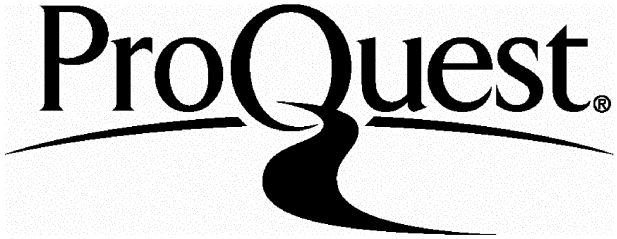
TRANSCRIPTIONAL REGULATION OF THE NEURON-RESTRICTIVE SILENCER FACTOR

Mohini Shashikant Mistry

A Thesis Presented for the Degree of Doctor of Philosophy

Wellcome Laboratory for Molecular Pharmacology
Department of Pharmacology
University College London



ABSTRACT

Neuron-restrictive silencer factor (NRSF), also known as RE1-silencing transcription factor (REST), is a DNA binding protein that silences the expression of multiple neurone-specific, terminal differentiation genes. During embryogenesis NRSF/REST is expressed at high levels in most non-neural tissues, but is downregulated during differentiation of neural progenitor cells into neurones. The proposed role of this transcription factor is to restrict the expression of neurone-specific genes to the nervous system. However NRSF/REST is also important in directing the proper spatial and temporal expression of NRSE/RE1-bearing genes within the developing peripheral and central nervous systems and this has implications for neuronal development. Furthermore in mature neurones the expression of low levels of NRSF/REST mRNA and the alteration of mRNA levels, in response to kainic acid, implies this transcription factor may also be involved in neuronal plasticity. To understand the mechanisms underlying transcriptional control of the NRSF/REST gene, the promoter region was isolated and characterised. A 6.5 kb genomic fragment of the mouse gene containing three exons (exon IA, exon IB and exon IC), that correspond to alternatively spliced 5' non-coding exons, was isolated. Reverse transcription-PCR identified splice variants in mouse fibroblast and neuroblastoma cell lines. Each splice variant contained one of the 5' non-coding exons spliced onto the first coding exon (exon II) resulting in transcripts with three different 5' untranslated regions (UTR). Splice variants containing type A, 5' UTR (corresponding to exon IA sequence) were by far the most abundant. Using 5' RACE PCR, RNase protection and *in vitro* transcription assays, two clusters of transcription initiation sites were identified within exon IA. Transient transfection analysis of the 6.5 kb genomic fragment, fused upstream of the luciferase reporter gene, showed this region was sufficient to direct cell-specific reporter gene expression in 3T3 fibroblast, Neuro2a neuroblastoma, PC12 pheochromocytoma cell lines and primary neuroepithelial cell cultures. Testing a series of 5' and 3' deletions within the 6.5 kb fragment in transient transfections led to the identification of several positive and negative *cis*-acting regulatory domains, some of which act in a cell-specific manner. An enhancer region, spanning the 3' end of exon IA and part of the first intron, was crucial for exon IA promoter activity. Characterisation of this proximal enhancer region revealed the presence of a sequence with similarity to the NRSE/RE1 motif, however NRSF/REST does not appear to bind. It is proposed that the differential levels of NRSF/REST gene expression in non-neural and neuronal cells are achieved by the combinatorial action of the cell-specific enhancer and repressor elements that are present within the mouse NRSF/REST gene.

DEDICATE

To my Mother,
Ranjan

ACKNOWLEDGEMENTS

Professor Noel Buckley, my principal supervisor, for giving me the opportunity to pursue these studies in his lab. Professor David Brown, for employment during the completing stages. Ian Wood for suggesting that I take up a Ph.D. and together with Avtar for many stimulating discussions over beer. Mariza and Stefania for giving me emotional support and generally for fussing over me. Patrick for making me laugh even when the going was tough and for all the support he has given during the completing stages. Tony for helping to deal with orders. All other people in the lab, including Fe, Miriea, Lucine, Gayle and Nicolas, who have all helped to make these Ph.D. studies a memorable journey.

All my family and friends who over the years have provided support. In particular Shashikant, my Dad, for his understanding. Veena and Ravi for being there. Finally Naresh, my husband, for allowing me to pursue my ambitions and being by my side.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

GLOSSARY OF ABBREVIATIONS

ac	*achaete*
Amp	ampicillin
AMV	avian myeloblastosis virus
AS-C	*achaete –scute* complex
AVP	arginine vasopressin
BAF	BRG1-associated factor
BDNF	brain–derived neurotrophic factor
bHLH	basic helix-loop-helix
bp	base pairs
cDNA	complementary DNA
ChAT	choline acetyltransferase
CNS	central nervous system
cRNA	complementary RNA
C-terminal	carboxyl terminal
DβH	dopamine β-hydroxylase
DMSO	dimethyl sulphate
dNTP	deoxynucleotide triphosphate
DRG	dorsal root ganglia
DTT	dithiothreitol
E(spl)-C	*Enhancer of split* Complex
E14	embryonic day 14
EDTA	ethylenediamine-tetraacetic acid
EMSA	electromobility shift assay
es	external sensory organs
GAP-43	growth-associated protein 43
GluR2	glutamate receptor subunit 2
HDAC	histone deacetylase
Inr	Initiator sequence
IPTG	isopropyl β-D-thiogalactopyransoside
KA	kainic acid
kb	kilobase; kilobase pairs
LB	Luria-Bertani medium
L1	L1 cell adhesion molecule
M_4	M_4 muscarinic acetylcholine receptor
M-MLV	moloney murine leukemia virus
Na type II	voltage dependent type II sodium channel
nAChR	neuronal nicotinic acetylcholine receptor
NE	neuroepithelial
Ng-CAM	neuron-glia cell adhesion molecule
NMDA R1	*N*-methyl-D-aspartate receptor subunit type 1
NRSE	neuron-restrictive silencer element
NRSF	neuron-restrictive silencer factor
$NRSF_{SF}$	truncated form of NRSF/REST
nt	nucleotide
N-terminal	amino terminal
N-tubulin	neuronal βIII tubulin
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline

PCR	polymerase chain reaction
PDGF	platelet-derived growth factor
PMSF	phenylmethylsulfonyl fluoride
PNS	peripheral nervous system
RACE	rapid amplification of cDNA ends
RE	restriction endonuclease
RE1	repressor element 1
REST	RE1-silencing transcription factor
rNTP	ribonucleotide triphosphate
RTase	reverse transcriptase
RT-PCR	reverse transcription–PCR
SANT	SWI13/ADA2/N-CoR/TFIIIB
sc	*scute*
SCG	superior cervical ganglion
SCLC	small cell lung cancer
SDS	sodium dodecyl sulphate
Su(H)	*Suppressor of Hairless*
SV40	simian virus 40
TdT	terminal deoxynucleotidyl transferase
TE	tris-EDTA buffer
TH	tyrosine hydroxylase
TSA	trichostatin A
UTR	untranslated region
X-gal	5-bromo-4-chloro-3-indolyl-B-d-galactopyranoside

Chapter 1

1.0 INTRODUCTION

1.1 Phenotypic diversity within the nervous system

The human nervous system contains approximately 10^{12} neurones. This large number of neurones consists of hundreds of distinct neuronal subtypes and furthermore the advent of monoclonal antibodies and DNA probes has revealed that neurones which appear morphologically identical can be molecularly distinct. Individual neurones differ from one another in their signalling capabilities, neuronal and hormonal responsiveness, neurotransmitter usage, expression of cell surface recognition molecules as well as the number and shape of their neurites and the neuronal connections they make. Furthermore the plasticity of mature neurones reflects regulated alterations in the expression of genes encoding, for example, neuropeptides and neurotransmitter-synthesising enzymes. One of the central challenges in neurobiology today is to understand the mechanisms underlying the morphogenesis of the mammalian nervous system. Genetic studies in invertebrates, such as *Drosophila melanogaster* and *Caenorhabditis elegans* has led to the identification of a large number of regulatory genes (encoding mainly transcription factors) that are involved in neurogenesis, and the identification of mammalian counterparts to these genes has greatly facilitated understanding neurogenesis in higher organisms. In addition, numerous studies have been performed identifying the *cis*-acting DNA elements and the *trans*-acting proteins that direct neurone-specific expression of terminal differentiation genes in mammals. As the links between transcription factors that regulate neuronal determination and those that are involved in neuronal differentiation are unravelled it has become apparent that neuronal development involves cascades of regulatory events in which transcriptional regulators are controlled by other transcriptional regulators acting in a hierarchical manner. Furthermore these hierarchies of transcription factors are themselves controlled

by signalling pathways that respond to cell-cell interactions and environmental signals. While both positive and negative regulation of transcription factors has been found to be crucial in the early stages of neurogenesis, the different classes of transcription factors that regulate the expression of neurone-specific, terminal differentiation genes have largely been found to be positive-acting factors. However, the importance of negative regulation in directing neurone-specific gene expression is becoming increasingly clear (He and Rosenfeld, 1991; Schoenherr and Anderson, 1995b; Simpson, 1995; Quinn, 1996; Goodman and Mandel, 1998; Thiel et al., 1999). In this regard a particular *cis*-regulatory element, called the neuron-restrictive silencer element (NRSE) (Mori et al., 1992), and also known as repressor element 1 (RE1) (Kraner et al., 1992) has been identified in multiple neurone-specific, terminal differentiation genes and has been the focus of much recent attention. The NRSE/RE1 interacts with the neuron-restrictive silencer factor (NRSF) (Schoenherr and Anderson, 1995a), also known as RE1-silencing transcription factor (REST) (Chong et al., 1995). The cell specificity of NRSF/REST gene expression during development and in adulthood is crucial to the role of this transcription factor in both restricting the expression of NRSE/RE1-bearing genes to the nervous system and modulating target gene expression in differentiated neurones. The questions are what are the factors that direct cell-specific expression of the NRSF/RST gene? And what are mechanisms that regulate NRSF/REST gene expression in different cell types?

1.2 Development of the nervous system

1.2.1 Neuronal development in vertebrates

In early vertebrate development, when the embryo consists of three layers: ectoderm, mesoderm and endoderm, chemical signals from the dorsal mesoderm trigger

a region of the ectoderm to thicken (neural plate) and indent, forming a neural groove on the surface of the embryo along what will become its anterior-posterior axis. The lips of the neural groove continue to swell, then meet and fuse to form the neural tube. The neuroepithelial cells that line the ventricular cavity of the neural tube are the precursors of neurones and glia in the adult brain and spinal cord. The post-mitotic progeny migrate from the ventricular zone towards outer layers and undergo terminal differentiation. During formation of the neural tube, a group of cells at its dorsal margin separates from it on either side to form the neural crest. All of the neurones forming the peripheral nervous system (PNS), including sensory and autonomic neurones arise from these cells. In vertebrates, the early steps of specification and differentiation of neural precursors have been studied in several models, including the spinal chord (Tanabe and Jessell, 1996) and the neural crest (Anderson et al., 1997).

Mammalian homologues of *Drosophila* proneural genes are involved in neuronal differentiation

Mutational genetic studies in both *Drosophila* and *C. elegans* have identified a large number of regulatory genes implicated in the determination of neuronal fate. However, the inability to perform large-scale genetic screens precludes such a systematic approach to gene identification in mammalian systems. Initial attempts at isolating regulatory genes has resorted to identifying genes that are mammalian counterparts of genes identified in invertebrate systems. Anderson and co-workers, who have studied the control of lineage commitment by cells in the neural crest (reviewed by Anderson et al., 1997) isolated *Mash1,* a mammalian homologue of the *Drosophila* proneural genes *ac-sc* (Johnson et al., 1990), which was the first such gene to be cloned in vertebrates. *Mash1* is able to efficiently complement the *ac-sc* loss-of-function

phenotype when expressed from a heat-shock promoter in *Drosophila* (cited in Anderson and Jan, 1997). Like its *Drosophila* counterparts, *Mash1* is expressed early in neurogenesis in restricted regions of the nervous system. For example in the mammalian PNS, *Mash1* is expressed by precursors of the autonomic, but not sensory, neurones (Lo et al., 1991; Guillemot and Joyner, 1993). Similarily, in the *Drosophila* PNS, *ac-sc* are expressed in precursor cells that give rise to external sensory (es) organs, but not those that generate chordotonal organs (Jan and Jan, 1994 and references therein). Targeted deletion of the *Mash1* gene in mice prevents the development of sub-populations of peripheral neurones, including the autonomic neurones, primary olfactory sensory neurones and a neuronal deficit is also seen in the enteric nervous system of *Mash-1 -/-* mice (Guillemot et al., 1993). These genetic data indicate that *Mash1*, like the *Drosophila ac-sc* proneural genes, is required for the genesis of particular subclasses of peripheral neurones. In *Drosophila* proneural genes play a role in cell determination, but their expression ceases at the stage of division of neural precursors. A more extensive analysis of the *Mash1* mutant phenotype *in vitro* indicated that MASH1 function is not essential until after commitment to neuronal differentiation (Sommer et al., 1995). In the absence of *Mash1,* arrested neuronal precursors die, thus the progression of these cells to fully differentiated neurones is dependent upon MASH1 function. The cellular function of MASH1 in the PNS is therefore distinct from that of ACHAETE-SCUTE in *Drosophila*, which is required at a relatively early stage in development. A major difference in vertebrates is that the nervous system forms as a result of induction from the dorsal mesoderm. Vertebrate proneural homologs are first expressed after formation of the neural plate and thus do not normally mediate choice between neural and epidermal fates, but are involved rather in the generation of neural precursor populations within the neural plate.

