The Neuron Restrictive Silencer Factor: A coordinate repressor of neuronal genes

Thesis by

Christopher J. Schoenherr

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1996

(Submitted October 3, 1995)

To my Mom and Dad and Uncle Mike

Acknowledgments

I would like to thank my advisor David Anderson for his constant support and enthusiasm. He gave me an opportunity at a critical time and for this I will always be grateful. I would also like to thank the members of my committee, Paul Patterson, Paul Sternberg, and Barbara Wold, for their support and patience.

I have been at Caltech for a long time, and I have met many people during my years here that have made this place seem like home.

In my first years here, I will never forget the good people I met in a bad situation. I have to thank Greg Wiederrecht for being the most funny and outrageous but still true to his roots Orange County Republican. I also thank him for all the biochemistry he taught me. It served me well. For my survival at Caltech, I owe much to Keith Harshman. He showed me the best burrito places in Pasadena. And the best french fries. Scot Moye-Rowley provided many an enthusiastic scientific discussion and many 'relaxing' parties. Special thanks go to Joanne Topol for making me see the world in a different light.

The second part of my tenure (actually a ninure) here involved a completely different group of people, all of whom I will miss. I thank Derek Stemple and Nirao Shah for putting up with my messy bench. They deserve a medal. I also enjoyed our many discussions of cell biology and foooood. I want to thank Andy Groves for his help with in situ hybridization and double entendres. Li-ching Lo was an invaluable source for antibody information. I tried to listen to her! Lukas Sommer deserves a special thank you for his patience in the last few weeks. But hurry up and give me those cells!!! A special I'm-tired-of-doing-those-type-of-experiments thank you goes to Alice Paquette. She made the last six months here infinitely more enjoyable. Kathy Zimmerman, Susan Birren, and Jane Johnson were always willing to listen to me ramble on. I owe many other people in the lab a thank you for discussions and support. Don't worry Hai soon our put orders will be too big.

Outside of the lab, I want to thank John Bowman for many good times and for introducing me to the great outdoors, Ultimate, and the Road Kill. I can't forget Zaven Kaprelian, who provided me with hours of entertaining conversation and almost as much advice. Special thanks go to Gregg Jongeward and Jonathan Bradley. Their knack for exquisite complaining always made me feel better about my life. My final thank you is for Katharine Liu. She has been a steady friend who has provided far more support than I can imagine.

Abstract

The transcriptional regulation of neuronal genes requires the combination of positive and negative control mechanisms. As a model neuronal gene, we have studied the neuron-specific gene, SCG10. The expression of SCG10 appears to be restricted to neurons by selective repression in non-neuronal cells. The upstream regulatory region of SCG10 contains a short sequence element that can repress, or silence, the activity of promoter fusion constructs in all non-neuronal cells assayed. In neuronal cells, this element has very little silencing activity. This neuron-restrictive silencer element (NRSE) was localized to about 21bp by deletional analysis. We have identified an NRSE binding protein that is present only in non-neuronal cell lines, but is absent from neuronal cell lines. This protein, the neuron-restrictive silencer factor (NRSF), is likely to mediate the repression activity of the NRSE as a double point mutation in the element that eliminates NRSF binding also eliminates silencing. Intriguingly, a similar element was identified in the type II sodium channel gene and shown to bind NRSF. Taken with its wide spread activity, this suggests that NRSF may be a coordinate regulator of neuronal gene expression in non-neuronal cells.

To determine the role of NRSF in neuronal gene regulation, we have isolated cDNA clones encoding a portion of human NRSF and the complete mouse homologue. NRSF is a novel protein with nine zinc fingers and several distinctive domains. Using in situ hybridization, expression of NRSF mRNA was detected in most non-neuronal tissues at several developmental stages, supporting the hypothesis that it functions as a near-global, sequence-specific repressor of neuronal gene expression. In the nervous system, NRSF mRNA was detected in neuronal progenitors, but not in postmitotic neurons. Its presence in precursor cells suggests that relief from NRSF-imposed repression may be an important event in the selection or execution of a neuronal differentiation program.

Further support for NRSF's role in neuronal gene regulation and development was provided by identification of potential NRSF target genes. Endogenous and recombinant

NRSF represses the activity of NRSE-containing reporter constructs and binds to consensus NRSEs in 14 other neuron-specific genes in addition to SCG10 and the type II sodium channel. At least seven additional neuronal genes were found to have sequences with significant similarity to the NRSE which are likely to represent functional binding sites for NRSF. These results suggest that one protein can coordinately repress many neuronal genes. Included amongst these genes are transcription factors that are implicated in the activation of neuronal differentiation, providing further evidence that NRSF may repress this process. Potential NRSEs also are found in non-neuronal genes which indicates that NRSF may have a function beyond the regulation of neuronal genes.

Table of contents

Chapter 1

Transcriptional regulation of eukaryotic genes

With few exceptions, all cells in a multicellular organism contain the same DNA. If the genetic material in each cell is the same, how then are the myriad of cell types in such organisms established and maintained? To answer this question, what distinguishes cell types must be determined. From developmental and genetic studies, we know that all cell types are products of their unique lineal histories and present environmental signals. And, from molecular and biochemical studies, it is clear that cell types can be characterized by the different proteins they express. It is thought, then, that a cell type is determined by the combined effects that lineage and environment have on the differential expression (and activity) of proteins. Thus, although it is an over-simplification, the question of cell type establishment and maintenance can be considered one of how differential protein expression is established and maintained.

Since we believe that the DNA content between cell types is the same, the regulation of differential protein expression must concentrate on RNA or the proteins themselves. While the nature of this regulation can be greatly influenced for extracellular signals, ultimately the control of protein expression must be performed by cell intrinsic factors. For RNA, these regulatory factors could focus on any of the steps required to convert genetic information into proteins, such as transcription, splicing, and translation. Proteins, on the other hand, could be subject to regulation by such methods as covalent modification, sequestration, and degradation. Much work in the field of molecular biology has focused on determining which of these processes are regulated and what relative role they play in defining different cell types.

This work has led to the conclusion that all of these processes are regulated to different degrees for each gene or protein. In fact, conventional wisdom suggests that if a process, or even a mechanistic step in that process, exists, then it will be regulated. This, however, should not imply that for a given gene each step is significantly regulated nor that each step is equally advantageous to regulate in all circumstances. In fact, overwhelming evidence indicates that for establishing and maintaining differential protein expression and,

thus, different cell types, transcription is the most pervasively regulated of all the potential target processes.

In this review, I want to discuss major concepts that are common to the regulation of transcriptional initiation of eukaryotic genes. Transcriptional regulation requires a highly complex orchestration of many interactive proteins that assemble into complicated structures. Understanding this multistep process requires a full integration of all the DNA sequences and proteins involved, from those that form chromatin, to the multitude of enhancer elements and factors that aid the general transcription factors required for all RNA polymerase II transcription. As an introduction, I will describe the major components of transcriptional regulation and give an overview of their properties. Then I will discuss mechanisms used by sequence specific factors to drive the initiation of transcription and how their arrangement in enhancers is critical to that activity. As they are of equal importance to activation mechanisms, methods of repressing transcription will also be detailed. Finally, mechanisms that address issues that may be distinct from classical enhancer models of transcription will be discussed.

THE COMPONENTS OF TRANSCRIPTIONAL REGULATION

The regulation of transcriptional initiation is accomplished predominantly by a combination of specific DNA sequences and the proteins that interact with them. Both the DNA sequences and the proteins can be broken down into two categories. The first category of each is required by almost all genes transcribed by RNA polymerase II. The second can differ from gene to gene.

The promoter

The focus of much of this regulation occurs at DNA sequences known as the promoter. For the purposes of this review a gene's promoter will be defined as sequences required for the assembly of the minimal complex necessary for RNA polymerase II to bind

and initiate transcription. Frequently, however, the term 'promoter' is used in a general manner to indicate sequences that are important for a basal level of transcription. Following the first definition, most promoters contain two distinct regions, a well conserved sequence (TATAA) known as the TATA box present at -30 and an Initiation sequence (CA) present at +1. The TATA box represents the high affinity binding site for the TATA binding protein (TBP) which is part of a large complex of proteins known as TFIID. These TBP-associated factors (TAFs) are essential for transcriptional activation by enhancer proteins. The binding of TFIID, followed by other protein complexes and RNA polymerase, is required before transcription can begin. Some genes, however, do not have an obvious consensus TATA box but appear to use the Initiation sequence (Inr) to recruit TFIID and thus the other components of the basic transcriptional machinery (Zawel and Reinberg, 1995). In fact, it has been proposed that the Inr represents the major nucleation site for TFIID, as there are many TATA-less promoters, and mutations in the Inr can abolish transcription whereas TATA box mutations only decrease transcription (Carcamo et al., 1991). In support of this idea, a recombinant TAF can recognize DNA containing an Inr and possibly can serve as the anchor for the TFIID complex (Verrijzer et al., 1994). An alternative possibility or a third route for nucleating TFIID involves Inr-binding proteins, such as the multifunctional YY1 (Roy et al., 1993; Seto et al., 1991).

Enhancers and repressors

A second set of DNA sequences are generally known as enhancers and repressors. Other names include UAS (for upstream activating sequences) and silencer elements. (For review see (Johnson, 1995; Mitchell and Tjian, 1989; Tjian and Maniatis, 1994)) These sequences are defined by their ability to increase (enhancers) or suppress (repressors) the rate of transcriptional initiation of a promoter on the same DNA molecule, i.e., in *cis*. They represent high affinity binding sites for the wide variety of sequence-specific DNA-binding proteins (transcription factors) which are responsible for their effect on transcription. Enhancers, as they were originally defined, are characterized by their ability to activate

transcription regardless of their orientation relative to the promoter and when located a considerable distance from a promoter (up to 50kb in some cases). Silencers, in their original definition, are also relatively orientation and distance independent in their ability to repress transcription (Brand et al., 1985). Both of these sequence elements typically are present in regions upstream of transcriptional initiation sites, but often can be found downstream of coding sequences and in introns. Furthermore, enhancers are usually comprised of several binding sites for different transcription factors that may interact with each other to define the overall characteristics of the enhancer. Individual enhancer elements can have different activities than the enhancer. Thus enhancers can also have silencer elements within them. Most genes contain multiple enhancers or silencers arranged in a particular manner that directs the cooperative and antagonistic interactions of DNA binding proteins that largely determine a gene expression characteristics.

Enhancer and silencer binding proteins

The proteins which recognize enhancer and silencer sequences make up many gene families. One count of the number of distinct families registered at least 12 distinct DNA binding domains (He and Rosenfeld, 1991). This is almost certainly an underestimation as several unique transcription factors are likely to have unidentified family members. Overall these DNA-binding proteins share certain characteristics important for transcriptional regulation. One theme common to almost all transcription factors is modular design. Almost all transcription factors have at least two domains: one responsible for DNA binding and another for modulating transcription. Most DNA binding domains contain alpha helices that interact with the major groove of DNA. A significant exception is the TATA binding protein. It uses a beta sheet to recognize the minor groove (Kim et al., 1993). Several other proteins families also recognize the minor groove (Tjian and Maniatis, 1994).

Although often considered to function solely as a tether for modulation domains, there is evidence that DNA binding is a dynamic process important to the activity of many

transcription factors. For example, some DNA binding domains induce bends in DNA that appear to be important for establishing contacts with other transcription factors (see below) (Natesan and Gilman, 1993; Tjian and Maniatis, 1994). DNA binding can also alter the effect a factor has on transcription. Members of the ligand-dependent nuclear receptor family, such as thyroid, retinoic acid, and estrogen receptors, bind to DNA as dimers with each monomer recognizing a 'half-site.' These half-sites can be separated by three to five nucleotides and, depending on the spacing, a bound factor will either activate or repress transcription (Naar et al., 1991; Umesono et al., 1991). This suggests that the DNA binding site acts as an allosteric effector to change the conformation of the transcription factor. The molecular basis for this binding site dependency is unknown.

The second portion of most transcription factors, modulation domains, usually referred to as activation or repression domains, are not well characterized either structurally or by primary amino acid sequence (Mitchell and Tjian, 1989). There are, however, shared characteristics that are used to classify activation or repression domains. Many of the first activation domains characterized were rich in acidic residues and were thought to have undefined secondary structure, thus giving rise to the term 'acid blob' (Sigler, 1988). Recent work, however, suggests that acidic activation domains of the yeast activators GAL4 and GCN4 may form β sheets (Leuther et al., 1993; Van Hoy et al., 1993). Several other activation domains that are commonly found in activator proteins have been characterized as glutamine rich, proline rich, or serine and threonine rich. Interestingly, mutagenesis studies on different activation domains suggest that the predominant amino acids that characterize these domains may not be the most important residues for activation (Cress and Triezenberg, 1991; Gill et al., 1994; Leuther et al., 1993). Instead, particular hydrophobic residues appear to be most important, as might be expected for a protein interaction domain. Several repression domains have been characterized and, just as for activation domains, there does not appear to be much sequence similarity other than an enrichment for certain amino acid residues (Cowell, 1994).

The modularity of DNA binding and modulation domains allows separate modification of DNA binding and transcriptional activity by posttranscriptional mechanisms In accordance with this possibility, these different domains are often on separate exons which creates opportunities for creating multiple proteins from one gene via differential splicing (Foulkes and Sassone-Corsi, 1992). This serves to generate diversity with out increasing the number of genes required. Other genes have taken advantage of this property by using alternative promoters or alternative translational initiation sites (Descombes and Schibler, 1991; Molina et al., 1993).

An important property of transcription factors is the ability of many of them to serve as either activators or repressors (Johnson, 1995). It is not clear whether all transcription factors have this property, but given the mechanisms used to achieve bifunctionality, it seems likely that many will. One mechanism involves binding of a protein to the transcription factor that changes its activity. For example, p53 becomes a repressor when the adenovirus E1B oncoprotein binds to it. In this case, the repression domain is encoded by E1B (Yew et al., 1994). Unliganded thyroid receptor acts as a direct repressor (see below) but changes to an activator upon ligand binding (Baniahmad et al., 1992). Recent experiments suggest that the receptor binds a protein that imparts repressor activity, and one role of the ligand is to release this activity (Baniahmad et al., 1995).

Requiring oligomerization for DNA binding, although not universal, is another feature common to many transcription factor families. These interactions can be either hetero- or homotypic and usually involve forming dimers. Often, one protein can dimerize with several different, but usually closely related proteins. For example, members of the basic helix-loop-helix (bHLH) family of transcription factors bind to DNA as dimers (Murre et al., 1989). The HLH domain, mediates the oligomerization, and the basic region is required for DNA binding (Davis et al., 1990; Murre et al., 1989; Voronova and Baltimore, 1990). Many members of this family do not appear to form functional homodimers, but must heterodimerize with a ubiquitously expressed bHLH protein

(Cabrera and Alonso, 1991; Lassar et al., 1991). A variation on this theme is seen in the basic leucine zipper family, in which members of the Jun family can form homodimers as well as heterodimers with several members of the Fos family (Hurst, 1994). This promiscuity can create oligomers with subtly or significantly different DNA binding or transcriptional activities and, thus, generates further levels of complexity and regulation. Furthermore, inhibitory subunits exist which when dimerized with a related factor prevent DNA binding (Benezra et al., 1990; Van Doren et al., 1991). The proposed utility of the multitude of interactions possible amongst related genes is to generate many different regulators with only a few genes (He and Rosenfeld, 1991).

Chromatin

In this section, I review some of the basic aspects of chromatin structure. There is increasing evidence that chromatin is a significant factor in the regulation of gene expression. Beyond the evidence that packaging of DNA into chromatin can repress transcription (Felsenfeld, 1992; Paranjape et al., 1994) in a non-specific manner, some studies suggest that it may be the agent of at least some sequence-specific repression mechanisms. On the other hand, recent work has also established a link between activation of transcription and chromatin. It is possible that in some cases activaton depends on the fact that DNA in eukaryotes in packaged into chromatin. Thus, it seems that to fully understand gene regulation we must understand the structure of chromatin.

In nuclei of eukaryotes, DNA is packaged with proteins into chromatin (for review see(Paranjape et al., 1994). This packaging provides for condensing the approximately two meters of DNA into a typical 5μm diameter eukaryotic nucleus. This packaging is characterized by at least two major levels of organization. The lowest level is the nucleosome, a small section of DNA (about 200bp) wrapped around a histone octamer core with the linker histone H1 usually present. Nucleosomes typically are present every 200bp and overall (linearly) compact the DNA roughly 7-fold. This 10nm fiber is further

compacted (again roughly 7-fold) into a structure known as the 30nm fiber. The structure of this fiber is unknown but clearly involves further coiling of the 10nm fiber. Most of the cell's DNA during interphase is present in the 30nm fiber. There is considerable evidence that organizing DNA into chromatin can have a negative effect on transcription in general and may represent a mechanism to maintain the normally inactive state of most genes in a cell. There is evidence, however, that the 30nm fiber is a dynamic structure that probably unfolds during transcription. Direct evidence for unfolding comes from electron microscopy of actively transcribed polytene chromosomes, and unfolding is inferred from the indirect evidence of increased sensitivity to DNaseI of transcriptionally active regions. DNase I sensitivity is commonly used as a measure of general chromatin structure with low sensitivity implying inactive chromatin and moderate sensitivity implying 'open' or actively transcribing chromatin. Regions of chromatin are also found that are hypersensitive to digestion. These sites are thought to represent the locations of regulatory regions.

While it is widely believed that it is the combination and arrangement of enhancers and silencers that determine a gene expression pattern, other mechanisms that appear to be distinct from these elements also contribute significantly to controlling transcriptional regulation. These mechanisms involve DNA domains and associated binding proteins that behave differently than enhancers and silencers do in standard transcriptional assays. Some of these domains, such as locus control regions, address the issue of independent domains of transcriptional regulation that may be necessary to prevent the repression of chromatin or to prevent the regulatory apparatus of adjacent genes from interacting (see below). There is, however, some controversy concerning the nature of these mechanisms. It is not clear yet if they are performing a function distinct from classically defined enhancers. On the other hand, these elements may alter the definition of enhancers to include their activities. In other cases, DNA domains known as matrix attachment regions or insulators are postulated to form boundaries that prevent cross-talk between regulated regions. Their actual function and their mechanism of action, however, are still

undetermined (see below). These areas are being actively explored and the understanding of their contribution to transcriptional regulation should increase.

MECHANISMS OF TRANSCRIPTIONAL ACTIVATION

To understand how transcriptional activators and repressors perform their duties, it is clear we need to understand the mechanism of transcriptional initiation. RNA polymerase II cannot recognize promoters on its own, but requires the stepwise assembly of several multiprotein complexes, known as general transcription factors, into an even larger complex before it can accurately initiate transcription (for review see, (Zawel and Reinberg, 1995). The first step in this assembly involves the binding of TFIID at the TATA box. TFIID is a general transcription factor that consists of the TATA binding protein (TBP) and several other TBP associated factors (TAFs). After TFIID has bound, a protein, TFIIB, binds to the complex. Then, TFIIF in association with RNA polymerase joins the complex. This is followed by TFIIE, TFIIH and TFIIJ to form the complete initiation complex. Although there is evidence for alternative assembly pathways for the initiation complex, the components are thought to be the same (Thompson et al., 1993). For the purposes of this review, however, it is important to consider that assembly of a complex takes place and transcription factors could alter the rate or the outcome of that assembly.

Assembly of this complex alone, however, is not sufficient to drive initiation. Some of these components have enzymatic activity essential for initiation. For example, the largest subunit of TFIIH has a helicase activity that is postulated to unwind the DNA helix over the start site to allow the polymerase to begin transcription. TFIIH has another subunit that phosphorylates heptapeptide repeats which characterize the C-terminal domain of RNA polymerase II (Lu et al., 1992). This phosphorylation appears to be crucial for the

disassociation of polymerase from the initiation complex to continue elongation of the RNA (Zawel and Reinberg, 1995).

In principle, anyone of the steps described above could be a target for activators (or repressors). In practice, however, evidence points to two main ideas for how enhancer proteins work. While it is widely considered that activators work by increasing the rate of initiation complex assembly, their mechanism of action is still not clear. One school of thought supports the idea of direct interactions between modulation domains and general transcription factors that potentiate their assembly into a functional complex. The other school of thought suggests that the main role of transcription factors is to relieve chromatin mediated repression of initiation complex assembly. These two models (discussed below) are not mutually exclusive, and it seems likely that at least some transcription factors will have both types of activity (Paranjape et al., 1994).

Activator-initiation factor interactions

A number of interactions between activator proteins and different general transcription factors have been detected (reviewed in (Tjian and Maniatis, 1994; Zawel and Reinberg, 1995). Direct physical interactions have been demonstrated between activators with acidic activation domains and TBP , TFIIB, and TFIIH (Ingles et al., 1991; Lin et al., 1991; Xiao et al., 1994). In some cases, there is a correlation between mutations that inactivate these domains and an inability to interact with these general factors. Conversely, mutations in TFIIB that are unresponsive to activators in in vitro transcription assays are also unable to bind acidic activators (Roberts et al., 1993). In some cases, the same activator has been shown to interact with several initiation factors, suggesting multiple separate targets or combined interactions (Lin et al., 1991).

An important question is whether there is any cell type specificity to the general factors associated with the initiation complex. The possibility for cell type specificity to basic mechanisms of transcriptional activation is implied by the 'co-activator' hypothesis. It has been suggested that transcription factors may not necessarily directly interact with

basal transcriptional machinery but may work through adapter or bridging molecules that may be cell type specific. While cell type specific co-activators have not been discovered, several activators can interact with particular subunits of TFIID (Gill et al., 1994; Goodrich et al., 1993). And as, some TFIID subunits are substochiometric, it is possible that not all subunits are essential for general function but could be specific to certain activators (Tjian and Maniatis, 1994). Circumstantial evidence for specific coactivators is provided by experiments showing that in certain tumor lines MyoD cannot activate transcription of reporter constructs or endogenous genes (Tapscott et al., 1993).

More direct evidence for coactivators comes from a recent discovery of a protein that binds to the transcription factor CREB (cAMP response element-binding protein). In response to increased cAMP levels in a cell, CREB is phosphorylated, thereby increasing its ability to activate transcription (Yamamoto et al., 1988). The CREB binding protein (CBP) was found to interact with the phosphorylated form of CREB (Chrivia et al., 1993). Moreover, CBP was also found to interact with TFIIB and activate transcription (Kwok et al., 1994). Taken together, these results suggest that CBP binds phosphorylated CREB and is responsible for its increased activation capability. Thus, CBP appears to serve as a coactivator for CREB. Whether CBP is specific for CREB is unknown. Interestingly, it is related to another protein , p300, which is implicated in the regulation of many genes (Arany et al., 1994).

Activators and chromatin

The other school of thought suggests that a significant function of enhancers is to relieve repression enforced by chromatin structure. In support of this 'antirepression' mechanism, several groups have found that activators frequently have only a small effect on the basal transcription rates of in vitro reactions using naked DNA templates (Paranjape et al., 1994). If, however, the templates are packaged into chromatin, they see a marked effect of these factors. This suggests that activator proteins work by displacing an inhibitory effect of nucleosomal condensation. In vitro evidence showing that TFIID

cannot bind chromatin supports this notion (Adams and Workman, 1993). Alternatively, it is possible that they require nucleosomes for full activity, as in vitro activation has been difficult to achieve from distances greater than 1kb without using a nucleosomal template (Laybourn and Kadonaga, 1992).

The above examples suggest that chromatin is a dynamic structure that can respond to transcription factor binding and, in contrast to previous thought, even provide an optimal substrate for some factors. Recently, different avenues of research have provided evidence for active remodeling of chromatin that requires transcription factors and a large, multisubunit complex. The first avenue came from genetic studies in yeast of the control of diverse regulatory networks, such as mating type switching and catabolite repression. These studies identified a group of genes, known as *SWI/SNF*, that are involved in the transcriptional activation of many genes that were not thought to be regulated by a common mechanism (Peterson and Herskowitz, 1992). A Drosophila homologue of the*SWI2* gene, *brahma*, also regulates many different genes, most notably the homeotics (for review see (Tamkun, 1995). The *SWI/SNF* genes are not essential for basal transcription but instead are required for efficient transcriptional stimulation by a wide variety of transcriptional activators. A possible mechanism for assisting activator proteins was postulated to involve counteracting the repressive effects of chromatin on transcription, as suggested by genetic interactions between *SWI/SNF* genes and chromatin proteins such as histones (Winston and Carlson, 1992).

This possibility was examined in vitro using a purified 2MDa complex containing products of the *SWI/SNF* genes or related complexes isolated from mammalian cells (Tamkun, 1995). As chromatin can prevent DNA binding of transcription factors, this complex was examined for its ability to stimulate nucleosomal binding by derivatives of the GAL4 transcription factor or TBP. The SWI/SNF complex was shown to stimulate binding of either protein to nucleosomal DNA by at least 10-fold but has no effect on binding to naked DNA. Furthermore, nucleosomal structure is altered by the complex even

in the absence of transcription factors. This reaction requires ATP hydrolysis catalyzed by the SWI2 protein and most likely involves a partial unwinding of the nucleosome (Cote et al., 1994; Imabalzano et al., 1994). Thus, it appears that this complex may enhance transcriptional activity by providing activators or TBP access to chromatin. Taken with the genetic evidence, this implies that disruption of chromatin structure is essential for at least some cases of transcriptional activation. These experiments, however, do not address whether the main function of transcriptional activators is to alter chromatin structure by targeting the SWI/SNF complex to disrupt specific chromatin domains or to enhance assembly of initiation complexes once the SWI/SNF complex has provided them access to DNA. Furthermore, it is important to note that not all genes in yeast are affected by the *SWI/SNF* genes (Winston and Carlson, 1992). This suggests that not all genes use this mechanism to relieve chromatin repression. Thus, *SWI/SNF* system may represent another level of differential gene regulation that recognizes a property common to a diverse set of genes.

CONCEPTS IN ENHANCER FUNCTION

Enhancer modules

As discussed above, transcriptional regulatory regions of some genes can be divided into discrete modules that confer a spatial or cell type specific expression. One example is the Drosophila homeodomain gene *fushi tarazu*. This gene is normally expressed in seven stripes of cells (segments) during Drosophila embryogenesis and in the nervous system later in development. Using fusion constructs in transgenic flies, it was shown that the striped pattern of expression and the neuronal expression could be conferred by two distinct, small sections of the *ftz* 5' flanking sequences (Hiromi and Gehring, 1987; Hiromi et al., 1985). These modules were separable and could work independently of each other. A similar but even more complex situation is seen in the regulation of another

Drosophila homeodomain gene *even skipped.* Like *ftz*, *eve* is also expressed in seven stripes in the early Drosophila embryo. Unlike *ftz*, however, whose stripe pattern is determined by one independent enhancer module, each stripe of eve expression is driven by an enhancer module specific to one stripe (Jiang and Levine, 1993). Thus, artificial constructs containing different combinations of enhancers can recreate a subset of the seven stripe pattern.

This modularity is not confined to Drosophila genes but is also seen in vertebrates genes. Modular enhancers have been found in mouse Hox genes (Puschel et al., 1991; Whiting et al., 1991). These enhancers drive expression in different tissues, and they have positional specificity. For an example of a non-regulatory gene, the intermediate filament protein, nestin, has independent enhancers for expression in muscle precursors and in neural progenitor cells, one in its first intron and the other in the second intron (Zimmerman et al., 1994). There are also well-described tissue specific enhancers that drive expression in pancreas, lymphocytes, and liver (Kruse et al., 1995; Liu et al., 1988; Staudt and Lenardo, 1991). Such enhancers often exist as functional entities as opposed to being made of elements scattered around the transcription unit.

The modularity also suggests something about the evolution of transcriptional regulation. It seems likely that transcriptional regulation could evolve in a manner similar to proteins, with modular enhancers behaving as exons. This analogy suggests many possible modes of enhancer evolution such as duplication and divergence within a gene or across genes. Duplication and translocation of an enhancer near a new promoter could now confer additional regulation upon the gene, often times without disrupting previous regulation. While this may reflect a simplistic view of transcriptional regulation, it seems likely that the modularity of enhancers would allow such evolution.

Cooperative interactions

The c-*fos* gene has been widely used as a model for the molecular mechanisms involved in transcriptional induction in response to extracellular signals. The c-*fos* gene is

rapidly and transiently induced by many different extracellular signals, such as calcium influx and growth factor signaling (for review see (Morgan et al., 1991). Extensive characterization of its proximal regulatory sequences has defined binding sites for four regulatory activities necessary for induction. As assayed by transient transfections, each of these elements responds almost exclusively to separate signaling pathways. In fact, this analysis lead to a model in which individual elements are thought to act independently to activate transcription in response to different signals (Gilman, 1988; Sheng et al., 1988). Recent evidence, however, has suggested that the transcription factors that bind to these elements cannot act independently but require the presence of all four binding proteins.

Using transgenic mice, Robertson et al. showed that c-*fos* promoter fusion genes containing the four binding sites were properly regulated in the brain (Robertson et al., 1995). Constructs with point mutations in any one of the four elements defined in transfection studies, however, did not respond properly to any inducing signals. This result suggests that the factors bound to all four sites are required for c-*fos* to respond, regardless of the type of signal. These results contradict some findings from transient transfection experiments and suggest a more cooperative model of transcriptional regulation. The results are supported by *in vivo* footprinting of the c-*fos* promoter region which shows occupancy of all four binding sites during induction (Herrera et al., 1989).

One model for transcriptional activation, described above, has individual transcription factors making contact with the initiation complex and facilitating its assembly. In this model, the cooperative enhancement seen with multiple factors is due to multiple separate contacts with members of the initiation complex. The authors suggest that this model cannot explain the extent of cooperativity they see for c-*fos* transcription. The authors propose that the concerted interaction of the four transcription factors forms a nucleation site for an 'interdependent transcription complex' (ITC). The ITC may contain adaptor proteins and initiation factors that bind cooperatively to form a functional unit capable of activating transcription. It is not known, however, which step the cooperativity

in c-*fos* is acting on. It is possible that cooperative DNA binding is responsible for the results seen, which is consistent with the first model.

One gene that clearly illustrates the need for cooperative DNA binding is Interleukin-2 (IL-2). Detailed analysis of the inducible IL-2 promoter illustrates several concepts that appear to hold true for many genes, whether they are induced transiently or developmentally regulated. A 300bp region of the IL-2 upstream sequences has been the subject of intensive investigations and have identified at least six binding sites necessary for proper IL-2 induction (for review see (Jain et al., 1995). Unlike c-*fos*, some but not all of these sites were found to completely eliminate activation when mutated. It should be pointed out, however, that most of these conclusions are based on data obtained from transient transfections, which may give misleading results as suggested for c-*fos*. Furthermore, sequences just distal to the first 300bp have regulatory activity, suggesting additional complexity (Novak et al., 1990). This may be the case as only 1 of 17 transgenic mice made with this region of the IL-2 promoter express properly (Brombacher et al., 1994).

IL-2 is a model of hierarchical cooperative interactions. At the first level, two of the six sites are actually composite elements that require the cooperative binding of different transactivators. One of these sites is composed of adjacent sites for a lymphoid specific factor known as NFAT (for nuclear factor of activated T-cells) and the Fos/Jun complex, AP-1 (Jain et al., 1993). When assayed individually, only NFAT can bind this binary site in vitro. The AP-1 site is very different from a consensus site and does not bind AP-1 alone. When assayed together however, an NFAT-AP1-DNA complex is formed indicating an interaction between the two transcription factors. This cooperative binding appears to require AP-1 to interact with DNA as mutations in the AP-1-like site abolish ternary complex formation. A similar situation is seen in another composite element that directs cooperative binding of Oct and AP-1 factors (Ullman et al., 1993). These composite element illustrate how different signaling pathways, such as the protein kinase C

and the calcium-dependent pathways, converge to induce IL-2 transcription. This interaction of AP-1 and NFAT and Oct proteins is postulated to explain the complex signaling required to activate IL-2 expression.

The second level of cooperative interactions takes place amongst the six major binding sites. For instance, if either of the composite sites described above are mutated such that AP-1 cannot bind, then IL-2 induction is eliminated. Similarly, induction is also eliminated if an apparently solo AP-1 site is mutated. This cooperativity was assayed directly by in vivo footprinting the IL-2 enhancer under different stimulatory conditions (Garrity et al., 1994). In unstimulated T-cells, the enhancer is unoccupied, even though some of the factors are present in the nucleus. Upon stimulation, coordinate occupation of all sites is observed. Further evidence of cooperativity was shown by the complete loss of a footprint caused by inhibitors of IL-2 induction even though they interfere with only one signaling pathway. Thus, just as in c-*fos*, cooperative interactions can decide the fate of the entire complex. In contrast to c-fos, the complex is not preformed but must assemble after stimulation.

Other studies have shown that a complex with a 'stereospecific' structure is required for activation. The mouse T-cell receptor α (TCRα) gene is driven by a minimal enhancer that can direct T-cell specific transcription. This enhancer contains binding sites for at least three different transcription factors, all of which are required for full enhancer function . Two of these factors are lymphocyte specific, but the other is expressed in many different cell types (Tjian and Maniatis, 1994). No one factor can activate TCRα transcription alone, but all three factors are required. Furthermore, not only is the binding of all three factors required, but the relative position of each binding site in the enhancer is essential for proper function. To explain this puzzle, the DNA binding and protein interaction properties of LEF-1 were examined.

LEF-1 is a lymphocyte specific HMG class protein known to induce a significant bend in DNA upon binding (Giese et al., 1992). Furthemore, LEF-1 appears to directly

interact with the other factors bound to the enhancer (Giese and Grosschedl, 1993). These results suggest that an LEF-1-induced bend, in the proper orientation, and protein interactions are required to form a tightly associated complex of proteins with DNA wrapped around them. Importantly, this indicates that a 'stereospecific complex' is required for activation. Previous to the characterization of the TCRα enhancer (and others), it was thought that the relative positions of individual binding sites was not essential for activation. The ability of artificial and rearranged promoters to activate transcription had suggested that position and even orientation of individual elements were largely irrelevant to enhancer function. It is not yet known how universal the need for a particular arrangement of enhancer elements is, although other examples are known (Natesan and Gilman, 1993; Thanos et al., 1993)

CONCEPTS IN TRANSCRIPTIONAL REPRESSION

Interactive repression

As with positive regulation, repression of transcription could also focus on the assembly of initiation complexes. However, since appreciable transcription requires specific activation, negative regulators can also be directed toward the activators themselves. One class of transcriptional repressors bind activator proteins and form complexes that are unable to bind DNA efficiently and, therefore, cannot activate transcription. Frequently this mechanism is seen between related transcription factor family members and is implicated in the regulation of numerous developmental decisions One system in which the genetic and molecular evidence is relatively complete is the development of the Drosophila peripheral nervous system.

In Drosophila, members of the bHLH family of transcription factors play an important role in the development of peripheral sensory organs (Jan and Jan, 1994). Specifically, loss-of-function mutations in bHLH genes *daughterless* (*da*) and members of

the achaete-scute complex (AS-C) result in the loss of particular sensory organs. On the other hand, loss-of-function mutations in the bHLH genes *extramacrochaete* and *hairy* (*emc* and *h*) result in the opposite phenotype, extra sensory organs (Van Doren et al., 1992). Thus, genetically *emc* and *h* are negative regulators of AS-C and *da*.

Isolation of the *emc* gene suggested a molecular model for this negative regulation. The emc protein contains an HLH dimerization domain but lacks the basic domain necessary for DNA binding. Thus, emc should be able to heterodimerize with other bHLH proteins, but such a complex should not be capable of binding DNA. This model was confirmed by showing that complexes of emc and AS-C members or da did not interact with a high affinity binding site (Van Doren et al., 1991). A homologous family of bHLH inhibitors, the Id genes, have been identified in vertebrates (Benezra et al., 1990). These factors inhibit DNA binding and transcriptional activation in the same manner as emc and also are implicated in regulating the differentiation of several cell lineages. Importantly, emc and Id can inhibit many different bHLH family members, most likely through interaction with the universal subunit, da or E12, respectively (Murre et al., 1989).

Other transcription factor gene families that bind DNA as dimers have negative regulators that work analogously to emc and Id. For example, in the basic -leucine zipper (bZIP) family, the C/EBP-homologous protein (CHOP) can inactivate the CAAT/enhancer binding protein (C/EBP) by dimerization and preventing DNA binding (Ron and Habener, 1992). In this case, the DNA-binding basic domain of CHOP is not deleted but apparently disrupted by the insertion of two proline residues. The function of CHOP is unknown, but it is postulated to antagonize some aspect of C/EBP-driven differentiation of adipocytes. Such repressors are also seen in the POU family of transcription factors (Ingraham et al., 1990). The I-POU protein has two lysine residues deleted from a highly conserved basic region and can form non-functional heterodimers with Cf1a, an activator of the dopa decarboxylase gene in Drosophila (Treacy et al., 1992). A different type of inhibitory partner is seen in the bZIP family. An alternatively spliced form of FosB that lacks an

activation domain, but not the DNA binding domain, forms heterodimers with Jun-related proteins that can bind DNA but not transactivate, presumably due to the absence of an activation domain (Wisdom et al., 1992).

While it is clear that the potential for forming nonfunctional complexes between closely related proteins is well exploited, unrelated transcriptional *activators* also can interact to form non-functional complexes and, thus, mutually inhibit each other. For example, both c-Jun and c-Fos can interact with the glucocorticoid receptor and form complexes that are unable to activate transcription (Diamond et al., 1990; Jonat et al., 1990; Yang-Yen et al., 1990). This interaction explains the many examples of glucocorticoid inhibition of c-Fos/c-Jun mediated inductions and, conversely, the inhibition of glucocorticoid mediated inductions by agents that increase levels of c-Fos/c-Jun. One physiologically relevant example of this mutual inhibition occurs with the collagenase gene. This gene is induced in fibroblasts of people with rheumatoid arthritis and is partially responsible for the destruction of tissues seen in this disease. Glucocorticoids have long been used as antiarthritic drugs and can lower collagenase levels through inhibition of an AP-1 site (see references in (Jonat et al., 1990).

Mutual interactive inhibition between unrelated factors may also play a role in myogenesis. Terminal differentiation and proliferation of myoblasts are apparently mutually exclusive processes. Terminal differentiation of myoblasts can be directed by increases in activity of members of myogenic family of bHLH genes, such as MyoD (Weintraub, 1993). Proliferation, on the other hand, can be maintained by the expression of transforming genes such as *jun* and *myc* (Miner and Wold, 1991; Su et al., 1991). These results suggested that a direct interaction between MyoD and nuclear proto-oncogenes would explain the exclusivity of differentiation and proliferation in myoblasts. In support of this idea, overexpression of MyoD and c-Jun was shown to mutually inhibit the activities of both proteins. Furthermore, these proteins could be cross-linked together and co-immunoprecipitated (Bengal et al., 1992). Whether this direct interaction is

necessary in development is unknown. As was the case for the positive integration of signaling pathways seen on the IL-2 gene, the regulatory interactions between Jun and both the glucocorticoid receptor and MyoD indicate that negative integration can also be important for determining the effects of an extracellular signal.

Competitive repression

Another mechanism of indirect repression is competition between activators and repressors for a DNA binding site. In a strict competition model, the repressor protein should only interfere with binding of an activator and not directly influence the initiation complex. Competition has been identified as a potential contributor to the spatial regulation of two segmentation genes in Drosophila. In the first example, a 730bp enhancer in the *Krüppel (Kr)* gene drives transcription in response to the homeodomain anterior determinant, *bicoid* (Hoch et al., 1992). This enhancer has six binding sites for bicoid as well as several binding sites for two genetically defined negative regulators of *Krüppel*, *knirps* and *tailless*. The knirps and tailless binding sites overlap with the bicoid sites, suggesting that these proteins may repress *Kr* expression by competing with bicoid for occupancy of the enhancer. Supporting this model, knirps can occlude DNA binding by bicoid to a DNA element with overlapping binding sites for the two proteins. Furthermore, knirps can inhibit bicoid-mediated activation of a reporter construct driven by this same DNA element.

Quenching

Another mode of repression involves interference with the activity of bound enhancer proteins. This mechanism, referred to as quenching, can be mediated by DNA bound repressors or soluble factors that interact with activators. In contrast to competition repression, quenching does not require overlap of binding sites but may require direct protein-protein interaction between activator and repressor. Although operationally similar to repression that acts on the initiation complex, repressors that function solely by activator

interference should have no effect on basal transcription. Of course, it is entirely possible that some repressors can work in multiple ways.

An excellent example of quenching occurs in an enhancer from the *rhomboid* gene that directs expression in the neuroectoderm of Drosophila (Gray et al., 1994). This restricted activity of the enhancer requires the zinc finger repressor, snail. In sna$^-$ embryos, this enhancer now drives expression in ventral mesoderm in addition to the neuroectoderm. Similarly, when all snail binding sites in the enhancer are mutated, expression also expands into the ventral mesoderm. Wild type activity can be restored to the mutant enhancer by placing synthetic snail binding sites 50 to 100bp away from the nearest activator elements. However, when these sites are placed 150bp away, repression is drastically reduced. This repression does not appear to be directed at the initiation complex as snail sites that are close to the TATA box but greater than 150bp away from the enhancer still do not repress. This mode of repression may be a general phenomenon as parallel experiments performed with *Kr* binding sites gave similar results. Thus, *snail* and *Kr* can only interfere with activators bound within a short distance. Further work should determine the nature of the protein-protein interactions that mediate this repression.

While sna and Kr must be bound to DNA to function, some repressors bind directly to activators and "mask" their activation domains without interacting with DNA. The best characterized example of masking is seen in the regulation of galactose catabolism in *S. cerevisiae* (reviewed in (Herschbach and Johnson, 1993) The genes necessary for galactose catabolism are upregulated by the activator protein GAL4 in the presence of galactose (Johnston, 1987). This protein is made constitutively but only activates transcription in the presence of galactose. In the absence of inducer, GAL4 is inactivated by the binding of the GAL80 protein. This does not occur by inhibiting DNA binding as GAL80 can interact with GAL4 when it is bound to DNA. This suggests that GAL80 makes the activation domain of GAL4 inaccessible for further protein-protein interactions. In support of this idea, GAL80 interacts with a subset of amino acids present in the

activation domain of GAL4. Inducer, however, does not cause GAL80 to disassociate but most act in a more subtle manner to relieve the repression (Leuther and Johnston, 1992).

The existence of competitive repression and quenching illustrates an important aspect of specific negative regulation. As many genes have enhancers with individual functions, there may be a need to repress them separately. Thus, repression by short range mechanisms allows for autonomous activity of different enhancers within an elaborately regulated gene. For instance, *eve* expression is regulated by separate enhancer modules for each stripe. The *eve* stripe 2 enhancer is repressed by *Kr* in the same cells that the stripe 3 enhancer is active(). Thus, short range repression of the stripe 2 enhancer is essential to proper *eve* expression. Presumably repressors such as *snail* or GAL80 could work similarly.

Direct repression

Repressors can also act directly on the initiation complex. In that sense, they are similar to activator proteins and may have similar target proteins, but with opposite effects. The hallmark of direct repression, in contrast to quenching, is the ability to repress unactivated (basal) transcription. This criterion makes direct repression difficult to distinguish from quenching and, perhaps, is best addressed in a purified in vitro transcription system. With this caveat, however, it is believed that many repressors identified will act directly on the initiation complex (Johnson, 1995). This appears to be true for even-skipped and the thyroid hormone receptor; the repression appears to act early in the assembly pathway, as complete initiation complexes are unaffected by the repressors (Fondell et al., 1993; Johnson and Krasnow, 1992). Further work on the many other examples of repressors will have to be done before their mode of action can be discerned.