Negative regulation of vertebrate homologues of proneural genes

Studies of the development of the *Drosophila* peripheral nervous system reveals that transcriptional activators are crucial in early neural development but negative regulation also plays an essential role in determining neural cell fate. Several mammalian homologues of genes that negatively regulate proneural genes in *Drosophila* have been identified. *Id* genes (Benezra et al., 1990), which are homologues of *extramacrochaete* are expressed in the ventricular zone of the spinal cord (Duncan et al., 1992). Four genes homologous to the Drosophila *hairy* and *Enhancer of split* Complex (*E(spl)-C*), termed *Hes1-Hes3* and *Hes5*, have been cloned in mouse (Akazawa et al., 1992; Sasai et al., 1992; Ishibashi et al., 1993). Genetic analysis has shown that *Hes1* and *Hes5* in particular have important roles in neurogenesis (Ohtsuka et al., 1995; Ishibashi et al., 1995). For example, forced expression of HES1 from a retroviral vector in neural progenitors prevented cells from migrating out of the ventricular zone and inhibited overt neurogenesis (Ishibashi et al., 1994) whereas targeted disruption of *Hes1* caused premature neurogenesis in the central nervous system (CNS) (Ishibashi et al., 1995). Intriguingly, premature neurogenesis in *Hes1* -/- mice was also accompanied by a premature induction of *Mash1* (Ishibashi et al., 1995) supporting the notion that *Hes* genes function in mammals as they do in *Drosophila*, to repress or delay the expression of positive-acting basic helix-loop-helix (bHLH) genes during neurogenesis. At present it is unknown if MASH-1 represents the critical transcription target for HES1-mediated repression of neuronal development, and/or whether HES-1 acts as a passive repressor by titrating out positive bHLH proteins by forming heterodimers *in vivo*. Human HES1 has been shown to bind a CACGCA sequence in the promoter of the *Hash1* gene and thereby repress its transcription in carcinoma cells (Chen et al., 1997). However wild-type HES1 and a non-DNA-binding mutant of HES1 were shown to repress MASH1

dependent transcriptional activation suggesting a passive repression mechanism is possible (Castella et al., 1999). In *Drosophila* the Notch signalling pathway activates expression of the *E(spl)-C* genes (Bailey and Posakony, 1995; LeCourtois and Sweisguth, 1995). Thus the Notch signalling pathway, which mediates cell-cell interactions, is a critical component of neurogenesis in invertebrates. Components of the Notch signalling pathway are conserved in vertebrates (reviewed in Lewis, 1996). Activation of *RBP-Jκ*, mammalian homologue of *Suppressor of Hairless* (*Su(H)*), by Notch signalling leads to increased expression of *Hes1* (Jarriault et al., 1995; Jarriault et al., 1998). Experimental studies in *Xenopus* embyonic CNS have suggested that the Notch signalling pathway plays a role in vertebrate neurogenesis (Chitnis et al., 1995). Furthermore targeted mutation of *Notch1* and *RBP-Jκ* in mice have also indicated a role for the Notch signalling pathway in mammalian neurogenesis (de la Pompa et al., 1997). Although HES1 expression was not altered in these mice, suggesting the *Hes1* gene can also be activated through a Notch-independent pathway, HES-5 expression was decreased in *Notch1* -/- mice and absent in *RBP-Jκ* -/- mice, and expression domains of *Mash1* and *Delta-1* were both increased in both mutants (de la Pompa et al., 1997).

1.2.2 Regulation of genes encoding transcription factors that directly regulate neurone-specific, terminal differentiation genes

The search for vertebrate counterparts of *Drosophila* proneural genes has been fruitful in obtaining a molecular understanding of neuronal differentiation in vertebrates. However, a great deal of work has also been devoted to defining the *cis*-acting DNA elements and *trans*-acting factors involved in directing expression of neurone-specific, terminal differentiation genes (reviewed in Mandel and McKinnon, 1993) and a large array of positive-acting transcription factors have been identified, including bHLH

(Guillemot, 1999); LIM domain (Tsuchida et al., 1994) and POU domain (Ryan and Rosenfeld, 1997) transcription factors amongst many others (reviewed by He and Rosenfeld, 1991; Struhl, 1991). Members of these transcription factor families often display specific patterns of expression and activity in developing and mature neurones. In vertebrates the relationship between transcription factors that act to determine neuronal fate and those that specify neuronal identity has begun to be understood. One of the best examples comes from the study of the regulation of noradrenergic phenotype in sympathetic neurone development in vertebrates (reviewed by Anderson et al., 1997; Guillemot, 1999). Recent years has seen the emergence of the idea that different components of neuronal identity are assembled piece by piece through the action of distinct subprograms of transcription factors that are controlled by various environmental signals (reviewed by Anderson and Jan, 1997). A substantial body of work both *in vivo* and *in vitro* has implicated that the neural tube, notochord and cells derived from the somites are involved in differentiation of sympathetic neurones (reviewed by Groves and Anderson, 1996 and see references therein). In chick embryos, that lack both a notochord and ventral neural tube structures, a loss of certain sympathetic characteristics is observed. Expression of noradrenergic biosynthetic enzymes such as tyrosine hydroxylase (TH) and dopamine β-hydroxylase (DβH) in neural crest cells aggregating in sympathetic ganglion adjacent to the dorsal aorta is abolished. The expression of PHOX2a, a paired homeodomain transcription factor that is an essential determinant of noradrenergic phenotype, and which may be a direct transcriptional regulator of DβH is also prevented by this manipulation. However, expression of CASH-1 (the chick homologue of MASH-1) is unaffected, as is expression of the pan-neuronal marker SCG10. These data suggest that expression of neurotransmitter identity and other pan-neuronal aspects of sympathetic identity is

Developmet vs Plasticity

- changes in [TF] in response to signals from environment or other cells.

- DRG requires NGF to survive

separately controlled by different subprograms of transcription factors. Pan-neuronal marker, SCG10, is controlled by MASH-1 and is independent of floor plate/notochord signals. By contrast, expression of neurotransmitter identity is controlled by PHOX2a and is dependent on signals from the notochord/floor plate. Although induction of MASH-1 in sympathetic neurones appears to be independent of signals from the floor plate/notochord, this does not mean *Mash-1* expression occurs cell-autonomously. *In vitro*, MASH-1 expression is induced in undifferentiated neural crest cells by BMP2, which also induces autonomic neuronal differentiation (Lo et al., 1997). *In vivo*, BMP2 mRNA is expressed by endothelial cells lining the dorsal aorta at the time when migrating neural crest cells arrive at this structure. Taken together these data suggest induction of MASH-1 in neural crest cells may be promoted by BMP2, derived from dorsal aorta, while induction of PHOX2a, DβH and TH occurs in response to signals from both the dorsal aorta and the floor plate/notochord. Although determination of pan-neuronal and neurotransmitter components of sympathetic identity can be experimentally uncoupled these are not parallel independent processes. In *Mash-1* -/- mice, sympathetic neurones do not express PHOX2a, TH or DβH. In fact MASH-1 is thought to indirectly or directly activate the expression of PHOX2a and couple the expression of pan-neuronal genes to neurotransmitter identity (Lo et al., 1998).

1.2.3 Concluding remarks on neuronal development

In summary the question of how cells acquire neuronal identity has been approached first by genetic analysis in invertebrates and more recently in vertebrates. Studies in *Drosophila* have identified many transcriptional regulatory proteins that are expressed at progressive stages of neuronal determination and differentiation. Many of these regulatory proteins are conserved in vertebrates, and have formed the basis of

identifying regulatory proteins involved in neuronal development in vertebrates. Also, the role of cell-cell interactions in the determination of neuronal fate, which was first recognised in invertebrates has been conserved in vertebrates. Moreover, the molecular analysis of neuronal differentiation in mammalian models, of which the work carried out by Anderson and colleagues on the neural crest cells and Jessell and co-workers on the spinal cord are arguably the best characterised, shows that both the sequential and combinatorial expression of different transcription factors are important in defining different types of neurones. Moreover subprograms of transcription factor cascades involved in neuronal differentiation are rarely controlled cell-autonomously but rely on environmental signals.

1.3 The NRSE/RE1 sequence restricts the expression of multiple terminal differentiation genes to within the nervous system

Studies into the regulation of neurone-specific gene expression has revealed that many classes of transcriptional activators are essential. However studies of two neurone-specific genes encoding the growth associated protein, superior cervical ganglion (SCG)-10 and rat type II sodium channel (Na type II), respectively, revealed that negative transcriptional regulation is important for directing neuronal gene expression (Maue et al., 1990; Mori et al., 1990). Transient transfection experiments revealed that an upstream negative regulatory domain in the SCG10 gene was required to repress transcription in HeLa cells but had no effect in PC12 cells (Mori et al., 1990). Moreover generation of transgenic mice harbouring a chimeric gene, consisting of the SCG10 gene 5' flanking sequence fused to a reporter gene, and containing a deletion of the upstream negative regulatory domain, showed ectopic expression of the transgene in non-neuronal tissue (Wuenchell et al., 1990). Data from transient transfection analysis of the Na type

II gene, which is expressed in neurones but not in muscle (muscle expresses other Na channel genes), revealed that three negative regulatory domains in the 5' flanking sequence were required to repress transcription of Na type II reporter genes in L6 myoblasts and had no effect in PC12 cells (Maue et al., 1990). DNase footprinting assays of the negative regulatory region within Na type II gene led to the identification of a 28 bp motif, that was termed RE1 (Kraner et al., 1992). The RE1 motif, fused to the Na type II minimal promoter was able to repress reporter gene expression in L6 skeletal muscle and B50 neural cells, but not in PC12 cells (Kraner et al., 1992). Similarly, delineation of the negative regulatory region in the SCG10 gene revealed a 62 bp domain, that silenced expression directed by the SCG10 proximal promoter region in HeLa cells but not in PC12 cells, and this region was found to contain a 21 bp sequence, termed NRSE, that was highly similar to the RE1 sequence in the Na type II gene as well as sequences present in the human and rat synapsin I genes (Mori et al., 1992). The NRSE/RE1 sequences in the SCG10 and NaII genes were shown to be orientation- and relatively distance-independent (Kraner et al., 1992; Mori et al., 1992) and therefore acted as a repressor element. Electromobility shift assays (EMSAs) revealed that cell extracts of non-neural cells, in which the Na type II and SCG10 reporter constructs showed no activity in transient transfection assays, contained NRSE/RE1 binding activity (Kraner et al., 1992; Mori et al., 1992). In contrast NRSE/RE1 binding activity was not detected in neuronal cell lines, including PC12 cells, that express the endogenous Na type II and SCG10 genes (Kraner et al., 1992; Mori et al., 1992). Interestingly, the presence of NRSE/RE1 binding activity in neuronal B50 cell extracts, together with the fact that reporter constructs containing the NRSE/RE1 motif fused to the Na type II minimal promoter were silenced in B50 cells, led to the suggestion that the NRSE/RE1 motif also silences Na type II gene expression in neural cells that do not

express the endogenous Na type II gene (Kraner et al., 1992). Several studies identifying NRSE/RE1 sequences in other neurone-specific genes had also began to emerge. For example, a study of the transcriptional regulation of the chicken neuron-glia cell adhesion molecule (Ng-CAM) gene revealed five tandem NRSE/RE1 motifs in the first intron that were important for silencing the Ng-CAM proximal promoter in non-neural 3T3 cells and these did not silence expression in Neuro2a neuroblastoma cells (Kallunki et al., 1995). In addition synapsin I (Li et al., 1993), brain-derived neurotrophic factor (BDNF) (Timmusk et al., 1993), choline acetyltransferase (ChAT) (Lönnerberg et al., 1995) and the β2-subunit of neuronal nicotinic acetylcholine receptor (nAChR β2 subunit) (Bessis et al., 1995) were also found to contain sequences showing high similarity to the NRSE/RE1 motif. Work from the Anderson laboratory, using a composite NRSE/RE1 sequence derived from the NRSE/RE1 sequences in the SCG10, Na type II, synapsin and BDNF genes to perform a search of the Genbank database, identified an additional 22 candidate genes containing sequences with two or fewer matches to the composite sequence and of these genes, 17 are expressed primarily in neurones (Schoenherr et al., 1996). Although NRSE/RE1 sequences were found in several non-neuronal genes, including actin, neuronal genes outnumbered non-neuronal genes by 4:1 (Schoenherr et al., 1996). The bias towards neuronal genes in the database search was consistent with NRSE/RE1-mediated repression restricting the expression of neurone-specific genes to the nervous system. The majority of NRSE/RE1-containing genes were found to encode proteins that contribute to different aspects of mature neuronal phenotype including neurotransmitter receptors, ion channels, neurotransmitter-synthesising enzymes, neuropeptides, cell adhesion molecules, synaptic vesicle proteins and cytoskeletal components. However, as more promoters are analysed, the repertoire of NRSE/RE1 containing genes has increased further and now also

includes the rat M_4 muscarinic receptor (M_4) (Wood et al., 1995; Wood et al., 1996; Mieda et al., 1996), *N*-methyl-D-aspartate receptor subunit type 1 (NMDA R1) (Bai et al., 1998; Okamoto et al., 1999), GluR2 (Myers et al., 1998), arginine vasopressin (AVP) and preprotachykinin (Coulson et al., 1999), L1 cell adhesion molecule (L1) (Kallunki et al., 1997) and neuronal βIII tubulin (N-tubulin) (Chen et al., 1998) genes. Mutation and deletion of the NRSE/RE1 sequence within the promoters of several of these reporter genes were carried out and showed that NRSE/RE1-mediated regulation is required to silence expression in non-neural cell lines and produces no change in reporter gene activity in neuronal cell lines (Mori et al., 1992; Kraner et al., 1992; Li et al., 1993; Lönnerberg et al., 1996; Schoch et al., 1996; Mieda et al., 1997; Kallunki et al., 1997; Bai et al., 1998; Myers et al., 1998). In the case of the human synapsin I gene, deletion of the NRSE/RE1 motif abolished neurone-specific expression entirely and led to same levels of expression in both non-neural cells and neuronal cell lines (Schoch et al., 1996). Importantly, some NRSE/RE1-containing genes, for example the rat M_4 receptor gene, are only expressed in subsets of neurones only indicating that NRSE/RE1-mediated regulation is not limited to pan-neuronal genes (see Wood et al., 1996 for discussion). Furthermore, the presence of an NRSE/RE1 in at least one positive transcriptional activator, P-LIM (Schoenherr et al., 1996) suggests that NRSE/RE1 mediated repression may also suppress expression of positive regulators of neuronal genes. However, an NRSE/RE1 sequence has also been found in the *Hes3* gene (Schoenherr et al., 1996), which encodes a cerebellar Purkinge cell-specific transcription actor that negatively regulates transcription (Hirata et al., 2000) and is expressed in Purkinje cells soon after birth and continues until adulthood (Sasai et al., 1992).