CHROMATIN STRUCTURE AND ACTIVATION

Role of chromatin-related mechanisms

The possibility that transcription could be positively regulated by proteins other than classically defined enhancer factors was first suggested by studies of gene regulation in transgenic mice. Transgenes are inserted in an apparently random fashion into the mouse genome and, before their insertion, frequently are ligated in a head-to-tail fashion to form multicopy concatamers. It was found, however, that many promoter fusion transgenes exhibited variable levels of expression that did not correlate with number of copies present in the genome (Palmiter and Brinster, 1986). Furthermore, independent insertion events of a given transgene gave widely variable levels of expression. This phenomenon appeared to depend on the position a transgene inserted into the genome and is assumed to be a function of nearby DNA sequences that can deregulate transgene expression.

At first, the influence of position effects on transgene expression was considered detrimental to elucidating the factors required for proper transcriptional control; it is now seen as an excellent assay system for defining sequences that can impart position independent and copy number dependent control of a transgene. Elements that can impart such activity are known as locus control regions (LCR) (Orkin, 1995). Several regulatory regions have been identified that have LCR-type activity (Bonifer et al., 1990; Grosveld et al., 1987; Palmiter et al., 1993), and many other regions have been implicated in such regulation. Important work for the future involves determining the mechanism of LCRs and distinguishing them from enhancers.

The defining LCR was first identified in the upstream region of the human β-globin locus (Grosveld et al., 1987). The β-globin locus is a cluster of five different globin isoforms that comprises approximately 60kb of DNA. Previous to the discovery of the globin LCR, experiments using transient transfections and transgenic mice identified regions both 5' and 3' of the gene that were important for the proper developmental and tissue specific expression of the β-globin gene (for review see (Orkin, 1995). However, transgenes containing these regions gave highly variable levels of expression that was independent of copy number. Furthermore, the levels of expression were much lower than

the endogenous gene. This classic position effect suggested that significant portions of the globin regulatory sequences were missing from the transgenes examined.

Using a construct that contains the globin LCR, Grosveld et al. (1987) was able to obtain, for the first time in a transgenic mouse, a transgene that showed copy number dependent and position independent levels of transcription (Grosveld et al., 1987). Other LCR's have been identified in the chicken lysozyme gene and the metallothionein gene (Bonifer et al., 1990; Palmiter et al., 1993). Thus, the function of the LCR appears distinct from the classically defined enhancer. However, several enhancer factors have been shown to be necessary for LCR activity, and one portion of the LCR does have enhancer activity in transient transfection assays (Orkin, 1995).

The LCR is widely considered to create a active chromatin domain (Orkin, 1990). Initial evidence for this hypothesis came from measuring the DNase I sensitivity (see above) of the β-globin locus. These studies identified four hypersensitive regions located about 40kb upstream from the globin cluster. This group of hypersensitive sites marks the globin LCR, thus linking chromatin structure to copy number-dependent, position-independent gene regulation. The hypothesis that the LCR can function autonomously as a 'chromatin opener' was directly examined in transgenic mice. By comparing the DNase I sensitivity of a construct containing the chicken β-globin LCR, enhancer, and promoter to one containing only the LCR and enhancer, Reitman et al. (1993) showed that the LCR is unable to open chromatin by itself but requires cooperation with a promoter (Reitman et al., 1993). This experiment, however, does not rule out that the LCR is necessary for opening chromatin.

While the LCR may not open chromatin on its own, genetic evidence for the necessity of open chromatin is provided by the recent cloning of a gene involved in an α-thalassemia (Gibbons et al., 1995). (α-globins do not appear to have a classic LCR but do have hypersensitive regions that are crucial for expression.) This gene, designated XH2, encodes a helicase similar to the brahma and SWI2 proteins described above that are

involved in chromatin/activator protein interactions. Mutations in this gene selectively reduce α-globin expression. One interpretation of the specificity of this mutation for α-globin is, that in the absence of a strong LCR, additional machinery is required to establish an open chromatin domain. Future work should determine whether this model or the equally interesting possibility of gene specific helicases is correct.

Long range and long term repression

A form of transcriptional repression that may be analogous but opposite to LCR activity is mediated by group of genes identified as negative regulators of homeotic genes (for review see (Pirrotta, 1995). In Drosophila, the Polycomb group (Pc-G) of genes work in concert to provide a mechanism of negative regulation that controls the activity of large chromosomal regions. More specifically these genes maintain the repression of certain genes that was established initially by transient regulatory factors. In Pc-G mutants, the initial expression pattern of genes such as *Ultrabithorax* and *Antennapedia* is normal, but later in development these patterns expand ectopically. Their patterns of expression is established by positive interactions with segmentation genes such as *ftz* and *engrailed* and repression by gap genes such as *hunchback* (Qian et al., 1993). Soon after the pattern is established, the specific repressors are no longer expressed, and the gene is activated ectopically. Thus, it appears that Pc-G proteins function to maintain the pattern of repression set up by transient regulatory molecules and, thus, provide a mechanism for propagating regulatory states established early in development.

The exact molecular mechanism for propagating repression is unknown but involves the formation of a complex made up of several Pc-G proteins that interacts with specific regions of DNA in the regulatory regions of many genes. There are twelve Pc-G genes identified genetically, seven of which have been characterized molecularly. Mutations in some Pc-G mutations are haploinsufficient and combinations of mutations have synergistic effects, suggesting that they must form a stochiometric complex for function. The possibility that Pc-G proteins form a complex on DNA was tested using

antibody staining of salivary gland polytene chromosomes. These experiments revealed that each Pc-G protein can be found at between 80-100 chromosomal positions, some of which correspond to known Pc-G regulated genes. Importantly, the sites for several proteins overlap almost completely, while only partially for others. Furthermore, mutations in different Pc-G genes almost eliminates the chromosomal interactions of other members of the group.

An important step was taken recently with the discovery of a Pc-G response element (PRE) (Chan et al., 1994). The first PRE was identified as a portion of the *Ubx* regulatory domain that could maintain the repressed state of a reporter construct bearing an enhancer that directed proper parasegmental expression. This 1.5kb element has no enhancer activity on its own and is dependent on wild type Pc-G activity. It is located 24kb from the *Ubx* promoter and is postulated to interact with enhancers up to 70kb away. Direct repression of the promoter, however, has not been ruled out. Notably, specific binding sites for Pc-G proteins have not yet been identified in this element.

Several mechanisms have been postulated to explain the long range and long term regulation imposed by Pc-G proteins. One hypothesis, based on cross-linking experiments that show a Pc-G protein associated with many regions in repressed genes (Orlando and Paro, 1993), suggests that Pc-G complexes extend over long regions of DNA, thereby occluding any activator proteins.

Matrix attachment regions

Eukaryotic chromatin appears to be organized into domains with an average length of 50-100kb (Sippel et al., 1993). These domains are postulated to represent independently regulable regions of the genome. They are thought to be established by binding to a protein structure termed the nuclear matrix or scaffold. The nuclear matrix refers to the protein structure remaining after isolated nuclei are treated with nucleases and extracted with various agents to remove proteins. A small amount of DNA is tightly associated with the matrix and is known as matrix-associated regions (MAR) or scaffold-

attachment regions (SAR) (Sippel et al., 1993). These MARs were found to be short (250-3000bp) A+T rich genomic fragments. No obvious consensus sequence, however, could be identified. It is believed that these MARs represent attachment sites to the nuclear matrix that create separate chromatin domains. In support of this idea, some MARs have been mapped to the boundaries of identified active chromatin domains. Furthermore in some cases, transgenes containing MARs have a significantly lower frequency of position effects than without such elements. In some cases, MARs have enhancer activity, suggesting a link between the two functions. It has been difficult, however, to show that MAR activity is distinct from enhancer activity or if the insulator type activity is dependent on matrix binding. With the increasing use of PCR for genomic footprinting, this issue could be addressed more directly than in the past.

CONCLUSION

Transcriptional regulation requires a complex, step-wise assembly of many interactive proteins into complicated structures. Moreover, these structures are likely to have enzymatic activity that is essential for the initiation process. Thus, a complete understanding requires not only a full integration of all the sequences and proteins involved in the process, but their respective activities as well. It is also clear that the process of initiating transcription can be broken down into many steps that can be regulated in parallel and in sequence. Furthermore, both positive and negative regulation of transcription are important for establishing the final pattern of gene expression. Finally, what is abundantly clear is that sex is more fun than science.

REFERENCES

Adams, C. C., and Workman, J. L. (1993). Nucleosome displacement in transcription. Cell *72*, 305-308.

Arany, Z., Sellers, W. R., Livingston, D., and Eckner, R. (1994). E1A-associated p300 and CREB-associated CBP belong to a conserved family of coactivators. Cell *77*, 799-800.

Baniahmad, A., Kohne, A. C., and Renkawitz, R. (1992). A transferable silencing domain is present in the thyroid hormone receptor, in the v-erbA oncogene product and in the retinoic acid receptor. EMBO J. *11*, 1015-1023.

Baniahmad, A., Leng, X. H., Burris, T. P., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (1995). The tau-4 activation domain of the thyroid-hormone receptor is required for release of a putative corepressor(s) necessary for transcriptional silencing. Mol. Cell. Biol. *15*, 76-86.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., and Weintraub, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell ***61***, 49-59.

Bengal, E., Ransone, L., Scharfmann, R., Dwarki, V. J., Tapscott, S. J., Weintraub, H., and Verma, I. M. (1992). Functional antagonism between c-Jun and MyoD proteins- a direct physical association. Cell *68*, 507-519.

Bonifer, C., Vidal, M., and Sippel, A. E. (1990). Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. EMBO J. *9*, 2843-2848.

Brand, A. H., Breeden, L., Abraham, J., Sternglanz, R., and Nasmyth, K. (1985). Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. Cell *41*, 41-48.

Brombacher, F., Schafer, T., Weissenstein, U., Tschopp, C., Anderson, E., Burki, K., and Baumann, G. (1994). IL-2 promoter-driven *lacZ* expression as a monitoring tool for IL-2 expression in primary T cells of transgenic mice. Intl. Immun. *6*, 189-197.

Cabrera, C., and Alonso, M. (1991). Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products of *Drosophila*. EMBO J *10*, 2965-2973.

Carcamo, J., Buckbinder, L., and D., R. (1991). The Initiator directs the assembly of a transcription factor-IID-dependent transcription complex. Proc. Natl. Acad. Sci. *88*, 8052-56.

Chan, C. S., Rastelli, L., and Pirrotta, V. (1994). A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. EMBO J. *13*, 2553-2564.

Chrivia, J. C., Kwok, R. P. S., Lamb, N., Hagiwara, M., Montminy, M. R., and Goodman, R. H. (1993). Phosphorylated CREB binds specifically to the nuclear-protein CBP. Nature *365*, 855-859.

Cote, J., Quinn, J., Workman, J. L., and Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. Science *265*, 53-60.

Cowell, I. G. (1994). Repression versus activation in the control of gene transcription. Trends Biochem. Sci. *19*, 128-132.

Cress, W. D., and Triezenberg, S. J. (1991). Critical structural elements of the VP16 transcriptional activation domain. Science *251*, 87-90.

Davis, R. L., Cheng, P. F., Lassar, A. B., and Weintraub, H. (1990). The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. Cell *60*, 733-746.

Descombes, P., and Schibler, U. (1991). A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same messenger-RNA. Cell *67*, 569-579.

Diamond, M. I., Miner, J. N., Yoshinaga, S. K., and Yamamoto, K. R. (1990). Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. Science *249*, 1266-1272.

Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. Nature *355*, 219-224.

Fondell, J. D., Roy, A. L., and Roeder, R. G. (1993). Unliganded thyroid-hormone receptor inhibits formation of a functional preinitiation complex- implications for active repression. Genes & Dev. *7*, 1400-1410.

Foulkes, N. S., and Sassone-Corsi, P. (1992). More is better- activators and repressors from the same gene. Cell *68*, 411-414.

Garrity, P. A., Chen, D., Rothenberg, E. V., and Wold, B. J. (1994). Interleukin-2 transcription is regulated in-vivo at the level of coordinated binding of both constitutive and regulated factors. Mol. Cell. Biol. *14*, 2159-2169.

Gibbons, R. J., Pickettes, D. J., Villard, L., and Higgs, D. R. (1995). Mutations in a putative global transcriptional regulator cause X-linked mental retardation with α-thalassemia (ATR-X syndrome). Cell *80*, 837-845.

Giese, K., Cox, J., and Grosschedl, R. (1992). The HMG domain of lymphoid enhancer factor-I bends DNA and facilitates assembly of functional nucleoprotein structures. Cell *65*, 185-195.

Giese, K., and Grosschedl, R. (1993). LEF-1 contains an activation domain that stimulates transcription only in a specific context of factor-binding sites. EMBO J. *12*, 4667-4676.

Gill, G., Pascal, E., Tseng, T. H., and Tjian, R. (1994). A glutamine-rich hydrophobic patch in transcription factor-SP1 contacts the DTAF(II)110 component of the Drosophila TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. *91*, 192-196.

Gill, G., Pascal, E., Tseng, Z., and Tjian, R. (1994). A glutamine-rich hydrophobic patch in transcription factor-Sp1 contacts the DTAF(II)110 component of the Drosophila TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. *91*, 192-196.

Gilman, M. Z. (1988). The c-fos serum response element responds to protein kinase C-dependent and -independent signals but not to cyclic AMP. Genes & Dev. *2*, 392-402.

Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A., and Tjian, R. (1993). Drosophila TAF(II)40 interacts with both a VP16 activation domain and the basal transcription factor- TFIIB. Cell *75*, 519-530.

Gray, S., Szymanski, P., and Levine, M. (1994). Short-range repression permits multiple enhancers to function autonomously within a complex promoter. Genes & Dev. *8*, 1829-1838.

Grosveld, F., Blom van Assendelft, G., Greaves, D. R., and Kollias, G. (1987). Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell *51*, 975-985.

He, X., and Rosenfeld, M. G. (1991). Mechanisms of complex transcriptional regulation: implications for brain development. Neuron *7*, 183-196.

Herrera, R. E., Shaw, P. E., and Nordheim, A. (1989). Occupation of the c-fos serum response element in vivo by a multi-protein complex is unaltered by growth factor induction. Nature *340*, 68-70.

Herschbach, B. M., and Johnson, A. D. (1993). Transcriptional repression in eukaryotes. Annu. Rev. Cell Biol. *9*, 479-509.

Hiromi, Y., and Gehring, W. J. (1987). Regulation and function of the Drosophila segmentation gene *fushi tarazu*. Cell *50*, 963-974.

Hiromi, Y., Kuriowa, A., and Gehring, W. J. (1985). Control elements of the Drosophila segmentation gene *fushi tarazu*. Cell *43*, 603-613.

Hoch, M., Gerwin, N., Taubert, H., and Jackle, H. (1992). Competition for overlapping sites in the regulatory region of the *Drosophila* gene *Kruppel*. Science *256*, 94-97.

Hurst, H. C. (1994). Transcription factors: BZIP proteins- introduction. Protein Profile *1*, 123-168.

Imabalzano, A. N., Kwon, H., Green, M. R., and Kingston, R. E. (1994). Facilitated binding of TATA-binding protein to nucleosomal DNA. Nature *370*, 481-485.

Ingles, C. J., Shales, M., Cress, W. D., Triezenberg, S. J., and Greenblatt, J. (1991). Reduced binding of TFIID to transcriptionally compromised mutants of VP16. Nature *351*, 588-590.

Ingraham, H. A., Albert, V. R., Chen, R., Crenshaw, E. B. I., Elsholtz, H. P., He, X., Kapiloff, M. S., Mangalam, H. J., Swanson, L. W., Treacy, M. N., and Rosenfeld, M. G. (1990). A family of POU-domain and pit-1 tissue-specific transcription factors in pituitary and neuroendocrine development. Annu. Rev. Physiol. *52*, 773-791.

Jain, J., Loh, C., and Rao, A. (1995). Transcriptional regulation of the IL-2 gene. Curr. Op. Immun. *7*, 333-342.

Jain, J., McCaffrey, P. G., Miner, Z., Kerppola, T. K., Lambert, J. N., Verdine, G. L., Curran, T., and Rao, A. (1993). The T-cell transcription factor NFATp is a substrate for calcineurin and interacts with Fos and Jun. Nature *365*, 352-355.

Jan, Y. N., and Jan, L. Y. (1994). Genetic control of cell fate specification in Drosophila peripheral nervous system. Annual Review of Genetics *28*, 373-393.

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.

Johnson, A. D. (1995). The price of repression. Cell *81*, 655-658.

Johnson, F. B., and Krasnow, M. A. (1992). Differential regulation of transcription preinitiation complex assembly by activator and repressor homeo domain proteins. Genes & Dev. *6*, 2177-2189.

Johnston, M. (1987). A model fungal gene regulatory mechanism: the *GAL* genes of *S. cerevisiae*. Microbiol. Rev. *51*, 458-576.

Jonat, C., Rahmsdorf, H. J., Park, K.-K., Cato, A. C. B., Gebel, S., Ponta, H., and Herrlich, P. (1990). Antitumor promotion and antiinflammation: Down-modulation of AP-1 (Fos/Jun) activity by glucocorticoid hormone. Cell *62*, 1189-1204.

Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature *365*, 520-527.

Kruse, F., Rose, S. D., Swift, G. H., Hammer, R. E., and Macdonald, R. J. (1995). Cooperation between elements of an organ-specific transcriptional enhancer in animals. Mol. Cell. Biol. *15*, 4385-4394.

Kwok, R. P. S., Lundblad, J. R., Chrivia, J. C., Richards, J. P., Bachinger, H. P., Brennan, R. G., Roberts, S. G. E., Green, M. R., and Goodman, R. H. (1994). Nuclear-protein CBP is a coactivator for the transcription factor CREB. Nature *370*, 223-226.

Lassar, A. B., Davis, R. L., Wright, W. E., Kadesch, T., Murré, C., Voronova, A., Baltimore, D., and Weintraub, H. (1991). Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. Cell *66*, 305-315.

Laybourn, P. J., and Kadonaga, J. T. (1992). Threshold phenomena and long-distance activation of transcription by RNA polymerase-II. Science *257*, 1682-1685.

Leuther, K. K., and Johnston, S. A. (1992). Non-dissociation of GAL4 and GAL80 gene regulatory proteins in vivo after galactose induction. Science *256*, 1333-1335.

Leuther, K. K., Salmeron, J. M., and Johnston, S. A. (1993). Genetic evidence that an activation domain of GAL4 does not require acidity and may form a β sheet. Cell *72*, 575-585.

Lin, Y. S., Ha, I., Maldonado, E., Reinberg, D., and Green, M. R. (1991). Binding of general transcription factor TFIIB to an acidic activating region. Nature *353*, 569-571.

Liu, J. K., Bergman, Y., and Zaret, K. S. (1988). The mouse albumin promoter and a distal upstream site are simultaneously DNaseI hypersensitive in liver chromatin and bind similar liver-abundant factors *in vitro*. Genes & Dev. *2*, 528-541.

Lu, H., Zawel, L., Fisher, L., Elgy, J.-M., and Reinberg, D. (1992). Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature *358*, 641-645.

Miner, J. H., and Wold, B. J. (1991). c-myc inhibition of MyoD and myogenin-initiated myogenic differentiation. Mol. Cell. Biol. *11*, 2842-2851.

Mitchell, P., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371-378.

Molina, C. A., Foulkes, N. S., Lalli, E., and Sassone-corsi, P. (1993). Inducibility and negative autoregulation of CREM - an altenative promoter directs the expression of ICER, an early response repressor. Cell *75*, 875-886.

Morgan, L., Jessen, K. R., and Mirsky, R. (1991). The effects of cyclic AMP on differentiation of cultured Schwann cells: progression from an early phenotype (O4) to a myelin phenotype (Po^+, $GFAP^-$, $N\text{-}CAM^-$, $NGF\text{-}receptor^-$) depends on growth inhibition. J. Cell Biol. *112*, 457-467.

Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobin enhancer binding, *daughterless, MyoD* and *myc* proteins. Cell *56*, 777-783.

Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B., Weintraub, H., and Baltimore, D. (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell *58*, 537-544.

Naar, A., Boutin, J.-M., Lipkin, S. M., Yu, V. C., Holloway, J. M., Glass, C. K., and Rosenfeld, M. G. (1991). The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. Cell *65*, 1267-1279.

Natesan, S., and Gilman, M. Z. (1993). DNA bending and orientation-dependent function of YY1 in the c-fos promoter. Genes & Dev. *7*, 2497-2509.

Novak, T. J., White, P. M., and Rothenberg, E. V. (1990). Regulatory anatomy of the murine Interleukin-2 gene. Nucleic Acids Res. *18*, 4523-4533.

Orkin, S. H. (1990). Globin gene regulation and switching: Circa 1990. Cell *63*, 665-672.

Orkin, S. H. (1995). Regulation of globin gene expression in erythroid cells. Eur. J. Biochem. *231*, 271-281.

Orlando, V., and Paro, R. (1993). Mapping *Polycomb*-repressed domains in the bithorax complex using *in vivo* formaldehyde cross-linked chromatin. Cell *75*, 1187-1198.

Palmiter, R. D., and Brinster, R. L. (1986). Germ-line transformation of mice. Ann. Rev. Genet. *20*, 465-499.

Palmiter, R. D., Sandgren, E. P., Koeller, D. M., and Brinster, R. L. (1993). Distal regulatory elements from the mouse metallothionein locus stimulate gene expression in transgenic mice. Mol. Cell. Biol. *13*, 5266-5275.

Paranjape, S. M., Kamakaka, R. T., and Kadonaga, J. T. (1994). Role of chromatin structure in the regulation of transcription by RNA polymerase II. Annu. Rev. Biochem. *63*, 265-297.

Peterson, C. L., and Herskowitz, I. (1992). Characterization of the SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell *68*, 573-583.

Pirrotta, V. (1995). Chromatin complexes regulating gene expression in *Drosophila*. Curr. Op. Gen. Dev. *5*, 466-472.

Puschel, A., Balling, R., and Gruss, P. (1991). Separate elements cause lineage restriction and specify boundaries of Hox-1.1 expression. Development *112*, 279-287.

Qian, S., Capovilla, M., and Pirotta, V. (1993). Molecular mechanisms of pattern formation by the BRE enhancer of the *Ubx* gene. EMBO J. *12*, 3865-3877.

Reitman, M., Lee, E., Westphal, H., and Felsenfeld, G. (1993). An enhancer/locus control region is not sufficient to open chromatin. .

Roberts, S. G. E., Ha, I., Maldonado, E., Reinberg, D., and Green, M. R. (1993). Interaction between an acidic activator and transcription factor-TFIIB is required for transcriptional activation. Nature *363*, 741-744.

Robertson, L. M., Kerpola, T. K., Vendrell, M., Luk, D., Smeyne, R. J., Bocchiaro, C., Morgan, J. I., and Curran, T. (1995). Regulation of *c-fos* expression in transgenic mice requires multiple interdependent transcription control elements. Neuron *14*, 241-252.

Ron, D., and Habener, J. F. (1992). CHOP, a novel developmentally regulated nuclear-protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene-transcription. Genes & Dev. *6*, 439-453.

Roy, A. L., Malik, S., Meisterenst, M., and Roeder, R. G. (1993). An alternative pathway for transcription initiation involving TFII-I. Nature *365*, 355-359.

Seto, E., Shi, Y., and Shenk, T. (1991). YY1 is an Initiator sequence-binding protein that directs and activates transcription in vitro. Nature *354*, 241-244.

Sheng, M., Dougan, S. T., McFadden, G., and Greenberg, M. E. (1988). Calcium and growth factor pathways of c-fos transcriptional activation require distinct upstream regulatory sequences. Mol. Cell. Biol. *8*, 2787-2796.

Sigler, P. B. (1988). Acid blobs and negative noodles. Nature *333*, 210-212.

Sippel, A. E., Schäfer, G., Faust, N., Saueressig, H., Hecht, A., and Bonifer, C. (1993). Chromatin domains constitute regulatory units for the control of eukaryotic genes. Cold Spring Harbor Symp. Quant. Biol. *58*.

Staudt, L. M., and Lenardo, M. J. (1991). Immunoglobulin gene transcription. Annu. Rev. Immunol. *9*, 373-398.

Su, H. Y., Bos, T. J., Monteclaro, F. S., and Vogt, P. K. (1991). Jun inhibits myogenic differentation. Oncogene *6*, 1759-1766.

Tamkun, J. W. (1995). The role of brahma and related proteins in transcription and development. Curr. Opin. Genet. Dev. *5*, 473-477.

Tapscott, S. J., Thayer, M. J., and Weintraub, H. (1993). Deficiency in rhabdomyosarcomas of a factor required for MyoD activity and myogenesis. Science *259*, 1450-1453.

Thanos, D., Du, W., and Maniatis, T. (1993). The high-mobility group protein HMG I(Y) is an essential structural component of a virus-inducible enhancer complex. Cold Spr. Harb. Symp. *58*, 73-81.

Thompson, C. M., Koleske, A. J., Chao, D. M., and Young, R. A. (1993). A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. Cell *73*, 1361-1375.

Tjian, R., and Maniatis, T. (1994). Transcriptional activation: A complex puzzle with few easy pieces. Cell *77*, 5-8.

Treacy, M. N., Neilson, L. I., Turner, E. E., He, X., and Rosenfeld, M. G. (1992). Twin of I-POU Homeodomain Distinguishes an Activator from an Inhibitor of Transcription. Cell *68*, 491-505.

Ullman, K. S., Northrup, J. P., Admon, A., and Crabtree, G. R. (1993). Jun family members are controlled by a calcium-regulated, cyclosporin A-sensitive signaling pathways in activated T-lymphocytes. Genes & Dev. *7*, 188-196.

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 D receptors. Cell *65*, 1255-1266.

Van Doren, M., Ellis, H. M., and Posakony, J. W. (1991). The *Drosophila* extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. Development *113*, 245-255.

Van Doren, M., Powell, P. A., Pasternak, D., Singson, A., and Posakony, J. W. (1992). Spatial regulation of proneural gene activity: auto- and cross-activation of *achaete* is antagonized by *extramacrochaete*. Genes & Dev. *6*, 2592-2605.

Van Hoy, M., Leuther, K. K., Kodadek, T., and Johnston, S. A. (1993). The acidic activation domains of the GCN4 and GAL4 proteins are not α helical but form β sheets. .

Verrijzer, C. P., Yokomori, K., Chen, J.-L., and Tjian, R. (1994). Drosophila TAFII150: Similarity to yeast gene TSM-1 and specific binding to core promoter DNA. Science *264*, 933-941.

Voronova, A., and Baltimore, D. (1990). Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. Proc. Natl. Acad. Sci. USA *87*, 4722-4726.

Weintraub, H. (1993). The MyoD family and myogenesis: redundancy, networks and thresholds. Cell *75*, 1241-1244.

Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W. J., Stott, D., and Allemann, R. K. (1991). Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. Genes & Dev. *5*, 2048-2059.

Winston, F., and Carlson, M. (1992). Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. Trends Genet. *8*, 387-391.

Wisdom, R., Yen, J., Rashid, D., and Verma, I. M. (1992). Transformation by FosB requires a transactivation domain missing in FosB2 that can be substituted by heterologous activation domains. Genes & Dev. *6*, 667-675.

Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J. L., Triezenberg, S. J., Reinberg, D., Flores, O., Ingles, C. J., and Greenblatt, J. (1994). Binding of

basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. Mol. Cell. Biol. *14*, 7013-7024.

Yamamoto, K. K., Gonzalez, G. A., Biggs, W. H., and Montminy, M. R. (1988). Phosphorylation-induced binding and transcriptional efficacy of nuclear factor CREB. Nature *334*, 494-498.

Yang-Yen, H. F., Chambard, J.-C., Sun, Y.-L., Smeal, T., Schmidt, T. J., Drouin, J., and Karin, M. (1990). Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction. Cell *62*, 1205-1215.

Yew, P. R., Liu, X., and Berk, A. J. (1994). Adenovirus E1B oncoprotein tethers a transcriptional repression domain to p53. Genes & Dev. *8*, 190-202.

Zawel, L., and Reinberg, D. (1995). Common themes in assembly and function of eukaryotic transcription complexes. Annu. Rev. Biochem. *64*, 533-561.

Zimmerman, L., Lendahl, U., Cunningham, M., McKay, R., Parr, B., Gavin, B., Mann, J., Vassileva, G., and McMahon, A. (1994). Independent regulatory elements in the nestin gene direct transgene expression to neural stem cells or muscle precursors. Neuron *12*, 11-24.

Chapter 2

A common silencer element in the SCG10 and type II Na^+ channel genes binds a factor present in nonneuronal cell but not in neuronal cells

Nozomu Mori, Christopher Schoenherr, David J. Vandenbergh, and David J. Anderson

Neuron, Vol. 9, 45–54, July, 1992, Copyright © 1992 by Cell Press

A Common Silencer Element in the SCG10 and Type II Na^+ Channel Genes Binds a Factor Present in Nonneuronal Cells but Not in Neuronal Cells

Nozomu Mori,*† Christopher Schoenherr,*
David J. Vandenbergh,*‡ and David J. Anderson*§
*Division of Biology
§Howard Hughes Medical Institute
California Institute of Technology
Pasadena, California 91125

Summary

We have localized a cell type–specific silencer element in the SCG10 gene by deletion analysis. This neural-restrictive silencer element (NRSE) selectively represses SCG10 expression in nonneuronal cells and tissues. The NRSE contains a 21 bp region with striking homology to a sequence present in a silencer domain of the rat type II sodium channel (NaII), another neuron-specific gene. We have identified a sequence-specific protein(s) that binds the SCG10 NRSE, as well as the homologous element in the NaII gene. A point mutation in the NRSE that abolishes binding of this neural-restrictive silencer-binding factor (NRSBF) in vitro also eliminates silencing activity in vivo. NRSBF is present in nuclear extracts from nonneuronal cells but not in extracts from neuronal cells, suggesting that the neuron-specific expression of SCG10 reflects, at least in part, the absence or inactivity of this protein. These data identify the NRSE as a potentially general DNA element for the control of neuron-specific gene expression in vertebrates.

Introduction

The molecular mechanisms that generate cellular diversity in the developing vertebrate nervous system remain largely unknown. Experiments in invertebrate systems amenable to genetic analysis have suggested that the development of particular types of neurons involves a series of operations (Ghysen and Dambly-Chaudiere, 1989; Jan and Jan, 1990). These operations include the choice between a neuronal and a nonneuronal fate, and the choice of neuronal subtype. One approach to the problem of neural cell type specification in vertebrates is to clone homologs of invertebrate neurogenic regulatory genes and subsequently determine their function (Coffman et al., 1990; Johnson et al., 1990). Another, more systematic approach is to isolate regulatory proteins that are required for the transcription of genes specifically expressed in neurons or their precursors (Bodner et al., 1988; Ingraham et al., 1988).

We have studied the regulation of expression of a neuron-specific gene, SCG10, that was originally identified as a marker for sympathetic neurons derived from the neural crest (Anderson and Axel, 1985). SCG10 is expressed by most or all developing neurons in the embryo and is one of the earliest markers of neuronal differentiation (Stein et al., 1988a). In addition, SCG10 is up-regulated by nerve growth factor and fibroblast growth factor and is repressed by glucocorticoid in PC12 cells (Stein et al., 1988b). The regulation of this gene is therefore likely to be relevant to the decision between neuronal and nonneuronal fates, rather than to the selection of a particular neuronal phenotype. SCG10 encodes a membrane-associated protein that accumulates in the processes and growth cones of developing neurons (Stein et al., 1988a). It is highly homologous to a family of more widely expressed phosphoproteins (Doye et al., 1989; Shubart et al., 1989), suggesting that SCG10 is a kinase substrate as well. These features of expression pattern, phosphorylation, and subcellular localization are similar to those of GAP-43 (for review, see Benowitz and Routtenberg, 1987) and suggest that SCG10 may be another GAP. However, there is no sequence homology between SCG10 and GAP-43 (Basi et al., 1987; Karns et al., 1987; Stein et al., 1988a). The function of SCG10 remains unknown, although its properties suggest that it may play a role in growth cone extension (Stein et al., 1988a).

Studies of SCG10 regulation in transfected cell lines and in transgenic mice revealed, unexpectedly, that the expression of this gene is restricted to neuronal cells and tissues by a differential repression mechanism. The promoter-proximal region of the SCG10 gene is active in both neuronal and nonneuronal cell types (Mori et al., 1990), suggesting that it contains a constitutive enhancer (although in transgenic mice this enhancer is slightly more active in neuronal tissues [Wuenschell et al., 1990]). Distal to this proximal region lies a silencer element that represses the activity of the constitutive SCG10 promoter–enhancer, as well as that of the heterologous thymidine kinase promoter (Mori et al., 1990). This repression is exerted in nonneuronal cells and tissues but is abrogated in neuronal cells (Vandenbergh et al., 1989; Wuenschell et al., 1990). Thus, in contrast to many other tissue-specific genes whose specificity of expression is achieved by selectively expressed positive-acting factors (for review, see Maniatis et al., 1987), the neural specificity of SCG10 expression is achieved in large part by differential repression. A similar observation has been made for the rat type II sodium channel (NaII) gene (Maue et al., 1990), suggesting that selective repression may be a general mechanism used by at least a subset of neuron-specific genes.

To understand the role of the silencer in SCG10 expression and in the development of neural progenitor cells, we have sequenced the SCG10 upstream

†Present address: Andrus Gerontology Center, University of Southern California, University Park, MC 0191, Los Angeles, California 90089.

‡Present address: NIDA/Addiction Research Center, PO Box 5180, Baltimore, Maryland 21224.

Figure 1. Genomic Sequence of the SCG10 Upstream Region

The data include the sequence of the promoter-proximal region (−541) previously published (Mori et al., 1990). A sequence appearing in an SCG10 cDNA clone (10-8; Stein et al., 1988a) is underlined. A TATA box and a CCAAT box are indicated in bold letters. The names of different deletion and addition constructs (see Figure 2) are indicated above the sequence, with a dot to specify the precise endpoint of the construct. The silencer-containing region (−1493 to −1460) is shown in italics and is underlined. A sequence showing similarity with other neuron-specific genes (not shown) is underlined in the proximal region.

region and further delineated the neural-restrictive silencer element (NRSE) by deletion and addition experiments. Furthermore, we identified a neural-restrictive silencer-binding factor (NRSBF) that interacts with this element in a sequence-specific manner. The interaction of this NRSBF with the SCG10 silencer is competed by a homologous element from the NaII silencer region. NRSBF activity is present in extracts from nonneuronal cells but not from neuronal cells, consistent with the distribution of silencing activity as determined by transfection assays. This suggests that the SCG10 silencer and its associated binding factor may be part of a general mechanism to repress the expression of neuron-specific genes outside of the nervous system.

Results

Delineation of the Silencer Element

Earlier, crude deletion experiments had indicated that the silencer element lies between approximately −2 kb and −0.5 kb, relative to the SCG10 transcription start sites (Mori et al., 1990; Wuenschell et al., 1990). To determine the number of silencer elements in this region and to localize them more precisely, we first sequenced the SCG10 upstream region, through −2.2 kb (Figure 1). This sequence extends approximately 1.6 kb upstream of the previously published sequence of the SCG10 promoter-proximal region (Mori et al., 1990). Next, a series of 5′ deletions of the chloramphenicol acetyltransferase (CAT)–SCG10 promoter fusion construct CAT16 were generated within the upstream region (Figure 2). The endpoints of these deletions are mapped on the sequence shown in Figure 1. Subsets of these constructs were then transfected into HeLa and PC12 cells to assay silencer activity, in a series of three separate experiments. Both cell lines were transfected to ensure that any silencer element identified showed the correct cell type specificity, i.e., that it repressed transcription strongly in HeLa cells (a nonneuronal line) but weakly or not at all in PC12 cells (a neuroendocrine line).

Deletion of approximately 130 bp from the 5′ end of CAT16 (Mori et al., 1990) yielded a construct (CAT15; Figure 2) that retained significant silencing activity (Figure 3A). (The absolute extent of silencing in HeLa cells for a given construct varied from experiment to experiment as the result of varying transfection efficiencies; see Experimental Procedures.) A series of ten deletions between approximately −1990 bp (CATd1; Figure 2) and −740 bp (CATd10; Figure 2) revealed an abrupt change in silencing activity between the construct CATd5 and CATd6 (Figure 3B), a region between −1510 bp and −1282 bp (Figure 1). Although the signal

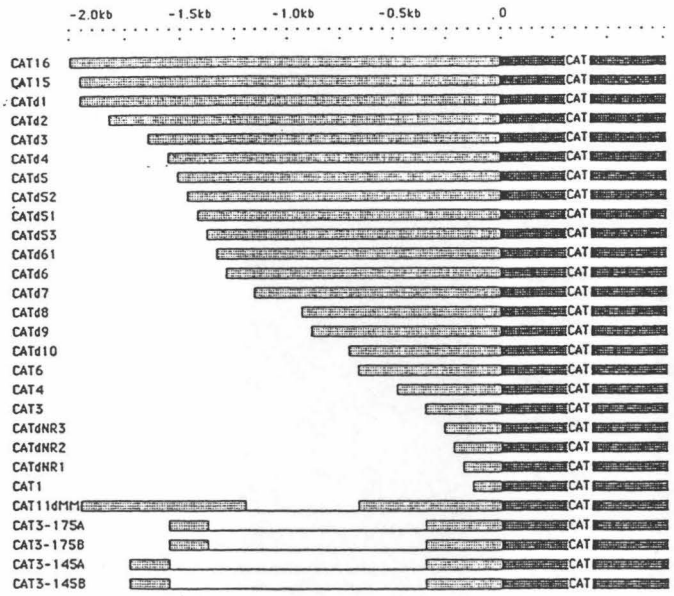


Figure 2. Schematic Diagram Illustrating SCG10-CAT Deletion and Addition Constructs

Thin lines indicate internal deletions. Constructs CATdNR1, 2, and 3 and CAT1 were made and examined in Mori et al., 1990, or in unpublished experiments and are included here for completeness only.

in HeLa cells in the experiment of Figure 3B was lower than usual because of reduced transfection efficiency (see figure legend), the silencing activity of the various deletion constructs is measured relative to the activity of the nonsilenced construct in the same cell line (see Experimental Procedures). Thus differences in transfection efficiency between HeLa and PC12 cells did not affect our ability to localize the silencer in such deletion experiments. Further deletions between −1510 and −1282 (Figure 3C) identified a silencer within a 62 bp domain (−1510 to −1448) defined by the endpoints between the constructs CATd5 and CATd52 (Figure 1; Figure 2). To determine whether this region was sufficient for silencing, or only necessary, addition constructs were made in which a 175 bp AluI fragment spanning the silencer was fused in either orientation to the promoter-proximal region contained in CAT3 (CAT3-175A,B; Figure 2). These addition constructs silenced efficiently in HeLa cells, but not in PC12 cells (CATa175A,B; Figure 3D). As a control, similar addition constructs generated from the adjacent 145 bp RsaI–AluI fragment (CAT3-145A,B; Figure 2) showed little or no silencing activity (CATa145A,B; Figure 3D).

Identification of a Homologous Silencer Element in the NaII Promoter

While this work was in progress, several studies appeared identifying silencer regions in other neuron-specific genes. In particular, analysis of the NaII gene indicated the presence of three separate regions containing silencing activity (Maue et al., 1990). Although the precise silencer elements were not identified in that study, we visually compared the sequences in the NaII silencer domains with those within the 62 bp SCG10 silencer region identified in the preceding deletion analysis. We identified a shorter sequence in which 17/21 bp were conserved between the two genes (Figure 4). The observation of a region of sequence identity within a silencer-containing domain of the NaII gene suggested that this region might define more precisely a silencer element, present not only in SCG10, but in other neuron-specific genes as well. We also noted that similar regions of homology were present in the promoter regions of both the human and rat synapsin I genes, which encode another neuron-specific protein (Sauerwald et al., 1990; Südhof, 1990) (Figure 4). Although neuronal cell type–specific expression is regulated by these synapsin I promoter regions, it is not yet clear whether they contain functional silencer elements (Sauerwald et al., 1990; Thiel et al., 1991).

We next asked whether the homology-containing region was necessary and sufficient for silencing activity. Deletion of a 25 bp fragment (−1487 to −1468) containing this sequence from the addition construct CAT3-175 (Figure 2; Figure 3) resulted in a virtual elimination of silencing activity (Table 1, CAT3-175Δ25), indicating that this shorter sequence was necessary for silencing. Moreover, an addition construct in which a 36 bp oligomer spanning the homologous region (−1503 to −1467) was fused to the CAT3 promoter-proximal construct showed silencing activity as well (Table 1, CAT3-S36(+)). Addition of a dimer of the 36 bp oligo (S36(++)) increased the extent of silencing by 3- to 5-fold (Table 1). Silencing was also observed when

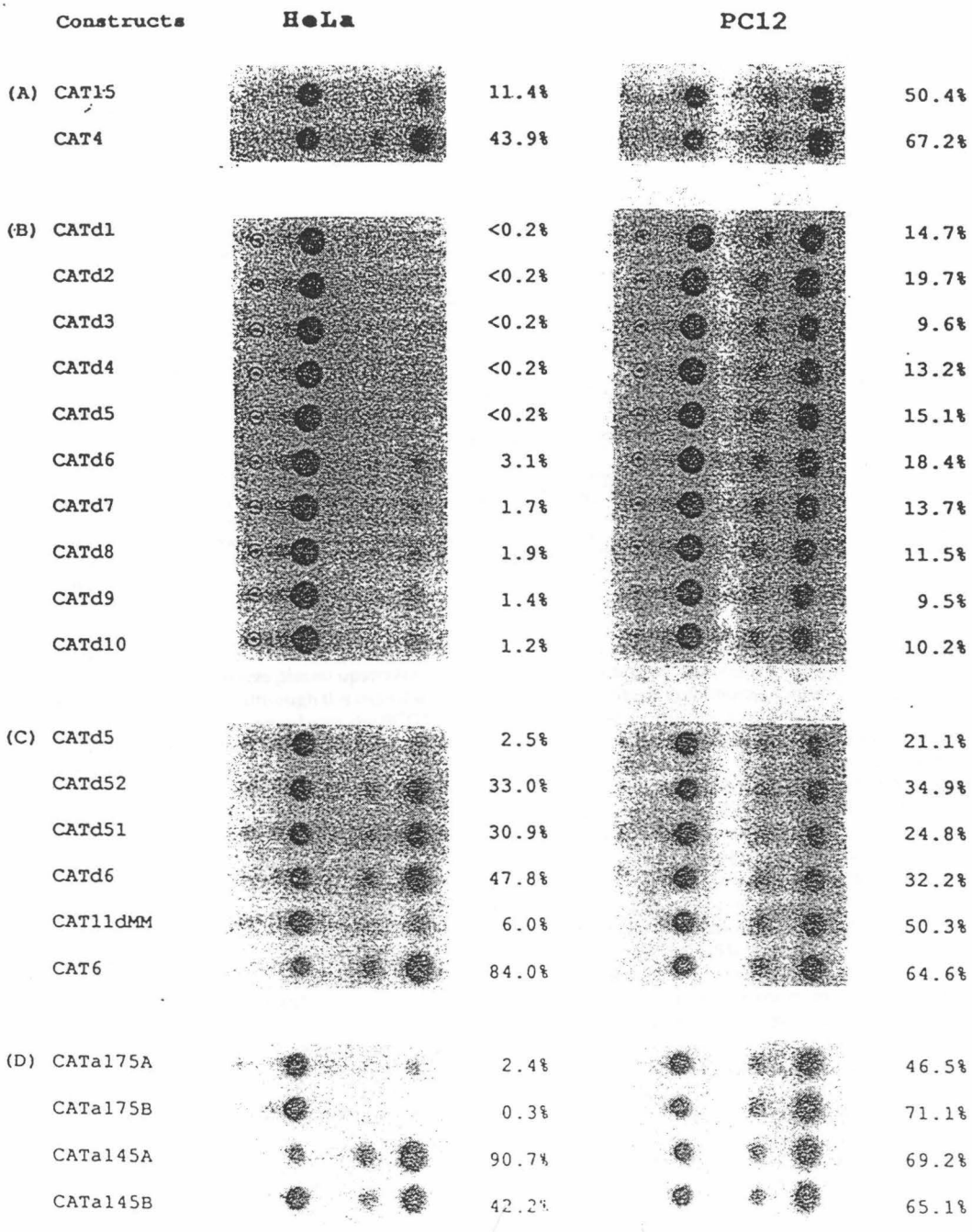


Figure 3. Localization of the SCG10 Silencer

Shown are a series of three separate experiments (A, B, and C/D) in which the constructs illustrated in Figure 2 were tested by parallel transient transfection into HeLa and PC12 cells. The percent conversions of chloramphenicol to acetylated chloramphenicol, indicated to the right of the autoradiograms, are not normalized and should not be directly compared between HeLa and PC12 cells. In the experiment of (B), the transfection efficiency in HeLa cells was lower than that in the other experiments because miniprep rather than CsCl-banded DNA was used (see Experimental Procedures) and because only 5 μg rather than 16 μg of plasmid per plate was used. However, since silencing activity is identified by comparing different constructs within the same cell line, an abrupt loss of silencing

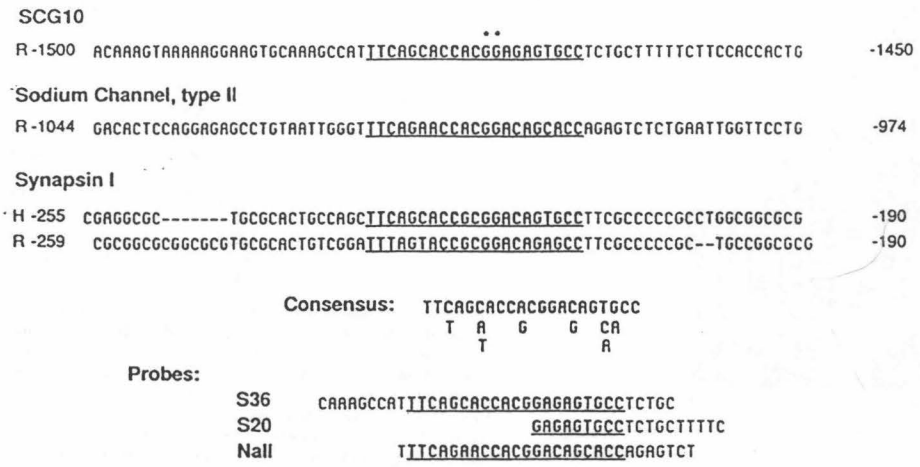


Figure 4. Alignment of NRSEs between SCG10 and Other Neuron-Specific Genes

R, rat; H, human. The sequence of the NaII gene is taken from Maue et al. (1990). The sequence of the synapsin I gene is from Sauerwald et al. (1990). This region of the synapsin I gene has not yet been demonstrated to contain a silencer. The region of maximum sequence identity is boxed and indicated in boldface, although not all residues within the box are perfectly conserved (see Consensus) Asterisks indicate the two guanines mutated to thymines in the Sm36 oligonucleotide. The probes and competitors used in the gel-shift experiments of Figures 5 and 6 are indicated below the consensus sequence.

the region of the NaII gene homologous to the SCG10 silencer (Figure 4, NaII) was placed upstream of CAT3 (Table 1, CAT3-NaII(+)), although the extent of silencing was less than that observed with the SCG10 oligomer. As in the case of the SCG10 silencer, a dimer of the NaII element increased the extent of silencing compared with monomer (Table 1, NaII(++)). Little or no silencing by either the 36 bp SCG10 element or the homologous NaII element was observed in PC12 cells (data not shown). Taken together, these data suggest that the region of sequence homology shared by the SCG10 and NaII genes contains a silencer that is both necessary and sufficient for the selective suppression of neuron-specific gene expression in nonneuronal cells. We have termed this silencer the neural-restrictive silencer element, or NRSE.

Identification of a Silencer-Binding Protein

To identify a protein or proteins that interact with the SCG10 silencer, we performed a series of electrophoretic mobility shift (EMS) assays on nuclear extracts from HeLa cells using the 36 bp silencer-containing oligonucleotide (S36; Figure 4) as a probe. Initial experiments revealed no shift, or only a very weak one (Figure 5, lane 1), with this probe. Because the affinity of many DNA-binding proteins is increased by dimerization of their sites, the EMS assays were repeated using a probe containing a dimer of the S36 fragment. Under these conditions, a much stronger shift was detected in the HeLa cell nuclear extracts (Figure 5, lane 4). The size of this complex was indistinguishable from that obtained with the monomeric probe, suggesting that the increased binding affinity was not due to cooperative binding of two of the silencer-binding factors. The stronger shift obtained with the S36 dimer in vitro is qualitatively consistent with the greater silencing activity observed for the S36 dimer compared with the S36 monomer in vivo (Table 1). However, it is curious that the absolute amount of binding activity obtained with the S36 monomer is so low given that this element is an effective silencer in transfection assays (Table 1). There are numerous possible explanations for this observation, including qualitative differences between in vitro and in vivo assay conditions. Studies with purified and/or cloned binding factor will be necessary to clarify this issue.

The specificity of the gel shift obtained with the S36 dimeric probe was established by a series of competition experiments. Specific competition for both monomer and dimer shifts was obtained with a 1200-fold molar excess of S36 monomeric DNA (Figure 5, lanes 2, 5, and 6), but not with a similar excess of irrelevant

between the constructs CATd5 and CATd6 in HeLa cells can still be detected. Silencing efficiencies also vary from experiment to experiment within HeLa cells (see Experimental Procedures), accounting for the differences in extent of silencing between the independent transfections of (A), (B), and (C/D). In (C) note the loss of silencing between the constructs CATd5 and CATd52. Note also that the addition construct CATa175 (CAT3-175 of Figure 2) contains potent silencing activity in both orientations (D), while an addition construct containing an adjacent fragment (CATa145, the same as CAT3-145 in Figure 2) does not. Little or no silencing is observed with these constructs in PC12 cells.

Table 1. Silencing Activity of NRSE-Containing DNA Fragments

	% Conversion ± SEM[a]	Fold Suppression[b]
CAT3	100 ± 11	1
CAT3-175	2.7 ± 0.2	37
CAT3-175Δ25	86 ± 6	1.2
CAT3-S36(+)	5.8 ± 0.7	17
CAT3-S36(++)	1.5 ± 0.1	67
CAT3-NaII(+)	13 ± 2	7.7
CAT3-NaII(++)	3.7 ± 0.3	27
CAT3-Sm36(+)	107 ± 16	0.9
CAT3-Sm36(++)	90 ± 8	1.1

[a] The percent conversion of chloramphenicol to acetylated chloramphenicol is shown for each construct, determined by liquid scintillation counting of spots excised from TLC plates. In these experiments, the HeLa cells were not split after transfection and glycerol shock was omitted. CAT3 activity is normalized to 100% for purposes of comparison to the other constructs. (+) indicates a monomeric site, and (++) indicates a dimeric site, all in the same orientation as the naturally occurring NRSE. The numbers represent the mean ± SEM of two independent experiments, each of which was performed in duplicate. In both experiments, the activity of the CAT constructs was normalized to that of pRSV-lacZ, a cotransfected internal control plasmid (Johnson et al., 1992).

[b] The fold suppression is calculated as described by Mori et al. (1990) for each of the constructs relative to CAT3, a construct containing the proximal promoter-enhancer.

octamer DNA (Figure 5, lanes 3 and 12) or with a 20 bp fragment (S20) (Figure 5, lane 11) partially overlapping the S36 region (Figure 4) that lacks silencer activity (data not shown). As expected from its sequence homology and functional similarity (Table 1), the NaII sequence (Figure 4, NaII) also competitively inhibited the formation of the S36 dimeric complex (Figure 5, lanes 7 and 8). However, the efficacy of competition was 2–5 times less than that observed with homologous S36 DNA (Figure 5, compare lanes 5 and 8). Consistent with this, the NaII silencer was 2–5 times less effective than the SCG10 silencer in transfection assays (Table 1). To provide further evidence for a relationship between the formation of the S36 complex and silencing activity, a mutation was introduced near the center of the element converting two adjacent guanines to thymines (CACGGAGA→CACTTAGA; see Figure 4). Preliminary footprinting experiments using methylation-interference indicate that the guanine residues changed in this mutation are in fact contacted by the NRSBF in vitro (C. Schoenherr, unpublished data). The mutant oligonucleotide (Sm36) was 50- to 100-fold less effective in competition than wild-type DNA (Figure 5, compare lanes 5 and 10). Moreover, in transient transfection assays, the same mutation abolished the silencing activity of both S36 monomer and dimer addition constructs (Table 1, CAT3-Sm36(+) and Sm36(++)). These data therefore suggest that the DNA-protein complexes revealed by the gel-shift assay in vitro reflect the activity of the factor involved in silencing SCG10 transcription in vivo.

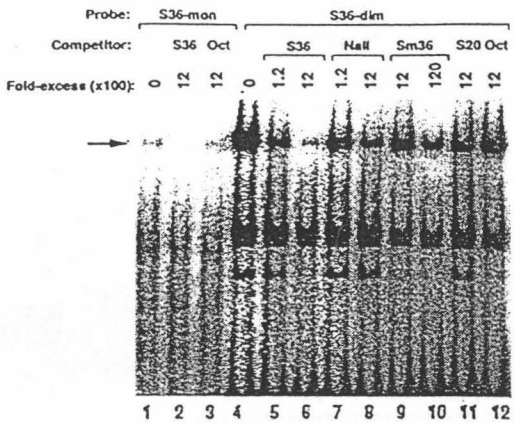


Figure 5. Identification of a Silencer-Binding Factor in HeLa Nuclear Extracts

EMS assays were performed using large scale nuclear extracts prepared as described in Experimental Procedures. Probes used were either the oligonucleotide S36 (Figure 4) containing the core silencer homology element (S36-mon), or a restriction fragment containing two tandem copies of S36 (S36-dim). Competitors used were S36, unlabeled S36; Oct, octamer-binding sequence (Muller et al., 1988); NaII, the NaII channel silencer homology element (see Figure 4); Sm36, mutant S36 (see text); S20, a 20 bp fragment of SCG10 overlapping but not containing the entire NRSE (Figure 4). The arrow indicates the complex that is specifically competed; the higher mobility complexes did not show specific competition.

The Silencer-Binding Factor Is Absent from Neuronal Cell Lines

The identification of an NRSBF for neural-specific genes such as SCG10 and NaII raises the question of how the cell specificity of silencing is achieved. A simple explanation for differential silencing is that nonneuronal cells contain active NRSBF, whereas neuronal cells do not. Alternatively, NRSBF could be present in both neuronal and nonneuronal cells, but its action might be compensated in neurons by positive-acting factors that bind elsewhere in the gene. As a first step toward distinguishing between these possibilities, we performed gel-shift assays using the S36 dimer probe on extracts from several different neuronal and nonneuronal cell lines. S36 binding activity was detected in nuclear extracts derived from three different nonneuronal cell lines able to silence SCG10-CAT constructs in transient transfection experiments (C. Schoenherr, unpublished data): human HeLa, mouse 10T1/2, and BALB/c 3T3 cells (Figure 6, lanes 1, 4, and 6). This activity was competed by an excess of S36 oligo (Figure, 6 lanes 2, 5, and 7), but not by irrelevant octamer oligo (lanes 3, 6, and 9). By contrast, little or no S36 dimer binding activity was detected in three different neuronal cell lines expressing endogenous and/or transfected SCG10: rat PC12 and MAH cells and SY5Y human neuroblastoma cells (Figure 6, lanes 10, 13, and 16). Control experiments showed that the nuclear extracts from all three neu-

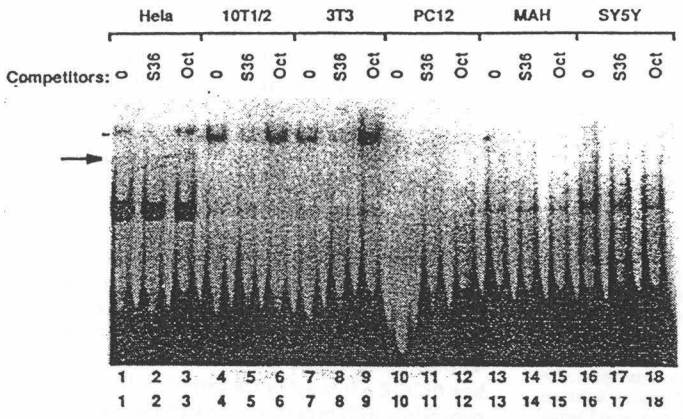


Figure 6. The NRSBF Is Present in Nuclear Extracts from Nonneuronal Cells but Not from Neuronal Cells

EMS assays were performed using small scale nuclear extracts made from the cell lines indicated above the lanes. Competitors are indicated and were used at a 1200-fold molar excess in each case. As the result of nonspecific nuclease activity in some of the cell lines (e.g., PC12), the probe is somewhat degraded in the absence of added cold competitor (lane 10); however, no specific complex is detected in the presence of the nonspecific octamer-binding sequence (Oct) competitor that eliminates the degradation (lane 12). This protection from probe degradation also accounts for the slightly higher amount of specific complex formed in HeLa, 10T1/2, and BALB/c cells in the presence of octamer competitor (lanes 3, 6, and 9). Control experiments using an octamer probe indicated that the PC12, MAH, and SY5Y nuclear extracts contained substantial amounts of octamer factor (data not shown). All three nonneuronal cell lines exhibited silencing of SCG10–CAT constructs in transient transfection assays (C. Schoenherr, unpublished data).

ronal lines contained substantial amounts of octamer-binding factor (data not shown), indicating that the failure to detect the NRSBF did not reflect a general inactivity of these extracts. These results therefore suggest that NRSBF activity is present in nonneuronal cell lines but absent from cell lines of neuronal or neuroendocrine origin.

Discussion

The neuron-specific expression of the SCG10 gene appears to be controlled by a silencer element that selectively represses transcription in nonneuronal cells and tissues (Mori et al., 1990; Wuenschell et al., 1990). This stands in contrast to many other tissue-specific genes from nonneuronal tissues, in which case specificity is achieved in large part by specifically expressed enhancer-binding factors (Maniatis et al., 1987). A similar mechanism of differential silencing has been observed for at least two other neuron-specific genes: the NaII channel (Maue et al., 1990) and choline acetyltransferase (Ibánez and Persson, 1991). This selective repression of several neuron-specific genes suggests that differential silencing could be a general mechanism controlling gene expression in the nervous system. As discussed previously (Mori et al., 1990), this may reflect the fact that many neuron-specific genes are members of multigene families which contain other genes expressed more broadly. All genes in such families might therefore contain nonselective enhancers, and the neural-specific genes may have evolved silencer elements to compensate for such enhancers in nonneuronal tissues. In the case of SCG10, for example, a closely related gene called P19/stathmin is expressed not only in the nervous system, but in many other cells and tissues as well (Doye et al., 1989; Shubart et al., 1989). Likewise, the NaII gene family contains channels expressed specifically in nonneuronal tissues such as muscle (Trimmer et al., 1989). However, although many neural-specific genes may use a common regulatory strategy, it is not yet clear whether the specific mechanism of silencing and the factors that mediate repression are shared by these genes.

We have narrowed the location of a silencer in the SCG10 gene to a 36 bp region located approximately 1.5 kb upstream of the promoter, by a series of deletion and addition experiments. This region contains a sequence of 21 bp, which we have termed the neural-restrictive silencer element (NRSE) that is highly similar to a sequence in the upstream regions of both the NaII and synapsin I genes. The present data suggest that the NRSE in SCG10 is both necessary and sufficient for silencing and that a similar sequence in the NaII gene possesses comparable activity. An independent deletion analysis of the NaII gene has localized a silencer element in the same region (residues −1017 to −996) that we identified by homology with SCG10 (Kraner et al., 1992).

Silencer elements have been identified in a number of genes expressed outside of the nervous system (for review, see Renkawitz, 1990). In some cases, these silencers interact with positive-acting elements to modulate quantitatively the extent of expression in different cells or tissues (Fujita et al., 1988; Camper and Tilghman, 1989; Baniahmad et al., 1990; Tada et al., 1991; Weissman and Singer, 1991). In other cases, the silencer elements contribute to lineage or tissue specificity, as in the cases of the collagen II (Savagner et al., 1990), cardiac myosin light chain 2 (Shen et al., 1991), immunoglobulin heavy chain (Weinberger et al., 1988), and T cell receptor (Winoto and Baltimore, 1989) genes. In all of these cases, however, the silencer elements are not the major determinant of lineage specificity, but rather achieve differential expression in closely related cell types of genes that also use

cell-specific enhancers. An exception is the CyIIIa actin gene of the sea urchin embryo, in which a negative-acting element appears to be the predominant determinant of the spatial specificity of expression (Hough-Evans et al., 1990). Proteins that interact with silencer elements have been cloned in only a small number of cases (Shore and Nasmyth, 1987; Diffley and Stillman, 1989; Kageyama and Pastan, 1989; Calzone et al., 1991; Höög et al., 1991). None of the silencer elements defined in these studies shows any substantial sequence similarity with the NRSE in SCG10 and NaII (data not shown).

Using a gel-shift assay, we were able to identify an apparent neural-restrictive silencer-binding factor that interacts with oligonucleotides containing the SCG10 and NaII NRSEs. There is a quantitative correlation between the relative activities of the SCG10 and NaII NRSE-containing sequences in in vitro and in vivo assays. These data provide a strong correlation between the presence of an NRSE in a DNA sequence and its ability to bind NRSBF and to silence transcription in vivo. This correlation suggests that the binding of NRSBF to the NRSE may be necessary for silencing in vivo, although this remains to be proven. Furthermore, the fact that a common factor (or family of related factors of similar size and sequence specificity) interacts with both the SCG10 and the NaII silencers suggests that the NRSBF could be a silencing protein used to repress a number of neuron-specific genes in nonneuronal cells and tissues.

NRSBF is present in nuclear extracts from several different nonneuronal cell lines but not in extracts from neuronal cell lines. There is thus a strong correlation between the presence of NRSBF, the capacity to silence transfected SCG10–CAT constructs, and the absence of endogenous SCG10 expression. The fact that silencer-containing SCG10–CAT constructs are not expressed in nonneuronal tissues of transgenic mice (Wuenschell et al., 1990), moreover, suggests that silencing activity is present in a wider variety of nonneuronal tissues than is represented by the cell lines we have used in transfection experiments. These data implicate NRSBF in silencing in vivo and suggest that it functions by continuously maintaining SCG10 repression, rather than by initiating a repression that is maintained, for example, by chromatin structure. Our results further suggest that SCG10 expression in neuronal tissues reflects the absence of NRSBF expression or activity. By extension, the initiation of SCG10 expression during neuronal development may involve a relief of repression maintained in precursors by NRSBF, i.e., specific derepression. As discussed previously (Mori et al., 1990), this derepression may occur by a loss of NRSBF or by the gain of a competing or inactivating anti-silencer protein. The elucidation of the mechanism of SCG10 derepression will require the isolation and cloning of NRSBF, a goal now made accessible by the identification of the NRSE. Such information should also clarify the issue of whether SCG10, NaII, choline acetyltransferase, and other neuron-specific genes are repressed by the identical silencer-binding proteins, or by a family of related proteins. Whatever the case, the results suggest that a mechanism for achieving cell-specific expression that previously was thought to be uncommon may be more the rule than the exception, at least in the nervous system.

Experimental Procedures

Constructions and Transfections

DNA sequencing, site-directed mutagenesis, and construction of addition and deletion constructs were carried out by standard molecular biological procedures. Deletion constructs of SCG10–CAT plasmids were generated according to Henikoff (1984). CAT37 (Mori et al., 1990) was first linearized with BglII, followed by digestion with exonuclease III (USB, 7000 U/ml) for up to 25 min. Aliquots were taken at 1 min intervals, added into a solution containing mung bean nuclease (NE Biolabs, 240 μ/ml), and incubated for 30 min at room temperature. The series of nuclease-digested DNAs were cleaved with XbaI or NdeI in order to cut out the far upstream region of the SCG10 genomic sequences, and then 5′ overhangs were filled in using the Klenow fragment of DNA polymerase I. Following transformation of this reaction, a series of plasmids with appropriately sized inserts were randomly chosen (CATd1–CATd10), and the boundaries around the deletion were sequenced. In a separate mutagenesis, deletion constructs CATd51–CATd69 were made to define the region between CATd5 and CATd6 more precisely. Constructs containing S36 wild-type and mutant and NaII oligonucleotides were made by restricting pCAT3 with HindIII and inserting their respective oligonucleotides. The sequence and orientation were confirmed by sequencing. The deletion in CAT3-175Δ25 was generated by site-directed mutagenesis using the Stratagene Mutator kit. The deletion removes the sequence from −1487 to −1468 and was sequenced to confirm the deletion. Other internal deletion and addition constructs were made by standard restriction and ligation procedures; details are available on request.

Conditions for transfection of HeLa and PC12 cells were as described previously (Mori et al., 1990), except that in some experiments, alkaline lysis miniprep rather than CsCl-banded plasmid DNA was used to facilitate the rapid analysis of multiple constructs (see Figure 3B). Control experiments indicated that the transfection efficiency using such miniprep DNA was 50% of that obtained with CsCl-banded DNA. In addition, internal controls using pRSV-lacZ (Johnson et al., 1992) indicated that the transfection efficiency in HeLa cells is lower than that in PC12 cells. However, within a given experiment the extent of silencing is calculated by comparing the activity of various mutant constructs with that of CAT3 or CAT4, the nonsilenced constructs, within the same cell line (Mori et al., 1990). In this way, differences in transfection efficiency between different cell lines do not influence the measurement of silencing activity. However, for a given construct, the extent of silencing relative to CAT3 or CAT4 varied from experiment to experiment. Plasmid titration experiments (N. Mori, unpublished data) have revealed that the extent of silencing decreases strongly as the amount of transfected plasmid is increased, probably reflecting saturation of the silencer-binding factor. Thus variations in the extent of silencing from experiment to experiment could reflect variations in transfection efficiency that affects the amount of DNA incorporated per cell.

Nuclear Extracts and Gel-Shift Assays

Large scale nuclear extracts (Figure 5) were prepared from 24 liters of HeLa cell suspension cultures grown in Dulbecco's modified Eagle's medium containing 10% calf serum. Extracts were made using 1 ml packed volume of HeLa cells, essentially as described by Harshman et al. (1988). Small scale nuclear extracts (Figure 6) were prepared from 10^8–10^9 cells according to Schreiber et al. (1989), with the exception that the nuclear extrac-

tion buffer contained 0.5 M NaCl instead of 0.4 M NaCl. EMS assays using large scale nuclear extracts were performed in 16 μl final volume of reaction buffer containing 20 mM HEPES (pH 7.6), 0.1% NP-40, 10% glycerol, 1 mM dithiothreitol, 2.5 mM $MgCl_2$, 250 mM KCl, and 125 μg/ml poly(dI-dC). Approximately 7 μg of the large scale HeLa extract was added to the reaction buffer and preincubated for 10 min at 4°C. Labeled DNA probe (1×10^4 to 2×10^4 cpm per reaction) and competitors were then added, followed by a 10 min incubation at room temperature. Electrophoresis was performed on a 4% polyacrylamide gel in 0.25× TBE, for 1.5–2 hr at 150 V. EMS assays using small scale extracts were similarly performed, except that 16 μg/ml HindIII-digested bacteriophage λ DNA was added and the KCl was replaced by 95 mM NaCl, contributed by the addition of the extracts. Positive controls for the activity of nuclear extracts were performed using an octamer factor-binding probe containing the core sequence ATTTGCAT (Muller et al., 1988). Gel shifts obtained using the NRSE probe were substantially weaker than those obtained using the control octamer probe.

Acknowledgments

We are grateful to Dr. Carl Parker for his advice and the use of his facilities for growing large scale HeLa cell cultures and preparing nuclear extracts. We thank Dr. Takashi Okazaki for assistance with some of the transfection experiments, Steven Padilla for plasmid preparations, and Dr. Susan Birren, Dr. Barbara Wold, and one of the reviewers for their helpful comments on the manuscript. This work was supported by National Institutes of Health grant NS23476. D. J. A. is an Assistant Investigator of the Howard Hughes Medical Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 USC Section 1734 solely to indicate this fact.

Received March 6, 1992; revised April 22, 1992.

References

Anderson, D. J., and Axel, R. (1985). Molecular probes for the development and plasticity of neural crest derivatives. Cell *42*, 649–662.

Baniahmad, A., Steiner, C., Köhne, A. C., and Renkawitz, R. (1990). Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. Cell *61*, 505–514.

Basi, G. S., Jacobson, R. D., Virág, I., Schilling, J., and Skene, J. H. P. (1987). Primary structure and transcriptional regulation of GAP-43, a protein associated with nerve growth. Cell *49*, 785–791.

Benowitz, L. I., and Routtenberg, A. (1987). A membrane phosphoprotein associated with neural development, axonal regeneration, phospholipid metabolism, and synaptic plasticity. Trends Neurosci. *10*, 527–532.

Bodner, M., Castrillo, J.-L., Theill, L. E., Deerinck, T., Ellisman, M., and Karin, M. (1988). The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. Cell *55*, 505–518.

Calzone, F. J., Hoög, C., Teplow, D. B., Cutting, A. E., Zeller, R. W., Britten, R. J., and Davidson, E. H. (1991). Gene regulatory factors of the sea urchin embryo. I. Purification by affinity chromatography and cloning of P3A2, a novel DNA-binding protein. Development *112*, 335–350.

Camper, S. A., and Tilghman, S. M. (1989). Postnatal repression of the α-fetoprotein gene is enhancer-dependent. Genes Dev. *3*, 537–546.

Coffman, C., Harris, W., and Kintner, C. (1990). *Xotch*, the *Xenopus* homolog of *Drosophila Notch*. Science *249*, 1438–1441.

Diffley, J. F. X., and Stillman, B. (1989). Similarity between the transcriptional silencer binding proteins ABF1 and RAP1. Science *246*, 1034–1038.

Doye, V., Soubrier, F., Bauw, G., Boutterin, M.-C., Beretta, L., Koppel, J., Vandekerckhove, J., and Sobel, A. (1989). A single cDNA encodes two isoforms of stathmin, a developmentally regulated neuron-enriched phosphoprotein. J. Biol. Chem. *264*, 12134–12137.

Fujita, T., Sakakibara, J., Sudo, Y., Miyamoto, M., Kimura, Y., and Taniguchi, T. (1988). Evidence for a nuclear factor(s), IRF-1, mediating induction and silencing properties to human IFN-β gene regulatory elements. EMBO J. *7*, 3397–3405.

Ghysen, A., and Dambly-Chaudiere, C. (1989). Genesis of the *Drosophila* peripheral nervous system. Trends Genet. *5*, 251–255.

Harshman, K. D., Moye-Rowley, W. S., and Parker, C. S. (1988). Transcriptional activation by the SV40 AP-1 recognition element in yeast is mediated by a factor similar to AP-1 that is distinct from GCN4. Cell *53*, 321–330.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene *28*, 351–357.

Hoög, C., Calzone, F. J., Cutting, A. E., Britten, R. J., and Davidson, E. H. (1991). Gene regulatory factors of the sea urchin embryo. II. Two dissimilar proteins, P3A1 and P3A2, bind to the same target sites that are required for early territorial gene expression. Development *112*, 351–364.

Hough-Evans, B. R., Franks, R. R., Zeller, R. W., Britten, R. J. and Davidson, E. H. (1990). Negative spatial regulation of the lineage specific CyIIIa actin gene in the sea urchin embryo. Development *108*, 41–50.

Ibánez, C. F., and Persson, H. (1991). Localization of sequences determining cell type specificity and NGF responsiveness in the promoter region of the rat choline acetyltransferase gene. Eur. J. Neurosci. *3*, 1309–1315.

Ingraham, H. A., Chen, R., Mangalam, H. J., Elsholtz, H. P., Flynn, S. E., Lin, C. R., Simmons, D. M., Swanson, L., and Rosenfeld, M. G. (1988). A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell *55*, 519–529.

Jan, Y. N., and Jan, L. Y. (1990). Genes required for specifying cell fates in *Drosophila* embryonic sensory nervous system. Trends Neurosci. *13*, 493–498.

Johnson, J. E., Birren, S. J., and Anderson, D. J. (1990). Two rat homologues of *Drosophila achaete-scute* specifically expressed in neuronal precursors. Nature *346*, 858–861.

Johnson, J. E., Birren, S. J., Saito, T., and Anderson, D. J. (1992). DNA binding and transcriptional regulatory activity of mammalian achaete-scute homologous (MASH) proteins revealed by interaction with a muscle-specific enhancer. Proc. Natl. Acad. Sci. USA *89*, 3596–3600.

Kageyama, R., and Pastan, I. (1989). Molecular cloning and characterization of a human DNA binding factor that represses transcription. Cell *59*, 815–825.

Karns, L. R., Ng, S.-C., Freeman, J. A., and Fishman, M. C. (1987). Cloning of complementary DNA for GAP-43, a neuronal growth-related protein. Science *236*, 597–600.

Kraner, S. D., Chong, J. A., Tsay, H.-J., and Mandel, G. (1992). Silencing the type II sodium channel gene: a model for neural-specific gene regulation. Neuron, this issue.

Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987). Regulation of inducible and tissue-specific gene expression. Science *236*, 1237–1245.

Maue, R. A., Kraner, S. D., Goodman, R. H., and Mandel, G. (1990). Neuron-specific expression of the rat brain type II sodium channel gene is directed by upstream regulatory elements. Neuron *4*, 223–231.

Mori, N., Stein, R., Sigmund, O., and Anderson, D. J. (1990). A cell type-preferred silencer element that controls the neural-specific expression of the SCG10 gene. Neuron *4*, 583–594.

Muller, M. M., Ruppert, S., Schaffner, W., and Matthias, P. (1988). A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B-cells. Nature *336*, 544–551.

Renkawitz, R. (1990). Transcriptional repression in eukaryotes. Trends Genet. *6*, 192–197.

Sauerwald, A., Hoesche, C., Oschwald, R., and Kilimann, M. W. (1990). The 5'-flanking region of the synapsin-I gene—a G+C-rich, TATA-less and CAAT-less, phylogenetically conserved sequence with cell type-specific promoter function. J. Biol. Chem. *265*, 14932–14937.

Savagner, P., Miyashita, T., and Yamada, Y. (1990). Two silencers regulate the tissue-specific expression of the collagen II gene. J. Biol. Chem. *265*, 6669–6674.

Schreiber, E., Matthias, P., Muller, M. M., and Schaffner, W. (1989). Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. Nucl. Acids Res. *17*, 6419.

Shen, R., Goswami, S. K., Mascareno, E., Kumar, A., and Siddiqui, M. A. Q. (1991). Tissue-specific transcription of the cardiac myosin light-chain 2 gene is regulated by an upstream repressor element. Mol. Cell. Biol. *11*, 1676–1685.

Shore, D., and Nasmyth, K. (1987). Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. Cell *51*, 721–732.

Shubart, U. K., Banerjee, M. D., and Eng, J. (1989). Homology between the cDNAs encoding phosphoprotein p19 and SCG10 reveals a novel mammalian gene family preferentially expressed in developing brain. DNA *8*, 389–398.

Stein, R., Mori, N., Matthews, K., Lo, L.-C., and Anderson, D. J. (1988a). The NGF-inducible SCG10 mRNA encodes a novel membrane-bound protein present in growth cones and abundant in developing neurons. Neuron *1*, 463–476.

Stein, R., Orit, S., and Anderson, D. J. (1988b). The induction of a neural-specific gene, SCG10, by nerve growth factor in PC12 cells is transcriptional, protein synthesis dependent, and glucocorticoid inhibitable. Dev. Biol. *127*, 316–325.

Südhof, T. C. (1990). The structure of the human synapsin I gene and protein. J. Biol. Chem. *265*, 7849–7852.

Tada, H., Lashgari, M. S., and Khalil, K. (1991). Regulation of JCV-L promoter function: evidence that a pentanucleotide "silencer" repeat sequence AGGGAAGGGA down-regulates transcription of the JC virus late promoter. Virology *180*, 327–338.

Thiel, G., Greengard, P., and Südhof, T. C. (1991). Characterization of tissue-specific transcription by the human synapsin I gene promoter. Proc. Natl. Acad. Sci. USA *88*, 3431–3435.

Trimmer, J. S., Cooperman, S. S., Tomiko, S. A., Zhou, J., Crean, S. M., Boyle, M. B., Kallen, R. G., Sheng, Z., Barchi, R. L., Sigworth, F. J., Goodman, R. H., Agnew, W. S., and Mandel, G. (1989). Primary structure and functional expression of a mammalian skeletal muscle sodium channel. Neuron *3*, 33–49.

Vandenbergh, D. J., Wuenschell, C. W., Mori, N., and Anderson, D. J. (1989). Chromatin structure as a molecular marker of cell lineage and developmental potential in neural crest-derived chromaffin cells. Neuron *3*, 507–518.

Weinberger, J., Pat, P: S., and Sharp, P. A. (1988). Localization of a repressive sequence contributing to B-cell specificity in the immunoglobulin heavy chain enhancer. Mol. Cell. Biol. *8*, 988–992.

Weissman, J. D., and Singer, D. S. (1991). A complex regulatory DNA element associated with a major histocompatibility complex class I gene consists of both a silencer and an enhancer. Mol. Cell. Biol. *11*, 4217–4227.

Winoto, A., and Baltimore, D. (1989). αβ lineage-specific expression of the a T cell receptor gene by nearby silencers. Cell *59*, 649–665.

Wuenschell, C. W., Mori, N., and Anderson, D. J. (1990). Analysis of SCG10 gene expression in transgenic mice reveals that neural specificity is achieved through selective derepression. Neuron *4*, 595–602.

GenBank Accession Number

The sequence of the SCG10 upstream region has been deposited in the GenBank database under the accession number M90489.

Chapter 3

NRSF: A zinc finger protein that mediates coordinate repression of multiple neuronal genes in non-neuronal cells

Christopher J. Schoenherr and David J. Anderson

(The main body of this chapter was published as a report in *Science* which is included as Appendix I.)

ABSTRACT

The expression of the *SCG10* gene is restricted to neurons by selective repression: the gene contains a neural restrictive silencer element (NRSE) that prevents transcription in non-neuronal cells. We have isolated cDNA clones encoding a novel zinc finger protein called the neural-restrictive silencer factor (NRSF) which can bind the SCG10 NRSE. NRSF binds to consensus NRSEs in three other neuron-specific genes besides *SCG10*, suggesting that it coordinately represses multiple target genes. Recombinant NRSF can also function to repress transcription in an NRSE-dependent manner in vivo. Expression of NRSF mRNA is detected in most non-neuronal tissues at several developmental stages, suggesting that it functions as a near-global, sequence-specific repressor of neuronal gene expression. In the nervous system, NRSF mRNA is expressed in neuronal progenitors, as well as in glial cells, but not in mature neurons. This suggests that relief from NRSF-imposed repression may be a central event in the selection or execution of a neuronal differentiation program.

INTRODUCTION

The molecular basis of neuronal determination and differentiation in vertebrates is not well understood. To elucidate this process, it is necessary to identify the cell-intrinsic and -extrinsic molecules involved. One systematic approach for identifying cell-intrinsic molecules involved in neuronal cell fate determination is the identification of transcriptional regulatory sequences in the promoters of neuron-specific genes, and the isolation of proteins that interact with these sequences. This approach has already proven successful in other mammalian lineages, yielding a number of cell-type specific transcriptional enhancer factors (for reviews, see (He and Rosenfeld, 1991; Johnson and McKnight, 1989; Maniatis et al., 1987; Mitchell and Tjian, 1989)). In several of these cases, moreover, inactivation of the genes encoding these regulators in mice leads to defects in the development of the specific cell types from which these genes were initially isolated (Corcoran et al., 1993; Li

et al., 1990; Pevny et al., 1991). These examples indicate that systematic promoter analysis of cell-type specific genes can lead to the identification of genetically essential regulators of lineage determination or differentiation.

To apply this approach to the development of neurons, we have examined the transcriptional regulation of a neuron-specific gene, *SCG10* (Anderson and Axel, 1985). SCG10 is a 22 Kd, membrane-associated protein that accumulates in growth cones and is transiently expressed by all developing neurons (Stein et al., 1988). Upstream regulatory sequences controlling SCG10 transcription have been analyzed using promoter fusion constructs, both in transient cell transfection assays and in transgenic mice (Mori et al., 1990; Wuenschell et al., 1990). The results of these studies indicated that the upstream region could be divided into two regulatory domains: a promoter-proximal region that is active in many cell lines and tissues, and a distal region (approximately 1.6 kb upstream) that selectively represses this transcription in non-neuronal cells. Deletion of this upstream region relieves the repression of SCG10 transgenes in non-neuronal tissues, such as liver, in transgenic mice (Vandenbergh et al., 1989; Wuenschell et al., 1990). Thus, this repressing region appears to be a major determinant of the lineage specific expression of SCG10. Furthermore, in transient cell transfection assays, this distal region could repress transcription from a heterologous promoter in an orientation and distance independent manner (Mori et al., 1990), satisfying the criteria for a silencer: a sequence analogous to an enhancer but with an opposite effect on transcription (Brand et al., 1985).

The finding that expression of a neuron-specific gene is controlled primarily by selective repression stands in contrast to most cases of previously studied mammalian tissue-specific genes, for which cell type specificity is achieved primarily by tissue- or lineage-specific enhancers (Maniatis et al., 1987; Mitchell and Tjian, 1989). While silencer elements have been detected in other cell type-specific genes, such as the T-cell receptor α, CD4, and certain globin genes (Gutman et al., 1992; Sawada et al., 1994; Stamatoyannopoulos et al., 1993; Winoto and Baltimore, 1989), these silencers are not

responsible for repression in all non-expressing tissues. Rather, they appear to work in conjunction with tissue-specific enhancers to extinguish expression in closely related but inappropriate cell types. In contrast, the silencing region of SCG10 appears to be a major determinant of neuronal specificity, acting to prevent the utilization of a broadly-active promoter-proximal domain in non-neuronal tissues (Mori et al., 1992; Mori et al., 1990).

A detailed analysis of the SCG10 silencer region identified a ca. 25 bp element necessary and sufficient for silencing (Mori et al., 1992). Interestingly, similar sequence elements were identified in two other neuron-specific genes: the rat type II sodium (NaII) channel and the human synapsin I genes (Kraner et al., 1992; Li et al., 1993; Maue et al., 1990; Mori et al., 1992). These sequence elements were shown to possess silencing activity in transfection assays as well. These data suggest not only that selective repression may be a common theme in the transcriptional regulation of several neuron-specific genes, but also that a common *cis*-acting silencer element may mediate repression of these genes. We have therefore named this element the neural restrictive silencer element (NRSE)(Mori et al., 1992); in the context of the NaII channel gene, it has also been called repressor element 1 (RE1) (Kraner et al., 1992).

Using electrophoretic mobility shift assays, the NRSEs in the SCG10, NaII channel and synapsin I genes were all shown to form complexes with a protein present in non-neuronal cell extracts, but absent in neuronal cell extracts (Kraner et al., 1992; Li et al., 1993; Mori et al., 1992). The cell type specificity of this binding activity detected in vitro thus paralleled that of the functional silencing activity exhibited by the NRSE in vivo. Both the SCG10 and the NaII channel NRSEs competed with similar efficacy for the SCG10 NRSE binding protein, suggesting that a common protein could bind both NRSEs (Mori et al., 1992). This protein(s), termed the neural restrictive silencer factor (NRSF), thus appears to be important for the lineage specific repression of at least two neuron-specific genes. Taken together with the broad activity of the NRSE as demonstrated by cell transfection and transgenic mouse assays, these data suggest that NRSF may be the first

sequence-specific silencer protein identified in vertebrates that represses cell type-specific gene expression in a near-global manner.

In this report we describe the isolation and characterization of cDNAs encoding NRSF. Recombinant proteins encoded by these cDNAs can bind to the SCG10 and NaII channel NRSE elements in vitro and can inhibit transcription in an NRSE-dependent manner in vivo. Furthermore, the recombinant protein can also bind to NRSE-homologous sequences in the synapsin I and brain-derived neurotrophic factor (BDNF) genes, suggesting that the same protein is responsible for silencing four distinct neuronal genes. Sequence analysis of NRSF cDNAs reveals that NRSF is a novel protein with eight zinc fingers and a domain exhibiting similarity to other known transcriptional repression domains. Examination of NRSF mRNA expression shows that it is widely transcribed in non-neuronal cells and tissues, and absent from (or expressed at low levels in) neuronal cells. In addition, NRSF is expressed in multipotent precursors of neurons and glia, and this expression is maintained in glial cells. These data are consistent with the idea that NRSF functions as a virtually universal repressor of neuron-specific gene expression, and that relief from NRSF-imposed repression may constitute a central event in the commitment to, or execution of, a neuronal program of differentiation.

MATERIALS AND METHODS

Isolation of NRSF cDNAs

A HeLa cell λgt11 cDNA expression library (the generous gift of Paula Henthorn) was screened according to methods of in situ detection of filter-bound DNA-binding proteins (Singh et al., 1988; Vinson et al., 1988). Briefly, the nitrocellulose filters which overlaid the phage plaques were treated with guanidine-HCl and probed as in Vinson et al. (1988) and washed as in Singh et al. (1988). Approximately 2 million phage were screened with a radiolabeled probe consisting of three tandem copies of Na33 (see below).

The probe was generated by restriction digest with EcoRI and XhoI of a plasmid containing three Na33 oligonucleotides inserted into the HindIII site of pBluescript and was end-labeled using [α-^{32}P] dATP, dTTP and Klenow fragment. The correct fragment was isolated by PAGE and was further purified using Elutip chromatography (Schleicher and Schuell). Probes containing two copies of the S36 or Sm36 were isolated in the same manner and were used to confirm the DNA-binding specificity of plaques that recognized the Na33 probe.

To obtain additional cDNAs, a HeLa cell λZAPII (Stratagene) and a Balbc/3T3 cell λEXlox (the generous gift of S. Tavtigian and B. Wold) cDNA library were screened using standard hybridization procedures. Four other cDNA libraries were screened including several size-fractionated for large inserts; however, no cDNAs longer than 2kb were isolated. The nucleotide sequence of both strands of each cDNA was determined by the dideoxy sequencing method using Sequenase version 2.0 (U.S. Biochemicals). The resulting sequences were assembled and analyzed using the GCG (Devereux et al., 1984) and BLAST programs (Altschul et al., 1990). The PROSITE data base (Bairoch, 1992) was used to search for protein sequence motifs. cDNAs for mouse NRSF were isolated from the Balbc/3T3 library to permit analysis of the expression pattern of NRSF mRNA in the mouse and the rat. Characterization of a full length mouse NRSF cDNA is described in Chapter 4.

Preparation of antisera to NRSF

The λH1 cDNA was inserted into the EcoRI site of pGEX-1, a prokaryotic glutathione S-transferase fusion expression vector (Smith and Johnson, 1988). GST-λH1 fusion protein was induced with IPTG and partially purified by isolation of inclusion bodies containing the fusion protein. The inclusion body preparation was subjected to SDS-PAGE, and gel slices containing the fusion protein were excised, mixed with adjuvant, and injected into mice. Sera from the injected mice were titered by Western blot analysis of the fusion protein. When the serum titer reached a sufficient level, a myeloma

was injected into the peritoneum of the mouse, and a tumor was allowed to develop for 10 days. The polyclonal ascites fluid (Ou et al., 1993) induced by this tumor was collected and clarified by centrifugation.