The NRSE/RE1 motif is one of the first repressor elements to be characterised that is important in directing the expression of neurone-specific genes to the nervous system by silencing expression in non-neural cells. Moreover, the NRSE/RE1 motif remains unique in the multiplicity of neurone-specific target genes that contain this *cis*-element. However transgenic studies have indicated that the silenced state of neurone-specific genes in the full spectrum of non-neural tissues is accomplished through the utilisation of multiple silencers, for example the first intron of the L1 gene contains a silencer element in addition to the NRSE/RE1 that is required to suppress expression in non-neural cells (Kallunki et al., 1997). Furthermore, neurone-specific genes whose expression is restricted to the nervous system in an NRSE/RE1-independent manner have been identified (reviewed by Schoenherr and Anderson 1995b). For example, negative regulatory regions in the DβH and Na,K ATPase α3 subunit genes that contain sequences showing limited similarity with the NRSE/RE1 sequence, and which restrict expression to neuronal cells (Ishiguro et al., 1993; Pathak et al., 1994) do not appear to interact with NRSF/REST in EMSAs (Schoenherr et al., 1996) suggesting these genes do not contain an NRSE/RE1 repressor element. In addition the promoter activity of the growth-associated protein 43 (GAP-43) gene is restricted to neuronal cells by a novel repressor element located immediately downstream of the TATA box (Weber and Skene, 1997). The GAP-43 repressor element, termed the SNOG element is highly position-dependent and is present in a similar position relative to the TATA box in two other neurone-specific genes (Weber and Skene, 1997). However in contrast to the NRSE/RE1 motif, factors that bind the SNOG element are present in both neuronal and non-neuronal cell extracts suggesting this element does not bind a tissue-specific transcription factor.

1.4 Discovery of the NRSE/RE1-binding protein (NRSF/REST)

In the Mandel laboratory, workers using a yeast one-hybrid approach isolated a partial cDNA clone encoding the DNA binding domain of an NRSE/RE1 binding protein from a HeLa cell cDNA library (Chong et al., 1995). This was subsequently used to screen a λZAP HeLa cDNA library and led to the isolation of a cDNA clone encoding the entire coding region of the NRSE/RE1 binding protein which was termed REST (Chong et al., 1995). In addition to a cluster of eight zinc finger motifs that formed the DNA binding domain, a lone zinc finger motif was identified at the C-terminus (Chong et al., 1995). All zinc fingers were found to be of the Cys_2His_2 type and exhibited homology with the GL1-Krüppel family of zinc fingers (Chong et al., 1995). The 5' untranslated region (UTR) of human NRSF/REST mRNA was determined by 5' RACE to be 326 nucleotides long, however Northern blot analysis indicated that the mRNA is 7.6 kb in length indicating the presence of a long 3' UTR (Chong et al., 1995). Workers in the Anderson laboratory simultaneously reported the isolation of a partial cDNA clone (λHZ4) from a HeLa cell cDNA library, encoding part of an NRSE/RE1 binding protein which they termed NRSF (Schoenherr and Anderson, 1995a). NRSF was found to contain a cluster of eight zinc finger motifs, and the interfinger sequences placed this protein in the GL1-Krüppel family of zinc fingers. The cDNA clones encoding NRSF and REST, corresponding to NRSE/RE1-binding proteins identified independently by two research groups, in fact encode the same protein. EMSAs were used to show that *in vitro* translated λHZ4 and full length REST protein both form a DNA-protein complex with the NRSE/RE1. Furthermore the addition of antibodies directed against recombinant λHZ4 and full length REST produced a supershift with native NRSF/REST present in HeLa and muscle cell nuclear extracts indicating that the NRSE/RE1 binding activity detected in these extracts corresponds to NRSF/REST protein (Chong et al.,

1995; Schoenherr and Anderson, 1995a). Finally, cDNA clones encoding partial (λHZ4) and full length NRSF/REST proteins were sufficient to repress the SCG10 promoter and Na type II promoter reporter genes, respectively, in pheochromocytoma PC12 cells (Chong et al., 1995; Schoenherr and Anderson, 1995a). Furthermore, when overexpressed in L6 myoblast cells, the N-terminal, eight zinc finger cluster (forming the DNA binding domain) behaved as a dominant negative, blocking the silencing activity of endogenous NRSF/REST (Chong et al., 1995). Despite further screens performed by different laboratories NRSF/REST is the only NRSE/RE1-binding protein that has been identified thus far. Analysis of the NRSF/REST protein structure showed that adjacent to the DNA-binding domain lies a highly basic region, followed by a proline rich region and a lone Cys_2His_2 zinc finger is found at the C-terminal end of NRSF/REST (Chong et al., 1995). Two distinct repression domains have been functionally mapped and are present within the N-terminal 83 amino acids and the C-terminal ninth zinc finger motif (Tapia-Ramirez et al., 1997; Thiel et al., 1998). Each repression domain is sufficient to repress the Na type II promoter as well as NRSE/RE1-containing heterologous promoter constructs (Tapia-Ramirez et al., 1997; Thiel et al., 1998). In transient transfections, overexpression of NRSF/REST in neuronal cell lines was shown to repress, but not silence, the activity of several NRSE/RE1 containing promoters including the M_4 muscarinic receptor (Mieda et al., 1997), ChAT (Lönnerberg et al., 1996), L1 (Kallunki et al., 1997) GluR2 receptor (Myers et al., 1998), NMDA R1 receptor (Okamoto et al., 1999) and arginine vasopressin (Coulson et al., 1999) reporter genes. However the effect of NRSF/REST on reporter gene expression was affected by the presence of other *cis*-regulatory regions. For example, the presence of a distal cholinergic enhancer region in ChAT reporter constructs suppressed NRSE/RE1 activity in cholinergic cells even after NRSF/REST was overexpressed (Lönnerberg et al., 1996),

CAN YOU DRAW DIAGRAM

of how PRSF binding

would act to release

gene expression!

TELL ME ABOUT PC12 stable &

trans.?

and reporter gene expression directed by the arginine vasopressin minimal promoter in small cell lung cancer (SCLC) cell lines required a greater excess of recombinant NRSF/REST to see a functional effect in cell lines displaying high levels of endogenous AVP gene expression (Coulson et al., 1999).

1.4.1 Mechanisms of transcriptional repression by NRSF/REST

The mechanism by which NRSF/REST represses gene expression has been the subject of several recent publications. The two distinct repressor domains of NRSF/REST both interact with distinct co-repressor complexes (Leichter and Thiel, 1999). Transcriptional repression via the recruitment of histone deacetylase (HDAC) has been characterised for many transcription factors (reviewed in Pazin and Kadonaga, 1997; Struhl, 1998) and in most cases these repressors recruit HDAC by binding the co-repressor mSin3 or NCoR/SMRT. The N-terminal repression domain has been shown to recruit a mSin3/HDAC complex (Huang et al., 1999; Naruse et al., 1999; Roopra et al., 2000; Grimes et al., 2000). In accord with these findings, the chromatin structure of the NRSE/RE1 and promoter region of the endogenous M_4 receptor gene was shown to be remodelled in PC12 cells, which do not contain NRSE/RE1 binding activity (Kraner et al., 1992; Mori et al., 1992), that had been stably transfected with λHZ4 (consisting of the N-terminal repression domain and the entire DNA binding domain of NRSF/REST) (Wood et al., 2001). Repression by the C-terminal zinc finger domain occurs through a different mechanism (Leichter and Thiel, 1999). A novel co-repressor named CoREST interacts with the C-terminal repression domain of NRSF/REST (Andrés et al., 1999). This co-repressor contains two SANT (SWI13/ADA2/N-CoR/TFIIIB) domains (Aasland et al., 1996) separated by a highly charged region, an arrangement also found in NCoR/SMRTe co-repressors for steroid hormone receptors (Hörlein et al., 1995; Park et

al., 1999). The mechanism used by CoREST to repress target genes is not known, however as transcriptional repression is not relieved by trichostatin A (TSA) (an inhibitor of histone deacetylases) the mechanism appears to be independent of the action of histone deacetylases (Huang et al., 1999; Naruse et al., 1999). Recently, a new protein, termed BAF57v, has been shown to interact with a domain of CoREST required for expression (Battaglioli et al., 2000). BAF57v corresponds to a splice variant of BAF57 (BRG1-associated factor) (Wang et al., 1998), an HMG protein component of the mammalian SWI/SNF chromatin remodeling complex. However, the pattern of CoREST gene expression compared to mSin3A gene expression is more restricted in early development (Grimes et al., 2000) and it has been suggested that mSin3A is required for NRSF/REST-mediated repression, in general, whereas CoREST may be recruited for more specialised repressor functions (Grimes et al., 2000).

1.4.2 NRSF/REST is expressed in a tissue-specific manner throughout vertebrate development and adulthood

NRSF/REST is widely expressed in non-neural tissue and CNS progenitor cells during embryogenesis

Early studies showed that NRSE/RE1 binding activity is present in non-neural cell extracts, of HeLa and rat L6 skeletal muscle cells, but not in neuronal cell extracts, of PC12 and NG-108 cells (Kraner et al., 1992; Mori et al., 1992; Li et al., 1993; Wood et al., 1996; Mieda et al., 1997). Northern blot and RNase protection analysis revealed that a high level of NRSF/REST mRNA is present in glial and fibroblast cell lines, but mRNA levels in neuronal PC12 and MAH cells were below the level of detection (Schoenherr and Anderson, 1995a). In accord with these data, Western blot analysis revealed a high level of NRSF/REST protein is present in rat L6 cells whereas a

relatively low level is present in PC12 cells (Chong et al., 1995). These data indicated that the absence of NRSE/RE1 binding activity in neuronal cells was due to the absence of NRSF/REST and not due to functional inactivation of this protein. Interestingly, NRSF/REST mRNA and activity are also absent in neuroendocrine cell lines which express many neuronal genes including those encoding NMDA receptor R1 and Na type II (Atouf et al., 1997). *In situ* hybridisation analysis showed that NRSF/REST mRNA is expressed ubiquitously in early development around E8.5 (Chen et al., 1998; Grimes et al., 2000). Later, at E11.5 and E13.5 NRSF/REST mRNA is still present at high levels in most non-neural tissues (Chong et al., 1995; Schoenherr and Anderson, 1995a) however only low levels of NRSF/REST mRNA are found within the developing nervous system. In the developing nervous system NRSF/REST mRNA is restricted mainly to the ventricular zone of the neural tube, which contains multi-potential progenitors of neurones and glia. In contrast, NRSF/REST mRNA expression is not detected in the adjacent marginal zone of the neural tube where differentiating neurones are present (Chong et al., 1995; Schoenherr and Anderson, 1995a). However low levels of NRSF/REST mRNA were detected in the developing dorsal root ganglia (DRG) at E11.5 and E13.5 (Chong et al., 1995; Schoenherr and Anderson, 1995a) and the presence of NRSF/REST in DRGs was proposed to silence Na type II expression within sub-populations of sensory neurones (Chong et al., 1995). The reciprocal expression of NRSF/REST gene and its target genes, including those encoding SCG10 and Na type II channel, supported the hypothesis that the *in vivo* role of NRSF/REST is to suppress neurone-specific gene expression in non-neural tissue (Chong et al., 1995 and Schoenherr and Anderson, 1995a). Furthermore, the presence of low levels of NRSF/REST mRNA in CNS progenitor cells in the ventricular zone, and the absence in the adjacent marginal zone, containing differentiating neurones, suggested NRSF/REST

may serve to prevent expression of neurone-specific genes in neural progenitor cells up to the time of neuronal differentiation (Chong et al., 1995 and Schoenherr and Anderson, 1995a). In accord with these data, the differentiation of a number of neuroblastoma cell lines produces an apparent decrease in NRSF/REST mRNA levels (Palm et al., 1999). In fact NB-OK-1 cells (a human neuroblastoma cell line) which show a marked decrease in NRSF/REST mRNA following differentiation, induced by staurosporine and cyclic AMP, show a concomitant increase in mRNA levels of synapsin I and other NRSE/RE1 containing genes (Nishimura et al., 1996; I. C. Wood and N. J. Buckley, unpublished observations). Furthermore the levels of synapsin I mRNA have been found to be inversely proportional to the level of NRSF/REST mRNA in a number of neuroblastoma cell lines that have been derived from neural crest cells at various stages of neuronal development (Lietz et al., 1998).