EMSAs

To generate recombinant protein, the λH1 insert was subcloned into the EcoRI site of pRSET B (Invitrogen), which provided an in-frame start codon, a poly-histidine tag, and a T7 promoter. Recombinant λH1 was produced by in vitro transcription from linearized plasmid and in vitro translation using a rabbit reticulocyte lysate according to manufacturer's protocol (Promega). Mobility shift assays were carried out in a 15μl reaction mixture containing 20mM Hepes (pH 7.6), 200mM KCl, 2.5mM MgCl2, 10% glycerol, 2μg of poly(dI-dC)·poly(dI-dC), 0.5μg supercoiled plasmid, 10μg of BSA and a titrated amount of reticulocyte lysate or 4μg of HeLa cell nuclear extract (prepared as described (Harshman et al., 1988)). For supershift experiments, ascites fluid was included during this incubation. This mixture was incubated for 10 minutes on ice. Labeled probe (0.3ng) in 1μl of 20mM Tris pH 7.5, 1M NaCl, 1mM EDTA and unlabeled competitors were then added to the reaction, followed by a 10 minute incubation at room temperature. Probes were labeled and isolated as described above, and unlabeled competitors were single copy, double-strand oligonucleotides added at the indicated molar excess. Electrophoresis was performed on a 4% polyacrylamide gel (30: 0.8% acrylamide:bis) in 0.25X TBE and electrophoresed for 2hr at 10V/cm at room temperature. In supershift experiments, all electrophoresis conditions were the same, except that the gel composition was 80:1% acrylamide:bis.

Transient transfections

To express NRSF in transient transfection experiments, we inserted the λHZ4 cDNA into the EcoRI site of pcDNA3-ATG, a modified form of pcDNA3 (Invitrogen), a mammalian expression vector containing the cytomegalovirus enhancer and an oligonucleotide which provides a start codon in-frame with λHZ4 and a stop codon in all

three reading frames. Transient transfections were performed using the calcium phosphate precipitation method (Wigler et al., 1979). PC12 cells ($4x10^5$ cells) cultured in a 60mm dish in DMEM containing 10% fetal calf serum, penicillin and streptomycin were cotransfected with 10μg of total plasmid for 6 hr. Each cotransfection included 5μg of a reporter plasmid (pCAT3 or pCAT3-S36++), the expression plasmid (pCMV-λHZ4+) at the concentrations indicated, pcDNA3-ATG to control for non-specific vector effects, 2μg of pRSV-lacZ to normalize transfections, and pBluescript to bring the total plasmid up to 10μg. Cells were harvested 48hr after transfection and processed for CAT and β-galactosidase assays as previously described (Mori et al., 1990), except CAT assays were quantified using a Molecular Dynamics Phosphor Imager.

In situ hybridization

The morning of the day of detection of a vaginal plug was designated as embryonic day 0.5. Fixation, embedding, sectioning, preparation of digoxygenin-labeled cRNA probes and in situ hybridization with nonradioactive detection were performed as described (Birren et al., 1993). Both sense and antisense probes for NRSF were generated from a 1.5kb mouse cDNA containing plasmid (pM5) excised from a λEXlox phage using a Cre recombinase system (Novagen). The antisense SCG10 probe has been described elsewhere (Stein et al., 1988).

RNase protection assays

RNase protections were performed as previously described (Johnson et al., 1992) with minor modifications as indicated. The mouse NRSF riboprobe was created using T7 polymerase and a linearized subclone of the EcoRI-Eco47 III fragment from pM5 subcloned into the EcoRI and SmaI sites of pBluescript-KS. A rat β-actin riboprobe (gift of M-J. Fann and P. Patterson) was included in each reaction as a control for the amount and integrity of the RNA. Total cellular RNA was isolated using the acid phenol method (Chomczynski and Sacchi, 1987).

Oligonucleotides

The sequence of the top strand of the oligonucleotides used for library screening and EMSAs are given below. The upper case sequences represent actual genomic sequence, the lower case sequences were used for cloning purposes.

S36: agctGCAAAGCCATTTCAGCACCACGGAGAGTGCCTCTGC;

Na33: agcttATTGGGTTTCAGAACCACGGACAGCACCAGAGTa;

Syn: agcttCTGCCAGCTTCAGCACCGCGGACAGTGCCTTCGCa;

BDNF: agcttAGAGTCCATTCAGCACCTTGGACAGAGCCAGCGGa;

Ets: agcttGCGGAACGGAAGCGGAAACCGa

RESULTS

Isolation of NRSF-encoding cDNA clone

In previous work, NRSF binding activity was detected in nuclear extracts from non-neuronal cell lines, such as HeLa cells (Mori et al., 1992). Therefore, to isolate a cDNA clone encoding NRSF, we screened a HeLa cell λgt11 cDNA expression library according to the methods of Singh and Vinson (see Experimental Procedures) for in situ detection of filter-bound DNA binding proteins (Singh et al., 1988; Vinson et al., 1988). The DNA probes used for screening the library are referred to as S36, Na33, and Sm36 (see Experimental Procedures for sequences). S36 and Na33 are the NRSE elements present in the SCG10 and NaII channel genes, respectively. Both of these elements have previously been shown to be sufficient to confer silencing activity and are bound by NRSF. The Sm36 sequence contains two point mutations in the S36 sequence and has an approximately 100-fold lower affinity for NRSF (Mori et al., 1992). Approximately 2 million plaques were screened initially using a radiolabeled probe consisting of three tandemly arrayed copies of the Na33 sequence. Positive plaques from this screen were tested further for sequence specific DNA-binding by an additional screen with probes containing the S36 or the mutated NRSE, Sm36. One phage was identified, λH1, that

bound both the S36 and the Na33 probes but not the control Sm36 probe. As this DNA-binding pattern was similar to that of native NRSF, we chose to examine the protein encoded by λH1 in more detail.

DNA-binding specificity of recombinant NRSF

As an additional test of the authenticity of the cDNA clone, we compared the DNA-binding specificity of the protein encoded by λH1 to that of NRSF present in a HeLa cell nuclear extract using an electrophoretic mobility shift assay (EMSA). To produce recombinant λH1 protein, the phage insert was subcloned into a T7 expression vector (pRSET, see Experimental Procedures) that provided an in-frame start codon, and RNA synthesized from this plasmid was used to program an in vitro translation reaction. The results (Fig. 1) of the EMSA comparing the λH1-encoded protein (lane 1, large arrowhead to left of panel) and native NRSF (lane 9, small arrowhead to right of panel) showed that both proteins form complexes with the S36 probe. No complexes were formed by an in vitro translation reaction to which no RNA had been added (data not shown). The faster mobility of the λH1-encoded protein:DNA complex most likely reflects a difference in molecular weight between the fusion protein and the endogenous factor, as the λH1 cDNA does not encode the full-length protein (see below). The sequence specificity of these complexes was tested by competition experiments using unlabeled, double-stranded oligonucleotide binding sites. The SCG10 (S36) and the NaII channel genes (Na33) NRSEs showed similar ability to compete both the λH1-encoded and the native protein:DNA complexes (Fig. 1, compare lanes 2-5 and 10-13). These complexes, however, were only poorly competed by the mutated NRSE (Sm36, lanes 6,7 and 14,15), and no competition was seen with an oligonucleotide containing an Ets factor binding site (lanes 8 and 16) (Lamarco et al., 1991). The data suggest that the protein encoded by λH1 and native NRSF have similar DNA-binding specificities as measured in this assay.

Immunological relatedness of recombinant and native NRSF

Pilot experiments indicated that the abundance of NRSF in HeLa nuclear extracts was too low to permit purification of sufficient quantities to obtain amino acid sequence for comparison to the λH1-encoded protein. As an alternative strategy, we pursued an immunological approach to obtain independent evidence for a relationship between native and recombinant NRSF proteins. A mouse polyclonal antibody (see Experimental Procedures) was generated as an ascites fluid against a glutathione-S-transferase-λH1 fusion protein (GST-λH1) and was tested for its ability to bind and supershift native NRSF and λH1-encoded protein:DNA complexes in an EMSA. Figure 2 (lower panel; bracket, lanes 1-4) shows that this antibody was able to supershift a portion of the λH1-encoded protein:DNA complex. The supershifted complex was competed by unlabeled NRSE oligonucleotide (Fig. 2 lower panel, lane 5), further indicating that it contained the λH1-encoded protein. A control ascites made in a similar manner against an unrelated protein (Fig. 2 lower panel, lanes 6-8), as well as several other ascites made against irrelevant antigens (data not shown), was unable to supershift the complex.

The preceding results showed that the antibody generated against the GST-λH1 fusion protein can recognize specifically the λH1 portion of the fusion protein, a result confirmed by immunoprecipitation of [^{35}S] methionine-labeled in vitro translation products (data not shown). We next tested the anti-GST-λH1 antibody for its ability to supershift the complex formed by native NRSF. Figure 2 (upper panel; bracket, lanes 1-4) shows that the antibody can supershift a portion of native NRSF complex. No supershift was seen with the control ascites (lanes 6-8) nor with several other control ascites (data not shown). Furthermore, the supershift complex was competed by excess unlabeled NRSE (Fig. 2 upper panel, lane 5), indicating that it contained NRSF. To show that the anti-GST-λH1 ascites was not itself the source of the supershift complex, a reaction containing this ascites alone was performed, and no complex was detected (lane 10). Therefore, the anti-GST-λH1 antibody can specifically recognize at least a component of native NRSF. Our inability to obtain a quantitative supershift leaves open the possibility that HeLa nuclear

extracts contain multiple NRSE-binding proteins. Nevertheless, the antigenic similarity of the λH1-encoded protein and native NRSF provides further evidence that λH1 encodes NRSF, or at least an immunologically-related protein of similar DNA-binding specificity.

NRSF interacts with NRSEs in multiple neuron-specific genes

NRSF-encoding cDNA clones were identified by virtue of their ability to bind to two independently-characterized functional NRSEs, one in the SCG10 gene, the other in the NaII channel gene. To determine whether NRSF also interacts with NRSE-like sequences identified in other neuron-specific genes, we performed EMSAs using probes containing potential NRSEs from the synapsin I and brain-derived neurotrophic factor (BDNF) genes. In the case of synapsin I, the NRSE-like sequence has been shown to function as a silencer by cell transfection assays (Li et al., 1993). In the case of BDNF, the element was identified by sequence homology but has not yet been tested functionally (Timmusk et al., 1993). Although BDNF is expressed both in neurons and in non-neuronal cells, this expression is governed by two sets of promoters which are separated by 16 kb; one set of the promoters is specifically utilized in neurons (Timmusk et al., 1993). Native NRSF from HeLa cells yielded a specific complex of similar size using probes from all four genes (Figure 3, lanes 1-4). The complexes obtained with the SCG10 and NaII channel NRSEs appeared to have slightly different mobilities than the other two elements, suggesting the possible existence of multiple NRSE-binding proteins in HeLa cells. However, at least a portion of all four of these complexes could be supershifted by the anti-GST-λH1 ascites, and the SCG10 NRSE complex could be competed by oligonucleotides containing NRSEs from the other three genes (data not shown). Furthermore, all four probes also generated specific complexes with the λH1-encoded protein (Fig. 3, lanes 5-8). These data therefore indicate that both native and recombinant NRSF are able to interact with consensus NRSEs in multiple neuron-specific genes.

Characterization of additional NRSF cDNA clones

The foregoing data demonstrate that the λH1 cDNA contains the DNA-binding domain of NRSF. However, Northern blot analysis of polyA$^+$ RNA isolated from HeLa cells using the λH1 cDNA insert as a probe revealed an mRNA species of approximately 7.5 kb; in contrast the λH1insert is only 1.1 kb indicating that it represents a partial cDNA (data not shown). This cDNA is unlikely to contain the entire NRSF coding sequence, since gel renaturation experiments suggested an approximate molecular weight of 200 Kd for native NRSF, while the λH1-encoded protein has an apparent molecular weight of 60 Kd (data not shown).

In an attempt to isolate a full length cDNA for NRSF, multiple cDNA libraries were screened by hybridization with the λH1 clone (see Experimental Procedures). Although many cDNAs were isolated and characterized, none of the inserts exceeded approximately 2kb. Moreover, attempts to isolate additional cDNA sequence by the 5' RACE procedure were unsuccessful. Therefore, the two longest clones isolated from a HeLa λZAPII cDNA library were characterized further. The sequence of the longest clone, λHZ4 (2.04 kb), is shown in Figure 4. λHZ4 has an open reading frame throughout its length with no candidate initiating methionine and no stop codon, indicating that the cDNA does not contain the full protein coding sequence for NRSF. Conceptual translation of the DNA sequence revealed that it contains a cluster of eight zinc fingers of the C_2H_2 class with interfinger sequences which place NRSF in the GLI-Krüppel family of zinc finger proteins (Fig. 5A, B) (Ruppert et al., 1988; Schuh et al., 1986). C-terminal to the zinc fingers is a 174 amino acid domain rich in lysine (26%; 46/174) and serine/threonine (21%; 37/174; Fig. 5A). A database search using the BLAST program did not reveal any sequences identical to λHZ4, indicating that NRSF represents a novel zinc finger protein (Altschul et al., 1990). While these searches did reveal many zinc finger genes with similarity to NRSF, none of these sequences had the same configuration of zinc fingers nor any significant homology outside of the zinc finger domain. The database searches did uncover significant similarity to two different 'expressed sequence tags' or ESTs defined by random

sequencing of human cDNA libraries. One EST, present in a human brain cDNA library, had 86% identity with nucleotides 315-496 (182bp) of NRSF. The other, present in a human bone cDNA library, had 93% identity with nucleotides 1938-2040 (103bp) of NRSF. We believe that these sequences represent partial NRSF cDNAs, as the ESTs were sequenced on only one strand and may contain a significant number of sequencing errors.

Repression of transcription by NRSF

To determine if the longest NRSF partial cDNA encoded a protein with transcriptional repressing activity, we cotransfected PC12 cells with reporter plasmids and a mammalian expression plasmid containing the λHZ4 clone, pCMV-HZ4 (see Experimental Procedures). One reporter plasmid (pCAT3-S36++) contained two copies of the NRSE inserted upstream of the proximal SCG10 promoter which was fused to the bacterial chloramphenicol acetyltransferase (CAT) gene. The control reporter plasmid (pCAT3) contained only the proximal SCG10 promoter fusion . In transient co-transfection experiments with pCAT3-S36++ and increasing amounts of pCMV-HZ4, we observed that the reporter plasmid activity was repressed from 11 to 32 fold (Fig. 6A; Table I). In parallel transfections performed with pCAT3 as the reporter plasmid, only a modest decrease (1.5 fold at maximum pCMV-HZ4 concentration) in activity was seen with increasing amounts of pCMV-HZ4 (Fig. 6B; Table I). Furthermore, in additional control experiments performed with a reporter construct containing two copies of the mutated NRSE element (Sm36) (Mori et al., 1992), only minimal repression was seen in the presence of pCMV-HZ4 (data not shown). These results indicated that the λHZ4 clone contained at least a portion of the domain required for transcriptional repression and provides further evidence that we have isolated a cDNA encoding a functional NRSF.

NRSF is expressed in neural progenitors but not in neurons

Previous work indicated that NRSE-dependent silencing activity and NRSE-binding activity are present only in non-neuronal cell lines and are absent from cell lines of

neuronal origin (Kraner et al., 1992; Maue et al., 1990; Mori et al., 1992; Mori et al., 1990). The absence of these activities in neuronal cells could reflect a lack of NRSF gene expression; alternatively, NRSF might be expressed but be functionally inactive in neuronal cells. To distinguish between these possibilities, we first performed RNase protection assays on several rodent neuronal and non-neuronal cell lines, using a portion of a mouse NRSF cDNA (see Experimental Procedures) as probe. No NRSF transcripts were detectable in two rat neuronal cell lines, MAH and PC12 cells (Fig. 7, lanes 4 and 5; rNRSF). In contrast several rat cell lines of glial origin (RN22, JS-1, NCM-1, and C6) expressed NRSF mRNA (Fig. 7, lanes 6-9). In addition, two non-neural cell lines, Rat-1 fibroblasts (rNRSF, lane 3) and mouse C3H10T1/2 fibroblasts (mNRSF, lane 2), expressed similar levels of NRSF transcripts. (The size difference between NRSF protected products of the mouse and rat most likely reflects a species difference in the sequence of the target mRNA, resulting in incomplete protection of the mouse probe by the rat transcript.) The absence of NRSF mRNA in several clonal neuronal cell lines and its presence in several non-neuronal cell lines is consistent with its proposed role as a negative regulator of neuron-specific gene expression in non-neuronal cells. Furthermore, the data imply that the absence of NRSF activity in neuronal cells is not due to functional inactivation of NRSF, but simply to the lack of NRSF expression.

The preceding RNase protection assays indicated that NRSF transcripts could be detected in glial but not in neuronal cell lines. In many parts of the embryonic nervous system, neurons and glia derive from multipotent progenitor cells (McConnell, 1991; McKay, 1989; Sanes, 1989). To determine whether such progenitor cells also express NRSF, we performed in situ hybridization experiments on mouse embryos. In transverse sections of E12.5 mouse embryos, NRSF hybridization was detected in the ventricular zone of the neural tube (Fig. 8A, arrow), a region containing mitotically active multipotential progenitors of neurons and glia (Leber et al., 1990) which do not express SCG10 mRNA (compare Fig. 8B, arrow). In contrast, the adjacent marginal zone of the

neural tube which contains SCG10 positive neurons (Fig. 8B) is largely devoid of NRSF expression (Fig. 8A). A similar complementarity of NRSF and SCG10 expression in the neural tube was detected at E13.5 (Fig. 8 C, D; arrows), when the marginal zone has expanded. NRSF mRNA was also detected in the ventricular zone of the forebrain (Fig. 10B, arrowhead).

In the peripheral nervous system, NRSF mRNA was absent or expressed at low levels in sympathetic and dorsal root sensory ganglia (DRG) at E13.5 (Fig. 8C, small and large arrowheads) whereas these ganglia clearly expressed SCG10 mRNA (Fig. 8D, small and large arrowheads). At E12.5, the DRG appeared to express higher levels of NRSF mRNA than the marginal zone of the neural tube (Fig. 8A, arrowheads). This NRSF expression may derive from undifferentiated neural crest cells that are present in DRG at these early developmental stages. These data suggest that NRSF is not expressed by differentiated ($SCG10^+$) neurons in vivo, but is expressed by their undifferentiated precursors. The expression of NRSF in multipotent neural precursors but not in neurons supports the idea that the induction of neuronal differentiation involves, at least in part, a relief from NRSF-imposed repression of neuron-specific gene transcription.

Widespread expression of NRSF in non-neural tissues

Previous work suggested that the NRSE is required to repress SCG10 expression in many, if not all, non-neural tissues throughout development (Wuenschell et al., 1990). To determine whether this broad requirement for the NRSE element is reflected in a broad expression of NRSF, we examined its expression in non-neuronal tissues by RNase protection and in situ hybridization experiments. In E13.5 embryos, NRSF mRNA was detected at variable levels in all non-neural tissues examined (Fig. 9). The highest levels of expression are detected in lung and limbs (Fig. 9, lanes 8 and 9), whereas much lower levels were found in heart and liver (lanes 5 and 6), a result supported by in situ hybridization data (Fig. 10A, B; arrows). In situ hybridization additionally revealed NRSF mRNA expression in other non-neural tissues such as the adrenal gland, aorta, genital

tubercle, gut, kidney, lung, ovaries, pancreas, parathyroid gland, skeletal muscle, testes, thymus, tongue, and umbilical cord (Fig. 10A, B and data not shown). NRSF mRNA was also detected in adult tissues, where its levels were more uniform than in embryonic tissues and comparable to that detected in the clonal cell line C3H10T1/2 (Fig. 9; 10T, lane 16), suggesting that a large percentage of the cells in these non-neural tissues express NRSF. As expected, all non-neural tissues contained higher levels of NRSF mRNA than brain (Fig. 9, lanes 4 and 10). The low level of expression in brain is likely to reflect expression in glial cells, as suggested by the detection of NRSF mRNA in several glial cell lines (Fig. 7). Taken together, these data indicate that NRSF is expressed by most non-neural tissues. This expression pattern is consistent with a role for NRSF as a near-global negative regulator of neuron-specific gene expression.

DISCUSSION

Although the molecular basis of cell type-specific transcriptional regulation has been intensively studied in many non-neuronal lineages, remarkably little is known about this process in the nervous system. Here we describe the isolation and characterization of cDNAs encoding a novel, zinc finger containing polypeptide that has the properties of a neural restrictive silencer factor (NRSF): it binds specifically to SCG10 and NaII channel silencer elements (NRSEs) in vitro and can repress transcription in an NRSE-dependent manner in vivo. Furthermore, we show that both the recombinant and native NRSFs can also bind to NRSEs present in the synapsin I and BDNF genes, suggesting that one polypeptide may negatively regulate at least four different neuron-specific promoters. The distribution of NRSF transcripts in the mouse indicates widespread expression in most non-neural tissues at several developmental stages, suggesting that NRSF is a near-global negative regulator of neuron-specific gene expression. In contrast, the expression of NRSF is low or undetectable in several neuronal populations from both the central and peripheral nervous systems. Taken together these data suggest that NRSF constitutes one of the first examples of a transcriptional regulator that functions as a major determinant of

cell-type specificity for multiple target genes by mediating sequence-specific repression in a virtually global manner.

NRSF is a novel zinc finger protein

Four lines of evidence indicate that the cDNA clones we isolated encode an authentic NRSF. First, the protein encoded by λH1 (the original cDNA recovered from the HeLa λgt11 library) and native NRSF have similar DNA binding specificities as measured in an EMSA. Second, antibodies generated against a GST-λH1 fusion protein are able to interact with native NRSF in an EMSA. Third, the presence of the putative NRSF mRNA detected in cell lines parallels both the silencing activity of the NRSE (as detected by transient transfection assays) and the in vitro DNA binding of native NRSF. Fourth, the longest NRSF cDNA clone (λHZ4) can repress NRSE-containing reporter constructs when cotransfected into PC12 cells. Therefore, while we cannot exclude the existence of other NRSF-like proteins, these data strongly suggest that we have isolated a cDNA encoding a functional fragment of NRSF.

The predicted amino acid sequence of NRSF has several notable features. First and foremost is the cluster of eight zinc fingers, all of which are members of the C_2H_2 class (Bairoch, 1992). The large number of zinc fingers is consistent with the large size of the NRSE (21-28bp) and suggests that most or all of the fingers are required for high affinity binding (Pavletich and Pabo, 1993; Pavletich and Pabo, 1991). The alignment of the eight zinc fingers and the interfinger sequences showed two features that differ from the canonical GLI-Krüppel zinc finger motif (Ruppert et al., 1988). First, the NRSF zinc fingers lack the phenylalanine or tyrosine residues that are usually found at position 10 in canonical zinc fingers. Instead, all eight fingers have a tyrosine at the 8th position. Second, six of the eight fingers lack the leucine typically found at position 16; rather, five of these fingers have an aromatic residue at this position (see Fig 5B). These differences suggest that the NRSF zinc fingers may have a slightly different secondary structure, and therefore a different set of DNA-binding characteristics than the canonical zinc finger. This

idea is consistent with structural studies of another non-canonical zinc finger present in ZFY, a mammalian Y-linked gene (Kochoyan et al., 1991). Whether this potential structural difference is important for NRSE binding by NRSF remains to be determined.

As the portion of NRSF we isolated encodes a protein with significant transcriptional repression activity, we compared the sequences N- and C-terminal to the zinc finger domain to those of known eukaryotic repressors. Although no extensive sequence similarities were found, the high lysine content of the C-terminal domain is reminiscent of a basic repression domain found in v-erbA (Baniahmad et al., 1992) (a repressor of the chick lysozyme gene (Baniahmad et al., 1990)) and with several artificial basic sequences found to repress transcription in yeast (Saha et al., 1993). Such similarities suggest that this basic region could be necessary for the silencing activity of NRSF. Given its basicity, this domain could repress transcription by blocking the interaction between acidic activation regions of enhancer-binding proteins and the general transcriptional machinery. Alternatively, the basic region could bind DNA and induce a bend that may be necessary for NRSF's repression activity. Such a mechanism has been postulated to explain YY1-mediated repression (Natesan and Gilman, 1993). The availability of a functional assay for NRSF provides the opportunity to delineate its functional domains as well as its mechanism of action.

NRSF is a near-global mediator of neuron-specific gene repression

In situ hybridization and RNase protection experiments indicate that NRSF transcripts are present in most non-neuronal cell types. These data are consistent with NRSF's proposed role as a near-global negative regulator of neuron-specific gene expression in non-neural tissues (Mori et al., 1990; Wuenschell et al., 1990). NRSF message, however, is absent from some non-neural tissues, such as embryonic heart and liver, indicating that it is not required in all non-neuronal cell types. Nevertheless, these cell types may acquire a dependence on NRSF function during later development, as adult heart

and liver express levels of NRSF mRNA comparable to other tissues and cell lines. In contrast to this developmentally-regulated pattern of expression, NRSF expression persists throughout development in other non-neural tissues, beginning at early embryonic stages and continuing into adulthood. This persistent expression of NRSF implies that it is necessary for the maintenance of neuronal gene silencing, rather than simply for its initiation. This widespread, continuous silencing by NRSF may involve specific protein-protein interactions with positive-acting transcription factors, or rather assembly of NRSE-containing regions of the genome into transcriptionally-inactive chromatin as suggested by the analysis of DNaseI hypersensitivity in SCG10 transgenes (Vandenbergh et al., 1989) (see below). The latter mechanism has been suggested to explain the action of certain silencers in yeast (for review, see (Rivier and Pillus, 1994)).

NRSF is a repressor of multiple neuron-specific genes

Functional NRSEs have been identified in three neuron-specific genes: SCG10, NaII channel, and synapsin I (Kraner et al., 1992; Li et al., 1993; Mori et al., 1992). We show that both native and recombinant NRSF can bind to all three of these NRSEs. In addition, NRSF binds to a consensus NRSE present in the BDNF gene. (While BDNF is expressed in both neurons and in selected non-neuronal cell types, its non-neuronal expression is dependent on a separate promoter which is 16 kb away from the neuron-specific promoter associated with the NRSE (Timmusk et al., 1993).) As our data for SCG10 and the NaII channel indicate a strong correlation between DNA binding by NRSF and susceptibility to NRSF-mediated silencing, we conclude that NRSF may repress at least four neuron-specific promoters. This establishes NRSF as a vertebrate silencer factor that coordinately regulates multiple lineage-specific genes. Such coordinate cell type-specific silencing suggests analogies to MATα2 in yeast, which coordinates repression of multiple **a**-specific genes in α cells (Herskowitz, 1989), and to the *Drosophila* Polycomb genes, which negatively regulate several homeotic genes (Paro, 1990). The identification

of NRSF suggests that coordinate repression of cell-type specific genes may be a more common mode of gene regulation in vertebrates than previously recognized.

Role of NRSF in neurogenesis

As a first step towards determining the role of NRSF in neurogenesis, we examined its expression pattern during embryonic development by in situ hybridization. These data indicated that NRSF is undetectable or expressed at low levels in neurons, but is expressed in regions of the embryonic CNS that contain neuronal precursors. Consistent with this, we have detected abundant expression of NRSF mRNA in undifferentiated P19 cells, a murine embryonal carcinoma cell line that can differentiate into neurons when cultured with retinoic acid (unpublished data). The presence of NRSF in neuronal precursors suggests that relief from NRSF-imposed repression could be important in either the initial selection or the execution of a neuronal program of differentiation. In either case, the absence of NRSF mRNA in neurons indicates that this derepression most likely occurs by an extinction of NRSF expression, rather than by its functional inactivation. Such a mechanism implies that neuronal precursors are actively prevented from differentiating until released from this repression by a signal that extinguishes NRSF expression. This idea has intriguing parallels to mechanisms recently shown to underlie neural induction in *Xenopus* embryos. In that system ectodermal cells are apparently actively prevented from adopting a neural fate by activin, and can undergo neural induction only after a relief from this repression by follistatin, an inhibitor of activin (Hemmati-Brivanlou et al., 1994; Hemmati-Brivanlou and Melton, 1994). It remains to be determined whether the action of follistatin is in any way related to the activity or expression of NRSF. In any case, the identification of NRSF provides an opportunity to further understand the control of an apparently central event in neurogenesis.

ACKNOWLEDGEMENTS

We thank Tom Kadesch and Paula Henthorn for advice and for sending cDNA libraries, Nozomu Mori for making the HeLa nuclear extract, Steven Padilla for technical assistance, Susan Ou for help in the preparation of antibodies, Sean Tavtigian and Barbara Wold for providing cDNA libraries, Liching Lo for advice on antibody production, and, along with Andrew Groves, for help with the in situ hybridization procedure, Ming-Ji Fann and Paul Patterson for providing actin cDNA probes, Lisa Banner for help with RNase protections, and Barbara Wold, Lukas Sommer and Scott Bultman for their comments on the manuscript. This work was supported by NIH grant NS23476. D.J.A. is an Associate Investigator of the Howard Hughes Medical Institute.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology *215*, 403-410.

Anderson, D. J., and Axel, R. (1985). Molecular probes for the development and plasticity of neural crest derivatives. Cell *42*, 649-662.

Bairoch, A. (1992). PROSITE: a dictionary of sites and patterns in proteins. Nucleic Acids Research *20*, 2013-2018.

Baniahmad, A., Kohne, A. C., and Renkawitz, R. (1992). A transferable silencing domain is present in the thyroid hormone receptor, in the v-erbA oncogene product and in the retinoic acid receptor. EMBO J. *11*, 1015-1023.

Baniahmad, A., Steiner, C., Kohne, A. C., and Renkawitz, R. (1990). Modular Structure of a Chicken Lysozyme Silencer: Involvement of an Unusual Thyroid Hormone Receptor Binding Site. Cell *61*, 505-514.

Birren, S. J., and Anderson, D. J. (1990). A v-myc-immortalized sympathoadrenal progenitor cell line in which neuronal differentiation is initiated by FGF but not NGF. Neuron *4*, 189-201.

Birren, S. J., Lo, L. C., and Anderson, D. J. (1993). Sympathetic neurons undergo a developmental switch in trophic dependence. Development *119*, 597-610.

Brand, A. H., Breeden, L., Abraham, J., Sternglanz, R., and Nasmyth, K. (1985). Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. Cell *41*, 41-48.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. *162*, 156-159.

Corcoran, L. M., Karvelas, M., Nossal, G. J. V., Ye, Z. S., Jacks, T., and Baltimore, D. (1993). Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival. Genes and Development *7*, 570-582.

Devereux, J. D., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Research *12*, 387-395.

Greene, L. A., and Tischler, A. S. (1976). Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. Proc. Natl. Acad. Sci. USA *73*, 2424-2428.

Gutman, P. D., Cao, S. X., Dave, H. P. G., Mittelman, M., and Schechter, A. N. (1992). Binding of erythroid and nonerythroid nuclear proteins to the silencer of the human epsilon-globin-encoding gene. Gene *110*, 197-203.

Harshman, K. D., Moye-Rowley, W. S., and Parker, C. S. (1988). Transcriptional activation by the SV40 AP-1 recognition element in yeast is mediated by a factor similar to AP-1 that is distinct from GCN4. Cell *53*, 321-330.

He, X., and Rosenfeld, M. G. (1991). Mechanisms of complex transcriptional regulation: implications for brain development. Neuron *7*, 183-196.

Hemmati-Brivanlou, A., Kelly, O. G., and Melton, D. A. (1994). Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. Cell *77*, 283-295.

Hemmati-Brivanlou, A., and Melton, D. A. (1994). Inhibition of activin receptor signaling promotes neuralization in Xenopus. Cell *77*, 273-281.

Herskowitz, I. (1989). A regulatory hierarchy for cell specialization, in yeast. Nature *342*, 749-757.

Johnson, J. E., Zimmerman, K., Saito, T., and Anderson, D. J. (1992). Induction and repression of mammalian achaete-scute homolog (MASH) gene expression during neuronal differentiation of P19 embryonal carcinoma cells. Development *114*, 75-87.

Johnson, P. F., and McKnight, S. L. (1989). Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. *58*, 799-839.

Kimura, H., Fischer, W. H., and Schubert, D. (1990). Structure, expression and function of a schwannoma-derived growth factor. Nature *348*, 257-260.

Kochoyan, M., Keutmann, H. T., and Weiss, M. A. (1991). Architectural rules of the zinc-finger motif: Comparative two-dimensional NMR studies of native and "aromatic-swap" domains define a "weakly polar switch." Proceedings of the National Academy of Science (USA) *88*, 8455-8459.

Kraner, S. D., Chong, J. A., Tsay, H. J., and Mandel, G. (1992). Silencing the type II sodium channel gene: a model for neural-specific gene regulation. Neuron *9*, 37-44.

Kumar, S., Huber, J., Pena, L. A., Perez-Polo, J. R., Werrbach-Perez, K., and de Vellis, J. (1990). Characterization of functional nerve growth factor receptor in a CNS glial cell line: Monoclonal antibody 217c recognizes the nerve growth factor-receptor on C6 glioma. J. Neurosci. Res. *27*, 408-417.

Lamarco, K., Thompson, C. C., Byers, B. P., Walton, E. M., and McKnight, S. L. (1991). Identification of ETS-related and Notch-related subunits in GA binding protein. Science *253*, 789-792.

Leber, S. M., Breedlove, S. M., and Sanes, J. R. (1990). Lineage, arrangement, and death of clonally related motorneurons in chick spinal cord. J. Neurosci. *10*, 2451-2462.

Li, L., Suzuki, T., Mori, N., and Greengard, P. (1993). Identification of a functional silencer element involved in neuron-specific expression of the synapsin I gene. Proc. Natl. Acad. Sci. U.S.A. *90*, 1460-1464.

Li, S., Crenshaw, E. B. I., Rawson, E. J., Simmons, D. M., Swanson, L. W., and Rosenfeld, M. G. (1990). Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene *pit*-1. Nature *347*, 528-533.

Lo, L.-C., Birren, S. J., and Anderson, D. J. (1990). V-*myc* immortalization of early rat neural crest cells yields a clonal cell line which generates both glial and adrenergic progenitor cells. Devel. Biol. *145*, 139-153.

Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987). Regulation of inducible and tissue-specific gene expression. Science *236*, 1237-1245.

Maue, R. A., Kraner, S. D., Goodman, R. H., and Mandel, G. (1990). Neuron-specific expression of the rat brain type II sodium channel gene is directed by upstream regulatory elements. Neuron *4*, 223-231.

McConnell, S. K. (1991). The generation of neuronal diversity in the central nervous system. Ann. Rev. Neurosci. *14*, 269-300.

McKay, R. D. G. (1989). The origins of cellular diversity in the mammalian central nervous system. Cell *58*, 815-821.

Mitchell, P., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371-378.

Mori, N., Schoenherr, C., Vandenbergh, D. J., and Anderson, D. J. (1992). A common silencer element in the SCG10 and type II Na^+ channel genes binds a factor present in nonneuronal cells but not in neuronal cells. Neuron *9*, 1-10.

Mori, N., Stein, R., Sigmund, O., and Anderson, D. J. (1990). A cell type-preferred silencer element that controls the neural-specific expression of the SCG10 gene. Neuron *4*, 583-594.

Natesan, S., and Gilman, M. Z. (1993). DNA bending and orientation-dependent function of YY1 in the c-fos promoter. Genes & Dev. *7*, 2497-2509.

Ou, S. K. H., Hwang, J. M. C., and Patterson, P. H. (1993). A modified method for obtaining large amounts of high-titer polyclonal ascites-fluid. Journal of Immunological Methods *165*, 75-80.

Paro, R. (1990). Imprinting a determined state into the chromatin of Drosophila. Trends in Genetics *6*, 416-421.

Pavletich, N. P., and Pabo, C. O. (1993). Crystal-structure of a 5-finger GLI-DNA complex - new perspectives on zinc fingers. Science *261*, 1701-1707.

Pavletich, N. P., and Pabo, C. O. (1991). Zinc finger DNA recognition - crystal-structure of a zif268-DNA complex at 2.1A. Science *252*, 809-817.

Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., Dagati, V., Orkin, S. H., and Costantini, F. (1991). Erythroid-differentiation in chimeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. Nature *349*, 257-260.

Pfeiffer, S. E., Betschart, B., Cook, J., Mancini, P. E., and Morris, R. J. (1978). Glial cell lines. In Glial cell lines, S. Federoff and L. Hertz, eds. (New York: Academic Press), pp. 287-346.

Rivier, D. H., and Pillus, L. (1994). Silencing speaks up. Cell *76*, 963-966.

Ruppert, J. M., Kinzler, K. W., Wong, A. J., Bigner, S. H., Kao, F. T., Law, M. L., Seuanez, H. N., O'Brian, S. J., and Vogelstein, B. (1988). The GLI-Kruppel family of human genes. Molecular and Cellular Biology *8*, 3104-3113.

Saha, S., Brickman, J. M., Lehming, N., and Ptashne, M. (1993). New eukaryotic transcriptional repressors. Nature *363*, 648-652.

Sanes, J. R. (1989). Analysing cell lineage with a recombinant retrovirus. Trends Neurosci. *12*, 21-28.

Sawada, S., Scarborough, J. D., Killeen, N., and Littman, D. R. (1994). A lineage-specific transcriptional silencer regulates CD4 gene expression during T lymphocyte development. Cell *77*, 917-929.

Schuh, R., Aicher, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Siefert, E., Nauber, U., Schroder, C., and Jackle, H. (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by *Kruppel* , a Drosophila segmentation gene. Cell *47*, 1025-1032.

Singh, H., LeBowitz, J. H., Baldwin Jr., A. S., and Sharp, P. A. (1988). Molecular cloning of an enhancer binding protein: isolation by screening of an expression library with a recognition site DNA. Cell *52*, 415-423.

Smith, D. B., and Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene *67*, 31-40.

Stamatoyannopoulos, G., Josephson, B., Zhang, J. W., and Li, Q. L. (1993). Developmental regulation of human gamma-globin genes in transgenic mice. Molecular and Cellular Biology *13*, 7636-7644.

Stein, R., Mori, N., Matthews, K., Lo, L.-C., and Anderson, D. J. (1988). The NGF-inducible SCG10 mRNA encodes a novel membrane-bound protein present in growth cones and abundant in developing neurons. Neuron *1*, 463-476.

Timmusk, T., Palm, K., Metsis, M., Reintam, T., Paalme, V., Saarma, M., and Persson, H. (1993). Multiple promoters direct tissue-specific expression of the rat BDNF gene. Neuron *10*, 475-489.

Vandenbergh, D. J., Wuenschell, C. W., Mori, N., and Anderson, D. J. (1989). Chromatin structure as a molecular marker of cell lineage and developmental potential in neural crest-derived chromaffin cells. Neuron *3*, 507-518.

Vinson, C. R., LaMarco, K. L., Johnson, P. F., Landschulz, W. H., and McKnight, S. L. (1988). In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes and Development *2*, 801-806.

Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G., and Chasin, L. (1979). DNA-mediated transfer of adenine phosphoribosyltransferase locus into mammalian cells. Proc. Natl. Acad. Sci. USA *76*, 1373-1376.

Winoto, A., and Baltimore, D. (1989). αβ lineage-specific expression of the α T cell receptor gene by nearby silencers. Cell *59*, 649-665.

Wuenschell, C. W., Mori, N., and Anderson, D. J. (1990). Analysis of SCG10 gene expression in transgenic mice reveals that neural specificity is achieved through selective derepression. Neuron *4*, 595-602.

Table I. Transcriptional Repression by λHZ4[a]

Reporter Plasmid	pCMV-HZ4	Percent CAT activity	Fold repression
pCAT3-S36++	0 μg	100	-
	1	8.3±0.6	11
	4	3.1±0.3	32
pCAT3	0	100	-
	1	77±0.8	1.3
	4	67±3.8	1.5

[a]PC12 cells were cotransfected with reporter plasmids and an expression plasmid containing λHZ4. The pCAT3 reporter plasmid consists of the SCG10 proximal region fused to the bacterial CAT enzyme; pCAT3-S36++ consists of pCAT3 with two tandem copies of the S36 NRSE inserted upstream of the SCG10 sequences. The NRSF expression plasmid (pCMV-HZ4) is derived from pCMV-ATG, a modified version of pcDNA3 (Invitrogen) that provides an initiating methionine and a stop codon for the λHZ4 cDNA. To control for non-specific promoter effects, each cotransfection is performed with a constant molar amount of expression plasmid consisting of differing amounts of pCMV-HZ4 and pCMV-ATG. An RSV-LacZ plasmid was included in all transfections to normalize for transfection efficiency. The activity of each reporter plasmid was normalized to 100% to compare the relative level of repression of each construct. The numbers represent the mean ± SD of two independent experiments performed in duplicate.

Figure 1. λH1 encoded protein has the same sequence specificity of DNA binding as native NRSF. Electrophoretic mobility shift assays were performed using a HeLa cell nuclear extract or the products of a rabbit reticulocyte lysate in vitro translation reaction programmed with RNA transcribed from a λH1 fusion construct. The probe was a radiolabeled restriction fragment containing two tandem copies of S36. Competitors used were the S36, Na33, and Sm36 oligonucleotides (see Experimental Procedures) and an oligonucleotide containing an Ets factor binding site (Ets) (Lamarco et al., 1991). The large arrowhead marks the λH1-encoded protein:DNA complex (lane 1), the small arrowhead marks the NRSF:DNA complex (lane 9).

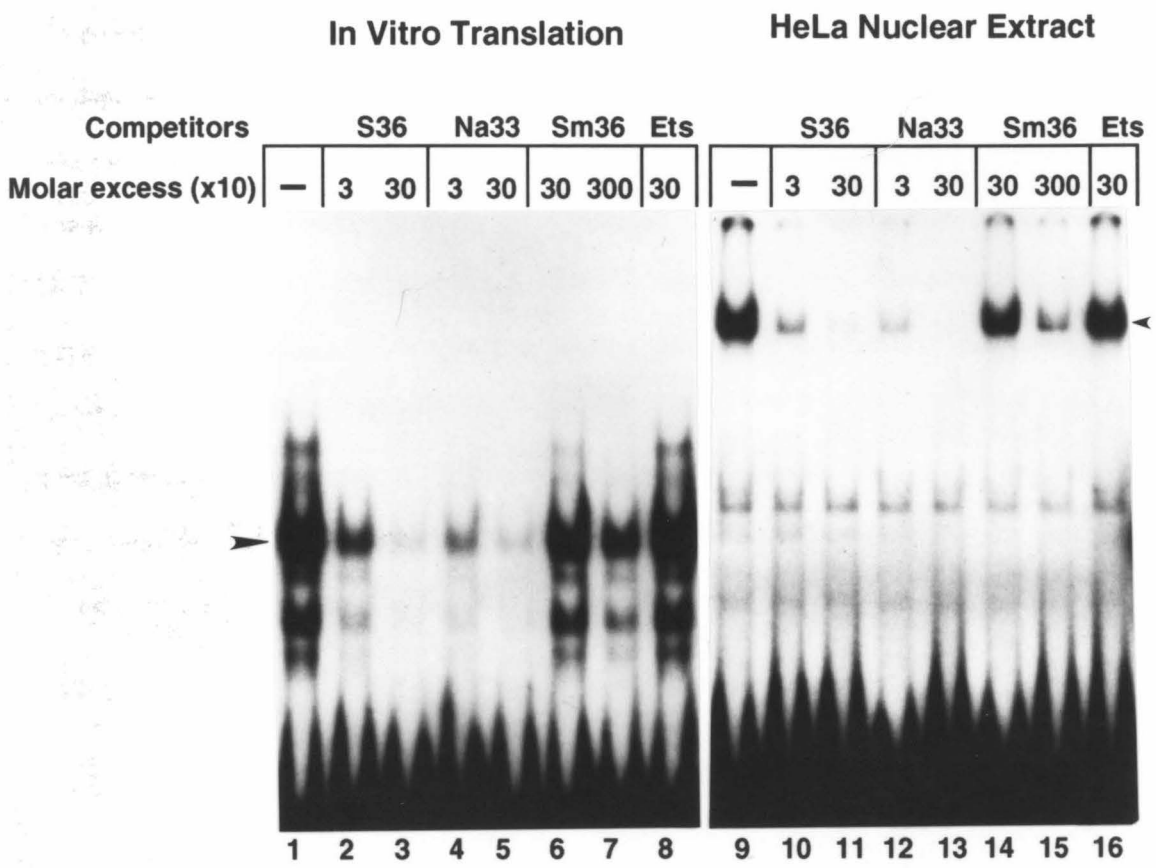
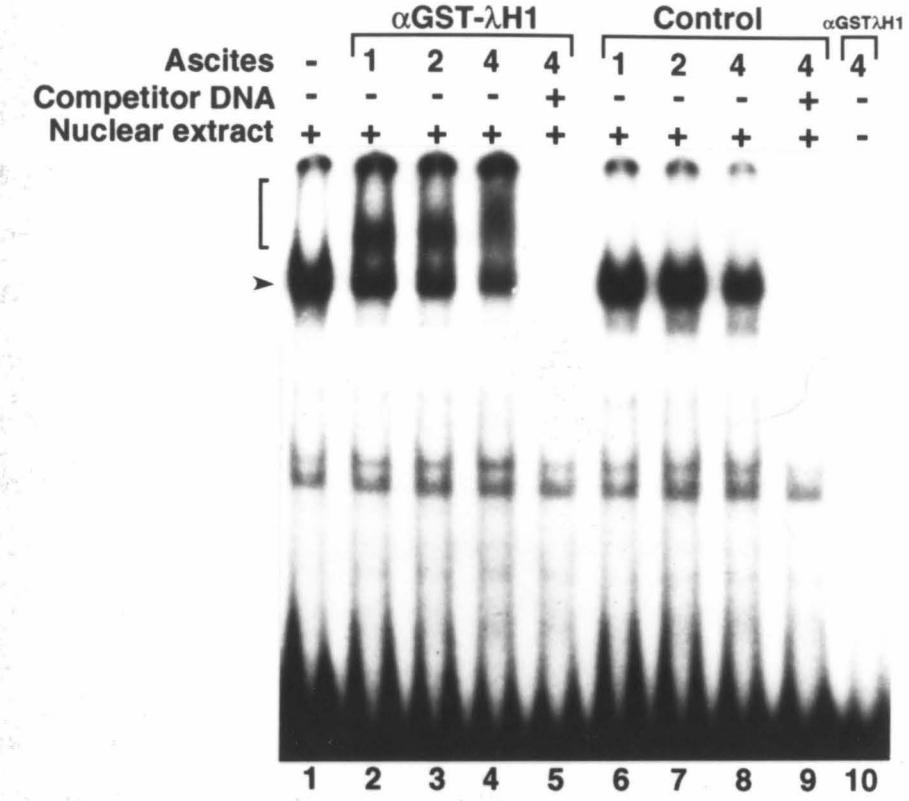

In Vitro Translation
HeLa Nuclear Extract
Competitors
S36
Na33
Sm36
Ets
Molar excess (x10)
—
3
30
3
30
30
300
30
—
3
30
3
30
30
300
30
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Figure 2. Antibodies against GST-λH1 recognize the native NRSF:DNA complex. Upper panel) The indicated amounts (in μl) of αGST-λH1 ascites or control ascites were added to a mobility shift reaction containing HeLa nuclear extract. The reactions were performed as in Figure 1 except that the acrylamide gel used for analysis had an 80:1 acrylamide to bis ratio instead of 30:0.8. The competitor was the S36 oligonucleotide present at 300 fold molar excess. The bracket indicates the supershifted NRSF:DNA complex, and the small arrowhead marks the NRSF:DNA complex. Lower panel) A mobility shift reaction using a rabbit reticulocyte reaction programmed with λH1 encoding RNA. The mobility shift reactions were performed and analyzed as in the upper panel. The bracket indicates the supershifted λH1-encoded protein:DNA complex, and the large arrowhead marks the λH1-encoded protein:DNA complex. Attempts to obtain a quantitative supershift using higher concentrations of antibody were precluded by the inhibition of DNA binding that occurred when the amount of ascites in the EMSA was increased.


αGST-λH1
Control
αGSTλH1
Ascites - 1 2 4 4 1 2 4 4 4
Competitor DNA - - - - + - - - + -
Nuclear extract + + + + + + + + + -
1 2 3 4 5 6 7 8 9 10

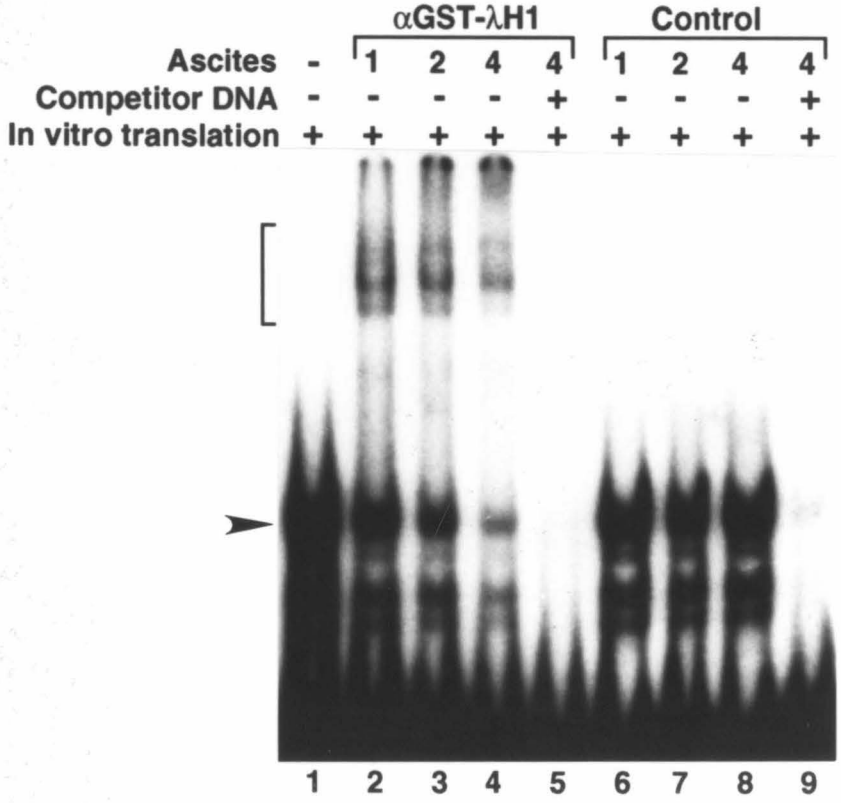

αGST-λH1
Control
Ascites - 1 2 4 4 1 2 4 4
Competitor DNA - - - - + - - - +
In vitro translation + + + + + + + + +
1 2 3 4 5 6 7 8 9

Figure 3. Native and recombinant NRSF recognize NRSEs in four different neuron-specific genes. Electrophoretic mobility shift assays were performed using either nuclear extract from HeLa cells (lanes 1-4), to reveal the activity of native NRSF, or using in vitro synthesized NRSF encoded by the λH1 cDNA (lanes 5-8). The labeled probes consisted of restriction fragments containing NRSEs (see Experimental Procedures) derived from the rat SCG10 gene (SCG10, lanes 1 and 5); the rat type II sodium channel gene (NaCh, lanes 2 and 6); the human synapsin I gene (Syn, lanes 3 and 7) or the rat brain-derived neurotrophic factor gene (BDNF, lanes 4 and 8). The large arrowhead indicates the specific complex obtained with recombinant NRSF; small arrowhead that obtained with native NRSF. Note that the complexes obtained with all four probes are of similar sizes. The complexes obtained using HeLa extracts were partially supershifted with antibody to recombinant NRSF (cf. Fig. 2) (data not shown).

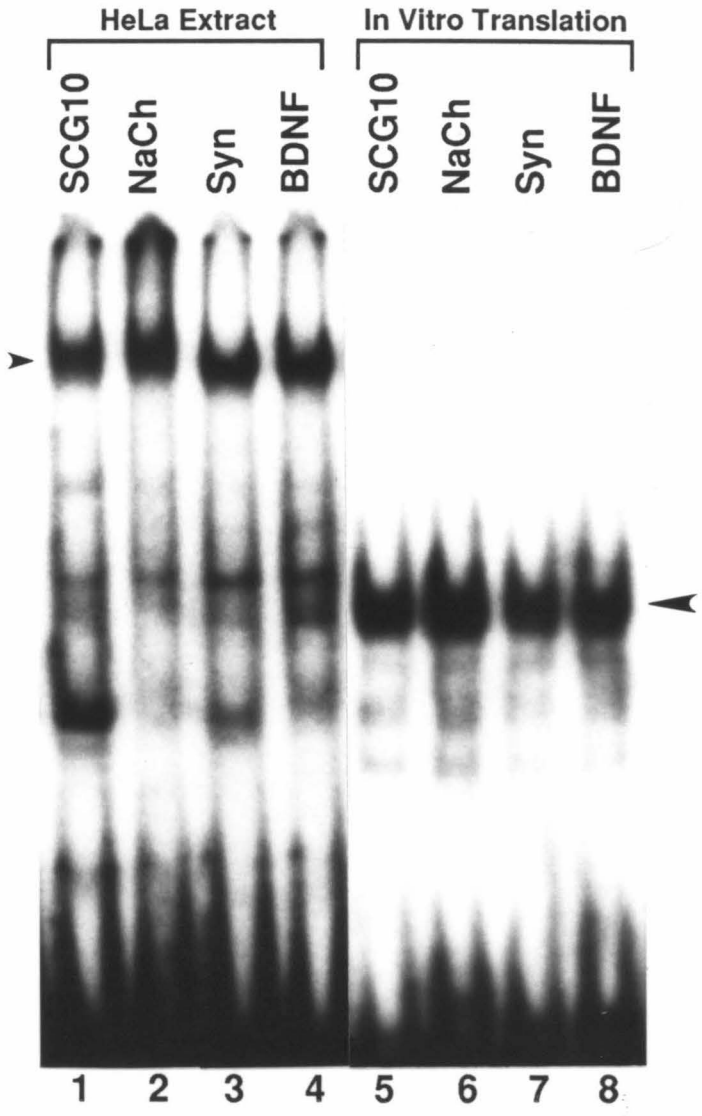

HeLa Extract
In Vitro Translation
SCG10
NaCh
Syn
BDNF
SCG10
NaCh
Syn
BDNF
1
2
3
4
5
6
7
8

Figure 4. Sequence analysis of NRSF. A) Nucleotide and predicted amino acid sequence of a partial cDNA (λHZ4) for human NRSF. The nucleotide sequence is numbered in standard type, and the amino acid sequence is in italics. The eight zinc fingers are underlined.

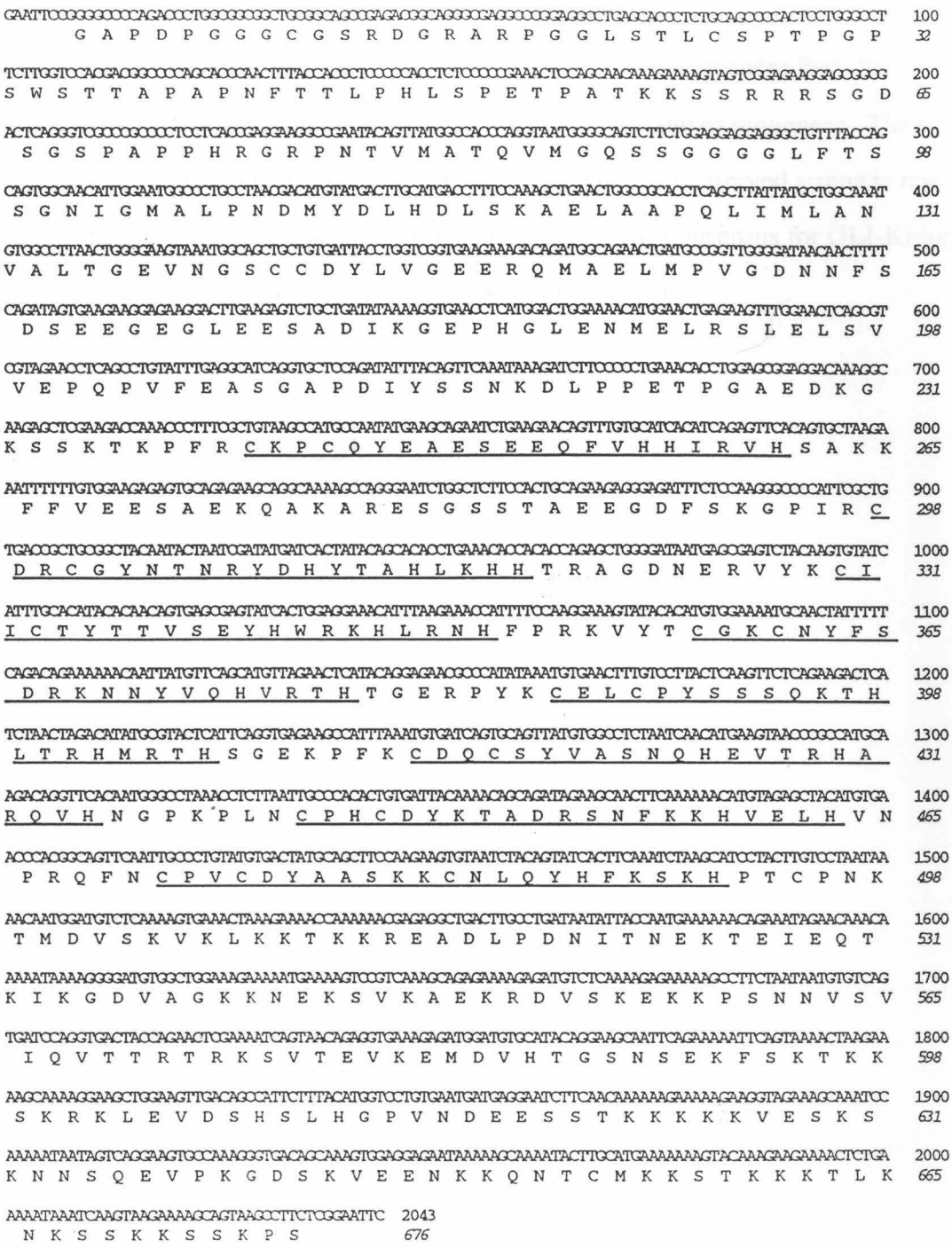
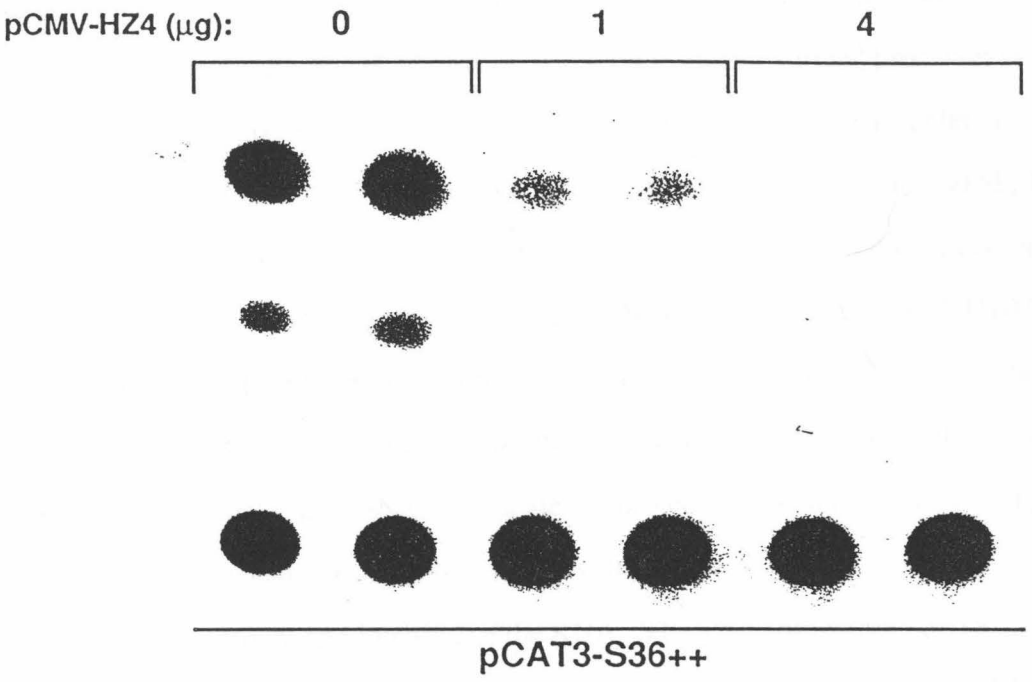


Figure 5. (A) Schematic diagram of the predicted amino acid sequences from the λHZ4 NRSF cDNA. (B) Alignment of NRSF zinc finger and interfinger sequences. The eight zinc fingers of human NRSF were aligned beginning with the conserved aromatic residue and including the interfinger sequences of fingers z2-7. The consensus for GLI-Krüppel zinc fingers and interfinger sequences is shown for comparison.

Figure 6. Repression of transcription by recombinant NRSF. (A) A representative autoradiogram of CAT enzymatic assays from cotransfection experiments in which increasing amounts of an expression plasmid (pCMV-HZ4) encoding a partial NRSF cDNA (clone λHZ4; see Fig. 5A) were cotransfected into PC12 cells together with a CAT reporter plasmid containing two tandem SCG10 NRSEs (pCAT3-S36++). (B) A similar experiment as in (A) except the CAT reporter plasmid (pCAT3) lacked NRSEs. See also Table I for quantification and further methodological details.

A.



B.

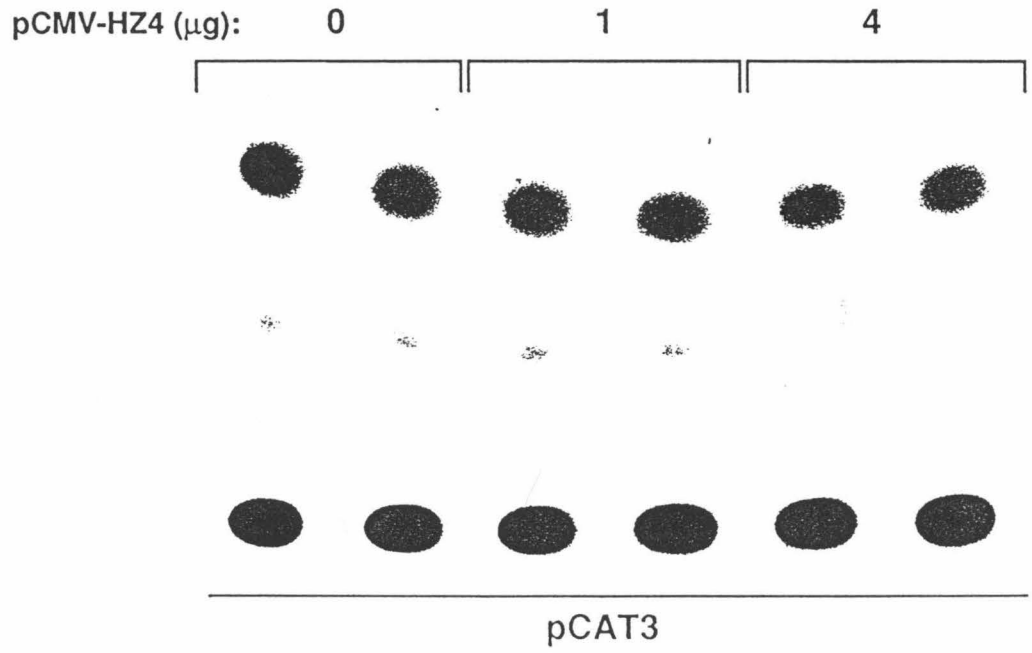


Figure 7. Analysis of NRSF message in neuronal and non-neuronal cell lines. RNase protections assays were performed on 10μg of total RNA form various cell lines. The two neuronal cell lines were MAH, an immortalized rat sympathoadrenal precursor (Birren and Anderson, 1990), and PC12, a rat pheochromocytoma (Greene and Tischler, 1976). The non-neuronal cell lines were: RN22 and JS-1, rat schwannomas (Kimura et al., 1990; Pfeiffer et al., 1978); NCM-1, an immortalized rat Schwann cell precursor (Lo et al., 1990); C6, a rat CNS glioma (Kumar et al., 1990); and Rat1 and mouse C3H10T1/2 (10T), embryonic fibroblast lines. A reaction containing yeast tRNA (tRNA) alone was performed as a negative control. The probes were derived from mouse NRSF and rat β-actin cDNAs. rNRSF and mNRSF indicate the protected products obtained using RNA from rat or mouse cell lines, respectively. β-actin probe was added to each reaction as control for the amount and integrity of the RNA. The autoradiographic exposure for the actin protected products was shorter than for NRSF. In this experiment, the RNase digestion was performed with RNase T1 only.

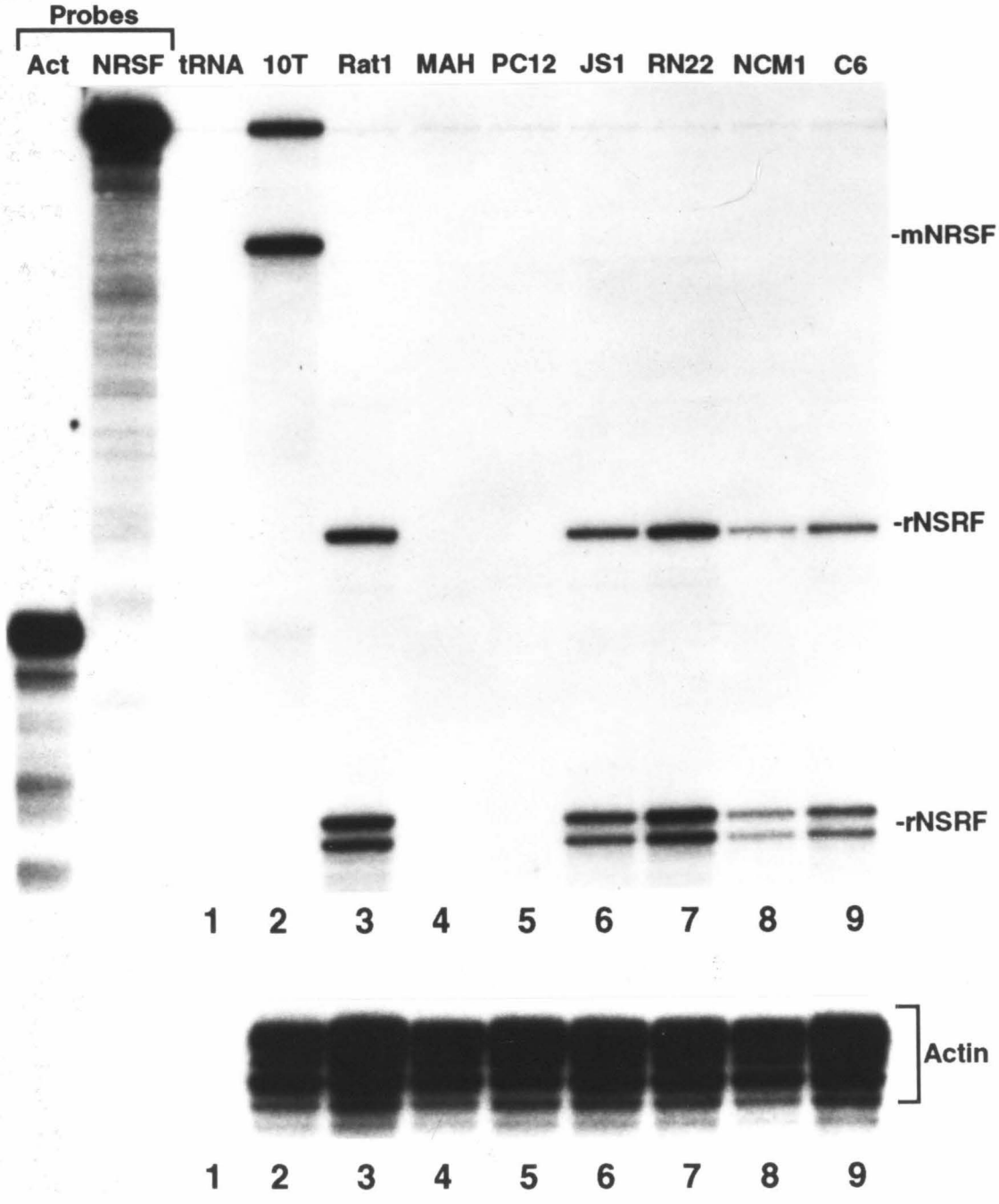

Probes
Act
NRSF
tRNA
10T
Rat1
MAH
PC12
JS1
RN22
NCM1
C6
-mNRSF
-rNSRF
-rNSRF
1 2 3 4 5 6 7 8 9
Actin
1 2 3 4 5 6 7 8 9

Figure 8. Comparison of NRSF and SCG10 mRNA expression by in situ hybridization. Adjacent transverse sections of E12.5 (A,B) and E13.5 (C,D) mouse embryos were hybridized with NRSF (A,C) or SCG10 (B,D) antisense probes. The arrows (A-D) indicate the ventricular zone of the neural tube. The large arrowheads (A-D) indicate the sensory ganglia and the small arrowheads, the sympathetic ganglia (C and D). Control hybridizations with NRSF sense probes revealed no specific signal (Fig. 10C and data not shown).

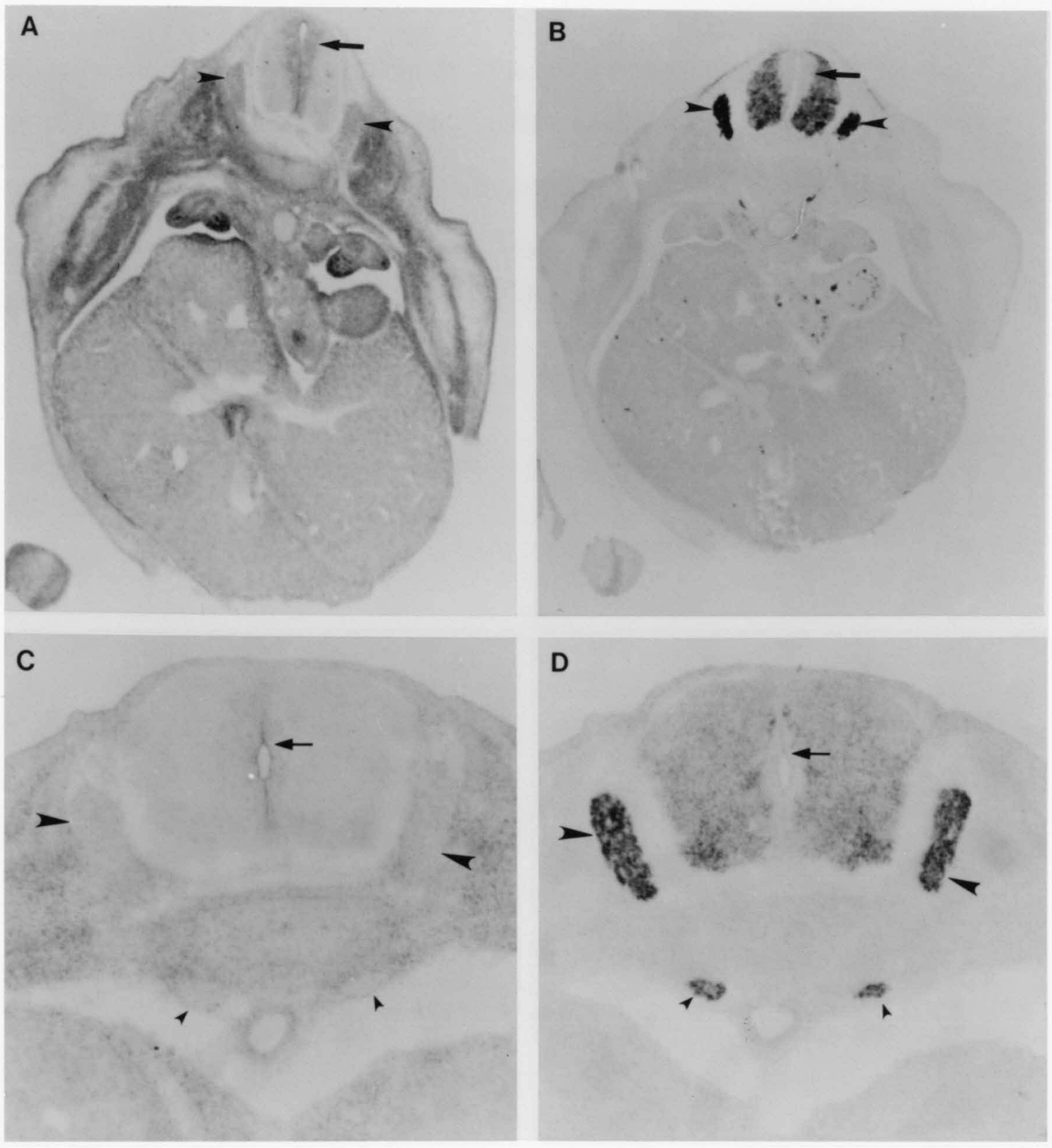

A
B
C
D

Figure 9. Analysis of NRSF message in tissues of the embryonic and adult mouse. RNase protection assays were performed on 10μg of total RNA isolated from various tissues of the embryonic day 13.5 and adult mouse. A reaction with tRNA alone (tRNA) was performed as a negative control. A low specific activity β-actin probe was added to each reaction as a control for the RNA. In this experiment, the RNase digestion was performed with RNase A and T1. Abbreviations: Br, Brain; He, heart; Ky, kidney; Li, liver; Lg, lung; Lm, limbs; NT, neural tube; Mu, muscle; 10T, C3H10T1/2, a fibroblast cell line.

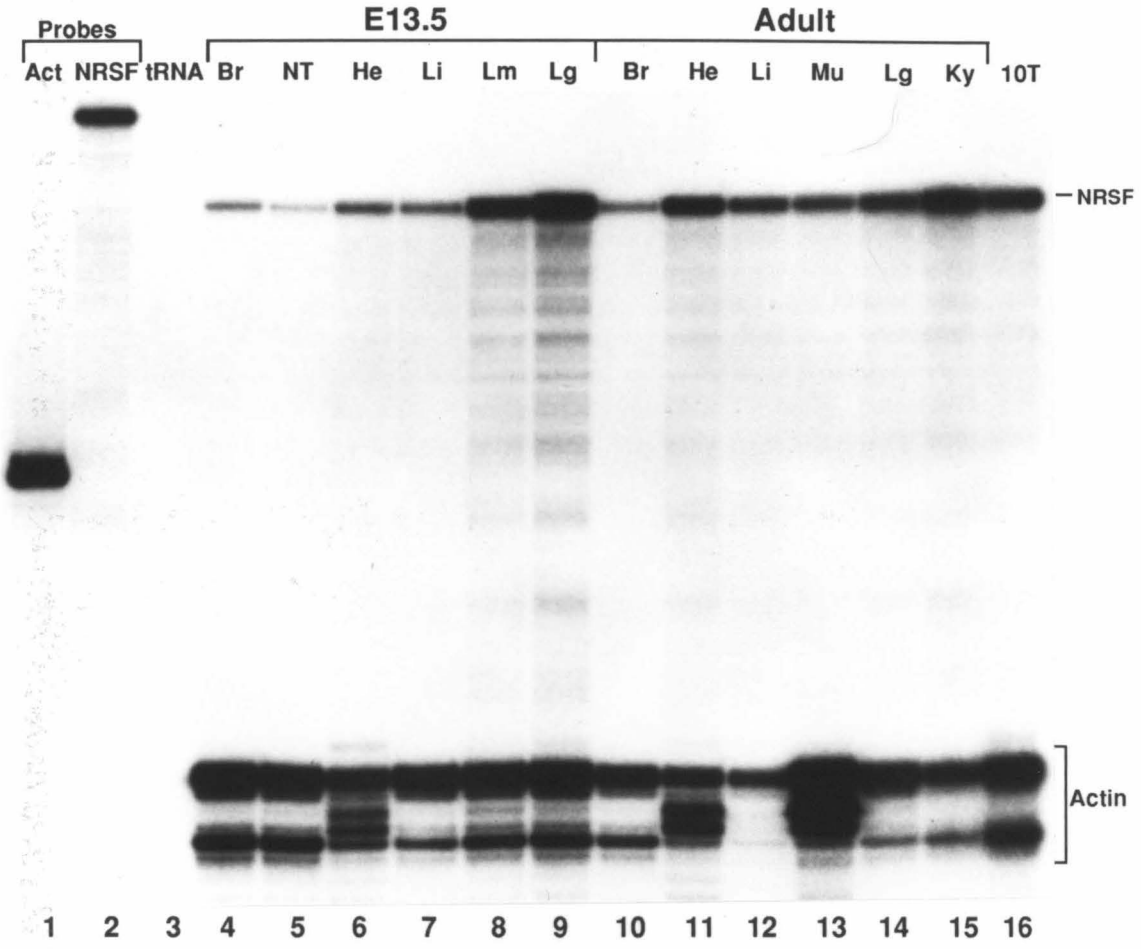
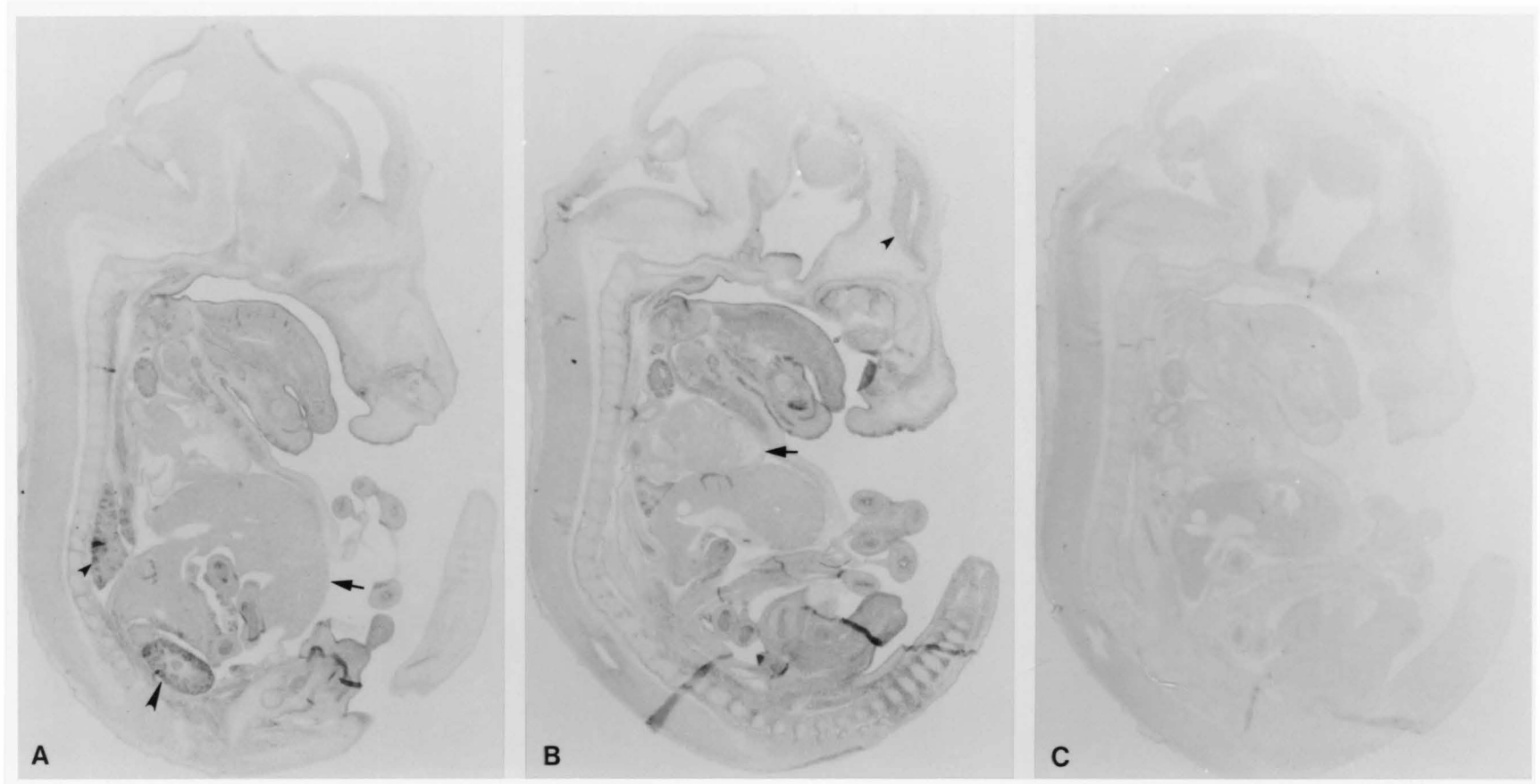

Probes
E13.5
Adult
Act
NRSF
tRNA
Br
NT
He
Li
Lm
Lg
Br
He
Li
Mu
Lg
Ky
10T
NRSF
Actin
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

Figure 10. Widespread expression of NRSF mRNA in non-neural tissues. In situ hybridization with an NRSF antisense probe (A,B) was performed on parasaggital sections of an E13.5 mouse embryo. (A) The arrowheads mark two positive tissues, the lung and the kidney; the arrow indicates the liver, which expresses much lower levels of NRSF mRNA (see also Fig. 9). (B) The arrowhead marks the ventricular zone in the telencephalon; the arrow indicates the heart. (C) An adjacent section to (B) was hybridized with an NRSF sense probe as a control for non-specific staining.


A
B
C

Chapter 4

Potential target genes for NRSF

Christopher J. Schoenherr, Alice J. Paquette, Vu Ngo, and David J. Anderson

ABSTRACT

Three neuronal genes, SCG10, type II sodium channel, and synapsin I, are negatively regulated by the same silencer factor, NRSF (Kraner et al., 1992; Li et al., 1993; Mori et al., 1992). NRSF represses the transcription of these genes in non-neuronal cells by binding to a conserved recognition sequence, the NRSE. In this report, we describe the isolation of a mouse NRSF cDNA and characterize a tissue specific alternative splicing event. We also provide evidence that there are many neuronal genes that have sequences with significant similarity to the NRSE and that these sequences are likely to represent functional binding sites for NRSF. Included in these genes are transcription factors that are implicated in the activation of neuronal differentiation, suggesting that NRSF may repress this process. Potential NRSEs also are found in non-neuronal genes which indicates that NRSF may have a more broad function than originally put forth.

INTRODUCTION

Transcriptional regulation of gene expression is an important mechanism in the development of neurons. Transcriptional regulatory proteins have been associated with many of the mutants affecting neuronal development in organisms such as *Drosophila* and *C. elegans* (Finney and Ruvkun, 1990; Jan and Jan, 1994; Way and Chalfie, 1988). In vertebrate organisms, overexpression of transcription factors in Xenopus oocytes and mutation of transcription factor genes in mice have also shown the importance of transcriptional regulation during neuronal determination and differentiation (Guillemot et al., 1993; Joyner and Guillemot, 1994; Korzh, 1994; Lee et al., 1995; Zimmerman et al., 1993). In parallel to these genetic studies, molecular analyses have identified transcription factors common to a subset of neurons that may be important for establishing and maintaining a particular neuronal phenotype (Bach et al., 1995; Ingraham et al., 1990;

Sasai et al., 1992; Tsuchida et al., 1994; Valarché et al., 1993). The regulatory networks that these transcription factors participate in, however, are not well characterized.

One of the most important steps in characterizing the regulatory network a transcription factor participates in is to determine the complement of genes this factor regulates. For example, in muscle development the MyoD family of basic helix-loop-helix (bHLH) transcription factors are known to autoregulate as well as cross regulate each other by directly activating their transcription (Olson and Klein, 1994). In addition, several muscle specific genes are known to be activated by MyoD family members (Weintraub et al., 1991). Combining this knowledge of target genes with the phenotypes of single and double mutations has led to detailed propositions of a regulatory cascade of transcription factors that begins with the earliest steps of myoblast determination and continues through differentiation to the maintenance of the muscle phenotype (Olson and Klein, 1994). While clearly incomplete, the knowledge of target genes directs and limits the choice of feasible regulatory models.

In the nervous system, target genes for several transcription factors known to be involved in neural development have been determined. In pituitary cells, Pit-1 is essential for proper pituitary development and is known to activate its own gene as well as other pituitary specific genes (Ingraham et al., 1990; Li et al., 1990). Similarly a target for mec-3 and unc-86, two proteins necessary for neurogenesis in *C.elegans*, is the *mec-3* gene itself (Finney and Ruvkun, 1990; Way and Chalfie, 1988; Xue et al., 1993). Recently, genes known to regulate the choice between a neural and epidermal fate were identified as target genes for the *Drosophila* proneural genes, *achaete* and *scute* (Singson et al., 1994; Van Doren et al., 1992). Thus, in each case, the beginning of a detailed regulatory cascade for neurogenesis is becoming clear.

Another transcription factor implicated in vertebrate neuronal development is the neuron-restrictive silencer factor (NRSF). NRSF was originally defined as a negative regulator of the neuron-specific gene, SCG10 (Mori et al., 1992). In addition, it was

determined that NRSF also negatively regulates the type II sodium channel and synapsin I genes (Kraner et al., 1992; Li et al., 1993). This factor binds to a conserved element, known as the neuron-restrictive silencer element (NRSE), that is present in all three genes. A fourth element that could bind NRSF was identified in the rat brain-derived neurotrophic (BDNF) gene, but its role in BDNF transcription is unknown. In contrast to most regulators of neuronal genes, NRSF activity and protein is not present in neuronal cells, but is found in many non-neuronal cells (Kraner et al., 1992; Mori et al., 1992). Thus, NRSF appears to prevent expression of certain neuronal genes in non-neuronal cells.

Recently the gene for human NRSF (also known as REST) was isolated and shown to encode a zinc finger transcription factor (Chong et al., 1995; Schoenherr and Anderson, 1995). In agreement with the proposed function for NRSF, its mRNA was detected in most non-neuronal cells but not in neuronal cells. Interestingly, NRSF mRNA is present in neuronal precursor cells, suggesting that NRSF may negatively regulate some aspect of neurogenesis (Chong et al., 1995; Schoenherr and Anderson, 1995). Given its ability to repress genes necessary for neuronal function, NRSF may be required to prevent precocious expression of the complete neuronal phenotype. In addition to repressing 'end-state' genes, NRSF could also inhibit neurogenesis by repressing the expression of activators of neuronal differentiation, such as transcription factors or growth factor receptors. Both of these models would be addressed if additional target genes of NRSF could be identified.

In this report, we describe the isolation of the full coding sequence for mouse NRSF and the characterization of a splicing variant. We also describe our attempts to identify NRSF target genes. Extensive DNA database searches using NRSF's recognition sequence, the 21bp NRSE, identified many genes that contain sequences with considerable similarity to the binding element. Most of the sequences with the highest similarity to the NRSE are found in neuronal genes and are located in regions associated with transcriptional regulation. Some of these sequences are conserved in several different

species, strongly supporting their functional relevance. More direct evidence for functional relevance of a subset of these sequences was provided by their ability to bind NRSF and to repress transcription of a heterologous promoter. Classification of the identified neuronal genes revealed that most of them are involved directly in neuronal function. However, at least one transcription factor implicated in activating neuronal differentiation may have a functional NRSE, thus providing a mechanism for NRSF to negatively regulate neurogenesis. Finally, several non-neuronal genes have NRSE-like sequences that can bind NRSF and repress transcription, raising the possibility that NRSF functions more broadly than in the repression of neuronal genes.

MATERIALS AND METHODS

DNA database searching

A consensus NRSE was determined by comparing the sequence of NRSEs in the rat SCG10, rat type II sodium channel, human synapsin I, and rat BDNF genes, all of which have been shown to bind NRSF [Schoenherr, 1995 #1594]. A residue was considered consensus if it was present in at least three of the four NRSE-like sequences. This consensus (see below) was used to search the Genbank DNA sequence database using the FASTA search program (Pearson, 1990). The parameters used were: word size 3, gap penalty 12.0, and gap extension penalty 4.0. Relevant groups of sequences in Genbank are divided into five different sections (invertebrates, other mammals, other vertebrates, rodents, and primates) and each of these sections was searched separately. Three hundred sequences were retrieved from each search. To limit the number of sequence alignments examined, a cutoff value of 54 for the 'optimized' similarity score (defined in Pearson, 1990) was chosen. The optimized similarity score is calculated by the FASTA program and reflects the relative quality of a gene segment's similarity to the consensus NRSE. Sequence alignments were not considered further if they contained gaps or a double point

mutation known to abolish NRSF binding (see Mori et al., 1992). Sequences also were removed from consideration if they were of unknown function (such as sequence tagged sites or pseudogenes) or if their mRNA expression pattern was unknown. Duplicate sequences from the same species were also removed. The remaining sequences were then divided into neuronal and non-neuronal categories. The same gene but from different species was counted as one gene. Similarly, potential NRSEs present in several members of closely related multigene families such as olfactory receptors and cytochrome P450s were counted as one gene.