Expression of NRSF/REST was not originally detected in differentiating neurones, however evidence is accumulating that NRSF/REST does play a role in directing gene expression within the nervous system. For example, deletion of the NRSE/RE1 in the first intron of the mouse L1 gene resulted in ectopic expression in non-neural tissues including the heart and kidneys, and expression of the mutated and wild type transgene were indistinguishable in the CNS (Kallunki et al., 1997). However, precocious and intense reporter gene expression was detected in the developing PNS (Kallunki et al., 1997). Furthermore analysis of post-natal transgenic mice showed increased reporter gene expression in the thalamus, whilst reduced expression was detected in the cortex, striatum and hippocampus (Kallunki et al., 1998). These data imply that the NRSE/RE1 acts as both a silencer and enhancer of L1 gene expression in the nervous system. However NRSE/RE1-mediated effects on the L1 promoter within

the CNS are linked to a developmental stage and a particular region of the nervous system. EMSAs performed using the L1 NRSE/RE1 sequence and brain extracts derived from post-natal mice indicated the presence of NRSE/RE1 binding activities (Kallunki et al., 1998). Two more transgenic studies, examining the expression of the nAChR β2 subunit and BDNF reporter genes, also pointed to a role for NRSE/RE1-mediated regulation in the nervous system (Bessis et al., 1997; Timmusk et al., 1999). A mutation of the NRSE/RE1 in the nAChR β2 subunit gene, that was shown to abolish protein binding, did not result in ectopic non-neural expression of the transgene but led to a loss of expression in certain brain regions, for example the thalamus and dorsal root ganglia, and a gain of expression in other neural structures for example the cortex (Bessis et al., 1997). By preparing promoter constructs in which the NRSE/RE1 was placed at different distances from an simian virus 40 (SV40) promoter and testing the activity of such constructs in neuroblastoma cells, containing low amounts of NRSE/RE1 binding activity, it was suggested that the NRSE/RE1 needed to be located less than 50 bp upstream of the TATA box to function as an enhancer (Bessis et al., 1997). However in the endogenous L1 gene, the NRSE/RE1 sequence is located approximately 10 kb downstream of the L1 promoter and still functions as an enhancer (Kallunki et al., 1998), indicating that the NRSE/RE1 does not necessarily have to be located proximal to the promoter to function as an enhancer in an endogenous gene. To determine if NRSF/REST is responsible for NRSE/RE1-mediated activation, Bessis et al. (1997) expressed NRSF/REST antisense in neuroblastoma cells and demonstrated a loss of NRSE/RE1-mediated activation, suggesting position-dependent activation is mediated by NRSF/REST. However a recent report showed that NRSF/REST alone is not sufficient to activate transcription in neuroblastoma cell lines irrespective of the location of the NRSE/RE1 (Thiel et al., 1998). Finally, transgenic analysis showed that the 5'

portion of the palindromic NRSE/RE1 sequence in the rat BDNF gene is important for controlling BDNF expression levels, in the brain, heart and lungs, and interestingly, the NRSE/RE1 is important in repressing kainic acid (KA) induced expression of the BDNF gene in hippocampal neurones (Timmusk et al., 1999). EMSAs have shown that the 5' portion of the BDNF NRSE/RE1 forms a complex with extracts of different regions of brain, including hippocampus (Timmusk et al., 1999). These data are interesting in the light of the fact that changes in BDNF expression in hippocampal and cortical neurones following KA-induced seizures is associated with long-term potentiation, neuroprotection and epilepsy (Timmusk et al., 1999). In summary, mutational studies in transgenic mice suggest that both NRSE/RE1-mediated repression and activation are important for directing the appropriate expression of the nAChR β2 subunit and L1 reporter genes in neurones within the developing nervous system and post-natal brain (Bessis et al., 1997; Kallunki et al., 1997; Kallunki et al., 1998). Also NRSE/RE1-mediated modulation is important for the response of BDNF gene expression to KA-induced seizures (Timmusk et al., 1999). More specifically the presence of low levels of NRSE/RE1-binding activity in neuroblastoma cell lines (Bessis et al., 1997; Kallunki et al., 1995), whole post-natal brain (Kallunki et al., 1998) and hippocampal (Timmusk et al., 1999) extracts is consistent with the expression of NRSF/REST in neuronal cells. In fact, low levels of NRSF/REST mRNA have been found in neuroblastoma cell lines (Lietz et al., 1998; Lönnerberg et al., 1996; Nishimura et al., 1996; Palm et al., 1999). More importantly a recent study has shown that low levels of NRSF/REST mRNA are present in spinal cord and all regions within adult rat brain that were examined (Palm et al., 1998). Moreover, *in situ* hybridisation shows NRSF/REST mRNA is localised within neurones (Palm et al., 1998). These data strongly imply that NRSF/REST is important not only in silencing the expression of neurone-specific genes in non-neural tissue and

neural precursors, but also in modulating the expression of neurone-specific genes in neurones; both during development and in adulthood.

NRSF/REST mRNA splice variants are expressed at low levels in differentiated neurones

Approximately 1% of all NRSF/REST mRNA transcripts expressed in brain consist of alternative mRNA splice variants, REST4 and REST5 (Palm et al., 1998; Palm et al., 1999), and these have been detected in embryonic and adult, rat and mouse brain (Palm et al., 1998; Palm et al., 1999). These neural-specific splice variants encode a truncated isoform, $NRSF_{SF}$, consisting of the N-terminal repression domain and five of the zinc finger motifs of the DNA binding domain (Palm et al., 1998; Palm et al., 1999). $NRSF_{SF}$ has been shown to bind the M_4 and Na type II NRSE/RE1 in EMSAs (Roopra et al., 2000) and possesses repressor activity (Palm et al., 1998). In one study EMSAs were performed using cell extracts containing different ratios of NRSF/REST isoforms, and $NRSF_{SF}$ was proposed to interact with full-length NRSF/REST to inhibit its binding to DNA (Shimojo et al., 1999). Further studies have shown that $NRSF_{SF}$ binds the cholinergic NRSE/RE1 motif as a monomer but 10-20 –fold less efficiently than full length NRSF/REST (Lee et al., 2000) and it is proposed that factors that perturb the equilibrium between monomeric and heteromeric transcription factors could alter the transcription of NRSE/RE1-bearing genes. This has been demonstrated using wild type PC12 cells which express mRNA encoding both full length NRSF/REST and $NRSF_{SF}$, in which the cholinergic gene is transcriptionally active (Shimojo et al., 1999). However, cholinergic gene expression is reduced in a protein kinase A deficient PC12 cell line which expresses mRNA encoding full length NRSF/REST only (Shimojo et al., 1999). The absence of REST4 in the protein kinase A deficient PC12 cell line is thought to lead

to repression of the cholinergic gene locus by endogenous full length NRSF/REST. In accord with these data both wild type and the mutant PC12 cells display similar low levels of NRSF/REST mRNA and protein, however while extremely low NRSE/RE1 binding activity was detectable using wild type PC12 cell extracts, prominent NRSE/RE1 binding activity was detected using mutant PC12 cell extracts (Shimojo et al., 1999). The presence of elevated levels of $NRSF_{SF}$ has been noted in several neuroblastoma cell lines, moreover REST4 mRNA levels increased following neuronal differentiation in several neuroblastoma cell lines (Palm et al., 1999). Northern blot analysis of rat brain shows the level of NRSF/REST mRNA and its splice variants increases in response to KA-induced seizures, and *in situ* hybridisation analysis reveals that increased NRSF/REST mRNA levels are present mainly in the hippocampus and cortex (Palm et al., 1998). However the physiological role of NRSF/REST truncated isoforms in neurones is not understood. The observed changes in NRSF/REST gene expression in neuroblastoma cell lines and brain tissues in response environmental cues suggests however that NRSF/REST may play a role in modulating gene expression in processes such as neuroprotection and synaptic plasticity within the nervous system.

1.4.3 Determination of the role of NRSF/REST in vertebrate development by loss-of-function and gain-of-function studies

Studies into early development of the vertebrate nervous system led to a proposal where vertebrate neural induction occurs through a default pathway (reviewed by Green, 1994). Implicit in this model is the idea that molecules, such as BMPs and related molecules, induce the epidermal fate in non-neural ectoderm (Sasai et al., 1995; Wilson and Hemmati-Brivanlou, 1995). Initially, NRSF/REST was proposed as a candidate that could fulfil the role of a master negative regulator of neuronal fate (Chong et al., 1995).

However the work by Anderson and colleagues, shows that loss of NRSF/REST function in mice and chicks, does not convert non-neural tissue into neurones (Chen et al., 1998) indicating that NRSF/REST is not involved in neuronal fate specification (Chen et al., 1998; reviewed in Hemmati-Brivanlou, 1998; Jones and Meech, 1999). However this transcription factor nevertheless must fulfil an essential role in early development as *Rest* -/- mice die by E11.5 (Chen et al., 1998). Histological examination at E9.25 revealed widespread cellular disorganisation, including the head mesenchyme and myotome. Expression of several NRSE/RE1 containing genes was examined including N-tubulin, SCG10, L1, synapsin I, calbindin and middle neurofilament. Among these only N-tubulin showed evidence of derepression. Expression was observed in non-neural tissues, where it is normally transiently expressed, for example increased expression of N-tubulin in myotome persisted until E10.5 by which time it is downregulated in wild type embryos, and *de novo* activation was observed in certain tissues, for example the epicardium (Chen et al., 1998). Thus the genetic evidence implies that NRSF/REST is required *in vivo* to repress neuronal gene expression in non-neural tissue.

Expression studies indicate that NRSF/REST gene expression in the ventricular zone is downregulated upon neuronal differentiation suggesting that relief from NRSF/REST-imposed repression may be important in the differentiation of neurones. However premature expression of NRSE/RE1-bearing genes was not detected within the CNS in *Rest* -/- mice, although onset of embryonic lethality preceeded the time these genes are normally expressed (Chen et al., 1998). Thus the generation of *Rest* -/- mice precluded a clear analysis of the role of NRSF/REST in neuronal differentiation. As an alternative means of inactivating NRSF/REST function, localised retrovirus-mediated

infection of chick embryos with a dominant negative form of NRSF/REST (dnNRSF) was carried out to enable analysis of gene expression at a later developmental time window then could be analysed in mouse (Chen et al., 1998). Infection of newly formed somites in chick embryo caused derepression of chick N-tubulin, Ng-CAM and SCG10. Injection of dnNRSF retrovirus in undifferentiated CNS progenitors and premigratory neural crest cells caused premature expression of SCG10 and Ng-CAM genes in the ventricular zone; and derepression and upregulation of SCG10 and Ng-CAM gene expression in Schwann precursor cells (Chen et al., 1998). Thus NRSF/REST-mediated repression is essential in controlling proper timing of neuronal gene expression in differentiating neurones in both the peripheral and central nervous systems.

However, the loss-of-function experiments did not address whether downregulation of NRSF/REST is essential for neuronal differentiation to proceed. Thus constitutive expression of NRSF/REST in the ventricular zone and sensory neurones of the DRG was examined by retroviral delivery of an expression construct encoding full length NRSF/REST into chick embryos (Paquette et al., 2000). Repression of N-tubulin gene expression in the ventricular zone was observed, however no decrease in expression was seen in the marginal zones. In contrast expression of Ng-CAM was decreased in the marginal zone. Therefore overexpression of NRSF/REST in differentiating neurones led to a downregulation of a subset of endogenous NRSE/RE1 bearing genes. Interestingly these same genes were upregulated in loss-of-function studies (Chen et al., 1998). Although overexpression of NRSF/REST did not prevent cells from attaining neuronal morphology, axons of commissural neurones showed a significant rise in pathfinding abnormalities (Paquette et al., 2000). Thus the downregulation of endogenous NRSF/REST expression that normally occurs during

neurogenesis appears important for those aspects of neuronal development that require fine tuning, such as axon-guidance. By extension this is similar to the proposed role of NRSF/REST expression in adult neurones in modulating gene expression underlying neuronal plasticity.

1.4.4 Deregulation of NRSF/REST gene expression in disease pathogenesis

The arginine vasopressin gene, which contains a putative NRSE/RE1 motif adjacent to the transcription start site, is frequently detected in small cell lung cancer and acts as an autocrine growth factor of these tumours (Coulson et al., 1999 and references therein). EMSAs performed with the AVP and SCG10 NRSE/RE1 sequences resulted in the detection of several DNA-protein complexes using SCLC cell extracts (Coulson et al., 1999). Intriguingly SCLC cell lines and tumours express high levels of a human NRSF/REST splice variant encoding $NRSF_{SF}$ (Coulson et al., 2000). The antagonistic role of rat $NRSF_{SF}$ in activating cholinergic gene expression in PC12 cells (Shimojo et al., 1999) and the abnormally high expression of the NRSF/REST mRNA splice variant in SCLC cell lines, led to the proposition that the AVP gene is derepressed in these tumours (Coulson et al., 2000). In another study, EMSAs performed using an NRSE/RE1 as a probe and nuclear extracts prepared from human medulloblastoma cell lines corresponding to cells arrested along the neuronal differentiation pathway resulted in the detection of a DNA:protein complex displaying similar migration to a complex obtained using HeLa cell extracts (Lawinger et al., 2000). Reporter gene expression driven by the Na type II promoter was shown to be silenced in an NRSE/RE1-dependent manner in medulloblastoma cell lines but not in a neural progenitor cell line, NT2 cells (Lawinger et al., 2000). Furthermore NRSE/RE1-binding activity using cell extracts of medulloblastoma cell lines and a subset of human medulloblastoma tumours was shown

to be more prominent than cell extracts of neuronal progenitor cells and fully differentiated neurones (Lawinger et al., 2000). Disruption in the hedgehog signalling pathway is postulated as a possible cause for increased NRSF/REST expression since a fraction of medulloblastoma patients show defects in the patched (*PTCH*) gene and *PTCH* is an integral component of the hedgehog signalling pathway (Lawinger et al., 2000 and references therein).

1.4.5 The importance of determining the mechanisms underlying the transcriptional regulation of NRSF/REST

The experiments outlined above demonstrate a role for NRSF/REST in (1) silencing the expression of neurone-specific genes in non-neural tissue; (2) controlling the proper timing of neuronal gene expression in the developing nervous system; and (3) modulating gene expression in mature neurones in response to extracellular stimuli. The essential function of NRSF/REST during development is further highlighted by the lethal phenotype of *Rest -/-* transgenic mouse lines. Moreover the multiplicity of NRSF/REST target genes, which includes upwards of 30 neurone-specific genes, implies that the tight control of NRSF/REST expression is of the utmost importance, however, little is known about its regulation. Many transcription factors are expressed in a developmental stage and/or tissue specific manner, while many others, particularly those that are expressed ubiquitously, use post-translational modifications to regulate their activity (reviewed in Maniatis et al., 1987; Mitchell and Tjian, 1989; Falvey and Schibler, 1991; Calkhoven and Ab, 1996). However, binding affinity, binding site context and interactions with other regulatory proteins all play an important role in focussing the action of transcription factors to their respective target promoters. In the case of NRSF/REST, differential levels of expression in non-neural and neuronal tissue

throughout development and adulthood appear to be important for directing this transcription factor to its target promoters. Genes encoding transcription factors, as for genes in general, display unique spatial and temporal expression patterns that are determined by a combination of *cis*-acting DNA elements. In higher eukaryotes, *cis*-elements are commonly found in the 5' flanking sequence but are also found in non-coding exons, introns, coding exons and even on transcribed RNA. Eukaroytic promoters can be divided into (1) the core promoter, which is recognised by the RNA polymerase II machinery and where transcription is initiated (2) distal regions which act to increase or decrease the level of basal transcription.

The identification of *cis*-acting regulatory elements in the NRSF/REST gene provides a powerful approach to identify the factors that act upstream of NRSF/REST during development and thus to allow determination of the relationship between NRSF/REST and other transcriptional regulators involved in neurogenesis. Furthermore the importance of understanding NRSF/REST gene regulation has been recently highlighted by reports showing the deregulation of NRSF/REST gene expression in two types of tumours: small cell lung cancer and medulloblastoma (Coulson et al., 2000; Lawinger et al., 2000). However, it is not yet clear if changes to NRSF/REST expression in these tumours are an effect or cause of pathogenesis.

1.5 Aims of the study

1) Isolate a genomic clone of the mouse NRSF/REST gene and determine the gene structure
2) Determine the positions of the transcription initiation sites to identify the NRSF/REST promoter region.

3) Perform transient transfection analysis of NRSF/REST genomic fragments directing luciferase gene expression in cell lines and identify the *cis*-regulatory domains that are important for cell-specific expression.

4) Identify a system to perform transient transfections that will allow the *cis*-regulatory elements that downregulates NRSF/REST gene expression during neuronal differentiation to be identified.

5) Isolate *trans*-acting factors that bind the NRSF/REST promoter and elucidate the mechanisms used to control the level of NRSF/REST gene expression in different cell types.

Chapter 2

2.0 METHODS AND MATERIALS

2.1 Cell culture

Cell lines were cultured in Nunclon™ 100mm plates (Life Technologies) in 5% CO_2 at 37°C. PC12 cells were grown in minimal essential alpha media containing 5% fetal calf serum and 5% horse serum. HEK 293, 3T3 and Neuro2a cells were cultured in Dulbecco's Modified Eagle Medium containing 10% fetal calf serum. All media were supplemented with 6g/L penicillin, 10g/L streptomycin and 2 mM L-glutamine and all reagents were obtained from Life Technologies. Primary neuroepithelial cultures (consisting of 70-80% neuroepithelial cells and 20-30% differentiated neurones) were prepared from embryonic day 14 (E14) rat cerebral cortex as described in Williams et al. (1997) and seeded in Nunclon™ 24-well multidish plates (10mm wells from Life Technologies).

2.2 Transient transfection analysis of NRSF/REST gene reporter constructs

2.2.1 Transient transfections

Cells were seeded one day before transfection in Nunclon™ 24-well multidish plates (10mm wells from Life Technologies). Each plasmid was transfected into triplicate wells therefore a mastermix was made from which equal aliquots were delivered to each well. At least three independent plasmid DNA preparations were used for each construct transfected. Equimolar amounts of each reporter plasmid were present in the transfection mix and to ensure differences in the amount of DNA in the transfection mix did not affect luciferase activity, Bluescript® II (Stratagene) carrier DNA was added; this has been shown to have no influence on luciferase activity (Myers et al., 1998). Cells were

co-transfected with pRL-CMV (Promega), which contains *Renilla* cDNA driven by the CMV promoter, at 1:250 ratio of pRL CMV to reporter plasmid for cell lines and 1:40 ratio for neuroepithelial cells. For cell lines, each well was transfected with 250 ng (Neuro2a), 500 ng (3T3) and 1 μg (PC12) of plasmid DNA using 4.5 μl of Tfx™-50 (Promega) per μg of DNA in 200 μl of Opti-MEM (Life Technologies). For HEK 293 cells each well was transfected with 1 μg DNA using 4.5 μl Tfx™-50 (Promega) per μg of DNA in 200 μl of Opti-MEM and volume of transfection mix was made up to 800 μl with Opti-MEM immediately before adding to cells. After 1-2 h, 1 ml of complete growth medium was added to each well without removal of transfection mix. For neuroepithelial primary cultures each well was transfected with 200 ng DNA, diluted in EC buffer (Qiagen) in a total volume of 60 μl, using 1.6 μl enhancer (Qiagen) and 2 μl Effectene™ transfection reagent (Qiagen) in 350 μl Opti-MEM. The transfection complex was replaced with complete growth media 6 h post-transfection. Control transfection experiments with β-galactosidase expression plasmid, pCMVβ (Clontech) indicated that transfection efficiencies ranged between 5% and 50%, depending on cell type, as judged by the percentage of stained cells after transfection.

2.2.2 Histochemical staining for β-galactosidase activity

Cells were washed in 1 x phosphate buffered saline (PBS), overlayed with fixative (see section 2.12) and incubated at room temp. for 5 min. Cells were washed three times with 1 x PBS, overlayed with β-galactosidase staining solution (see section 2.12) and incubated in the dark at 37°C for 2-4 h. The β-galactosidase staining solution was replaced with 1 x PBS once the blue stain had developed.

2.2.3 Luciferase reporter assays

Cells transfected in 24-multiwell dishes were washed with 1 x PBS, then 100 µl Passive Lysis Buffer (Promega) was added, the cells freeze/thawed at -80°C and room temp, respectively, and shaken at room temp. for 15 min. before assaying extracts. Cell extracts were assayed using Dual-Luciferase™ Assay system (Promega). All luciferase measurements were carried out in a Turner TD-20e luminometer. Firefly luminescence, driven from pGL3 Basic (Promega) derived reporter constructs, was normalised to *Renilla* luminescence, driven from pRL CMV (Promega). Normalised luminescence was expressed relative to normalised luminescence driven from promoterless pGL3 splice (see section 2.3.3).

2.3 Plasmid construction

2.3.1 Techniques in subcloning

Ligation

Ligation reactions were performed in 10 µl of 1 x T4 DNA ligase buffer (Boehringer Mannheim) containing 3:1 molar ratio of insert DNA:linearised vector backbone DNA and 1 unit T4 DNA ligase (Boehringer Manheim). Reactions were incubated at 14°C overnight, then the DNA precipitated and resuspended in 5 µl water.

Dephosphorylation of Vector Backbone

2 µg vector backbone was linearised using an appropriate restriction endonuclease (RE) in a 20 µl reaction and the enzyme heat inactivated. To dephosphorylate the 5' ends, 1 unit (1 µl) of shrimp alkaline phosphatase (Roche) was added, and the reaction incubated for 1 h at 37°C. The enzyme was heat inactivated at 65°C for 15 min before use.

Gel Purification of DNA

DNA was purified from 0.8-1.2% agarose gel slices using QIAquick Gel Extraction Kit (Qiagen) as per manufacturers instructions. In the case of extracting DNA from a polyacrylamide gel slice, the slice was crushedand the DNA eluted by incubation in 2 volumes of TE pH 7.4 at 37°C overnight. To remove the polyacrylamide the sample was subjected to centrifugation through VectaSpin filtration units (Whatman).

Cloning PCR products

PCR products to be cloned were either gel-purified; purified directly using High Pure™ PCR product purification kit (Roche) or used directly. PCR products generated in the presence of Taq extender PCR additive (Stratagene), or other Taq polymerases containing 3' to 5' exonuclease activity will have blunt ends and therefore dATP has to be added to the 5' end prior to cloning into pGEMT-Easy. Thus, immediately after PCR amplification, the reaction was heated to 95°C for 20 min and 15 µl of dATP (from a 2mM dATP stock) and 5 units of Taq DNA Polymerase (Promega) were added per 50 µl and the reaction incubated at 70°C for 15 min. Between 2µl and 4 µl of the PCR reaction were used to perform a ligation with pGEMT-Easy vector (Promega).

Blunt ended cloning

Blunt ends were generated from DNA restriction fragments containing 5' or 3' overhangs using 1 unit (0.5 µl) Klenow enzyme (Roche) in a 20 µl reaction containing 1 x buffer H (Roche) and 250 nM each dGTP, dCTP, dATP and dTTP. The reaction was allowed to proceed at room temp. for 15 min and heat inactivated before use in ligation reactions.

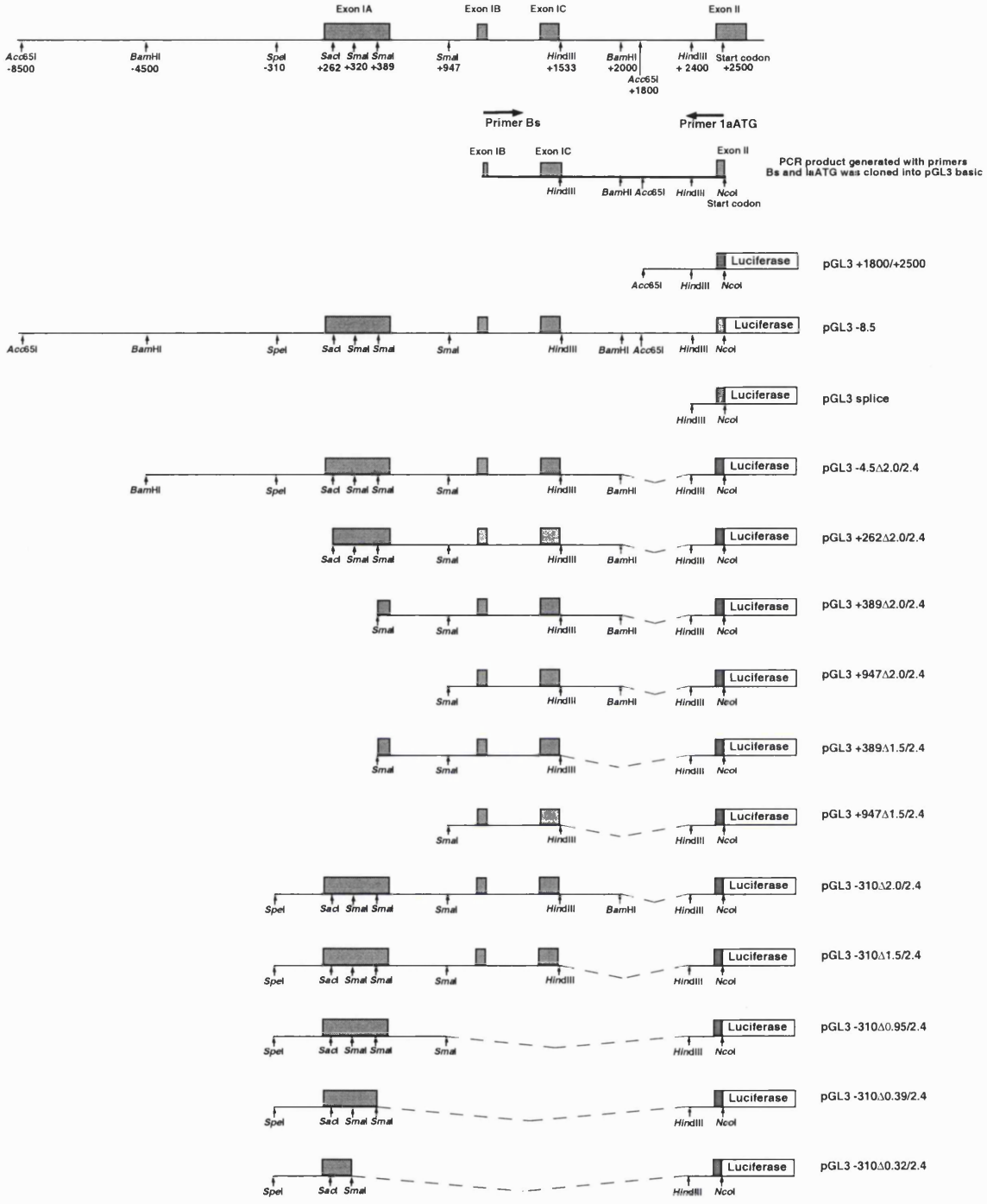
2.3.2 Subcloning the mouse NRSF/REST P1 clone

1 μg of P1 DNA was subjected to restriction digestion using 10 units of either *Hin*dIII, *Acc*65I or *Pst*I in a 20 μl reaction. Then 5 μl of the reaction was removed and used to set up a ligation with 50 ng dephosphorylated, linearised Bluescript® II SK(+) (Stratagene) that had been digested with the same restriction endonuclease. The ligation reaction was transformed into electrocompetent JS5 bacterial cells and plated on LB/Amp/X-gal/IPTG agar plates, and 100 white colonies were picked and streaked onto LB/Amp plates to generate an archive of P1 clone restriction fragments. Screening of the P1 clone restriction fragments was carried out by colony hybridisation using radiolabelled gene specific probes to allow isolation of DNA fragments that span the mouse NRSF/REST gene and flanking sequences (see Figure 1). Notably, attempts to isolate genomic clones corresponding to the region spanning intron ii, between exons II and III, repeatedly failed. Clones were subjected to restriction mapping and sequence analysis.