EMSAs and transient transfections

EMSAs were performed using native NRSF and in vitro translated human NRSF (λH1) essentially as described (Mori et al., 1992; Schoenherr and Anderson, 1995), except that in some experiments Klenow labeled double-stranded oligonucleotides were used as probes. The top strand of each oligonucleotide is provided below. Each oligonucleotide was synthesized with HindIII compatible ends for insertion into the SCG10 promoter fusion construct, CAT3 (Mori et al., 1990). Cell culture, transient transfections, and CAT assays were performed as described (Mori et al., 1992; Schoenherr and Anderson, 1995).

NRSE Oligonucleotides

One strand of each oligonucleotide probe used for NRSF binding assays is shown below. The upper case letters represent genomic sequence, and the lower case sequence was added for cloning purposes. The portion with similarity to the NRSE is underlined.

Rat SCG10:	agctGCAAAGCCATTTCAGCACCACGGAGAGTGCCTCTGC
Rat Na^+ channel, type II:	agcttATTGGGTTTCAGAACCACGGACAGCACCAGAGTa
Human Synapsin I:	agcttCTGCCAGCTTCAGCACCGCGGACAGTGCCTTCGCa
Rat BDNF:	agcttAGAGTCCATTCAGCACCTTGGACAGAGCCAGCGGa

Human glycine receptor:	agctt AGGCGTTTCAGCACCACGGAGAGCGTCCAGAa
Rat NMDA1 receptor:	agctt ACACGCTTCAGCACCTCGGACAGCATCCGCCa
Human ACh receptor β2:	agctt CGCGGCTTCAGCACCACGGACAGCGCCCCACa
Chicken β4 tubulin:	agctt CCGCCGTTCAGCACCGCGGACAGCGCCGCCTa
Chicken middle neurofilament:	agctt CGGGGTTTCAGCACCACGGACAGCTCCCGCGa

Isolation of NRSF cDNA

Initial mouse NRSF clones were isolated by screening a Balb/c 3T3 cDNA library in λEXlox (a generous gift of S. Tavtigian and B. Wold) with a human NRSF cDNA (Schoenherr and Anderson, 1995). Seven clones were characterized and found to have inserts no longer than 1.5kb. To obtain a full length cDNA, a random-primed cDNA library was constructed using polyA+ RNA isolated from CH310T1/2 cells using the FAST Track RNA Isolation System (Invitrogen). Two 5μg aliquots of RNA were converted to double stranded cDNA with attached EcoRI adaptors using the Stratagene cDNA Synthesis Kit and protocol, except that one aliquot of RNA was treated with methylmercury before first strand synthesis and both reverse transcriptions were performed at 50° for two hours. The cDNA was size selected for 1kb and above using a 1% agarose gel. The remaining cDNA from both reactions was mixed and ligated into λEXlox. Approximately 1.5 million phage from the unamplified library were screened by hybridization to a 1.5kb mouse NRSF clone. Positives were screened by PCR for the longest inserts. Eight clones were characterized by restriction mapping and sequencing. Clones were sequenced by a combination of automated fluorescent and manual dideoxynucleotide sequencing. The resulting sequences were assembled and analyzed using the GCG program (Devereux et al., 1984).

RT-PCR

Reverse transcription-PCR was performed on RNA isolated by the method of Chomczynski et al. 1987. First strand synthesis on 1μg of total RNA was performed with the Gibco Preamplification System. One-twentieth of the product was used in a standard PCR reaction for 40 cycles at 94°, 1min; 57°, 1min; and 72°, 2min in a Perkin Elmer Thermocycler. 5' primer: GGTCAAGAAGCAAAGATCCGCTTC; 3' primer: ATCTCACTCAGCAGGCTCAGCT. Primers were used at 1μM and began at nucleotide 2068 and 2793, respectively, according to Figure 1. Products were analyzed on a 1.5% agarose gel.

RESULTS

Isolation of mouse NRSF

To determine the structure and function of NRSF, it was necessary to isolate a cDNA that contains the entire coding sequence. The cloning of a partial cDNA for human NRSF is described in Chapter 3. As previous attempts to isolate complete human cDNAs were unsuccessful, the human NRSF was used as a probe to screen a mouse 3T3 cell library. Several positive clones were isolated, characterized, and shown to contain the mouse NRSF. One of these clones was used to screen another mouse cDNA library (made from CH310T1/2 cells) to obtain a full length cDNA clone. Two overlapping clones were isolated, sequenced, and spliced together. A comparison of this sequence to human NRSF revealed that the mouse NRSF cDNA was missing three of seven copies of a 16 amino acid proline rich repeated motif that were identified in human NRSF (Chong et al., 1995). (See Figure 3.) This result suggested that either these repeats are not present in mouse NRSF or that our cDNA clone represents a differential splicing event. Evidence for the latter was obtained by sequence analysis of mouse genomic fragments that contain an NRSF pseudogene. The sequencing identified a region in the pseudogene with homology to the repeats in human NRSF (Chen and Schoenherr, unpublished).

To isolate a mouse cDNA that contains these repeats, RT-PCR was performed on RNA from several cell types. A PCR product was obtained that matched the size obtained from amplifying the NRSF pseudogene. The PCR product was sequenced and conceptually spliced into the original mouse NRSF sequence. There is an open reading frame across the PCR product that retains the reading frame of the original cDNA. This indicates that there are least two splice variants of mouse NRSF. The sequence of this composite cDNA is shown in Figure 1. The sequence of the insertion is double underlined (Fig. 1). This composite cDNA is 4375bp long and has a long open reading frame of 1082 amino acids that begins with a methionine codon. This ATG is preceded by an in-frame stop codon 172bp upstream, indicating that it is likely to be the initiating codon. The TAG triplet at base pair 3602 is likely to be the stop codon as there are many stop codons in all three reading frames downstream of it.

Comparison of mouse and human NRSF

The conceptual translation of the composite cDNA was compared to human NRSF. The comparison of the NRSF homologs showed regions of high homology interspersed with ones of much lower homology (Fig. 2). The N-terminus, beginning with the initiating methionine up to the first zinc finger, is well conserved between mouse and human NRSF (84% identity). The eight zinc finger domain, which can bind NRSEs in isolation, is 96% identical between the two genes; the ninth finger, although not necessary for NRSE binding, is also well conserved (22/23 residues). A domain just C-terminal to the eighth zinc finger that is rich in lysines and serine/threonines shows only moderate conservation of primary sequence (64%) but the high basicity of the region is retained. A small acidic region near the ninth finger is almost identical between the two homologs (37/38 residues).

The proline rich repeats identified by Chong et al., 1995, however, are not well conserved between the two species. While four similar sequences could be found in the

mouse gene, they matched only about half of the 16 residues of the human repeat. Furthermore, a five residue repeat (consensus MEVAQ) that occurs 13 times in human NRSF does not occur with similar frequency in mouse NRSF. The location of both repeats in the human and mouse isoforms is shown in Figure 3. A significiant portion of this region is deleted in the smaller isoform of mouse NRSF although the four proline rich repeats remain. Although the overall conservation in this region is comparatively low (38%), the proline rich nature of this region is conserved with 21% proline residues in human NRSF and 26% in mouse. The function of this proline rich region is unknown.

Tissue-specific splicing variant of NRSF

To determine if the two splice variants of NRSF are expressed in a tissue-specific manner, RT-PCR was performed on RNA from cell lines and tissues taken from embryonic day 13.5 and adult mice. As shown in Figure 4, the larger splicing variant was found to predominate in adult muscle, heart, lung, and kidney, as well as the fibroblast line CH310T1/2 and the embryonic carcinoma, P19. The smaller variant predominates in E13.5 limb tissue. These results showed that differential splicing of NRSF is tissue specific. Adult liver did not give either expected product but did give a much smaller band at about 220bp. The nature of this product and the unidentified products derived from the other tissues is unknown. Furthermore, as the function of this region is unknown, the significance of this regulation remains to be determined.

Potential NRSF Target Genes

NRSF has been shown to repress the transcription of three neuronal genes in non-neuronal cells by binding to the NRSE (Kraner et al., 1992; Li et al., 1993; Mori et al., 1992). These results suggested that NRSF may regulate many different neuronal genes. As an initial step to determine if additional genes were regulated by NRSF, a consensus NRSE was determined (TTCAGCACCnCGGACCAnGCC) and was used to search the

Genbank DNA sequence database using the FASTA program (Pearson, 1990). This search should identify candidate NRSF-regulated genes by the presence of sequence elements similar to the NRSE. As expected from using a short sequence element, these searches revealed a large number of genes with sequences that have substantial similarity to the NRSE. To increase the likelihood that a gene may contain a functional NRSE and to limit the number of genes examined to a manageable number, genes with sequences that had an 'optimized' similarity score (calculated by the FASTA program) below an arbitrary cut-off (see below) were not considered further. Sequences also were not considered if they were similar to a known double point mutation that inactivates the NRSE (Mori et al., 1992).

Initial inspection of the remaining genes with potential NRSEs revealed that many of them are neuron-specific. However, many of the remaining genes are widely expressed or not known to be expressed in neurons. Thus, we wanted to address two questions: What portion of these potential NRSEs are likely to be functional repressor elements? And does NRSF interact with NRSEs in non-neuronal genes as well as in neuronal ones? As a first step, the number of neuronal and non-neuronal genes and their average similarity scores were determined. A gene was considered neuronal if it shows expression in neurons and no or limited expression elsewhere. Thus, genes such as atrial natriuretic peptide and calbindin which, in additions to neurons, are expressed in restricted types of non-neuronal cells were considered neuronal. While on the other hand, adenosine phosphoribosyl-transferase (aprt) is expressed in many cell types including neurons, and was considered non-neuronal. This analysis revealed that about 39 distinct neuronal genes have sequences with a similarity score at or above the cutoff score of 54; the average score is about 62 ± 7.3, with five genes having the perfect match score of 76. Two genes have more than one sequence that meets these criteria. For the non-neuronal genes, which numbered 52, the average score was 56 ± 4.7, much lower than the average neuronal gene score. If the cutoff score is raised to 62 (which corresponds to 2 mismatches from the

consensus), the results are even more dramatic: 21 of the NRSE-like sequences are in neuronal genes and only 6 are in non-neuronal genes. Importantly, the neuronal genes identified in this search included the three genes with functionally defined NRSEs. These results indicate a strong bias toward neuronal genes having NRSE-like sequences over non-neuronal genes. A list of 25 neuronal and 10 non-neuronal genes and their NRSE-like sequences is shown in Table Ia and Ib.

Potential NRSEs are preferentially located in regulatory regions

The functional relevance of the NRSE-like sequences was addressed further by noting their locations in their respective transcription units. While this method does not directly assess function, a locational pattern may exist that would lend credence to the relevance of these sequences. The locations in the transcription unit of the 91 sequences were determined and divided into five categories: Regulatory (5' or 3' non-transcribed), 5'UTR, intron, 3'UTR, and coding. A summary of this tabulation is shown in Table II. This analysis revealed that a greater percentage of the neuronal potential NRSEs with similarity scores 54 and above are located in non-coding regions than those from non-neuronal genes (26/41 versus 18/50). Furthermore, many more of the neuronal sequences compared to the non-neuronal ones (11/41 versus 2/50) are located in the 5'UTR. The comparison of locations was more striking when only sequences with similarity scores of 62 and above were considered. In this case, 18/21 (including 10 in the 5'UTR) neuronal sequences and 6/6 non-neuronal sequences are located in non-coding regions. Thus, most of the neuronal sequences with the highest similarity to the NRSE are preferentially located in non-coding regions with an additional bias toward the 5'UTR. While the bias towards 5'UTRs (versus regulatory and intronic regions) may reflect the bias in the database for cDNA sequences, the disproportionate number of these elements in this location could reflect a conserved, functionally relevant pattern. Considering that NRSF is a negative regulator, such a placement could maximize transcriptional repression. Thus, the NRSE-

like sequences in neuronal genes do appear to be preferentially located in non-coding regions when compared to those in non-neuronal genes. However, six non-neuronal genes have potential NRSEs with high similarity scores present in non-coding regions.

Evolutionary conservation of putative NRSEs

Additional support for the functional relevance of a portion of the potential NRSEs is derived from examining their sequence conservation in different species. Table II shows the species and the alignment of potential NRSEs that were conserved between species at least as distant as mouse and humans. Conserved sequences present in coding regions are not shown, as such conservation may reflect protein function rather than DNA function. Amongst the neuronal genes, several NRSE-like sequences and their location within the gene were found to be conserved in distantly related species. For example, a potential NRSE was found in the first intron of the corticotrophin-releasing factor gene of four species, from Xenopus to humans. Six other neuronal sequences were found in two or more different species. One NRSE-like sequence from a non-neuronal gene (skeletal actin) was conserved according to these criteria. Such sequence conservation over considerable evolutionary time strongly suggests that these sequences are functional.

Many potential NRSEs can bind NRSF and repress transcription

Although the above evidence suggests that some of these sequences are likely to be functional NRSEs, more direct evidence was obtained by determining their ability to bind NRSF and repress transcription. Although we considered it impractical to test all 91 potential NRSEs, we examined 24 of the sequences, 16 from neuronal genes and 8 from non-neuronal genes, for their ability to bind NRSF. Double stranded oligonucleotide probes representing the potential NRSEs were tested for their ability to bind in vitro translated human NRSF in an electrophoretic mobility shift assay (EMSA). Figure 4 shows that nine probes derived from neuronal genes could bind NRSF in a manner similar

to the NRSE originally defined in SCG10. Three sequences, from Na channel, synapsin I and BDNF, have previously been shown to bind NRSF (Schoenherr and Anderson, 1995). Furthermore, an additional six of seven probes derived from neuronal genes could also bind NRSF. We also tested eight probes from non-neuronal genes, of which five could bind NRSF (Paquette and Schoenherr, data not shown). All probes derived from sequences with a similarity score above 60 successfully bound NRSF. Of those with scores between 59 and 54, only two of six successfully bound NRSF. The qualitative aspects of these binding results were confirmed by EMSAs performed using each oligonucleotide as a competitor against the SCG10 NRSE probe. These competition EMSAs were performed with HeLa NRSF, indicating that native as well as recombinant NRSF could bind these sequences (Paquette and Schoenherr, data not shown). To test their ability to repress transcription, a subset of the sequences assayed for NRSF binding were placed upstream of the SCG10 promoter reporter construct, CAT3 (Mori et al., 1990), and introduced into CH310T1/2 cells. These experiments revealed a complete parallel between a sequence's ability to bind NRSF in vitro and its ability to repress transcription (Paquette and Schoenherr, data not shown). Thus, these results showed that most but not all of the identified neuronal sequences can bind NRSF and repress transcription. Additionally, some non-neuronal sequences behaved as functional NRSEs, supporting the possibility that NRSF-mediated repression may not be limited to neuronal genes.

DISCUSSION

Mouse NRSF

We have isolated a cDNA for mouse NRSF and identified two products of differential splicing. A comparison between mouse and human NRSF reveals regions of high sequence conservation and other regions with much less conservation. As expected,

the zinc finger DNA binding domains are nearly identical. The functions of the other regions of high homology are unknown, but their conservation suggests they are likely to be important for NRSF activity. The general characteristics of the low similarity regions, however, do seemed to be conserved. This type of conservation is reminiscent of what has been seen in the several classes of transcriptional activation domains that have been characterized (Mitchell and Tjian, 1989). Thus, any interpretation of the function of these divergent regions should take this into account. The existence of a differential splicing event in this region further suggests that the proline rich region may be significant. Functional analysis of NRSF should address these questions.

Potential NRSF-regulated genes

We believe that the majority of the NRSE-like sequences we found in neuronal genes will prove to be functional. First, over half of the sequences have no more than two base changes from the consensus; five are identical to the NRSE. Furthermore, the three original NRSEs were identified in these searches, verifying the search parameters. Also , most of the best candidate NRSEs are located in gene regions frequently associated with transcriptional activity, consistent with a regulatory role. Almost half of these elements are located in the 5'UTR, perhaps providing NRSF with an optimal position to repress transcription. Additional evidence suggesting the importance of these elements is provided by their evolutionary conservation. At least seven of these potential NRSEs are conserved over considerable evolutionary distance. The last circumstantial evidence supporting the repressor function of some of these elements comes from previous analyses of the regulation of two of the identified genes. For example, in the rat VGF gene, a 218bp fragment containing the NRSE-like sequence was shown to have repressor activity in non-neuronal but not in neuronal cells . In addition, Keegan et al. showed that an NRSE-containing 4kb portion of the rat corticotropin-releasing factor gene can repress transgene expression in non-neuronal cells (Keegan et al., 1994).

The most direct evidence that these potential NRSEs might be important in transcriptional regulation comes from the NRSF binding assays and transient transfection experiments. These experiments showed that most of the neuronal sequences tested could bind NRSF and repress transcription in a manner similar to the original SCG10 NRSE. In fact, all neuronal sequences we tested with similarity scores of 60 and above could bind NRSF and repress transcription, implying that most or all of the untested sequences with such scores would behave similarly. On the other hand, one of the two neuronal elements and three of four of the non-neuronal elements with scores below 59 performed poorly in both assays. This suggests that many of the sequences with scores of 59 or less will not be functional. Results of these experiments validated the database searches as a method for identifying sequences that can function as bona fide NRSEs. While some of these elements may not operate in the context of their normal transcription unit, it is difficult to argue that all these similar sequences that can repress transcription are non-functional and, thus, merely coincidental. Therefore, if we accept that all NRSE-like sequences with similarity scores 60 and above are probable silencer elements, then the number of NRSF-regulated neuronal genes identified climbs from 3 to at least 23.

The arguments above also apply to the potential NRSEs in non-neuronal genes. five of these elements showed high affinity binding and repressor activity. These sequences are located in regulatory regions and one shows evolutionary conservation. Thus it is probable that NRSF does regulate certain non-neuronal genes. Such an activity is not incompatible with its proposed functions in neuronal gene regulation. In fact, NRSF's widespread expression pattern would make a single regulatory function seem unlikely. Other transcription factors such as YY1, AP-1, and hormone receptors are widely expressed and have different functions depending on the gene being examined [Schüle, 1991 #812; Umesono, 1991 #1772; Natesan, 1993 #1750]. Thus, it is conceivable that NRSF may only partially repress or even activate transcription in the context of these non-neuronal genes. It will be interesting to determine what role, if any, these elements play in

regulating non-neuronal genes that are expressed in the same cells as NRSF, such as skeletal muscle actin. This includes some of the neuronal genes that have limited expression outside the nervous system.

For the most part we have ignored the potential NRSEs in coding regions. The average similarity score for both the neuronal and non-neuronal sequences in coding regions is much lower than those in non-coding regions, suggesting that many of these sequences will not bind NRSF. In agreement with this, only two of six EMSA probes made from sequences with similarity scores of 59 or below bound NRSF in vitro or repressed transcription. This does not mean that all of sequences in coding will prove to be non-functional. At least two of these elements, in the neuronal genes encoding Hes-3 and a subset of olfactory receptors, have high similarity scores (68) and are likely to bind NRSF. It would be interesting to examine whether NRSE-containing coding sequences can also repress transcription. Coding sequences, as well as 5'UTRs and introns, are usually excluded in studies of transcriptional regulation. Many of the best candidate NRSEs are present in these regions, which are often not included in reporter constructs. Our results suggest that inclusion of largely intact neuronal genes in transgenic and transient transfection experiments may be necessary to closely mimic normal transcriptional regulation.

Ideally, we would like derive a set of rules with which we could determine whether a given sequence will bind NRSF or not. Inspection of the sequences tested for NRSF binding, however, did not reveal such a paradigm. In fact, the results suggested that no single residue was critical for binding, perhaps due to the length of the NRSE (21bp). If true, this would make a systematic, single point mutation study uninformative. However, examination of sequences that did not or only weakly bound NRSF suggested that certain residues are more important than others. For example, two sequences that do not bind NRSF, from T-cell receptor beta and the myosin light chain, have only one additional mutation (C8->A or G12->T, respectively) that is not found in sequences that do bind.

This suggests that single mutations in these two residues may eliminate NRSF binding. On the other hand, the lower affinity of these elements may be due to the combined effects of their particular differences from the consensus NRSE. In the end, the binding data provide a limited 'template' which can be compared to potential NRSEs. If there is a match to one of the 24 sequences we have tested, then a particular sequence can be included or excluded.

NRSF and invertebrates

We also examined invertebrate sequences in Genbank for potential NRSEs. Interestingly, no neuronal genes were identified, and all eight non-neuronal genes found had similarity scores of 58 or below. In addition, preliminary experiments to identify NRSE-binding activity in a Drosophila cell nuclear extract or an NRSF-like mRNA by degenerate PCR have been unsuccessful (Paquette, unpublished). While these are negative results, they suggest that the NRSF regulation system is present only in vertebrate species.

Function of NRSF

At the outset of this study, we wanted to identify NRSF-regulated target genes as a way to address its role in neuronal differentiation. We found, as originally proposed, that NRSF probably regulates many genes involved in neuronal function. Moreover, we could see no pattern to the neuronal genes NRSF regulates. They included genes involved in virtually all aspects of the neuronal phenotype. This suggests that NRSF's role in differentiation is to directly prevent ectopic expression of the entire neuronal program. This may be important as a precursor cell proceeds toward neuronal determination and acquires transcription factors (or closely related family members) present in neurons. However, we did identify at least one neuronal transcription factor that is likely to be regulated by NRSF.

This gene, P-Lim (also known as mLIM-3 or lim3a), is a LIM homeodomain protein and an activator of pituitary specific genes [Bach, 1995 #1702; Seidah, 1994 #1703; Tsuchida, 1994 #1548]. P-Lim mRNA can be detected in pituitary neuroendocrine

precursors and in adult pituitary cells. It is also expressed transiently in motor neuron precursors. P-Lim also was shown to activate transcription of Pit-1, a gene required for proper pituitary development (Bach et al., 1995). These results suggest a regulatory cascade in which the postulated NRSF-mediated repression of P-Lim would prevent proper pituitary development by inhibiting an activator of Pit-1 expression. Then, once a signal to down-regulate NRSF expression was received, P-Lim could be transcribed and would activate Pit-1 and other pituitary specific genes. Thus, NRSF may negatively regulate one or more positive regulators of the neuronal phenotype.

REFERENCES

Bach, I., Rhodes, S. J., Pearse, R. V., Heinzel, T., Gloss, B., Scully, K. M., Sawchenko, P. E., and Rosenfeld, M. G. (1995). P-Lim, a Lim-homeodomain factor, is expressed during pituitary organ and cell commitment and synergizes with Pit-1. Proc. Natl. Acad. Sci. *92*, 2720-2724.

Chong, J. A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J. J., Zheng, Y., Boutros, M. C., Altshuller, Y. M., Frohman, M. A., Kraner, S. D., and Mandel, G. (1995). REST: A mammalian silencer protein that restricts sodium channel expression to neurons. Cell *80*, 949-957.

Devereux, J. D., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Research *12*, 387-395.

Finney, M., and Ruvkun, G. (1990). The *unc-86* gene-product couples cell lineage and cell identity in C.elegans. Cell *63*, 895-905.

Guillemot, F., Lo, L.-C., Johnson, J. E., Auerbach, A., Anderson, D. J., and Joyner, A. L. (1993). Mammalian achaete-scute homolog-1 is required for the early development of olfactory and autonomic neurons. Cell *75*, 463-476.

Ingraham, H. A., Albert, V. R., Chen, R., Crenshaw, E. B. I., Elsholtz, H. P., He, X., Kapiloff, M. S., Mangalam, H. J., Swanson, L. W., Treacy, M. N., and Rosenfeld,

M. G. (1990). A family of POU-domain and pit-1 tissue-specific transcription factors in pituitary and neuroendocrine development. Annu. Rev. Physiol. *52*, 773-791.

Jan, Y. N., and Jan, L. Y. (1994). Genetic control of cell fate specification in Drosophila peripheral nervous system. Annual Review of Genetics *28*, 373-393.

Joyner, A. L., and Guillemot, F. (1994). Gene targeting and development of the nervous system. Curr. Op. Neuro. *4*, 37-42.

Keegan, C. E., Karolyi, I. J., Knapp, L. T., Bourbonais, F. J., Camper, S. A., and Seasholtz, A. F. (1994). Expression of corticotropin-releasing hormone transgenes in neurons of adult and developing Mice. Mol. Cell. Neurosci. *5*, 505-514.

Korzh, V. P. (1994). Genetic control of early neuronal development in vertebrates. Curr. Op. Neuro. *4*, 21-28.

Kraner, S. D., Chong, J. A., Tsay, H. J., and Mandel, G. (1992). Silencing the type II sodium channel gene: a model for neural-specific gene regulation. Neuron *9*, 37-44.

Lee, J. E., Hollenberg, S. M., Snider, L., Turner, D. L., Lipnick, N., and Weintraub, H. (1995). Conversion of *Xenopus* extoderm into neurons by NeuroD, a basic helix-loop-helix protein. Science *268*, 836-844.

Li, L., Suzuki, T., Mori, N., and Greengard, P. (1993). Identification of a functional silencer element involved in neuron-specific expression of the synapsin I gene. Proc. Natl. Acad. Sci. U.S.A. *90*, 1460-1464.

Li, S., Crenshaw, E. B. I., Rawson, E. J., Simmons, D. M., Swanson, L. W., and Rosenfeld, M. G. (1990). Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene *pit*-1. Nature(*London*) *347*, 528-533.

Mitchell, P., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371-378.

Mori, N., Schoenherr, C., Vandenbergh, D. J., and Anderson, D. J. (1992). A common silencer element in the SCG10 and type II Na^+ channel genes binds a factor present in nonneuronal cells but not in neuronal cells. Neuron *9*, 1-10.

Mori, N., Stein, R., Sigmund, O., and Anderson, D. J. (1990). A cell type-preferred silencer element that controls the neural-specific expression of the SCG10 gene. Neuron *4*, 583-594.

Olson, E. N., and Klein, W. H. (1994). bHLH factors in muscle development: Dead lines and commitments, what to leave in and what to leave out. Genes & Dev. *8*, 1-8.

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. Meth. Enzm. *183*, 63-98.

Possenti, R., Di Rocco, G., Nasi, S., and Levi, A. (1992). Regulatory elements in the promoter region of *vgf,* a nerve growth factor-inducible gene. Proc. Natl. Acad. Sci. USA *89*, 3815-3819.

Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to *Drosophila hairy* and *enhancer of split*. Genes & Dev *6*, 2620-2634.

Schoenherr, C. J., and Anderson, D. J. (1995). The Neuron-Restrictive Silencer Factor (NRSF): A coordinate repressor of multiple neuron-specific genes. Science *267*, 1360-1363.

Schüle, R., and Evans, R. M. (1991). Cross-coupling of signal transduction pathways: Zinc finger meets leucine zipper. Trends Genet *7*, 377-381.

Seidah, N. G., Barale, J.-C., Marcinkiewicz, M., Mattei, M.-G., Day, R., and Chretien, M. (1994). The mouse homeoprotein mLIM-3 is expressed in cells derived from the neuroepithelium and persists in adult pituitary. DNA Cell Biol. *13*, 1163-1180.

Singson, A., Leviten, M. W., Bang, A. G., Hua, X. H., and Posakony, J. W. (1994). Direct downstream targets of proneural activators in the imaginal disc include genes involved in lateral inhibitory signaling. Genes & Dev. *8*, 2058-2071.

Tsuchida, T., Ensini, M., Morton, S. B., Baldassare, M., Edlund, T., Jessell, T. M., and Pfaff, S. L. (1994). Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. Cell *79*, 957-970.

Valarché, I., Tissier-Seta, J.-P., Hirsch, M.-R., Martinez, S., Goridis, C., and Brunet, J.-F. (1993). The mouse homeodomain protein Phox2 regulates NCAM promoter activity in concert with Cux/CDP and is a putative determinant of neurotransmitter phenotype. Development *119*, 881-896.

Van Doren, M., Powell, P. A., Pasternak, D., Singson, A., and Posakony, J. W. (1992). Spatial regulation of proneural gene activity: auto- and cross-activation of *achaete* is antagonized by *extramacrochaete*. Genes & Dev. *6*, 2592-2605.

Way, J. C., and Chalfie, M. (1988). *mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in C. elegans. Cell *54*, 5-16.

Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T. K., Turner, D., Rupp, R., Hollenberg, S., Zhuang, Y., and Lassar, A. (1991). The *myoD* gene family: nodal point during specification of the muscle cell lineage. Science *251*, 761-766.

Xue, D., Tu, Y., and Chalfie, M. (1993). Cooperative interactions between the *Caenorhabditis elegans* homeoproteins UNC-86 and MEC-3. Science *261*, 1324-1328.

Zimmerman, K., Shih, J., Bars, J., Collazo, A., and Anderson, D. J. (1993). *XASH-3*, a novel *Xenopus achaete-scute* homolog, provides an early marker of planar neural induction and position along the medio-lateral axis of the neural plate. Development *119*, 221-232.

Table Ia. Neuronal genes with NRSE-like sequences[a]

Gene Name	Sequence Comparison	Similarity Score
	Consensus TTCAGCACCnCGGACAGnGCC	
Rat SCG10	----------G------	69
Rat Type II Na Channel	-----A----------A--	62
Human Synapsin	--------------------	76
Rat BDNF	---------T----------	69
Rat NMDA Rec. I	----------------AT-	64
Human Nicotinic ACh Rec. ß2	--------------------	76
Chicken ß4-tubulin	--------------------	76
Chicken Middle Neurofilament (rev)	----------------T--	69
Human Glycine Rec. (rev)	----------G----T-	62
Rat Glycine Rec. (rev)	----------A------T--	62
Rat Synaptophysin	-C--------T----------	62
Human L1	---------G--------AA	61
Rat Atrial Natriuretic Peptide	----------------CG-	64
Mouse Calbindin	AG----------------	68
Rat GABA-A Rec.δ subunit (rev)	----------A----G---GA	54
Rat Nicotinic ACh Rec. α7[b]	AG----G------C-------	54
Mouse P-Lim	--------------------	76
Mouse Hes-3	GG----------------	68
Human CRF	--------------------	76
Human Olfactory Rec. (rev)	C---------A----------	65
Mouse Synaptotagmin	----------------A	72
Mouse AMPA rec. (rev)	--T----------------T	65
Rat VGF (rev)	--------G-T----------	62
Rat prodynorphin (rev)	--------A-----G------	62
Rat Cyto. Dynein Heavy Chain (rev)	AGTT----------------	60

[a]The first 16 sequences have been assayed for NRSF binding. The remaining sequences have not been tested.

[b]was unable to bind NRSF

Table Ib. Non-neuronal genes with NRSE-like sequences[a]

Gene Name	Sequence Comparison	Similarity Score
	Consensus TTCAGCACCnCGGACAGnGCC	
Rat APRT (rev)	A----------	72
Sheep Keratin	A----------G-	65
Mouse Skeletal Actin (rev)	GG----------C-------	61
Bovine P450 (rev)	----------A--------G-	62
Human Steroid ß-Hydroxylase (rev)	----------A-------AG-	57
Human T-cell Rec. ß-chain[b]	G------A--T----------	58
Human Myosin Light Chain (rev)[b]	C---------AT---------	55
Mouse Macrophage Prot.[b]	CC---------A-C-------	54
Rat Somatostatin Trans. Factor (rev)	----------T	72
Rat Choline Kinase	C----------AA	64

[a]The first eight sequences have been assayed for NRSF binding. The remaining sequences have not been tested.

[b]was unable to bind NRSF

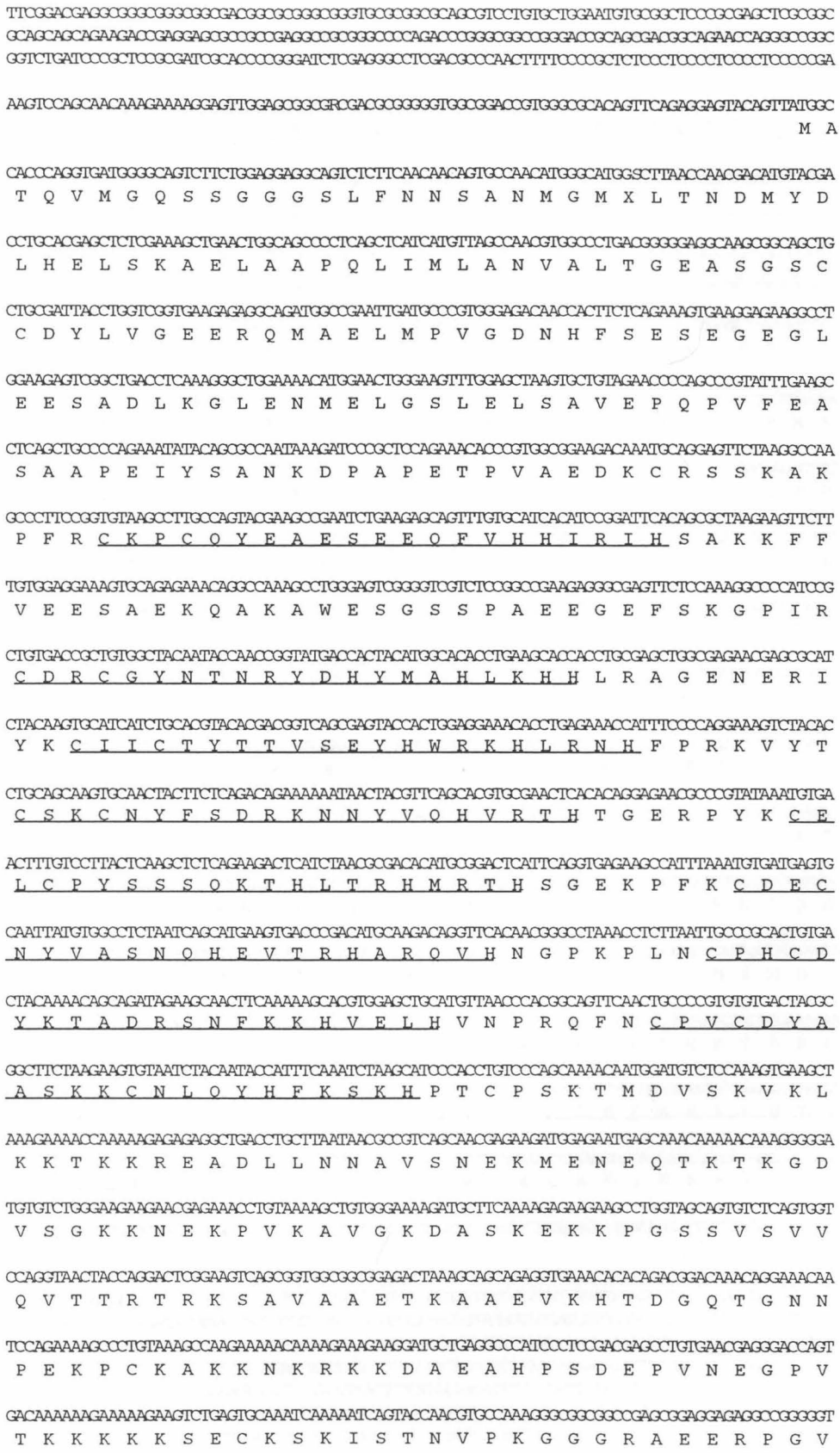
Table II. Location of potential NRSEs within transcription units

	Regulatory	5' UTR	Intron	3' UTR	Coding
Neuronal 54+	5 (13%)	11 (25%)	8 (20%)	2 (5%)	15 (37%)
Non-neuronal 54+	7 (13%)	2 (4%)	8 (17%)	1 (2%)	32 (65%)
Neuronal 62+	5 (22%)	9 (45%)	5 (22%)	0	2 (10%)
Non-neuronal 62+	2 (33%)	1 (17%)	2 (33%)	0	1 (17%)

Table III. Evolutionary conservation of potential NRSEs

Gene	Species	Comparison
		TTCAGCACCNCGGACAGNGCC
CRF	Sheep	--------T----------
	Xenopus	----------AA
	Human	--------------------
	Rat	----------T-
nACH ß-2	Human	--------------------
	Mouse	----------T
	Rat	----------T
NMDA receptor 1	Human	--------------------
	Rat	----------AT-
	Duck	----------G---G------
Synapsin I	Human	--------------------
	Rat	--T--T---------------
L1	Human	----------G--------AA
	Rat	-C----------AGA
Atrial Natr. peptide	Cow	----------T-------AG-
	Horse	----------T---AAA
	Human	----------T-------AGA
	Mouse	----------A-------CG-
	Guinea pig	----------AG-
	Rat	----------CG-
Calbindin	Chicken	G--------------------
	Human	AG--------A----------
	Rat	AG--------------------
	Mouse	AG--------------------
	(Consensus rev)	GGCNCTGTCCGNGGTGCTGAA
Skeletal actin	Cow	-------A----------GC
	Mouse	-------G----------CC

Figure 1. Nucleotide and predicted amino acid sequence of a composite cDNA for mouse NRSF. The nucleotide sequence is numbered in standard type beginning at the first nucleotide of the cDNA. The amino acid sequence is numbered in italics and begins with the first methionine. The nine zinc fingers are underlined. The spliced-in sequence of the alternative exon is marked with a double underline.



```
2071 CAAGAAGCAAAGCGCTTCCCTTAAGAAAGGCACAAAGAAGACGCCGCCCAAGACAAAGACAAGTAAAAAAGGTGGCAAACTTGCTCCAAA 2160
573  K  K  Q  S  A  S  L  K  K  G  T  K  K  T  P  P  K  T  K  T  S  K  K  G  G  K  L  A  P  K  602

2161 GGGGATGGGGCAGACAGAACCTTCTTCTGGGGCATTGGCTCAAGTGGGGGTGTCTCCAGACCCTGCCCTCATTCAGGCAGAGGTCACCGG 2250
603  G  M  G  Q  T  E  P  S  S  G  A  L  A  Q  V  G  V  S  P  D  P  A  L  I  Q  A  E  V  T  G  632

2251 GTCAGGATCTTCTCAGACAGAGCTTCCTTCACCCATGGATATTGCTAAGTCAGAGCCCGCCCAGATGGAGGTTTCCCTAACAGGGCCACC 2340
633  S  G  S  S  Q  T  E  L  P  S  P  M  D  I  A  K  S  E  P  A  Q  M  E  V  S  L  T  G  P  P  662

2341 TCCGGTGGAGCCTGCTCAAATGGAGCCATCGCCTGCGAAACCTCCCCAGGTAGAAGCACCCACTTACCCCCAGCCTCCCCAAAGGGGCCC 2430
663  P  V  E  P  A  Q  M  E  P  S  P  A  K  P  P  Q  V  E  A  P  T  Y  P  Q  P  P  Q  R  G  P  692

2431 TGCCCCTCCCACGGGGCCTGCCCCTCCCACGGGGCCTGCCCCTCCCACGGAGCCTGCCCCTCCCACGGGGCTTGCCGAGATGGAACCTTC 2520
693  A  P  P  T  G  P  A  P  P  T  G  P  A  P  P  T  E  P  A  P  P  T  G  L  A  E  M  E  P  S  722

2521 TCCCACGGAGCCTTCCCAGAAGGAACCACCTCCCAGTATGGAGCCTCCCTGCCCCGAGGAGCTGCCTCAGGCCGAGCCACCTCCTATGGA 2610
723  P  T  E  P  S  Q  K  E  P  P  P  S  M  E  P  P  C  P  E  E  L  P  Q  A  E  P  P  P  M  E  752

2611 GGATTGTCAGAAGGAGCTGCCTTCTCCCGTGGAGCCCGCTCAGATTGAGGTTGCTCAGACGGCCCCTACGCAGGTTCAGGAGGAGCCCCC 2700
753  D  C  Q  K  E  L  P  S  P  V  E  P  A  Q  I  E  V  A  Q  T  A  P  T  Q  V  Q  E  E  P  P  782

2701 TCCTGTCTCGGAGCCACCTGGGGTGAAGCCAACCAAAAGATCATCTCTCCGGAAAGACAGAGCAGAGAAGGAGCTGAGCCTGCTGAGTGA 2790
783  P  V  S  E  P  P  R  V  K  P  T  K  R  S  S  L  R  K  D  R  A  E  K  E  L  S  L  L  S  E  812

2791 GATGGCGCGGCAGGAGCAGGTCCTCATGGGGGTTGGCTTGGTGCCTGTTAGAGACAGCAAGCTTCTGAAGGGAAACAAGAGCGCCCAGGA 2880
813  M  A  R  Q  E  Q  V  L  M  G  V  G  L  V  P  V  R  D  S  K  L  L  K  G  N  K  S  A  Q  D  842

2881 CCCCCCAGCCCCACCGTCACCATCGCCAAAGGGAAACTCGAGGGAAGAGACACCCAAGGACCAAGAAATGGTCTCTGATGGGGAAGGAAC 2970
843  P  P  A  P  P  S  P  S  P  K  G  N  S  R  E  E  T  P  K  D  Q  E  M  V  S  D  G  E  G  T  872

2971 TATAGTATTCCCTCTCAAGAAAGGAGGACCAGAGGAAGCTGGAGAGAGTCCAGCTGAGTTGGCTGCTCTCAAGGAGTCTGCCCGTGTTTC 3060
873  I  V  F  P  L  K  K  G  G  P  E  E  A  G  E  S  P  A  E  L  A  A  L  K  E  S  A  R  V  S  902

3061 ATCCTCTGAACAAAACTCAGCCATGCCAGAGGGTGGAGCATCACACAGCAAGTGTCAGACTGGCTCCTCTGGGCTTTGTGACGTGGACAC 3150
903  S  S  E  Q  N  S  A  M  P  E  G  G  A  S  H  S  K  C  Q  T  G  S  S  G  L  C  D  V  D  T  932

3151 TGAGCAGAAGACAGATACTGTCCCCATGAAAGACTCCGCAGCAGAGCCAGTGTCCCCTCCTACCCCAACAGTGGACCGTGACGCAGGGTC 3240
933  E  Q  K  T  D  T  V  P  M  K  D  S  A  A  E  P  V  S  P  P  T  P  T  V  D  R  D  A  G  S  962

3241 ACCAGCTGTAGTGGCCTCCCCTCCTATCACGTTGGCTGAAAACGAGTCTCAGGAAATTGATGAAGATGAAGGCATCCATAGCCATGATGG 3330
963  P  A  V  V  A  S  P  P  I  T  L  A  E  N  E  S  Q  E  I  D  E  D  E  G  I  H  S  H  D  G  992

3331 AAGTGACCTGAGTGACAACATGTCTGAGGGGAGTGACGACTCAGGACTGCACGGGGCTCGGCCGACACCACCAGAAGCTACGTCAAAAAA 3420
993  S  D  L  S  D  N  M  S  E  G  S  D  D  S  G  L  H  G  A  R  P  T  P  P  E  A  T  S  K  N  1022

3421 TGGGAAGGCAGGGTTGGCTGGTAAAGTGACTGAGGGAGAGTTTGTGTGTATTTTCTGTGATCGTTCTTTTAGAAAGGAAAAAGATTATAG 3510
1023 G  K  A  G  L  A  G  K  V  T  E  G  E  F  V  C  I  F  C  D  R  S  F  R  K  E  K  D  Y  S  1052

3511 CAAACACCTCAATCGCCACTTGGTGAATGTGTACTTCCTAGAAGAAGCAGCTGAGGAGCAGGAGGAGCAGGAGGAGCGGGAGGAGCAGGA 3600
1053 K  H  L  N  R  H  L  V  N  V  Y  F  L  E  E  A  A  E  E  Q  E  E  Q  E  E  R  E  E  Q  E  1082

3601 GTAGCTGAGCCTCGGGAGAAGCACCGTGCAGACTTTGTGAGCATGCAATTTTAATTTGTAGACAAACGCAAGCTTGCTTTAATTAGTCTC 3690
1083 *

3691 CAAGGCTGAGTTTTCAGTAACATTCTTTTTCTTAGGACTGTACATCTATTTAGTGTTTGTTGCATAAATCTTAGCAAATCCTCGGGAGTT 3780
3781 AATGTAAGAGGACAGATATGTAACTAGCTCGTGCAGGCAGGTGCAAGGAGAAGGGTAAGATGGTGGAACACACCACTTGCCTTGTCTGCC 3870
3871 TACAACCTGTTGGGTTTTCTTTTCACGGTAGTTCCTAATTTTTAGTTACTTGTTTAGATCGATAAAAATTGGCTTAGTAAATTACTTGAA 3960
3961 GAATTTGCCTGCTTTATATAAATTAAGTTAGCACTTTACAGTTYCTTTAGAGATGAAAAAAAAGAGATTTTAATTGGAGAGAAATTCTCA 4050
4051 ACATTGGACATTGTATCTGTCCAGGTAATTGCTTCCTAACTTGCTATCAATATTTTGTGTTTATATGTTAATCGTTATAAAAAGTGATTT 4140
4141 TTGTTTTTTGGGTATTTTTTATTTTGGTGCTTTTCTGGCTTAAGATGTTGCACATGGTTCTTGTTTTTGTTTCTTTAACCTATGCAGTTA 4230
4231 ATCTCCCTTCCCCTGAAACAGCGTTGTGTTAAATAGTAACACTATACAGATATATGCATGGTTTTTTTTTTTGTTTGTTTGTTTGT 4320
4321 TTTTCCTTTTTGGAGGGATGCTTTTAGGCTTGTTTGCCTCGTSCCGAATTCGATA 4375
```

Figure 2. Alignment of predicted amino acid sequences of mouse and human NRSF. The alignment of mouse to human NRSF was performed using the Gap program (Devereux et al., 1984). The overall sequence identity is 79%.