2.3.3 Construction of NRSF/REST reporter plasmids

NRSF/REST genomic fragments were cloned into the luciferase reporter vector, pGL3 Basic (Promega). To produce pGL3 -8.5 a two step cloning strategy was used. First a 0.7 kb *Acc*65I/*Nco*I DNA restriction fragment, spanning the initiation codon and 700 bp of upstream intronic sequence in the mouse NRSF/REST gene, was obtained by restriction digestion of a PCR product generated using primers Bs and 1a ATG on mouse genomic DNA template. Primer 1a ATG introduces a point mutation into the sequence immediately surrounding the initiation codon to generate an *Nco*I restriction site. The 0.7 kb *Acc*65I/*Nco*I DNA restriction fragment was cloned into pGL3 Basic, cut with *Acc*65I/*Nco*I, to generate pGL3 +1800/+2500. A 10 kb *Acc*65I mouse NRSF/REST genomic fragment containing exon IA, exon IB, exon IC and flanking sequences was

cloned into pGL3 Basic +1800/+2500 to generate pGL3 -8.5. All subsequent NRSF/REST reporter plasmids were constructed using pGL3 splice. This reporter plasmid was produced by replacement of the linker region between the *Hin*dIII and *Nco*I sites in pGL3 Basic (Promega) with a 120 bp region of the mouse NRSF/REST gene containing the NRSF/REST initiation codon, the exon II splice acceptor sequence and some intronic sequence. A 6.5 kb *Bam*HI fragment, internal to the 10 kb *Acc*65I fragment, spanning the three non-coding exons and flanking sequences was cloned into the *Bgl*II site of pGL3 splice to generate pGL3 -4.5 Δ2.0/2.4. This plasmid was digested with either *Sac*I and religated resulting in pGL3 +262 Δ2.0/2.4, or *Sma*I and religated, resulting in pGL3 +389 Δ2.0/2.4 and pGL3 +947 Δ2.0/2.4. Subsequently pGL3 +389 Δ2.0/2.4 and pGL3 +947 Δ2.0/2.4 were digested with *Hin*dIII and religated to generate pGL3 +389 Δ1.5/2.4 and pGL3 +947 Δ1.5/2.4, respectively. The reporter plasmid pGL3 -4.5 Δ2.0/2.4 was digested with *Nhe*I/*Spe*I and religated to generate a 5' deletion resulting in the reporter construct pGL3 -310 Δ2.0/2.4. Subsequently pGL3 -310 Δ2.0/2.4 was digested with *Hin*dIII and religated to produce pGL3 -310 Δ1.5/2.4. Further internal deletions were generated by linearising pGL3 -310 Δ1.5/2.4 with *Hin*dIII (which cuts at the 3' end of the NRSF/REST genomic insert), performing partial digestions with *Sma*I, blunting the ends and religating. This procedure resulted in the generation of reporter constructs pGL3 -310 Δ0.95/2.4, pGL3 -310 Δ0.39/2.4 and pGL3 -310 Δ0.32/2.4. Heterologous promoter constructs were generated using the vector pGL3 Inr (Wood et al., 1999) which contains the initiator element of the adenovirus major late promoter cloned into the *Hin*dIII site of pGL3 Basic. A *Sma*I fragment spanning nucleotides +389 to +947 of the mouse NRSF/REST gene was cloned into the *Sma*I site of pGL3 Inr to generate pGL3 Inr +389/+947 and pGL3 Inr +947/+389. PCR amplification of pGL3 Inr +389/+947 using primers RV3 and GL2 resulted in a PCR



Schematic showing construction of the deletion series of mouse NRSF/REST gene reporter plasmids. A map of the 5' end of the mouse NRSF/REST gene and flanking region is shown. Boxes represent exonic sequence and lines represent introns. Vertical small arrows indicate positions of restriction sites which were used to generate deletions and these are numbered with respect to the 5' most transcription start site defined as +1. Horizontal thick arrows show positions of primers. Dashed lines indicate deleted portions of NRSF/REST genomic sequence.

product containing nucleotides +389 to +947 of the mouse NRSF/REST gene and some vector sequence. The PCR fragment was digested with *Sau*3AI and the resulting fragments cloned into the *Bgl*II site of pGL3 Inr to generate pGL3 Inr +394/+659; pGL3 Inr +659/+763 and pGL3 Inr +947/+763. Clones containing the NRSF/REST genomic fragment +763/+947, contained the insert in the reverse orientation only.

2.4 RNA extraction and analysis

2.4.1 Extraction and purification of RNA

Total RNA was extracted from confluent cultured cells growing in 10 cm^2 TC dishes using 1 ml of RNAzol™ B (Biotecx Laboratories, Inc.) as per manufacturers instructions. Poly A^+ RNA was purified from total RNA using the Oligotex™ kit (Qiagen).

2.4.2 RNase protection assay

RNase protection assays were performed using the RPA II kit (Ambion). Following synthesis of an internally radio-labelled cRNA probe (see section 2.8.3), approximately 10^5 cpm of cRNA probe and the target RNA (either total RNA or poly A^+ mRNA) were co-precipitated (see section 2.10). The resulting RNA pellet was resuspended in 20-30 µl of solution hybridisation buffer (see section 2.12), denatured at 90°C for 10 min followed by incubation overnight an appropriate annealing temp. Samples were subjected to RNase digestion using a 1:100 dilution of RNase A/T1 mixture (Ambion) at 37°C for 30 min, precipitated and resuspended in 8 µl of formamide gel loading buffer. Prior to loading, samples were denatured at 90°C for 10 min and electrophoresed through a 3.5-6% urea-acrylamide gel in 1 x TBE at 60 watts for approximately 2 h. Finally, the gel was dried and exposed to X-ray film.

2.4.3 Primer extension assay

An oligodeoxynucleotide, end radio-labelled by kinasing using γ-^{32}P ATP (Amersham) (see section 2.8.1), was co-precipitated with either poly A$^+$ mRNA or total RNA. The RNA pellet was resuspended in solution hybridisation buffer (see section 2.12), denatured at 90°C for 10 min, incubated at an appropriate annealing temperature overnight, reprecipitated and resuspended in 10 µl of 2 x Superscript II buffer (Life Technologies). Alternatively, the RNA pellet was resuspended directly in 10 µl of 2 x Superscript II buffer, denatured at 85°C for 10 min and incubated at an appropriate annealing temperature for 1 h. Following annealing, once the RNA and probe were in 2 x Superscript II buffer the temperature was adjusted to either 37°C or 42°C for M-MLV and AMV reverse transcriptase respectively, and 0.5 mM dATP, 0.5 mM dTTP, 0.5 mM dGTP, 0.5 mM dCTP, 10 mM DTT, 1 µl RNase Inhibitor (Roche) and 200 units M–MLV reverse transcriptase (Promega) or 10 units AMV reverse transcriptase (Promega) were added taking the final volume up to 20 µl. Incubation of the reverse transcription reaction was allowed to proceed for 1-2 hours; then 10 units (1 µl) RNase, DNase-free (Roche) was added and incubation continued at 37°C for 30 min. The reaction was stopped by phenol/chloroform extraction (see section 2.12), the samples precipitated and the DNA pellet resuspended in 8 µl water. An equal volume of formamide gel loading dye was added before electrophoresing through a 4% urea-acrylamide gel in 1 x TBE. The gel was dried and exposed to X-ray film for the detection of primer extension products by autoradiography.

2.4.4 *In vitro* transcription assay

100 ng of gel purified, linearised NRSF/REST genomic DNA fragments (excised from a plasmid), 3mM $MgCl_2$ and 8 units HeLa nuclear extract (Promega) were assembled on

ice. To this 1 µl [α-^{32}P] rCTP (3000Ci/mmol, 10mCi/ml) (Amersham Life Sciences) and 1 µl of rNTP mix (containing 10mM rATP, 10mM rUTP, 10mM rGTP and 0.4mM rCTP) were added and the reaction incubated at 30°C for 1 h The reaction was terminated with 175 µl of stop mix (Promega), phenol/chloroform extracted, precipitated and the RNA pellet resuspended in 5 µl H_20. An equal volume of formamide loading dye was added and samples electrophoresed through a 4% urea-acrylamide gel in 1 x TBE. The gel was dried and exposed to X-ray film for detection of transcription run-off products by autoradigraphy.

2.5 Polymerase chain reaction (PCR) analysis

2.5.1 Standard PCR

Reactions were performed in 25 µl of 1 x Taq Extender buffer (Stratagene) containing template DNA, 250 nM dGTP, 250 nM dCTP, 250 nM dTTP, 250 nM dATP, 5 nM of each forward and reverse primers, 1.25 units Taq polymerase (Promega) and 1.25 units Taq Extender PCR additive (Stratagene). In some cases PCR amplification only yielded products in the presence of dimethyl sulphate (DMSO). In these cases DMSO was added to the PCR cocktail at a concentration of 5% w/v in the final reaction. Standard cycling conditions were 1 cycle 94°C/3 min; 30 cycles 94°C/30 s, 55°C/30 s, 72°C/1 min; 1 cycle 72°C/5 min.

2.5.2 Reverse transcription-PCR (RT-PCR)

Traces of o contaminating genomic DNA were removed by incubation of 10 µg of total RNA with 10 units (1 µl) DNAse I, RNase-free (Roche) at 37°C for 30 min in 1 x transcription buffer (Roche). The reaction was extracted with phenol/chloroform and precipitated (see section 2.12) and the RNA resuspended in 5 µl of water and stored at

–20°C until use. For reverse transcription, 2 µl (4 µg) of DNase I -treated RNA was added to 8 µl water containing 250 ng oligo $d(T)_{15}$ primer (Promega) and denatured at 65°C for 2 min then immediately cooled on ice. First strand cDNA synthesis was performed in 1 x Superscript II buffer (Life Technologies) containing 10 mM DTT, 1 mM dGTP, 1 mM dCTP, 1 mM dATP and 1 mM dTTP, 1 µl RNase Inhibitor (Roche) and 200 units (1 µl) M-MLV reverse transcriptase (Promega) at 37°C for 1 h. Then, 2 µl of cDNA template was subjected to standard PCR analysis (see section 2.4.1) and amplified products were electrophoresed through 2% MetaPhor® agarose (FMC BioProducts).

2.5.3 5' RACE (rapid amplification of cDNA ends) PCR

5' RACE PCR was carried out using a modified procedure of Frohman et al. (1988). First strand cDNA was synthesised as described in section 2.5.2, purified using High Pure™ PCR purification kit (Roche) to remove excess dNTPs and eluted in 100 µl of water. A tailing reaction was set up, to add a dATP tail to the 3' end of the first strand cDNA, in 1 x terminal deoxynucleotidyl transferase (TdT) buffer (Promega) containing 1 mM dATP and ~50 units (2 µl) TdT (Promega). The tailing reaction was incubated at 37°C for 1 h and heat inactivated at 65°C for 10 min. PCR amplification was performed on tailed cDNA using a gene specific primer and a t17 adapter primer, RACE-1 under standard PCR cycling conditions. A second round of amplification was performed using 10% of the amplified product from the first round of PCR as template and a set of primers internal to the amplified product. These primers consisted of an internal gene specific primer and the adaptor primer, RACE-2. PCR products from the second round of amplification were cloned directly into pGEMT-Easy cloning vector (Promega) and resultant clones were verified by DNA sequence analysis.

2.5.4 Semi-quantitative PCR analysis

To determine the PCR conditions required to produce a signal that is proportional to the amount of starting template, a series of increasing amounts of plasmid standards were used for PCR amplification. The plasmid standards were generated by cloning the PCR products obtained from using gene specific primer pairs to amplify cDNA. Subsequently standard RT-PCR was carried out except the optimal number of cycles was defined by the minimal number of cycles required to obtain a signal.

2.6 DNA isolation and analysis

2.6.1 P1 clone DNA isolation

A single colony of mouse NRSF/REST P1 (in JS5 cells), streaked on LB agar plates containing 25μg/ml kanamycin, was used to inoculate a 5 ml overnight culture. The following day 50 ml of LB media containing 25 μg/ml kanamycin was seeded with 500 μl of this overnight culture and incubated at 37°C with shaking until A_{550} 1.3-1.5 was reached. Cells were harvested by centrifugation and P1 DNA was isolated using a Tip-20 anion exchange column provided with the Plasmid Mini kit (Qiagen) but with two modifications to the manufacturers' protocol: (1) the volumes of P1, P2 and P3 were increased to 3 mls and (2) the elution buffer was prewarmed to 50°C. Induction of P1 replication was carried out by addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM from freshly made, filter sterilised, 100 mM IPTG stock solution when the culture reached A_{550} of 0.15. Recovery of P1 DNA from JS5 bacterial strain was found to be >10-fold higher per ml of culture than the NS3529 bacterial strain in which the P1 clone was provided. A comparison between restriction digestion patterns of the P1 clone isolated from the original host NS3529 and new host JS5 confirmed that no deletions or recombinations had occured.

2.6.2 Transduction of P1 clone into a Cre⁻ bacterial strain

The mouse NRSF/REST genomic clone in the P1 vector backbone, p*Ad*10-*SacB*II, was provided in host strain NS3529 a derivative of *E.coli* DH5 containing the *lac*Iq and *Cre* genes. Difficulties were experienced in obtaining sufficient yields of P1 clone DNA from host strain NS3529 for subsequent restriction analysis and sub-cloning. Typical yields were 0.5 μg of P1 plasmid DNA per 10 ml of exponentially growing cells (3 x 10^9 cells) when high-copy-number P1 lytic replicon was induced with IPTG. Low P1 plasmid yields from host strain NS3529 have been attributed to the presence of cre recombinase and transfer to a Cre⁻ host has been reported to improve P1 DNA recovery by 3-10 fold (Sternberg et al., 1994). P1 phage-mediated transduction was carried out (as described by Sternberg et al., 1994) to transfer the P1 clone to *E. coli* JS5 $lacI^q$ cre⁻ host strain shown to produce higher yields of P1 plasmid DNA (Kroon and MacDonald, 1997).