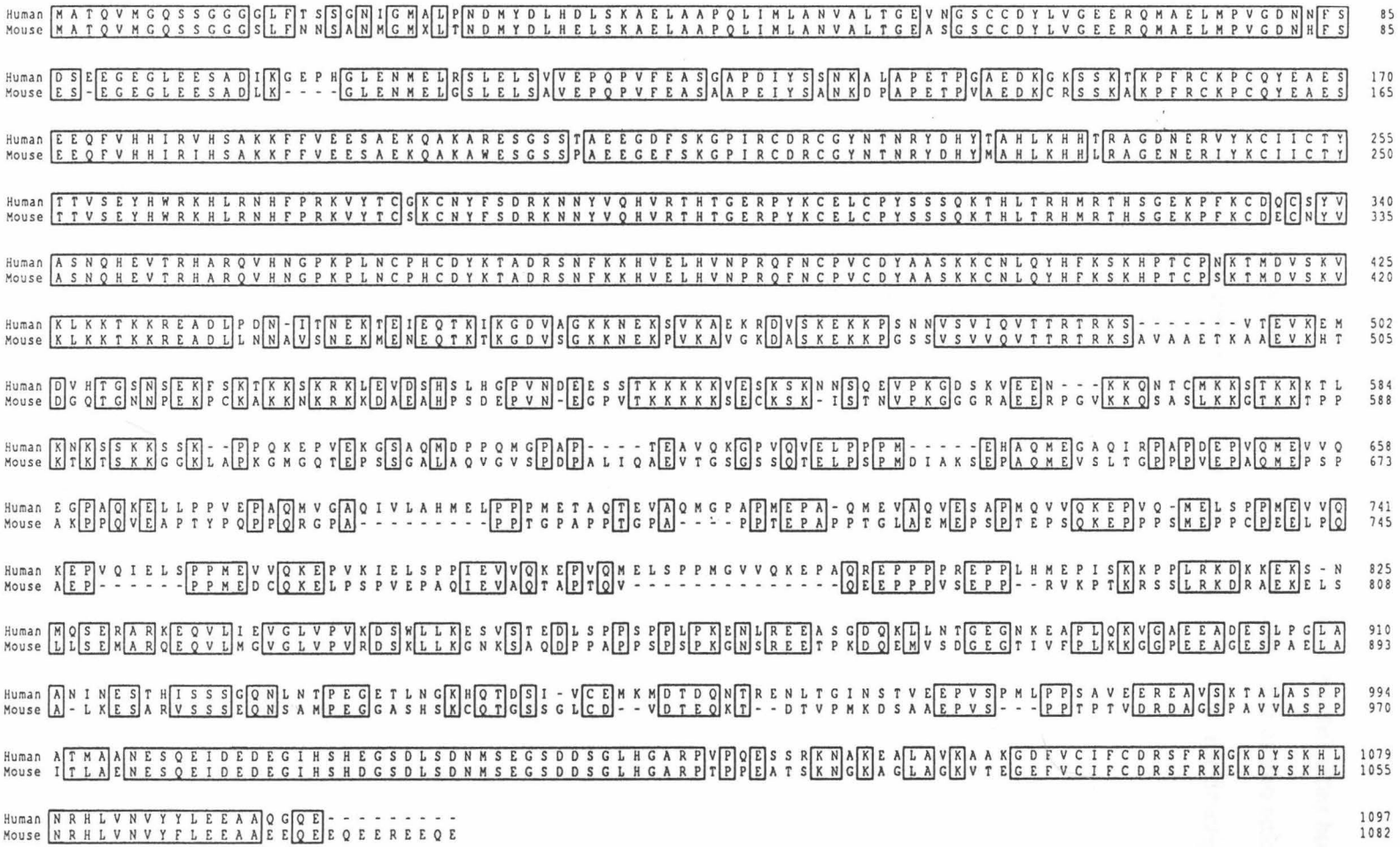

Human MATQVMGQSSGGGGLFTSSGNIGMALPNDMYDLHDLSKAELAAPQLIMLANVALTGEVNGSCCDYLVGEERQMAELMPVGDNNFS 85
Mouse MATQVMGQSSGGGSLFNNSANMGMXLTNDMYDLHELSKAELAAPQLIMLANVALTGEASGSCCDYLVGEERQMAELMPVGDNHFS 85
Human DSEEGEGLEESADIKGEPHGLENMELRSLELSVVEPQPVFEASGAPDIYSSNKALAPETPGAEDKGKSSKTKPFRCKPCQYEAES 170
Mouse ES-EGEGLEESADLK----GLENMELGSLELSAVEPQPVFEASAAPEIYSANKDPAPETPVAEDKCRSSKAKPFRCKPCQYEAES 165
Human EEQFVHHIRVHSAKKFFVEESAEKQAKARESGSSTAEEGDFSKGPIRCDRCGYNTNRYDHYTAHLKHHTRAGDNERVYKCIICTY 255
Mouse EEQFVHHIRIHSAKKFFVEESAEKQAKAWESGSSPAEEGEFSKGPIRCDRCGYNTNRYDHYMAHLKHHLRAGENERIYKCIICTY 250
Human TTVSEYHWRKHLRNHFPRKVYTCGKCNYFSDRKNNYVQHVRTHTGERPYKCELCPYSSSQKTHLTRHMRTHSGEKPFKCDQCSYV 340
Mouse TTVSEYHWRKHLRNHFPRKVYTCSKCNYFSDRKNNYVQHVRTHTGERPYKCELCPYSSSQKTHLTRHMRTHSGEKPFKCDECNYV 335
Human ASNQHEVTRHARQVHNGPKPLNCPHCDYKTADRSNFKKHVELHVNPRQFNCPVCDYAASKKCNLQYHFKSKHPTCPNKTMDVSKV 425
Mouse ASNQHEVTRHARQVHNGPKPLNCPHCDYKTADRSNFKKHVELHVNPRQFNCPVCDYAASKKCNLQYHFKSKHPTCPSKTMDVSKV 420
Human KLKKTKKREADLPDN-ITNEKTEIEQTKIKGDVAGKKNEKSVKAEKRDVSKEKKPSNNVSVIQVTTRTRKS-------VTEVKEM 502
Mouse KLKKTKKREADLLNNAVSNEKMENEQTKTKGDVSGKKNEKPVKAVGKDASKEKKPGSSVSVVQVTTRTRKSAVAAETKAAEVKHT 505
Human DVHTGSNSEKFSKTKKSKRKLEVDSHSLHGPVNDEESSTKKKKKVESKSKNNSQEVPKGDSKVEEN---KKQNTCMKKSTKKKTL 584
Mouse DGQTGNNPEKPCKAKKNKRKKDAEAHPSDEPVN-EGPVTKKKKKSECKSK-ISTNVPKGGGRAEERPGVKKQSASLKKGTKKTPP 588
Human KNKSSKKSSK--PPQKEPVEKGSAQMDPPQMGPAP----TEAVQKGPVQVELPPPM-----EHAQMEGAQIRPAPDEPVQMEVVQ 658
Mouse KTKTSKKGGKLAPKGMGQTEPSSGALAQVGVSPDPALIQAEVTGSGSSQTELPSPMDIAKSEPAQMEVSLTGPPPVEPAQMEPSP 673
Human EGPAQKELLPPVEPAQMVGAQIVLAHMELPPPMETAQTEVAQMGPAPMEPA-QMEVAQVESAPMQVVQKEPVQ-MELSPPMEVVQ 741
Mouse AKPPQVEAPTYPQPPQRGPA---------PPTGPAPPTGPA----PPTEPAPPTGLAEMEPSPTEPSQKEPPPSMEPPCPEELPQ 745
Human KEPVQIELSPPMEVVQKEPVKIELSPPIEVVQKEPVQMELSPPMGVVQKEPAQREPPPPREPPLHMEPISKKPPLRKDKKEKS-N 825
Mouse AEP------PPMEDCQKELPSPVEPAQIEVAQTAPTQV--------------QEEPPPVSEPP--RVKPTKRSSLRKDRAEKELS 808
Human MQSERARKEQVLIEVGLVPVKDSWLLKESVSTEDLSPPSPPLPKENLREEASGDQKLLNTGEGNKEAPLQKVGAEEADESLPGLA 910
Mouse LLSEMARQEQVLMGVGLVPVRDSKLLKGNKSAQDPPAPPSPSPKGNSREETPKDQEMVSDGEGTIVFPLKKGGPEEAGESPAELA 893
Human ANINESTHISSSGQNLNTPEGETLNGKHQTDSI-VCEMKMDTDQNTRENLTGINSTVEEPVSPMLPPSAVEEREAVSKTALASPP 994
Mouse A-LKESARVSSSEQNSAMPEGGASHSKCQTGSSGLCD--VDTEQKT--DTVPMKDSAAEPVS---PPTPTVDRDAGSPAVVASPP 970
Human ATMAANESQEIDEDEGIHSHEGSDLSDNMSEGSDDSGLHGARPVPQESSRKNAKEALAVKAAKGDFVCIFCDRSFRKGKDYSKHL 1079
Mouse ITLAENESQEIDEDEGIHSHDGSDLSDNMSEGSDDSGLHGARPTPPEATSKNGKAGLAGKVTEGEFVCIFCDRSFRKEKDYSKHL 1055
Human NRHLVNVYYLEEAAQGQE--------- 1097
Mouse NRHLVNVYFLEEAAEEQEEREEQE 1082

Figure 3. Schematic diagram of the predicted amino acid sequences from human NRSF and the two mouse NRSF isoforms. The black line represents the amino acid sequence of NRSF. Notable regions of the sequence are highlighted. (See text for description.)

Mouse NRSF Isoforms

Human NRSF

MEVAQ Repeat

Zinc Finger

Proline Rich Repeat

Lysine Ser/thr Region

Figure 4. Tissue-specific expression of two NRSF splice variants. Various tissues and cell lines were analyzed by RT-PCR for the presence of the two identified splicing products. The upper band at 725bp and the lower band at 400bp are the PCR products from the two splice variants. M indicates size standards in base pairs.

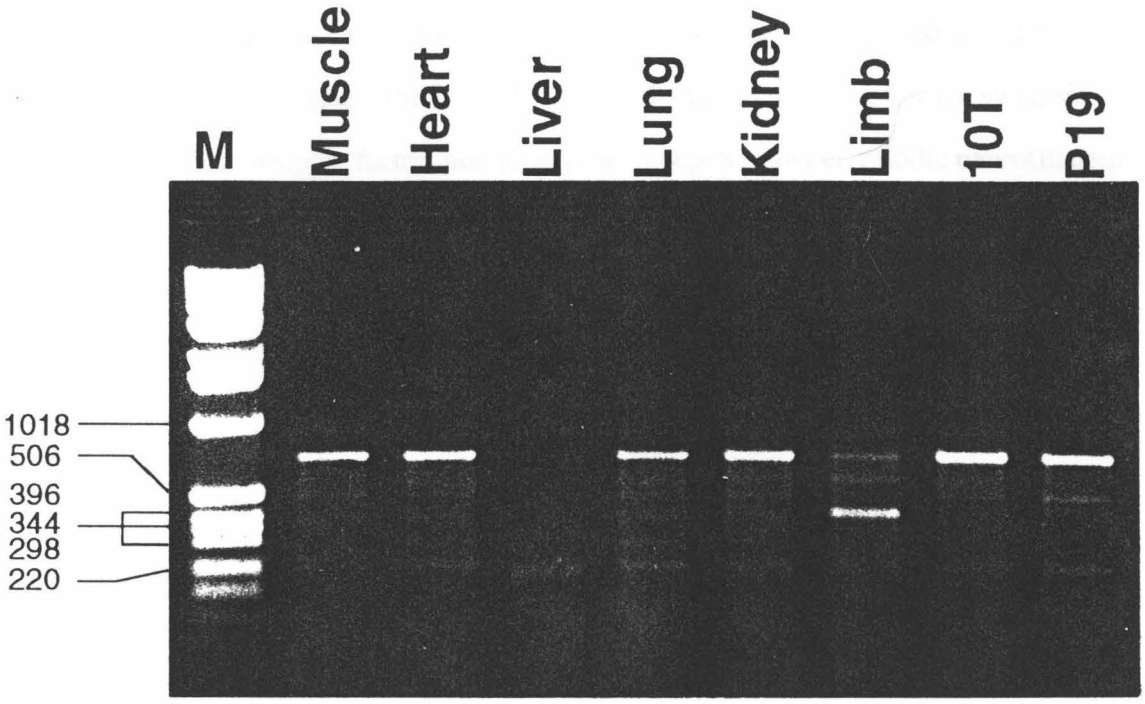
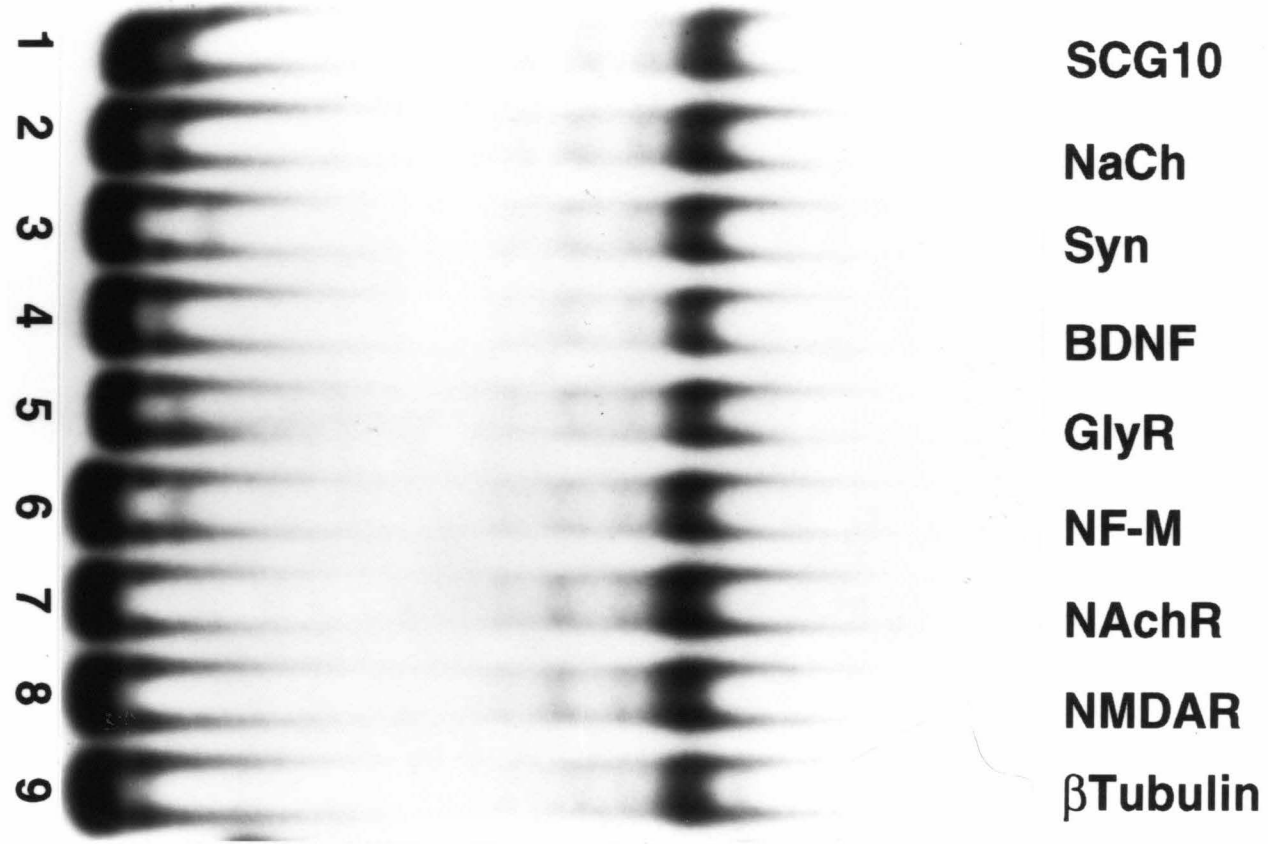

M
Muscle
Heart
Liver
Lung
Kidney
Limb
10T
P19
1018
506
396
344
298
220

Figure 5. Recombinant NRSF recognize NRSEs in nine different neuron-specific genes. Electrophoretic mobility shift assays were performed using in vitro synthesized human NRSF encoded by the λH1 cDNA (Schoenherr and Anderson, 1995). The labeled probes consisted of restriction fragments derived from the following genes (listed in order presented from left to right): rat SCG10, rat type II sodium channel, human synapsin I, rat brain-derived neurotrophic factor, human glycine receptor, chicken middle neurofilament, ß4 subunit of rat nicotinic acetylcholine receptor, rat NMDA1 receptor, and chicken ß4 tubulin.


1
2
3
4
5
6
7
8
9
SCG10
NaCh
Syn
BDNF
GlyR
NF-M
NAchR
NMDAR
βTubulin

Chapter 5

Summary and Future Directions

The neuron-restrictive silencer factor (NRSF) is a negative regulator of neuronal gene expression in non-neuronal cells. In this thesis, I have described the initial characterization of NRSF and its gene. NRSF is a large zinc finger protein with several distinctive protein domains of unknown function. It is capable of binding DNA and repressing transcription in a sequence-specific manner. As predicted from earlier work, its mRNA is present in many non-neuronal cells. Its mRNA is also in neuronal precursors but is absent from embryonic neurons, suggesting that NRSF may inhibit some aspect of neurogenesis. In support of its proposed general role in neuronal gene regulation, potential NRSEs were found in many neuron-specific genes. Included amongst these genes is a transcription factor (P-lim) implicated in activating neurogenesis, further supporting NRSF's proposed role in neuronal development.

NRSF appears to be a unique molecule. To the best of my knowledge, there is no other molecule that directly represses a large battery of tissue specific genes in cells that do not normally express them. There are, however, several indications that it may not stay unique. In yeast, the MATα2 repressor is expressed in α cells and represses a battery of **a** specific genes (Herskowitz, 1989). This gene is probably the most analogous to NRSF as it represses genes specific to a particular cell type. The Polycomb (Pc-G) genes and many early developmental factors in Drosophila also repress a large number of genes. However, Pc-G genes maintain a predetermined state of repression, while NRSF appears more likely to be involved in establishing as well as maintaining repression. Furthermore, their known targets are transcriptional regulators of early pattern formation (Chapter 1). While it may be inaccurate to suggest there is a distinction between regulating pattern formation genes and genes required to execute a cell's ultimate function (end-state genes), it is possible that the mechanisms of repression may be very different. In vertebrates, a negative regulatory system that may prove to be analogous to the NRSF system appears to prevent ectopic activation of a bHLH-regulated immunoglobulin genes (Genetta et al., 1994; Weintraub et al., 1994). In this system,

another zinc finger protein (ZEB; unrelated to NRSF) can inhibit the ectopic activation of an E box containing enhancer present in the immunoglobulin heavy chain gene (Genetta et al., 1994).

Our main interest, however, is in the role of NRSF in development of the nervous system. As I see it, NRSF could have two distinct and not mutually exclusive roles in neuronal differentiation. The first is a corollary to its proposed function in non-neuronal cells. Neuronal precursors express several neuronal markers (L. Sommer, personal communication). This is likely to be a consequence of expressing transcription factors that are the same or closely related to ones present in neurons (Guillemot et al., 1993; Lee et al., 1995). This neuronal character, however, does not usually express itself fully until terminal differentiation. Thus, the need for a repression mechanism specific to neuronal genes would be greatest in neuronal precursors. Therefore, NRSF may act as a fail-safe against precocious expression of end-state genes.

NRSF may also play a more direct role in inhibiting neurogenesis. In this sense, it may act along with other negative regulators of neurogenesis, such as *extra-macrochaete* and the *Notch /Delta* system, to inhibit activators of neuronal development. NRSF regulation, however, appears to be distinct from these pathways and could act in concert with these systems as part of a network of negative regulation that controls neurogenesis. Whether these three systems are overlapping or regulate completely distinct sets of genes remains to be determined.

An alternative model for potential NRSF functions has been suggested by another group studying NRSF. Chong et al. detected NRSF mRNA in adult DRG by Northern analysis and have suggested that NRSF, or REST (for repressor element 1 -silencing transcription factor) as they have named the protein, is expressed in some classes of neurons in the adult animal (Chong et al., 1995). The potential presence of NRSF in mature neurons is invoked to explain the down-regulation of some neuronal genes that occurs in different subpopulations of neurons as they mature. This model, however, is in

conflict with NRSF's regulation of many different neuronal genes. To address this issue, they propose that a whole family of NRSF-like factors might exist. How these factors discriminate between different NRSEs is not discussed. This problem could also be addressed by invoking differences in enhancer architecture between different NRSF-regulated genes that allows down-regulation of some genes with little or no effect on other. This possibility would require proteins that cancel NRSF's repressor activity.

While the evidence supporting this alternative model is indirect and the model itself has difficulty explaining the regulation of many neuronal genes with potential NRSEs, I do believe it should be addressed. The most direct experiment would be to determine if there are any fully mature neurons that express NRSF. Initial attempts to answer this question using monoclonal antibodies against NRSF were unsuccessful (Schoenherr, unpublished). In principle, such an approach should prove fruitful with the proper reagents.

Given the many models for NRSF function, future directions should be to determine which if any of the proposed models will turn out to be true. Our two main assertions about NRSF function are that it represses neuronal genes in non-neuronal cells and that it negatively regulates neuronal development, not just by simply regulating end-state genes, but by repressing genes that activate differentiation, such as transcription factors and growth factor receptors. In some respects, all the evidence supporting the role of NRSF in regulating neuronal genes is circumstantial. Thus, an important avenue of research would be to directly prove our assertions about NRSF function. This can be addressed in several ways.

The obvious first experiment that could address both assertions is to create a null mutation in NRSF using homologous recombination. It is interesting and constructive to consider what the likely phenotypes could be. Starting with the most difficult to analyze, an NRSF mutant could be lethal very early in development, perhaps before implantation. One explanation could be that neuronal genes are now being expressed in all cells, and

this is incompatible with normal development. Another explanation centers on the possibility that NRSF regulates some non-neuronal genes. Their deregulation could also be lethal. Similarly, one potential phenotype would reflect the absence of NRSF expression in embryonic heart and liver (Schoenherr and Anderson, 1995). It is possible that NRSF regulates the development of these organs, defects in which could cause death at a stage too early to see a neuronal phenotype. Such a phenotype would suggest the necessity of renaming NRSF.

Alternatively, it is equally plausible that NRSF mutant embryos will be normal except for a low level of ectopic neuronal gene expression, and we would have a simple confirmation of one of our propositions. Although this phenotype is not dramatic developmentally, it would make a significant contribution to models explaining how genes are kept silent in inappropriate cells. Most models of silencing suppose that there is no distinction between genes specific to one cell type or another. Thus, they focus on mechanisms that can repress all genes. NRSF, along with MATα2, Pc-G genes, and ZEB, suggest that some classes of genes may require specific repression to remain silent. Alternatively, they may represent the first examples of a general phenomenon, as we may find that most genes are under specific repression.

Assuming that NRSF mutant embryos survive to a stage when neurogenesis is taking place, a spectrum of effects on the nervous system could be seen. At one end of the spectrum, loss of NRSF may allow premature differentiation of neurons. This could result either in a significantly smaller number of neurons due to depletion of dividing precursors or problems with patterning caused by disturbing the timing of migration and axon outgrowth. At the other end of the spectrum, loss of NRSF may cause an expansion of neuronal populations by allowing all multipotent neural precursors to choose a neuronal fate. An intermediate phenotype, suggested by the possibility of NRSF-mediated repression of P-lim (Chapter 4), would show defects only in a subset of neuronal populations.

From the possibilities given for NRSF mutant phenotypes, it is clear that alternative methods may be required to assess NRSF's role in neuronal development. With that in mind, overexpression studies in cell culture, mouse embryos, and Xenopus oocytes are underway. Ideally, such studies will be complemented by loss of function experiments. For example, in vitro differentiation of ES stem cells may be an excellent system to study neurogenesis of NRSF deleted cells. Other approaches include antisense techniques and, perhaps, dominant negative perturbations. It seems likely that many different avenues will be required to fully elucidate what NRSF does. And, while my future directions lie elsewhere, it will be more than interesting to see where the future takes NRSF.

REFERENCES

Chong, J. A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J. J., Zheng, Y., Boutros, M. C., Altshuller, Y. M., Frohman, M. A., Kraner, S. D., and Mandel, G. (1995). REST: A mammalian silencer protein that restricts sodium channel expression to neurons. Cell *80*, 949-957.

Genetta, T., Ruezinsky, D., and Kadesch, T. (1994). Displacement of an E-box-binding repressor by basic helix-loop-helix proteins: implications for B-cell specificity of the immunoglobulin heavy chain enhancer. Mol. Cell. Biol. *14*, 6153-6163.

Guillemot, F., Lo, L.-C., Johnson, J. E., Auerbach, A., Anderson, D. J., and Joyner, A. L. (1993). Mammalian achaete-scute homolog-1 is required for the early development of olfactory and autonomic neurons. Cell *75*, 463-476.

Herskowitz, I. (1989). A regulatory hierarchy for cell specialization, in yeast. Nature *342*, 749-757.

Lee, J. E., Hollenberg, S. M., Snider, L., Turner, D. L., Lipnick, N., and Weintraub, H. (1995). Conversion of *Xenopus* extoderm into neurons by NeuroD, a basic helix-loop-helix protein. Science *268*, 836-844.

Schoenherr, C. J., and Anderson, D. J. (1995). The Neuron-Restrictive Silencer Factor (NRSF): A coordinate repressor of multiple neuron-specific genes. Science *267*, 1360-1363.

Weintraub, H., Genetta, T., and Kadesch, T. (1994). Tissue-specific gene activation by MyoD - determination of specificity by cis-acting repression elements. Genes & Dev. *8*, 2203-2211.

Appendix I

The Neuron-Restrictive Silencer Factor (NRSF): A Coordinate Repressor of Multiple Neuron-Specific Genes

Christopher J. Schoenherr and David J. Anderson

The Neuron-Restrictive Silencer Factor (NRSF): A Coordinate Repressor of Multiple Neuron-Specific Genes

Christopher J. Schoenherr and David J. Anderson*

The neuron-restrictive silencer factor (NRSF) binds a DNA sequence element, called the neuron-restrictive silencer element (NRSE), that represses neuronal gene transcription in nonneuronal cells. Consensus NRSEs have been identified in 18 neuron-specific genes. Complementary DNA clones encoding a functional fragment of NRSF were isolated and found to encode a novel protein containing eight noncanonical zinc fingers. Expression of NRSF mRNA was detected in most nonneuronal tissues at several developmental stages. In the nervous system, NRSF mRNA was detected in undifferentiated neuronal progenitors, but not in differentiated neurons. NRSF represents the first example of a vertebrate silencer protein that potentially regulates a large battery of cell type-specific genes, and therefore may function as a master negative regulator of neurogenesis.

The molecular basis of vertebrate neurogenesis is not well understood. To identify transcriptional regulators of neurogenesis we previously analyzed the transcriptional regulation of a neuron-specific gene, *SCG10* (*1*). The SCG10 5′ regulatory region can be dissected into two functional domains: a proximal region that is active in many cell lines and tissues, and a distal region that represses this transcription in nonneuronal cells (*2*, *3*). This distal region satisfies the criteria for a silencer: a sequence analogous to an enhancer but with an opposite effect on transcription (*4*).

A 24-bp (approximately) element is necessary and sufficient for silencing of SCG10 (*5*). Similar sequence elements with functional silencing activity have been identified in other neuron-specific genes: the rat type II sodium (NaII) channel, human synapsin I (*5*–*8*), and neuronal Na,K-ATPase subunit (*9*) genes. These data suggest that a common cis-acting silencer element may mediate the transcriptional repression of multiple neuron-specific genes. We have therefore named this element the neuron-restrictive silencer element (NRSE) (*5*); in the context of the NaII channel gene, it has been called repressor element 1 (RE1) (*7*). The NRSEs in the SCG10, NaII channel, and synapsin I genes all form complexes with a protein, the neuron-restrictive silencer factor (NRSF), present in nonneuronal cell extracts, but absent in neuronal cell extracts (*5*, *7*, *8*).

To isolate a complementary DNA (cDNA) clone encoding NRSF, we screened a HeLa cell λgt11 cDNA expression library (*10*, *11*) with a probe containing three copies of the NaII NRSE (*12*). One phage was identified, λH1, that like native NRSF bound both the S36 and the Na33 probes but not the control Sm36 probe (*5*, *12*). Competition experiments with unlabeled probes in an electrophoretic mobility shift assay (EMSA) confirmed that the sequence specificity of the λH1-encoded protein (*13*) was similar to that of native NRSF in HeLa cell nuclear extracts (Fig. 1, compare lanes 2 through 7 and 10 through 15). Further evidence for a relationship between native and recombinant NRSF was obtained with a mouse polyclonal antibody to recombinant NRSF (anti-NRSF) (*14*). This antibody specifically supershifted a portion of the λH1-encoded protein-DNA complex (Fig. 2B, lanes 1 to 4), as well as a portion of the native NRSF complex (Fig. 2A, lanes 1 to 4). No supershifts were seen with a control ascites (Fig. 2, A and B, lanes 6 to 8). The antigenic similarity of the recombinant and native NRSF proteins provides independent evidence that the cDNA clone encodes a portion of NRSF.

We performed parallel EMSAs with probes containing potential NRSEs from

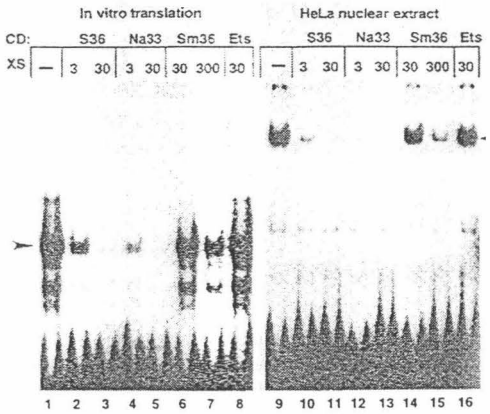


Fig. 1. λH1 encoded protein has the same DNA-binding specificity as native NRSF. EMSAs were performed using a HeLa cell nuclear extract or in vitro translated NRSF (*13*). The probe was a restriction fragment containing two copies of S36. Competitors used were the S36, Na33, and Sm36 oligonucleotides (*12*) and an Ets binding site oligonucleotide (Ets) (*30*). XS indicates molar excess of competitor DNA (CD). The large arrowhead marks the λH1-encoded protein-DNA complex (lane 1), the small arrowhead marks the NRSF-DNA complex (lane 16). The λH1 cDNA does not encode the full-length protein.

C. J. Schoenherr, Division of Biology 216-76, California Institute of Technology, Pasadena, CA 91125, USA.
D. J. Anderson, Howard Hughes Medical Institute, Division of Biology 216-76, California Institute of Technology, Pasadena, CA 91125, USA.
*To whom correspondence should be addressed.

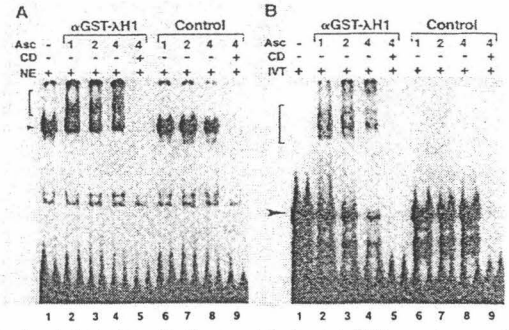


Fig. 2. Antibodies against GST-λH1 recognize the native NRSF-DNA complex. (**A**) The indicated amounts (in μl) of αGST-λH1 (*14*) or a control ascites (Asc) were added to an EMSA containing HeLa nuclear extract (NE). The competitor DNA (CD) was the S36 oligonucleotide present at 300-fold molar excess. (**B**) An EMSA with in vitro translated (IVT) NRSF (*13*). The EMSAs were performed as in (*A*), except that the acrylamide gel used for analysis had an 80:1 acrylamide to bis ratio. Brackets indicate the antibody-supershifted protein-DNA complexes, and the arrowheads the unperturbed complexes. No complexes were formed in a reaction containing Asc alone (*16*).

Table 1. Recombinant NRSF has repressor activity. PC12 cells were cotransfected with reporter plasmids and an expression plasmid containing λHZ4 (*22*). The pCAT3 reporter contains the SCG10 promoter (*2*) fused to the bacterial CAT enzyme; pCAT3-S36++ is pCAT3 with two tandem S36 NRSEs inserted upstream of the SCG10 sequences. The activity of each reporter plasmid in the absence of pCMV-HZ4 was normalized to 100%. The numbers represent the mean ± the standard deviation of two independent experiments performed in duplicate.

Reporter plasmid	pCMV-HZ4 (μg)	CAT activity (%)	Repression
pCAT3.S36−+	0	100	1
	1	8.3 ± 0.6	11.4
	4	3.1 ± 0.3	32
pCAT3	0	100	1
	1	77 ± 0.8	1.3
	4	67.5 ± 4	1.5

*Repression is calculated as 100 ÷ percent CAT activity at a given plasmid concentration.

the synapsin I and brain-derived neurotrophic factor (BDNF) (*8*, *15*) as well as the SCG10 and NaII channel genes. Native NRSF yielded similarly sized complexes with all four probes (Fig. 3, lanes 1 to 4). A portion of these four complexes could be supershifted by the antibody to NRSF (*16*), and all four probes bound recombinant NRSF (Fig. 3, lanes 5 to 8). Thus both native and recombinant NRSF were able to interact with putative NRSEs in multiple neuron-specific genes. Additional consensus NRSEs were identified in at least 14 other neuronal genes by a nucleotide database search (*17*).

To isolate longer NRSF cDNA clones, multiple cDNA libraries were screened using a λH1 probe (*18*). Although Northern blots indicated that the NRSF mRNA is 7 to 8 kb (*16*), we were unable to isolate NRSF cDNAs <2 kb, perhaps reflecting a strong stop to reverse-transcription. The sequence of the longest human clone obtained, λHZ4 (2.04 kb), has a continuous open reading frame (*19*) that encodes a novel protein containing eight zinc fingers of the C_2H_2 class with interfinger sequences that place NRSF in the GLI-Krüppel family of zinc finger proteins (Fig. 4, A and B) (*20*, *21*). However, these zinc fingers contain a conserved tyrosine residue absent from the canonical finger sequence (Fig. 4B, dashed box). COOH-terminal to the zinc fingers is a 174–amino acid domain rich in lysine (26%; 46 of 174) and serine or threonine (21 percent; 37 of 174; Fig. 4A).

To determine whether the longest NRSF cDNA encoded a protein with transcriptional repressing activity, we transfected an expression vector containing λHZ4 (pCMV-HZ4) into PC12 cells (which do not contain NRSF activity) together with various target plasmids (*22*). Increasing amounts of pCMV-HZ4 repressed transcription from an NRSE-containing target plasmid from 11 to 32 times (Table 1). In control transfections performed with a target plasmid lacking an NRSE or containing a mutated (*5*) NRSE, the repression was only 1.5 times at the maximal pCMV-HZ4 concentration (Table 1) (*16*). These results indicated that the λHZ4 clone contains at least a portion of the transcriptional repression domain and that this repression requires NRSE-binding.

The absence of NRSF activity in neuronal cells (*2*, *5*–*7*) could reflect a lack of NRSF gene expression or an inactivation of NRSF. To distinguish between these possibilities, we performed RNase protection assays (*23*) on several neuronal and nonneuronal cell lines. No NRSF transcripts were detectable in two neuronal cell lines, MAH and PC12 (Fig. 5, lanes 4 and 5; rNRSF). In contrast several glial and two fibroblast cell lines expressed NRSF mRNA (Fig. 5, lanes 6 to 9). These data indicated that the absence of NRSF activity in neuronal cells is due to a lack of NRSF expression, not to its functional inactivation.

Using a mouse NRSF cDNA clone (*16*)

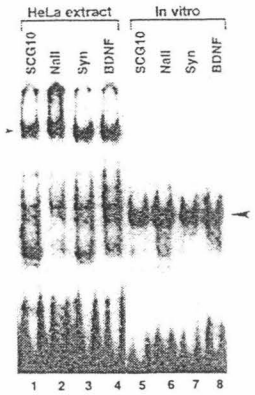


Fig. 3. Native and recombinant NRSF recognize NRSEs in four different neuron-specific genes. EMSAs were performed using either HeLa nuclear extract (lanes 1 to 4) or in vitro synthesized NRSF (lanes 5 to 8). The probes contained NRSEs from the SCG10 (lanes 1 and 5); NaII channel (lanes 2 and 6); synapsin I (lanes 3 and 7); or the BDNF (lanes 4 and 8) genes. The large and small arrowheads indicate the specific complexes obtained with recombinant and native NRSFs, respectively.

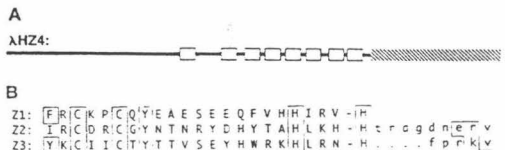


Fig. 4. (**A**) Schematic diagram of the predicted amino acid sequence (*19*) from the NRSF λHZ4 cDNA clone. Stippled boxes indicate the position of zinc fingers, cross-hatched region a domain rich in basic amino acids. (**B**) Alignment of NRSF zinc finger and interfinger sequences. The eight zinc fingers of human NRSF were aligned beginning with the conserved aromatic residue and including the interfinger sequences of fingers z2-7. The consensus (Cons) for GLI-Kruppel zinc fingers and interfinger sequences is shown for comparison.

as a probe, we next performed in situ hybridization experiments on mouse embryos (*24*). At E12.5, NRSF mRNA was detected in the ventricular zone of the neural tube (Fig. 6A, arrow), a region containing multipotential progenitors of neurons and glia (*25*), which do not express SCG10 mRNA (compare Fig. 6B, arrow). In contrast, the adjacent marginal zone of the neural tube which contains SCG10 positive neurons (Fig. 6B, open arrow) was largely devoid of NRSF expression (Fig. 6A, open arrow) NRSF mRNA was also detected in the ventricular zone of the brain (Fig. 6E, arrowhead). In the peripheral nervous system, NRSF mRNA was absent or expressed at low levels in sympathetic and dorsal root sensory ganglia (DRG) at E13.5 (Fig. 6C, small and large arrowheads), whereas these ganglia expressed SCG10 mRNA (Fig. 6D, small and large arrowheads). Thus, these data suggest that NRSF is expressed by undifferentiated neuronal progenitors but not by differentiated neurons in vivo.

The SCG10 NRSE is required to prevent expression in multiple nonneural tissues throughout development (*3*). This broad requirement for the NRSE was reflected in a broad expression of NRSF mRNA. The NRSF mRNA was detected in many embryonic nonneural tissues such as the adrenal gland, aorta, genital tubercle, gut, kidney, lung, ovaries, pancreas, parathyroid gland, skeletal muscle, testes, thymus, tongue, and umbilical cord (Fig. 6, E and F) (*16*). RNase protection revealed NRSF transcripts in many adult nonneuronal tissues, including heart and liver (*16*), which expressed little NRSF mRNA in embryos (Fig. 6). This broad expression pattern is consistent with a role for NRSF as a near-ubiquitous negative regulator of neuron-specific gene expression.

Four lines of evidence support the conclusion that our cDNA clones encode a functional fragment of authentic NRSF. First, recombinant and native NRSFs showed similar in vitro DNA binding specificities. Second, antibodies generated against recombinant NRSF bound to native NRSF. Third, the presence or absence of NRSF mRNA in cell lines paralleled both NRSE-dependent silencing activity and NRSF DNA-binding activity in nuclear extracts. Fourth, the longest NRSF cDNA clone repressed transcription in

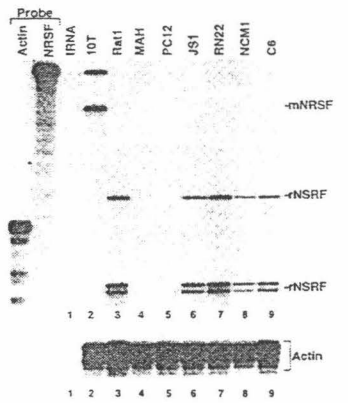


Fig. 5. Analysis of NRSF message in neuronal and nonneuronal cell lines. RNase protection assays (*23*) were performed on total RNA from various cell lines. The two neuronal cell lines were MAH (*31*) and PC12 (*32*). The glial lines were: RN22 (*33*), JS-1 (*34*), NCM-1 (*35*), and C6 (*36*); the fibroblast lines were Rat1 and mouse C3H10T 1/2 (10T). "tRNA" indicates a negative control. The probes were derived from mouse NRSF and rat β-actin cDNAs. rNRSF and mNRSF indicate the protected products obtained with RNA from rat or mouse cell lines, respectively. The size difference between mNRSF and rNRSF most likely reflects an incomplete protection of the mouse probe by the rat transcript.