Preparation of P1 vir lysate

A single colony of NS3529 containing the NRSF/REST P1 clone, streaked on LB agar containing 25μg/ml kanamycin was used to inoculate a 5 ml overnight culture. The following day 100 μl of this overnight culture was used to seed 5 ml of LB media containing 25 μg/ml kanamycin and the culture incubated at 37°C with shaking for 2 h. Bacteria were harvested by centrifugation at 6000 rpm for 5 min in a bench top microcentrifuge. Cells were resuspended in 100 μl LB media containing 5mM $CaCl_2$, followed by the addition of 10 μl (10^9) of P1 vir phage (Genome Systems, Inc.) and the cells were incubated at 37°C for 5 min. The cells were harvested by centrifugation and resuspended in 1 ml LB media containing 10 mM $MgCl_2$, incubated at 37°C with shaking for 2 h, then lysed by addition of 20 μl chloroform, vortexed and the cell debris

pelleted by centrifugation at maximum speed in a bench top microcentrifuge. The supernatant containing the P1 lysate was incubated at 37°C for 10 min to evaporate chloroform before storage at 4°C until required (stable for at least one year).

Transduction protocol

The Cre$^-$ bacterial strain into which the P1 was to be transferred was streaked on selective plates and a single colony used to inoculate 5 ml LB and grown overnight. The following day 5 ml of LB media was seeded with 100 µl of this overnight culture and incubated at 37°C, with shaking, for 2 h. 1 ml of this culture was harvested by centrifugation and the bacterial pellet resuspended in 100 µl LB media containing 5 mM $CaCl_2$. 70 µl of transducing phage (prepared as described above) was added and allowed to adsorb to cells at 37°C for 5 min. Cells were harvested by centrifugation and resuspended in 1 ml LB media containing 10 mM Na citrate and incubated at 37°C for 45 min after which cells were pelleted, resuspended in 100 µl LB media and plated onto LB agar plates containing 25 µg/ml kanamycin.

2.6.3 Plasmid DNA preparation

Plasmid DNA was routinely prepared from 5 ml bacterial overnight cultures using either Wizard Plus SV Miniprep DNA purification system (Promega) or Plasmid mini kit (Qiagen) as per manufacturers instructions. When higher quantities of plasmid DNA were required either 50 ml or 100 ml bacterial cultures were grown, harvested and the plasmid DNA isolated using Plasmid Midi or Maxi kit (Qiagen) respectively.

2.6.4 Restriction endonuclease digestion of DNA

Digestions of DNA with restriction endonucleases were performed in 20 µl of 1 x appropriate restriction enzyme buffer containing 0.3-1 µg DNA (depending on plasmid size) and usually 3-10 units/ug DNA of restriction endonuclease (obtained from Roche or Promega). Reactions were incubated for 1 h at the appropriate temperature for the restriction endonuclease and terminated by either heat inactivation or addition of 2 µl RE stop dye. Partial digestion of DNA with restriction endonuclease was carried out by assembling a 100 µl reaction containing 1.5-5 µg DNA (depending on plasmid size) in 1 x restriction enzyme buffer. The reaction mixture was divided into 5 tubes, that were placed on ice, such that tube 1 contained 30 µl, tubes 2 to 4 contained 20 µl and tube 5 contained 10 µl volumes. Restriction endonuclease was added to tube 1 and using a different pipette tip, 10 µl from tube 1 was added to tube 2. The serial dilution process was continued up to tube 5 and when completed all five tubes contained 20 µl volumes. All five tubes were incubated for 15 min at the temperature appropriate for the restriction endonuclease and reactions were terminated by addition of 2 µl of RE stop dye. Restriction mapping of the P1 clone was performed by partial restriction digestion analysis. The P1 clone DNA was fully digested with one restriction endonuclease, e.g. *Bam*HI and a series of partial digestions were performed with a second restriction endonuclease, e.g. *Xho*I, Southern hybridisation was performed with radio-labelled probes corresponding to the 5' and 3' end of fully digested fragments which allowed restriction sites internal to the fully digested fragments to be mapped.

2.6.5 Southern and colony hybridisation

To perform Southern hybridisation, DNA digested with a suitable restriction endonuclease was electrophoresed on a 0.8-1% agarose gel and transferred to Hybond

N+ nylon membrane (Amersham Life Science) by capilliary blotting as described (Sambrook et al., 1989). Colony lifts were prepared using Hybond N+ nylon membrane as described (Sambrook et al., 1989). DNA was fixed to Hybond N+ nylon membrane by crosslinking using a UV Stratalinker® 2400 (Stratagene). Pre-hybridisation was carried out by placing Southern blots/colony lifts in a plastic bag containing 5-10 ml of hybridisation buffer (see section 2.12) that was sealed and allowed to incubate in a shaking waterbath at 65°C for 1-4 h. Denatured probe was added into the pre-hybridisation buffer, the bag resealed and returned to the shaking water bath for incubation at an appropriate annealing temperature for at least 12 h. Following probe incubation, the membrane was washed twice in 1 x SSPE; 0.1% SDS at 65°C for 20 min each time and then exposed to X-ray film.

2.7 Protein preparation and analysis

2.7.1 Whole cell protein extracts

Whole cell protein extracts were routinely prepared from confluent 10 cm^2 dishes of cells. The cells were washed in 1 x PBS and harvested by scraping into 400 µl of whole cell extract buffer (see section 2.12). Cells were lysed using three rounds of freeze/thaw, each round consisting of freezing on dry ice, rapidly thawing in a 37°C waterbath and vortexing. Extracts were centrifuged at maximum speed in a bench top microcentrifuge to pellet the cell debris and aliquots of the supernatant were placed transferred to cryotubes (Nunclon, Life Technologies) for storage in liquid nitrogen.

2.7.2 Assaying protein concentration

The concentration of proteins was determined using DC Protein Assay (BIO-RAD) microassay procedure according to manufacturers instructions.

2.7.3 Electromobility shift assay (EMSA)

Approximately 500 μg of whole cell extract, prepared as described in section 2.7.1, was incubated on ice with or without competitor DNA for 20 minutes in 1 x whole cell extract gel shift buffer or 1 x gel shift buffer (see section 2.12) containing 1 μg sheared calf thymus DNA. Approximately 20,000 cpm of a radio-labelled DNA fragment was added to each reaction and incubation was continued at room temp. for a further 20 min Samples were electrophoresed through a 4% non-denaturing polyacrylamide in 0.25 x TBE, fixed, dried and exposed to X-ray film.

2.8 Labelling probes

2.8.1 Radio-labelling oligodeoxynucleotides

Tailing

Reactions were performed in 20 μl of 1 x TdT buffer (Roche), containing 1 pmole oligodeoxynucleotide, 4 μl [α^{32}-P] dATP (3000 Ci/mmol, 10mCi/ml) (Amersham Life Science) and 25 units (1 μl) TdT (Roche) and incubated at 37°C for 1 h. Reactions were terminated by heat inactivation and the probe purified by spin column chromatography (Sambrook et al., 1989) using Sephadex G-50 (Sigma) columns.

Kinasing

Reactions were performed in 20 μl of 1 x T4 polynucleotide kinase buffer (Roche) supplemented with 5 mM DTT containing 1 pmole oligodeoxynucleotide, 5 μl [γ^{32}-P] ATP (>5000 Ci/mmol; 10mCi/ml) (Amersham Life science) and 10 units (1 μl) T4 polynucleotide kinase (Roche). To produce probes with higher specific activity, 15 μl

[γ^{32}-P] ATP (>5000 Ci/mmol; 10mCi/ml) and 0.3 pmole of oligodeoxynucleotide were used instead. Reactions were incubated at 37°C for 1 h, then the reaction terminated by heat inactivation and the probe was purified by spin column chromatography.

2.8.2 Radio-labelling double stranded DNA fragments

Klenow fill in

Reactions were performed in 20 µl of 1 x Klenow buffer (Roche), containing 100 ng of a gel-purified DNA fragment, 0.1 mM each dCTP, dTTP and dGTP, 2 µl [α^{32}-P] dATP (3000 Ci/mmol; 10mCi/ml) (Amersham Life Science) and 2 units (1 µl) Klenow enzyme (Roche) and incubated at 37°C for 30 min. Following this 1 µl of dNTP mixture (5mM of each dCTP, dTTP, dGTP and dATP) was added and incubation continued at 37°C for a further 10 min. Reactions were terminated by heat inactivation and probes were purified by spin column chromatography.

Random primer labelling

Gel-purified DNA fragments were labelled using the Prime-It II random primer labelling kit (Stratagene). Labelled probes were purified using tbe QIAquick nucleotide removal kit (Qiagen).

2.8.3 Synthesis of cRNA probes

Internally radiolabelled cRNA probes were synthesised in 20 µl of 1 x transcription buffer (Roche) containing 500 ng of linearised DNA template, 125 mM each rATP, rGTP, rUTP and 12 µM rCTP, 2.5 µl of [α-^{32}P] rCTP (800 Ci/mmol; 10mCi/ml) (Amersham Life Science) 20 units (1 µl) T7 RNA polymerase (Roche) and 1 µl RNase Inhibitor (Roche) at 37°C for 1 h. The DNA template was removed by addition of 10

units (1 µl) DNase I, RNase-free (Roche) and continued incubation at 37°C for 15 min. Following DNase I digestion, an equal volume of formamide loading buffer was added and the probe electrophoresed through a 5% urea-acrylamide gel and the cRNA probe was gel purified as described in section 2.3.1. Alternatively 280 µl probe elution buffer (Ambion) was added and the probe could be used directly in RNase protection assays. Incorporation of the radiolabelled nucleotide was measured by adsorption to DE-81 filters (Sambrook et al., 1989).

2.9 Sequencing

2.9.1 Automated sequencing

Sequencing reactions were performed using BigDye™ terminator kit (ABI Prism®, PE Biosystems) in 10 µl reactions containing 4 µl terminator ready reaction mix, 400 ng double stranded plasmid DNA or 2 µl PCR product template and 3.2 pmole primer. Sequencing of DNA fragments corresponding to the 5' end of the mouse NRSF/REST gene were more effective with the inclusion of DMSO added to a final concentration of 5% w/v in the reaction. PCR cycling conditions used were 30 cycles 94°C/30 s, 50°C/15 s, 60°C/4 min. Samples were precipitated and resuspended in 3 µl fomamide before electrophoresis through a urea-acrylamide gel prepared using Sequagel Automatrix 4.5 solution (National Diagnostics) on an 377 automated DNA sequencer (Applied Biosystems).

2.9.2 Manual sequencing

Reactions were performed using Sequenase DNA sequencing kit (Amersham Life Science) using [α-^{35}S] dATP (1000Ci/mmol) as per manufacturers protocol.

2.10 Precipitation of nucleic acids

Precipitation of RNA

Stock solution of 5 M NH_4OAc was added to RNA and the concentration adjusted to 0.5 M, then 2.5 volumes of ethanol was added, the samples mixed and placed at -20°C for 15 mins. to precipitate. The RNA was pelleted at maximum speed in a bench top microcentrifuge at 4°C, the pellet washed in 75% ethanol before being air dried and resuspended in TE pH 7.4.

Precipitation of DNA

Stock solution of 2 M NaCl was added to DNA to a final concentration of 0.2 M, then 2 volumes ethanol was added and samples placed on ice for 10 min. The DNA was pelleted at maximum speed in a bench top microcentrifuge and the DNA pellets were washed with 70% ethanol before being air dried and resuspended in TE pH 7.4.

Phenol/chloroform extraction of RNA and DNA samples

The concentration of salt was adjusted using 2 M NaCl or 5 M NH_4OAc as for ethanol precipitation. An equal volume of phenol:chloroform:isoamyl alcohol mix (25:24:1) was added and the samples vortexed to form an emulsion. Samples were centrifuged at maximum speed for 5 min. in a bench top microcentrifuge and the supernatant transferred to a fresh tube, extracted with an equal volume of chloroform and precipitated.