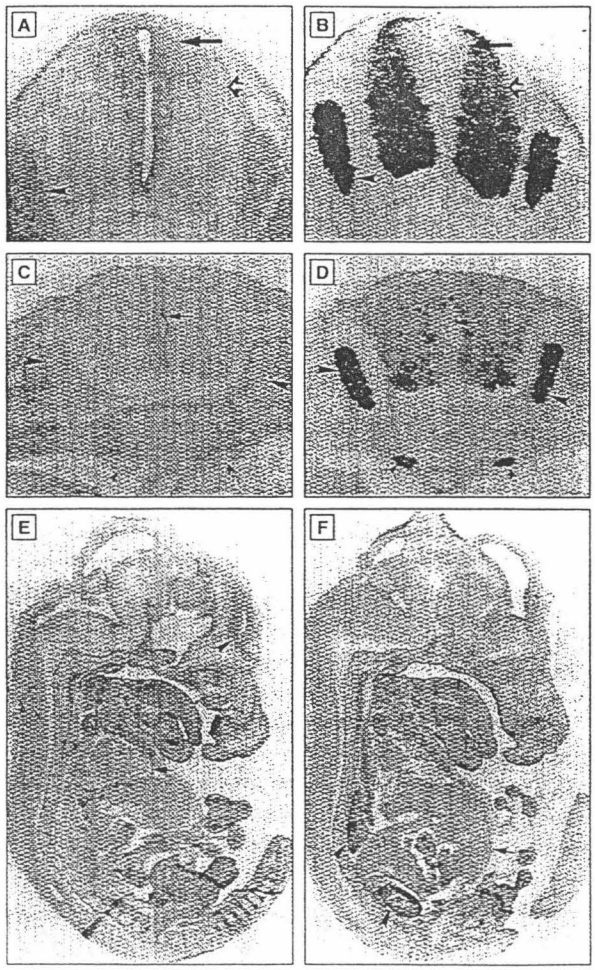


Fig. 6. (**A** to **D**) Comparison of NRSF and SCG10 mRNA expression by nonradioactive in situ hybridization (*24*). Adjacent transverse sections of E12.5 (A and B) and E13.5 (**C** and **D**) mouse embryos were hybridized with NRSF (A and C) or SCG10 (B and D) antisense probes. The solid and open arrows (A to D) indicate the ventricular and marginal zones of the neural tube, respectively. The large and small arrowheads (A to D) indicate the sensory and sympathetic ganglia, respectively. Control hybridizations with NRSF sense probes revealed no specific signal (*16*). (**E** and **F**) Widespread expression of NRSF mRNA in nonneural tissues. In situ hybridization with an NRSF antisense probe was performed on parasaggital sections of an E13.5 mouse embryo. Arrowheads mark several positive tissues, the arrows negative tissues.

vivo in an NRSE-dependent manner.

Functional NRSEs have been identified in four neuron-specific genes: SCG10, NaII channel, synapsin I (*5–8*) and neuronal Na,K-ATPase subunit (*9*), while 14 other neuronal genes contain consensus NRSEs (*17*). Although silencer function has not yet been demonstrated for these potential NRSEs, native and recombinant NRSF bound to six of these sequences (Fig. 3) (*16*), and previous data indicate a strong correlation between NRSF binding and silencing activity (*5*, *7*). We therefore conclude that NRSF may silence at least 18 neuron-specific promoters. Thus NRSF may be the first vertebrate silencer factor that coordinately represses a battery of cell type–specific genes. This would provide experimental support for the idea that the maintenance of the differentiated state involves active negative regulation of gene expression (*26*).

In other systems, positive-acting transcription factors that regulate multiple lineage-specific target genes have been shown to function as master regulators of cell type determination or differentiation (*27–29*). By analogy, NRSF may function as a master negative regulator of the neuronal phenotype. Specifically, the presence of NRSF in neuronal progenitors, together with its proposed coordinate negative regulation of many neuronal genes, suggests that relief from NRSF-imposed repression may be a key event in neurogenesis. The identification of NRSF therefore provides an opportunity to further understand the control of this event.

REFERENCES AND NOTES

1. R. Stein, N. Mori, K. Matthews, L.-C. Lo, D. J. Anderson, *Neuron* **1**, 463 (1988).
2. N. Mori, R. Stein, O. Sigmund, D. J. Anderson, *ibid.* **4**, 583 (1990).
3. C. W. Wuenschell, N. Mori, D. J. Anderson, *ibid.*, p. 595.
4. A. H. Brand, L. Breeden, J. Abraham, R. Sternglanz, K. Nasmyth, *Cell* **41**, 41 (1985).
5. N. Mori, C. Schoenherr, D. J. Vandenbergh, D. J. Anderson, *Neuron* **9**, 1 (1992).
6. R. A. Maue, S. D. Kraner, R. H. Goodman, G. Mandel, *ibid.* **4**, 223 (1990).
7. S. D. Kraner, J. A. Chong, H. J. Tsay, G. Mandel, *ibid.* **9**, 37 (1992).
8. L. Li, T. Suzuki, N. Mori, P. Greengard, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1460.
9. B. G. Pathak, J. C. Neumann, M. L. Croyle, J. B. Lingrel, *Nucleic Acids Res.* **22**, 4748 (1994).
10. H. Singh, J. H. LeBowitz, A. S. Baldwin Jr., P. A. Sharp, *Cell* **52**, 415 (1988).
11. C. R. Vinson, K. L. LaMarco, P. F. Johnson, W. H. Landschulz, S. L. McKnight, *Genes Dev.* **2**, 801 (1988).
12. S36 and Na33 are the NRSE elements present in the SCG10 and NaII channel genes, respectively. The Sm36 sequence contains two point mutations in the S36 sequence and has an approximately 100 times lower affinity for NRSF. For oligonucleotides, the sequence of the top strands of the oligonucleotides used for library screening and EMSAs are given below. The upper case sequences represent actual genomic sequence, the lower case sequences were used for cloning purposes: S36: agctGCAAAGCCATTTCAGCACCACGGAGAGTGCCTCTGC; Na33: agcttATTGGGTTTCAGAACCACGGACAGCACCAGAGTa; Syn: agcttCTGCCAGCTTCAGCACCGCGGACAGTGCCTTCGCa; BDNF: agcttAGAGTCCATTCAGCACCTTGGACAGAGCCAGCGGa; Ets: agcttGCGGAACGGAAGCGGAAACCGa.
13. The λH1 cDNA was subcloned into the Eco RI site of pRSET B (Invitrogen). Recombinant λH1 was produced by coupled in vitro transcription and translation with a rabbit reticulocyte lysate according to manufacturer's protocol (Promega). Mobility shift assays were performed as described [N. Mori, C. Schoenherr, D. J. Vandenbergh, D. J. Anderson, *Neuron* **9**, 45 (1992)], except that 0.5 μg supercoiled plasmid and 10 μg of BSA were included in each reaction.
14. The λH1 cDNA was inserted into the Eco RI site of pGEX-1, a prokaryotic glutathione *S*-transferase fusion expression vector [D. B. Smith and K. S. Johnson, *Gene* **67**, 31 (1988)]. Gel slices containing the fusion protein were used as antigen to produce a polyclonal mouse ascites as described [S. K. H. Ou, J. M. C. Hwang, P. H. Patterson, *J. Immunol. Meth.* **165**, 75 (1993)].
15. T. Timmusk *et al.*, *Neuron* **10**, 475 (1993).
16. C. S. Schoenherr and D. J. Anderson, unpublished observations.
17. A search of the Genbank database with three different algorithms [J. D. Devereux, P. Haeberli, O. Smithies, *Nucleic Acids Res.* **12**, 387 (1984); S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990)] identified 14 additional neuronal genes that show, on average, 93 percent sequence similarity to the consensus NRSE. These genes include NMDA, ACh, and glycine receptor subunits, neurofilament and neuron-specific tubulin. These database searches also revealed NRSE-like sequences in several nonneuronal genes. The average similarity was only 84%, however, compared to 93% for the neuronal genes.
18. Five different cDNA libraries, derived from human, mouse, and rat tissues were screened by plaque hybridization. The selection of libraries included those made with inserts size-selected for length greater than 4 kb. No cDNA isolated from any library extended beyond the 5′ end of clone λHZ4.
19. The sequence of the λHZ4 human NRSF cDNA has been submitted to Genbank, under accession no. U13879.
20. R. Schuh, *et al.*, *Cell* **47**, 1025-1032 (1986).
21. J. M. Ruppert *et al.*, *Molecular and Cellular Biology* **8**, 3104-3113 (1988).
22. The λHZ4 cDNA was cloned into the Eco RI site of a modified pcDNA3 (Invitrogen) to generate pCMV-λHZ4. Transient transfections of PC12 cells were performed essentially as described [N. Mori, C. Schoenherr, D. J. Vandenbergh, D. J. Anderson, *Neuron* **9**, 45 (1992)]. Each cotransfection included 5 μg of a reporter plasmid (pCAT3 or pCAT3-S36++), pCMV-λHZ4 or pcDNA3 at the concentrations indicated, 2 μg of pRSV-lacZ to normalize transfections, and pBluescript to a total of 10 μg of DNA. Cells were harvested 48 hours after transfection and processed for CAT and β-galactosidase assays as described [N. Mori, R. Stein, O. Sigmund, D. J. Anderson, *Neuron* **4**, 583 (1990)], except CAT assays were quantified with a Phosphor Imager (Molecular Dynamics).
23. Ribonuclease protections were performed essentially as described [J. E. Johnson, K. Zimmerman, T. Saito, D. J. Anderson, *Development* **114**, 75 (1992)]. Total cellular RNA was isolated with the acid phenol method [P. Chomczynski and N. Sacchi, *Anal. Biochem.* **162**, 156 (1987)].
24. Fixation, embedding, sectioning, preparation of digoxygenin-labeled cRNA probes, and in situ hybridization with nonradioactive detection were performed as described [S. J. Birren, L. C. Lo, D. J. Anderson, *Development* **119**, 597 (1993)] with the use of sense and antisense probes for mouse NRSF.
25. S. M. Leber, S. M. Breedlove, J. R. Sanes, *J. Neurosci.* **10**, 2451 (1990).
26. H. M. Blau, *Annu. Rev. Biochem.* **61**, 1213 (1991).
27. S. Li *et al.*, *Nature* **347**, 528 (1990); L. Pevny *et al.*, *ibid.* **349**, 257 (1991).
28. L. M. Corcoran *et al.*, *Genes Dev.* **7**, 570 (1993).
29. H. Weintraub *et al.*, *Science* **251**, 761 (1991).
30. K. Lamarco, C. C. Thompson, B. P. Byers, E. M. Walton, S. L. McKnight, *ibid.* **253**, 789 (1991).
31. S. J. Birren and D. J. Anderson, *Neuron* **4**, 189 (1990).
32. L. A. Greene and A. S. Tischler, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2424 (1976).
33. S. E. Pfeiffer, B. Betschart, J. Cook, P. E. Mancini, R. J. Morris, in *Glial Cell Lines*, S. Federoff and L. Hertz, Eds. (Academic Press, New York, 1978), pp. 287–346.
34. H. Kimura, W. H. Fischer, D. Schubert, *Nature* **348**, 257 (1990).
35. L.-C. Lo, S. J. Birren, D. J. Anderson, *Dev. Biol.* **145**, 139 (1990).
36. S. Kumar *et al.*, *J. Neurosci. Res.* **27**, 408 (1990).

37. We thank T. Kadesch and P. Henthorn for cDNA libraries, N. Mori for HeLa nuclear extract, S. Padilla for technical assistance, S. Ou for antibody preparation, S. Tavtigian and B. Wold for cDNA libraries, L. Lo, A. Groves, and L. Banner for valuable advice, M.-J. Fann and P. Patterson for actin probes, and B. Wold, L. Sommer, and S. Bultman for comments on the manuscript. Supported by NIH grant NS23476 and an Associate Investigatorship of the Howard Hughes Medical Institute (D.J.A.).

31 October 1994; accepted 1 February 1995

Appendix II

Silencing is golden: Negative regulation in the control of neuronal gene transcription

Christopher J. Schoenherr and David J. Anderson

Summary: Recent work has identified negative-acting DNA regulatory elements that function to prevent the expression of neuronal genes in non-neuronal cell types or in inappropriate neuronal subtypes. In some cases, the protein factors that interact with these silencer elements have been isolated and characterized. These data suggest that negative regulation plays a major role in determining the diverse patterns of gene expression within the nervous system.

INTRODUCTION

The study of the transcriptional regulation of neuronal genes is of fundamental importance to understanding the differentiation, diversification, survival and plasticity of neurons. During neurogenesis progenitor cells must select between neuronal vs. non-neuronal fates; in the former case they must also select among a large repertoire of neuronal subtype fates. These developmental decisions require the action of transcriptional regulatory proteins, as indicated by the genetic analysis of neurogenesis in both vertebrates and invertebrates (Jan and Jan, 1994; Joyner and Guillemot, 1994). Transcriptional regulation is important in the function of neurons as well as in their genesis. For example, some forms of neuromodulation such as LTP and sensitization are dependent on the transcriptional induction and, possibly, repression of certain genes (Silva and Giese, 1994). Thus, to fully understand the molecular mechanisms that underly not only neuronal development but also neuronal plasticity, an analysis of the transcriptional regulatory components of these processes is essential.

Over the last decade, a great deal of work has been done to define the *cis*-acting DNA elements and *trans*-acting protein factors involved in the transcriptional regulation of tissue-specific genes in many non-neuronal cell types, such as erythrocytes, lymphocytes and liver cells (Johnson and McKnight, 1989; Maniatis et al., 1987). Most of these studies have emphasized the role of cell type- or tissue-specific positively-acting transcription factors in controlling gene expression. Such positively-acting factors have also been

identified in the study of neuronal gene expression (He and Rosenfeld, 1991). However, recent studies have revealed that negative regulation plays a significant role in the control of neuron-specific gene expression as well. This review, therefore, will focus on the transcriptional regulation of neuronal genes by negatively-acting DNA sequences and their associated regulatory proteins.

Negative regulatory elements that repress neuronal genes in non-neuronal cells

Because there are few cell lines that accurately recapitulate patterns of neuronal gene expression, particularly in the CNS, DNA regulatory elements in neuronal genes are most reliably assayed in transgenic mice. A number of such studies have reported that deletion of certain regulatory domains from reporter constructs exhibiting neuron-specific expression resulted in the ectopic expression of the reporter gene in non-neuronal tissues. Genes presenting this type of result include the growth-associated proteins SCG10 (Vandenbergh et al., 1989; Wuenschell et al., 1990) and GAP-43 (Vanselow et al., 1994); a neuron-specific a subunit of the Na-K ATPase (Pathak et al., 1994); hypoxanthine phosphoribosyltransferase (HPRT) (Rincón-Limas et al., 1994); and corticotropin-releasing hormone (CRH) (Keegan et al., 1994). Taken together, such results suggest the existence of negative-acting regulatory domains that function to repress transcription of neuronal genes in non-neuronal cells, and imply that such a repressive mechanism is common to many neuronal genes.

Evidence for negative regulation of neuronal genes has also been obtained from transient transfection experiments in cell lines. For example, the upstream negative regulatory domain in the *SCG10* gene was shown to repress transcription in HeLa but not in PC12 cells. Moreover, this region was shown to be orientation-independent, relatively position-insensitive and able to repress transcription from a heterologous promoter as well as from the SCG10 promoter (Mori et al., 1990). Similar data were obtained from analysis of the type II sodium channel (NaII) gene, which is expressed in neurons but not in muscle

(Maue et al., 1990). These features are a defining characteristic of silencers, elements first defined in yeast which behave similarly as enhancers but which have an opposite effect on transcription (Brand et al., 1985). More recently, the analysis of several other neuronal genes using transient transfection assays has revealed evidence of negative regulatory elements that restrict expression to neurons. These include the alpha1-chimaerin (Dong et al., 1995), VGF (Possenti et al., 1992), DBH (Ishiguro et al., 1993; Shaskus et al., 1995) and rat growth hormone genes (Guérin et al., 1993).

Negative regulatory elements that repress neuronal genes in neuronal subtypes

The experiments mentioned above reveal a common theme in the regulation of a number of neuronal genes: expression is restricted to neurons, at least in part, by negative-acting sequences that repress transcription in non-neuronal cell types. There is also evidence that negative-acting *cis*-elements may function to restrict the expression of neuron-specific genes amongst different neuronal subtypes. Such genes can include those encoding neurotransmitter-synthetic enzymes, ion channels, receptors and neuropeptides. Perhaps the most exemplary case derives from analysis of the regulatory domains in the dopamine-ß-hydroxylase (DBH) gene (Kapur et al., 1991; Mercer et al., 1991). DBH is normally expressed by noradrenergic neurons, but not dopaminergic neurons, in both the CNS and PNS. Analysis of DBH regulatory regions in transgenic mice provided evidence for two types of negative-acting *cis*-elements: those that repress expression in dopaminergic neurons, and those that repress expression in non-catecholaminergic neurons in the CNS (Hoyle et al., 1994). The requirement for a negative-acting element to repress expression in dopaminergic neurons was suggested to reflect the existence of shared positive-acting elements in genes expressed in dopaminergic and noradrenergic neurons, such as DBH and TH (Hoyle et al., 1994).

Negative regulatory elements in the DBH gene have also been identified in transient transfection assays (Ishiguro et al., 1993; Shaskus et al., 1995). These elements were

shown to repress expression in non-neuronal cell lines, but neuronal cell lines of non-catecholaminergic origin were not examined. However, taken together with the transgenic experiments, the data suggest that the DBH gene contains negative-acting elements that repress transcription in non-neuronal cell types as well as in inappropriate neuronal cell types; whether these elements are one and the same remains to be determined. In the case of the choline acetyltransferase (ChAT) gene, a negative regulatory domain repressing expression in both non-neuronal and in non-cholinergic neuronal cell lines was demonstrated in transient transfection assays (Ibáñez and Persson, 1991; Lönnerberg et al., 1995). Again, whether distinct silencers for each function can be separated within this domain remains to be determined.

Trans-acting factors that repress neuronal genes in non-neuronal cells: identification of NRSF/REST

A fine-structural analysis of negative-acting *cis*-elements is a prerequisite to identify the proteins that interact with these elements. Such an analysis has been performed for the silencers in the SCG10 (Mori et al., 1992) and NaII channel (Kraner et al., 1992) genes. Surprisingly, a comparison of the silencer elements delineated by deletional analysis in these two genes revealed considerable similarity (18/21 identity), suggesting that they might bind the same protein (Mori et al., 1992). This idea was supported by the results of in vitro DNA binding assays, which revealed that both silencers bound a factor present only in extracts from nonneuronal cells, and absent from neuronal cell extracts, parallelling the presence or absence of repressing activity. As the silencer element in the SCG10 and Na II channel genes appeared to bind the same factor and to have the same function, it was named the neuron-restrictive silencer element (NRSE) (Mori et al., 1992) or repressor element-1 (RE-1) (Kraner et al., 1992). The protein that binds to this element was named the neuron restrictive silencer factor (NRSF) or repressor element 1 silencing transcription factor (REST). Further work identified an NRSE/RE-1-like element in the synapsin I gene; this element also behaved like a silencer and bound a protein specifically in non-neuronal

cells (Li et al., 1993). These data suggested that NRSF/REST might bind a similar sequence in three different neuronal genes.

To obtain further information about the structure, function and regulation of NRSF/REST, it was important to isolate the gene encoding this protein. cDNAs encoding NRSF/REST were isolated independently in two different laboratories (Chong et al., 1995; Schoenherr and Anderson, 1995). The predicted protein encoded by the apparently full-length human cDNA has an molecular weight of 116 Kd and contains nine zinc fingers (Chong et al., 1995), with eight fingers clustered near the amino terminus and one at the carboxyl terminus. These zinc fingers are related, but not identical, to the consensus sequence for zinc finger proteins in the Gli-Kruppel family. The ninth zinc finger is apparently not required for NRSE/RE1 binding as truncated versions of NRSF/REST lacking this ninth finger still bind with high affinity and have silencing activity in transfected cells (Schoenherr and Anderson, 1995). The human protein also has two other notable features: a lysine- and serine/threonine- rich region just C-terminal to the zinc finger domain, and six tandem repeats of a novel protein sequence (Chong et al., 1995). The functions of these domains are unknown but they may represent regions of the protein that function in repression, analagous to the transcriptional activation domains present in enhancer-binding factors. Cloned NRSF/REST binds to NRSE-like elements in the SCG10, Na II channel and synapsin I as well as BDNF genes (Schoenherr and Anderson, 1995), confirming that these elements in four different neuronal genes could interact with a common protein.

To determine if any other neuronal genes might be regulated by NRSF/REST, an extensive search of a DNA sequence database was performed using a consensus sequence drawn from the SCG10, Na II channel, BDNF and synapsin I NRSE/RE1s. This search identified at least 20 neuronal genes, including channels, receptors, cytoskeletal and synaptic proteins, with NRSE/RE1-like sequences in regulatory regions (Schoenherr and Anderson, 1995). The majority of these sequences are able to interact with native and

recombinant NRSF, as well to silence transcription from an heterologous promoter (C. Schoenherr, A.J. Paquette and D.J. Anderson, unpublished data). Taken together with the data from the SCG10, Na II channel and synapsin I NRSEs, these results suggest that NRSF/REST can regulate a large number of neuronal genes, and therefore that silencing may represent a general mechanism for restricting neuronal gene expression to the nervous system.

Possible biological functions of NRSF/REST

Previous studies of the SCG10 gene in transgenic mice indicated that the NRSE-containing distal regulatory domain was required to repress inappropriate expression in most or all non-neuronal tissues examined (Mori et al., 1990). This raised the question of whether NRSF/REST is responsible for such global negative regulation. In situ hybridization analysis of NRSF/REST mRNA in sections of mouse embryos revealed detectable expression in most nonneuronal cell types examined; conversely, NRSF/REST mRNA was absent from neuronal cells examined at these stages (E11.5 - E13.5) (Chong et al., 1995; Schoenherr and Anderson, 1995). At present, the earliest stage of embryogenesis at which NRSF/REST mRNA is first detectable has not yet been identified. Nevertheless, the fact that many nonneuronal tissues continue to express NRSF/REST mRNA from embryogenesis into adulthood (C. Schoenherr and D.J. Anderson, unpublished data) suggests that NRSF/REST is involved in the maintenance, and not just the initiation, of neuronal gene repression in non-neuronal tissues. This, taken together with the apparently large battery of neuronal genes that contain NRSE/RE1-like elements (see above), suggests that NRSF/REST represents a global negative regulator of neuronal gene expression in vertebrates. Not all nonneuronal cell types express NRSF/REST mRNA, however; both embryonic heart and liver appear to contain little or no transcripts, although these tissues do express NRSF/REST in the adult. Since embryonic heart and liver do not express neuronal genes, this result suggests that NRSF/REST is not required

in all non-neuronal cell types at all times in development to repress neuronal gene expression.

It has been suggested that NRSF/REST is responsible not only for silencing neuronal genes in nonneuronal cells but also for repressing the type II Na channel gene in maturing peripheral neurons (Chong et al., 1995). This raises the question of how pan-neuronal genes which contain NRSE/RE-1s, such as SCG10 and synapsin I, would escape repression by NRSF/REST in these sensory neurons. Chong et al. (Chong et al., 1995) suggest that this may reflect the existence of multiple NRSF/REST-like proteins, but this remains to be demonstrated. Finally, the presence of NRSF/REST in the ventricular zone of the neural tube (Chong et al., 1995; Schoenherr and Anderson, 1995), where undifferentiated neuronal progenitors are located, suggests a possible function for this negative regulator in neurogenesis. Specifically, relief from NRSF/REST-imposed repression may be important in either the determination or the differentiation of neurons. Loss- or gain-of-function perturbations in NRSF/REST will be required to test this hypothesis.

De-repression of neuronal genes by NGF in differentiating PC12 cells

While it appears that transcriptional repression is important in achieving cell type-specific gene expression in the nervous system, repression has also been implicated in the modulation of transcription within a given cell type, by environmental signals such as neurotrophins. For example, many neuron-specific genes, such as peripherin and SCG10, are expressed at a low level in undifferentiated PC12 cells, and become up-regulated three- to five-fold upon NGF treatment (Leonard et al., 1987; Stein et al., 1988). In the case of peripherin, this up-regulation is due, in part, to relief from repression in uninduced cells imposed by a negative regulatory element (NRE - not to be confused with the NRSE) present in the peripherin gene (Thompson et al., 1992). Interestingly, the NRE interacts with proteins in both uninduced and induced PC12 cell extracts. However, the protein:DNA complex obtained from uninduced cells is larger than that from induced cells.

Cloning of the gene encoding an NRE-binding protein revealed it to be a member of the CCAAT transcription factor/nuclear factor-1 (CTF/NF-1) transcription factor family, called NF1-L (Adams et al., 1995). The size of the complex formed by recombinant NF1-L is similar to that obtained from NGF-treated PC12 cell nuclear extracts. Addition of extract from uninduced cells to recombinant NF1-L converted the complex to a size comparable to that detected in uninduced cell extracts (Adams et al., 1995). This suggests that uninduced extracts contain a protein that interacts with NF1-L and converts it to a negative regulator of peripherin transcription; NGF treatment would then cause dissociation of this subunit from NF1-L, allowing full induction of peripherin transcription. Interestingly, NF1-L has also been recently identified as a silencer-binding protein for the rat growth hormone gene (Roy and Guérin, 1994).

Negative regulatory factors in search of target genes

A number of studies have identified transcription factors with repressor activity that are expressed in the nervous system, although in most cases the target genes regulated by these repressors are not yet clearly established. For example, several helix-loop-helix proteins related to *hairy* and *enhancer of split*, two negative regulators of neurogenesis in *Drosophila*, have been cloned (Akazawa et al., 1992; Feder et al., 1993; Sasai et al., 1992) and one of these, Hes-3, is expressed specifically in Purkinje cells (Sasai et al., 1992). Other examples of negative factors present only in subsets of neurons come from the POU family of transcription factors (Ingraham et al., 1990). While most POU-family proteins appear to activate transcription, recent work has ascribed a repressor function to neuron-specific splice variants of Oct-2. The neuronal isoforms, Oct 2.4 and Oct 2.5, can inhibit transcriptional activation mediated by the product of a third alternatively-spliced transcript from this gene, Oct 2.1, an activating isoform originally found in B-cells (Lillycrop et al., 1994). Another POU protein expressed in sensory neurons, Brn3b, was shown to repress transcription from reporter plasmids containing appropriate binding sites, as well as to antagonize activation by closely related POU family members (Morris et al., 1994). One

case in which a biologically-relevant target of repression by a POU-domain protein has been identified concerns SCIP/Tst-1/Oct-6, which has been suggested to function as a transcriptional repressor of myelin-specific genes (such as P_0) in glial precursor cells (Monuki et al., 1990). The mechanism of repression may involve a promoter context-dependent "quenching" interference with normal transactivators of these genes (Monuki et al., 1993; Monuki et al., 1993). The biological signficance of this repression is not completely clear, but may reflect a requirement to delay expression of the myelination program in proliferating and immature glial progenitors, since an extinction of SCIP/Tst-1/Oct-6 expression in these cells is always tightly correlated with the initiation of myelination.

CONCLUSION

Negative transcriptional regulation is clearly emerging as an important theme in understanding the control of gene expression in the nervous system, perhaps to a greater extent than it has in the study of cell type-specific transcriptional regulation in other vertebrate tissues. This may reflect, at least in part, the use of overlapping sets of positive-acting transcription factors to generate the enormous diversity of cell type-specific patterns of gene expression in the brain (Struhl, 1991). If such combinatorial mechanisms are operative, many neuronal genes will share binding sites for common positive factors even if these genes are not expressed in the same neuronal cell type. Specific negative regulation would therefore be required to prevent expression of some genes in the wrong kinds of neurons. Similarly, if neuronal genes shared positive-acting factors with other genes normally expressed in non-neuronal tissues, it could explain why silencer elements are required to repress neuronal gene transcription outside of the nervous system. Whether NRSF/REST-like mechanisms are unique to the nervous system or used more broadly will become apparent as binding proteins are identified for silencers used in other tissues and cell types (Sawada et al., 1994).

Although this article has emphasized recent advances in our understanding of negative regulation, it should not be taken to imply that positive regulation is any less important in the regulation of neuronal gene expression. In the end, the promoter of a neuronal gene is likely to function analagously to the axon hillock region of a neuron: it integrates the sum total of positive and negative influences acting on the gene (Davidson, 1994) (through enhancers and silencers rather than through excitatory and inhibitory synapses), and then "computes" a frequency of transcriptional initiation, analagous to a frequency of firing action potentials. The challenge for the future is to identify the players in the network of molecular "connections" that impinge on a given neuronal gene, and understand how their influence is integrated.

ACKNOWLEDGEMENTS

Portions of the work cited from the authors' laboratory was supported by NIH grant NS23476. D.J.A. is an Associate Investigator of the Howard Hughes Medical Institute.

REFERENCES

Adams, A. D., Choate, D. M., and Thompson, M. (1995). NF1-L Is the DNA-binding component of the protein complex at the peripherin negative regulatory element. J. Biol. Chem. *270*, 6975-6983.

Akazawa, C., Sasai, Y., Nakanishi, S., and Kageyama, R. (1992). Molecular characterization of a rat negative regulator with a basic helix-loop-helix structure predominantly expressed in the developing nervous system. J. Biol. Chem., 21879-21885.

Brand, A. H., Breeden, L., Abraham, J., Sternglanz, R., and Nasmyth, K. (1985). Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. Cell *41*, 41-48.

Chong, J. A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J. J., Zheng, Y., Boutros, M. C., Altshuller, Y. M., Frohman, M. A., Kraner, S. D., and Mandel, G. (1995). REST: A mammalian silencer protein that restricts sodium channel expression to neurons. Cell *80*, 949-957.

Davidson, E. H. (1994). Molecular biology of embryonic development: how far have we come in the last ten years? BioEssays *16*, 603-615.

Dong, J.-M., Smith, P., Hall, C., and Lim, L. (1995). Promoter region of the transcriptional unit for human α1-chimaerin, a neuron-specific GTPase-activating protein for p21rac. Eur. J. Biochem. *227*, 636-646.

Feder, J. N., Jan, L. Y., and Jan, Y. N. (1993). A rat gene with sequence homology to the *Drosophila* gene *hairy* is rapidly induced by growth factors known to influence neuronal differentiation. Mol. Cell. Biol. *13*, 105-113.

Guérin, S., Anzivino, M. J., Roy, R. J., and Moore, D. D. (1993). Expression of the rat growth-hormone gene is under the influence of a cell-type-specific silencer element. Eur. J. Biochem. *213*, 399-404.

He, X., and Rosenfeld, M. G. (1991). Mechanisms of complex transcriptional regulation: implications for brain development. Neuron *7*, 183-196.

Hoyle, G., Mercer, E. H., Palmiter, R. D., and Brinster, R. L. (1994). Cell-specific expression from the human dopamine ß-hydroxylase promoter in transgenic mice is controlled via a combination of positive and negative regulatory elements. J. Neurosci. *14*, 2455-2463.

Ibáñez, C. F., and Persson, H. (1991). Localization of sequences determining cell type specificity and NGF responsiveness in the promoter region of the rat choline acetyltransferase gene. Eur. J. Neurosci. *3*, 1309-1315.

Ingraham, H. A., Albert, V. R., Chen, R., Crenshaw, E. B. I., Elsholtz, H. P., He, X., Kapiloff, M. S., Mangalam, H. J., Swanson, L. W., Treacy, M. N., and Rosenfeld,

M. G. (1990). A family of POU-domain and pit-1 tissue-specific transcription factors in pituitary and neuroendocrine development. Annu. Rev. Physiol. *52*, 773-791.

Ishiguro, H., Kim, K. T., and Kim, K.-S. (1993). Neuron-specific expression of the human dopamine-ß-hydroxylase gene requires both the cAMP-response element and a silencer region. J. Biol. Chem. *268*, 17987-17994.

Jan, Y. N., and Jan, L. Y. (1994). Neuronal cell fate specification in *Drosophila*. Curr. Op. Neuro. *4*, 8-13.

Johnson, P. F., and McKnight, S. L. (1989). Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. *58*, 799-839.

Joyner, A. L., and Guillemot, F. (1994). Gene targeting and development of the nervous system. Curr. Op. Neuro. *4*, 37-42.

Kapur, K. P., Hoyle, G. W., Mercer, E. H., Brinster, R. L., and Palmiter, R. D. (1991). Some neuronal cell populations express human dopamine ß-hyroxylase-*lacZ* transgenes transiently during embryonic development. Neuron *7*, 717-727.

Keegan, C. E., Karolyi, I. J., Knapp, L. T., Bourbonais, F. J., Camper, S. A., and Seasholtz, A. F. (1994). Expression of corticotropin-releasing hormone transgenes in neurons of adult and developing Mice. Mol. Cell. Neurosci. *5*, 505-514.

Kraner, S. D., Chong, J. A., Tsay, H. J., and Mandel, G. (1992). Silencing the type II sodium channel gene: a model for neural-specific gene regulation. Neuron *9*, 37-44.

Leonard, D. G. B., Ziff, E. B., and Greene, L. A. (1987). Identification and characterization of mRNAs regulated by nerve growth factor in PC12 cells. Mol. Cell. Biol. *7*, 3156-3157.

Li, L., Suzuki, T., Mori, N., and Greengard, P. (1993). Identification of a functional silencer element involved in neuron-specific expression of the synapsin I gene. Proc. Natl. Acad. Sci. U.S.A. *90*, 1460-1464.

Lillycrop, K. A., Estridge, J. K., and Latchman, D. S. (1994). Functional interaction between different isoforms of the Oct-2 transcription factor expressed in neuronal cells. Biochemistry Journal *298*, 245-248.

Lönnerberg, P., Lendahl, U., Funakoshi, H., Ärhlund-Richter, L., Persson, H., and Ibáñez, C. F. (1995). Regulatory region in choline acetyltransferase gene directs developmental and tissue-specific expression in transgenic mice. Proc. Nat. Acad. Sci. U.S.A. *92*, 4046-4050.

Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987). Regulation of inducible and tissue-specific gene expression. Science *236*, 1237-1245.

Maue, R. A., Kraner, S. D., Goodman, R. H., and Mandel, G. (1990). Neuron-specific expression of the rat brain type II sodium channel gene is directed by upstream regulatory elements. Neuron *4*, 223-231.

Mercer, E. H., Hoyle, G. W., Kapur, R. P., Brinster, R. L., and Palmiter, R. D. (1991). The dopamine ß-hydroxylase gene promoter directs expression of E. coli *lacZ* to sympathetic and other neurons in transgenic mice. Neuron *7*, 703-716.

Monuki, E. S., Kuhn, R., and Lemke, G. (1993). Cell-specific action and mutable structure of a transcription factor effector domain. Proc. Natl. Acad. Sci. U.S.A. *90*, 9978-9982.

Monuki, E. S., Kuhn, R., and Lemke, G. (1993). Repression of the myelin P0 gene by the POU transcription factor SCIP. Mech. Devel. *42*, 15-32.

Monuki, E. S., Kuhn, R., Weinmaster, G., Trapp, B. D., and Lemke, G. (1990). Expression and activity of the POU transcription factor SCIP. Science *249*, 1300-1303.

Mori, N., Birren, S. J., Stein, R., Stemple, D., Vandenbergh, D. J., Wuenschell, C. W., and Anderson, D. J. (1990). Contributions of cell-extrinsic and cell-intrinsic factors to the differentiation of a neural-crest-derived neuroendocrine progenitor cell. In Cold Spring Harbor Symposium on Quantitative Biology, pp. 255-263.

Mori, N., Schoenherr, C., Vandenbergh, D. J., and Anderson, D. J. (1992). A common silencer element in the SCG10 and type II Na^+ channel genes binds a factor present in nonneuronal cells but not in neuronal cells. Neuron *9*, 1-10.

Mori, N., Stein, R., Sigmund, O., and Anderson, D. J. (1990). A cell type-preferred silencer element that controls the neural-specific expression of the SCG10 gene. Neuron *4*, 583-594.

Morris, P. J., Theil, T., Ring, C. J. A., Lillycrop, K. A., Moroy, T., and Latchman, D. S. (1994). The opposite and antagonistic effects of the closely related POU family transcription factors Brn-3a and Brn-3b on the activity of a target promoter are dependent on differences in the POU domain. Mol. Cell. Biol. *14*, 6907-6914.

Pathak, B. G., Neumann, J. C., Croyle, M. L., and Lingrel, J. B. (1994). The presence of both negative and positive elements in the 5'-flanking sequence of the rat Na,K-ATPase α3 subunit gene are required for brain expression in transgenic mice. Nuc. Acids Res. *22*, 4748-4755.

Possenti, R., Di Rocco, G., Nasi, S., and Levi, A. (1992). Regulatory elements in the promoter region of *vgf*, a nerve growth factor-inducible gene. Proc. Natl. Acad. Sci. USA *89*, 3815-3819.

Rincón-Limas, D. E., Geske, R. S., Xue, J.-J., Hsu, C. Y., Overbeek, P. A., and Patel, P. I. (1994). 5'-Flanking sequences of the human HPRT gene direct neuronal expression in the brain of transgenic mice. J. Neurosci. Res. *38*, 259-267.

Roy, R. J., and Guérin, S. L. (1994). The 30-kDa rat liver transcription factor nuclear factor 1 binds the rat growth-hormone proximal silencer. Eur. J. Biochem. *219*, 799-806.

Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to *Drosophila hairy* and *enhancer of split*. Genes & Dev *6*, 2620-2634.

Sawada, S., Scarborough, J. D., Killeen, N., and Littman, D. R. (1994). A lineage-specific transcriptional silencer regulates CD4 gene expression during T lymphocyte development. Cell *77*, 917-929.

Schoenherr, C. J., and Anderson, D. J. (1995). The Neuron-Restrictive Silencer Factor (NRSF): A coordinate repressor of multiple neuron-specific genes. Science *267*, 1360-1363.

Shaskus, J., Zellmer, E., and Lewis, E. J. (1995). A negative regulatory element in the rat dopamine β-hydroxylase gene contributes to the cell type specificity of expression. J. Neurochem. *64*, 52-60.

Silva, A. J., and Giese, K. P. (1994). Plastic genes are in! Curr. Op. Neuro. *4*, 413-420.

Stein, R., Orit, S., and Anderson, D. J. (1988). The induction of a neural-specific gene, SCG10, by nerve growth factor in PC12 cells is transcriptional, protein synthesis dependent, and glucocorticoid inhibitable. Dev. Biol. *127*, 316-325.

Struhl, K. (1991). Mechanisms for diversity in gene expression patterns. Neuron *7*, 177-181.

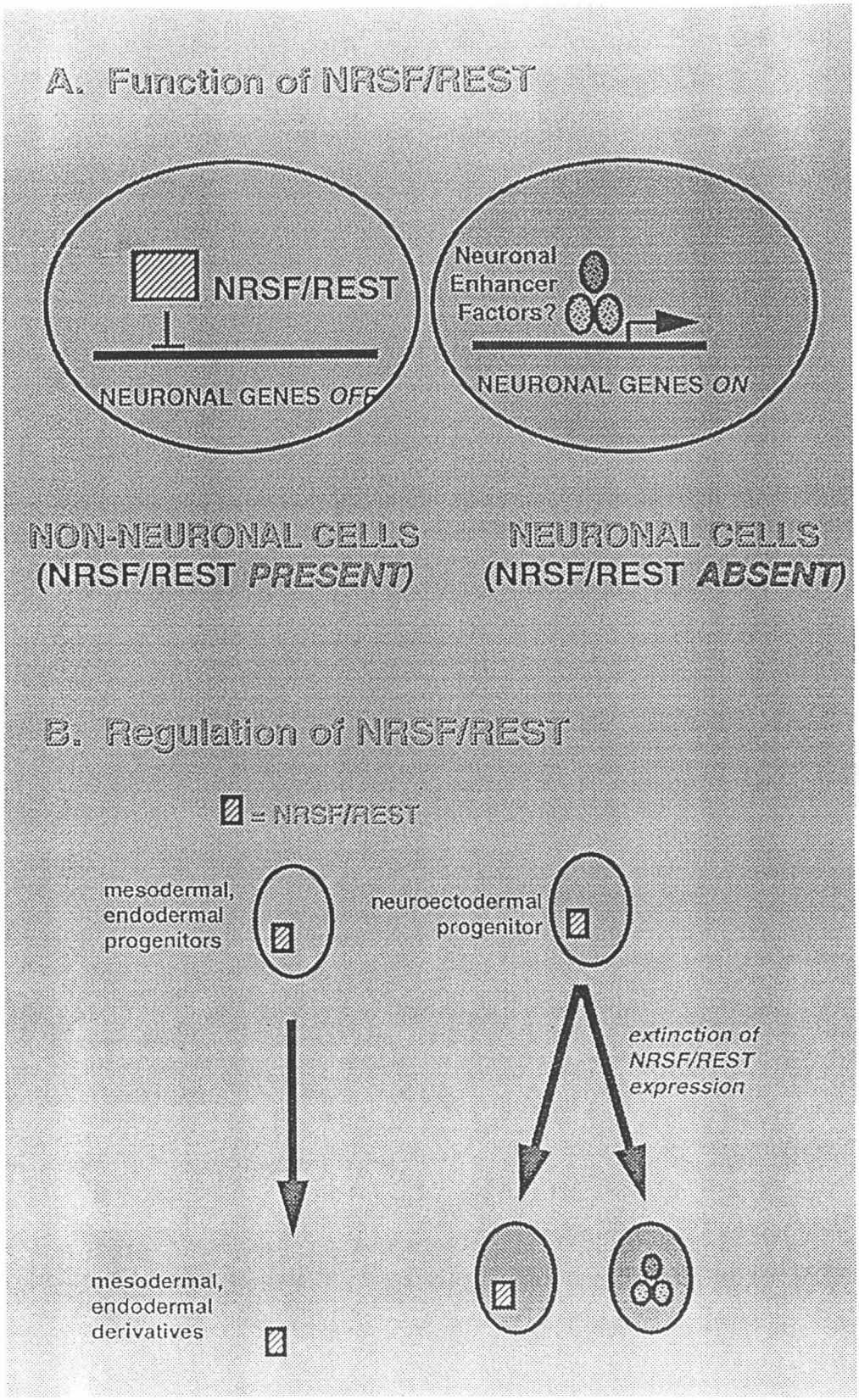
Thompson, M. A., Lee, E., Lawe, D., Gizang-Ginsberg, E., and Ziff, E. B. (1992). Nerve Growth Factor-induced derepression of peripherin gene expression Is associated with alterations in proteins binding to a negative regulatory element. Mol. Cell. Biol. *12*, 2501-2513.

Vandenbergh, D. J., Wuenschell, C. W., Mori, N., and Anderson, D. J. (1989). Chromatin structure as a molecular marker of cell lineage and developmental potential in neural crest-derived chromaffin cells. Neuron *3*, 507-518.

Vanselow, J., Grabczyk, E., Ping, J., Baetscher, M., Teng, S., and Fishman, M. C. (1994). GAP-43 transgenic mice: dispersed genomic sequences confer a GAP-43-like expression pattern during development and regeneration. J. Neurosci. *14*, 499-510.

Wuenschell, C. W., Mori, N., and Anderson, D. J. (1990). Analysis of SCG10 gene expression in transgenic mice reveals that neural specificity is achieved through selective derepression. Neuron *4*, 595-602.

Figure 1. Function and regulation of NRSF/REST. (A) NRSF/REST is present in non-neuronal cells, where it represses transcription of neuronal genes (left). NRSF/REST is absent in neuronal cells, permitting the transcription of neuronal genes in response to positive-acting enhancer factors, some of which may be neuron-specific (right). This illustration is an oversimplification in that some types of neurons may express NRSF/REST, which could function in those cells to repress some neuronal genes but not others (Chong et al., 1995). (B) NRSF/REST is expressed in most or all non-neural tissues and their progenitors (left). In embryonic neural tissue, NRSF/REST is expressed by multipotent progenitors of neurons and glia but is then selectively extinguished in those cells that adopt a neuronal fate and maintained in those cells that adopt a glial fate (right). The illustration is again an oversimplification in that NRSF/REST may not be expressed in some types of glial cells, and is not expressed in some embryonic non-neural tissues such as heart and liver.


A. Function of NRSF/REST
NRSF/REST
NEURONAL GENES OFF
Neuronal Enhancer Factors?
NEURONAL GENES ON
NON-NEURONAL CELLS (NRSF/REST PRESENT)
NEURONAL CELLS (NRSF/REST ABSENT)
B. Regulation of NRSF/REST
= NRSF/REST
mesodermal, endodermal progenitors
neuroectodermal progenitor
extinction of NRSF/REST expression
mesodermal, endodermal derivatives