2.11 Oligodeoxynucleotides sequences

1700s	GGG ATT AAC GGC CAT AGC CCA C
1500s	GGT AAG GCT GGG GCT AAC TTC TGC T
1570s	CCG CCC CTT TAA TGA GCC TAA CGC
1270s	CGC GCG CGG ACG GCG AGG GGG CGT GT
As	GCA GCA GCA GAC CGA GCG CCG CCG A
Bs	TAC AGG CCC GAT CGG ACG CCG GTA
Cs_2	GGT GGC AGC GGC ACA TAG GAA TT
1Aa	CTT TGT TGC CGG ACT TTC GGG GGA GGG
$1Aa_2$	CGG GAG CCG CAC ATT CCA GC
$1Aa_4$	ATT CCA GCA CAG GAC GCT GCG
69a	GTC GTT AGG CAG GGC CAT TCC
137a	CCA CGT TGG CTA ACA TGA TGA GC
406s	CAT CCG CTG TGA CCG CTG
934a	CAG TTG AAC TGC CGT GGG
1a ATG	TGG GTG GCC ATG GCT GTA CTC TG
NRSF RE1s	GTA CCG CTG CTG GTG ACC GCG GTC CTG AAA CTT C
NRSF RE1a	GTA CGA AGT TTC AGG ACC GCG GTC ACC AGC AGC G
NaII RE1s	GGT GCT GTC CGT GGT TCT GAA
NaII RE1a	TTC AGA ACC ACG GAC AGC ACC
RV3	CTA GCA AAA TAG GCT GTC CC
GL2	CTT TAT GTT TTT GGC GTC TTC CA
RACE-1	GAC TCG AGT CGA CAT CGA TTT TTT TTT TTT TTT TT
RACE-2	GAC TCG AGT CGA CAT CG
Sp1s	GAT CCC GAT CGG GGC GGG GCG
Sp1a	GAT CCG CCC CGC CCC GAT CGG

2.12 Solutions

Hybridisation buffer
5 x SSPE
5 x Denhardt's solution
0.5% SDS
100 μg/ml SS DNA, sonicated and denatured before addition

Solution hybridisation buffer (1x)
80% formamide,
0.4 M NaCl,
1 mM EDTA
0.04 M PIPES pH 6.5

Whole cell extract buffer
20 mM HEPES pH 7.9
450 mM NaCl
0.4 mM EDTA
25% glycerol
0.5 mM DTT added just prior to use
0.5 mM PMSF added just prior to use

Gel shift buffer (2x)
40 mM HEPES pH 7.9
16% glycerol
10 mM $MgCl_2$
200 mM KCl

Whole cell extract gel shift buffer (5x)
25 mM HEPES pH 7.9
25 mM $MgCl_2$
500 mM KCl
TBS (5.5 mM Tris pH 7.4, 137 mM NaCl)

β-galactosidase staining solution
(Made up in 1 x PBS)
1 mg/ml X-gal
5mM potassium ferricyanide $(K_3Fe(CN)_6)$
5mM potassium ferrocyanide $K_4Fe(CN_6).3H_2O$
2mM $MgCl_2$

Fixative (made up in 1 x PBS)
2% formaldehyde,
0.2% gluteraldehyde

Chapter 3

Page

3.0 RESULTS

3.1 Introduction

The expression pattern of NRSF/REST, showing higher abundance in non-neural than neuronal tissues throughout development and adulthood (Chong et al., 1995; Schoenherr and Anderson, 1995a), is consistent with the proposed role of this transcription factor in silencing neurone-specific gene expression in non-neural tissue. However, NRSF/REST is also expressed within the developing nervous system, in the ventricular zone of the neural tube, although its expression is downregulated in differentiating neurones (Chong et al., 1995; Schoenherr and Anderson, 1995a). Both the inactivation and overexpression of NRSF/REST in the ventricular zone caused alterations in the expression patterns of three NRSE/RE1-bearing genes: Ng-CAM, N-tubulin and SCG10 (Chen et al., 1998; Paquette et al., 2000) and although overt neurogenesis was not affected, sub-populations of neurones displayed axon guidance abnormalities (Paquette et al., 2000). Thus the expression of NRSF/REST in neural precursor cells, and its downregulation upon neuronal differentiation, is important for controlling the proper expression of neurone-specific genes during neurogenesis which in turn is crucial for certain aspects of neuronal development (Paquette et al., 2000). Low levels of NRSF/REST mRNA were recently detected in sub-populations of differentiated neurones in the adult brain (Palm et al., 1998), moreover an increase in NRSF/REST mRNA within hippocampal and cortical neurones is detected following KA-induced seizures (Palm et al., 1998). Transgenic studies have shown that abrogation of the NRSE/RE1 in the L1 and nAChR β2 subunit transgenes causes alterations to reporter gene expression in neuronal tissue, as well as non-neural tissue, during embryonic and post-natal development (Bessis et al., 1997; Kallunki et al., 1997; Kallunki et al., 1998). These experiments imply that NRSF/REST also controls gene expression in neurones. Despite the accumulating evidence that NRSF/REST plays a

crucial role in silencing the expression of multiple neurone-specific genes in non-neural tissue and modulating their expression in developing and mature neurones, the molecular mechanisms that control cell-specific expression of the NRSF/REST gene are currently unknown. The aims of this study were to obtain a genomic clone of the mouse *Rest* gene and to identify the *cis*-acting genomic elements, and the *trans*-acting factors, that govern tissue-specific expression of the *Rest* gene. During the course of this study, the rat *Rest* gene structure was described and shown to contain three 5' non-coding exons (exon I, exon II and exon III) each of which are alternatively spliced onto the first coding exon (exon IV) resulting in mRNA containing three different 5' UTRs designated type A, type B and type C, respectively (Palm et al., 1998). Expression of mRNA transcripts containing type A, 5' UTR constitutes the majority of rat NRSF/REST transcripts in all non-neural and neuronal tissues analysed (Palm et al., 1998). The structure of the 5' end of the mouse NRSF/REST gene was also recently described and shown to have similar organisation to the 5' end of rat gene (Koenigsberger et al., 2000). In addition, the transcription start sites of mRNA initiated at each of the three alternative 5' non-coding exons in mouse were mapped, and the 5' proximal regions of each of the alternative exons were shown to have promoter activity (Koenigsberger et al., 2000).

3.2 Characterisation of the mouse NRSF/REST gene structure

The *cis*-acting genomic elements that regulate transcription of all eukaryotic, protein-coding genes, can be arrayed within several hundred base pairs of the transcription initiation site, but are usually found over larger distances ranging between 1-50 kb. To ensure the genomic clone could be large enough to contain genomic elements required to direct tissue-specific expression of the *Rest* gene, a P1 bacteriophage-based mouse genomic clone was obtained. The P1 genomic clone was characterised and since the *Rest* gene structre was still unknown when the study was initiated, the gene structure was elucidated.

3.2.1 Restriction mapping of a P1 genomic clone containing the mouse NRSF/REST gene

Bacteriophage P1 based genomic clones have several distinct advantages compared to many other vectors used to construct genomic DNA libraries, including the ability to contain inserts of up to 100 kb, a very low frequency of DNA rearrangements owing to maintenance of the plasmid as a single copy and the ease of recovering and purifying plasmid DNA by standard protocols. A P1 genomic clone of the mouse NRSF/REST gene was obtained from Genome Systems, Inc. upon provision of primers: 1s (containing sequence from the first intron) and 137a (containing sequence from exon II), which spanned the most 5' portion of mouse NRSF/REST sequence available at the time. The copy number of the P1 vector backbone, pAd10-*Sac*BII, developed by Sternberg et al. (1992) can be increased by induction of the *Lac*I^q-repressed lytic replicon with IPTG. Nonetheless the yield of the P1 genomic clone was very low from bacterial host strain NS3529 used for the generation of the P1-based genomic library by Genome Systems, Inc. Thus, P1-phage mediated transduction (as described by Sternberg

et al., 1994) was employed to transfer the P1 genomic clone to Cre⁻ host strain JS5 resulting in a >10-fold increase in P1 DNA yield. The P1 genomic clone was restricted with *Pst*I, *Hin*dIII and *Acc*65I and subcloned into pBluescript® II SK(+) (Stratagene). Subsequently P1 restriction fragments spanning the NRSF/REST gene and flanking sequences were identified and sequenced (Figure 1). Furthermore *Bam*HI and *Xho*I sites within the mouse NRSF/REST gene locus were mapped as shown in Figure 1.

3.2.2 Determination of the mouse NRSF/REST gene structure

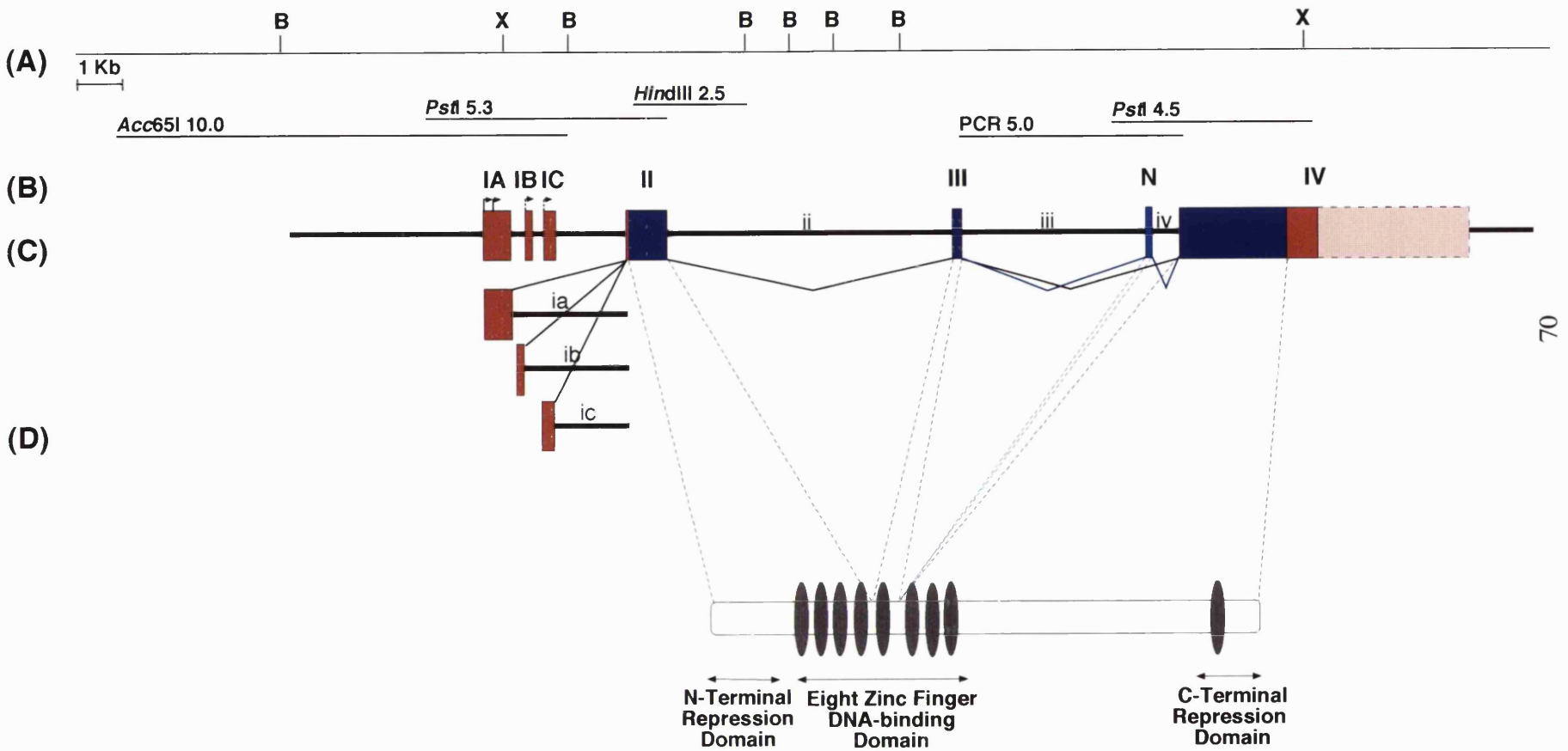
Identification of multiple 5' UT exons within the mouse NRSF/REST gene

Since full length cDNA sequence of mouse NRSF/REST was unavailable, sequence corresponding to the 5' UTR was obtained by performing 5' RACE PCR on RNA extracted from 3T3 cells, a mouse NRSF/REST expressing fibroblast cell line (Figure 5). The longest 5' RACE PCR clone contained 115 bp of sequence upstream of the adenine of the initiation codon. The 5' RACE sequence showed 77% homology with 116 bp of the 3' portion of rat type A, 5' UTR (accession number AF037200) and 70% homology with 118 bp of the 3' portion of human 5' UTR (accession number U13879 and HSU22680) indicating this sequence corresponds to at least part of the 5' UTR. When aligned to mouse NRSF/REST genomic sequence, the 5' RACE sequence diverged 9 bp upstream of the adenine of the initiation codon, indicating that the remaining 106 bp of 5' RACE sequence was derived from an upstream exon(s). By performing Southern hybridisation, a radio-labelled oligodeoxynucleotide probe, Aa, corresponding to 5' RACE sequence between positions +442 and +468 (numbering relative to transcription initiation site at +1 as shown in Figure 8) was shown to hybridise to a 2.0 kb *Hin*dIII restriction fragment located approximately 1.2 kb upstream of the initiation codon (data not shown). Sequence analysis of the 2.0 kb *Hin*dIII restriction fragment (Figure 8) revealed the 106 bp corresponding to entire 5' RACE

sequence was present. Furthermore alignment with rat type A, 5' UTR sequence (accession number AF037200) showed that 312 bp of the mouse NRSF/REST genomic sequence was 85% homologous with rat type A, 5' UTR sequence and this region was designated exon IA (Figure 1).

During the course of these studies, rat NRSF/REST mRNAs with three distinct 5' UTRs, designated type A, type B and type C, were identified which are transcribed from alternative 5' UT exons: exon I, exon II and exon III, respectively within the rat NRSF/REST gene (Palm et al., 1998). To determine if the mouse gene also contained alternative 5' UT exons, the sequences of rat type B and type C, 5' UTRs were aligned to mouse NRSF/REST genomic sequence. A 79 bp region located 545 bp downstream of exon IA showed 99% homology with rat type B, 5' UTR sequence (accession number AF037201) and was therefore designated exon IB (Figure 1). Furthermore a 145 bp region, located 216 bp downstream of exon IB and 1.2 kb upstream of exon II, showed 89% homology with rat type C, 5' UTR sequence (accession number AF037202) and was therefore designated exon IC (Figure 1). Comparison of the structure of the 5' portion of the mouse and rat NRSF/REST genes revealed that the sizes of the intronic regions between exon IB and exon IC and between exon IC and exon II are similar in the mouse and rat genes. In an attempt to measure the size of the intron separating exon I (exon IA in mouse) from the first coding exon in the rat gene (exon II in mouse), Palm et al., (1998) used exonic primers to amplify rat genomic DNA. Their failure to generate any specific amplified products was taken to imply the presence of a large (>15kb) intron. However, the size of this intron in mouse is, in fact, approximately 2.2 kb. Difficulty was also experienced in amplifying DNA across this intron but this was

Figure 1 **Genomic organisation of the mouse NRSF/REST gene.** **(A)** Restriction map of the mouse NRSF/REST gene locus with *Xho*I, X and *Bam*HI, B sites indicated. **(B)** Subcloned P1 clone restriction fragments and PCR fragments. **(C)** Linear representation of the exon/intron structure of the mouse NRSF/REST gene showing the relative positions of exons (shown as boxes) and introns (shown as thick, horizontal lines). The ORF is shown as blue filled boxes. The neural-specific exon (N) is filled lighter blue. The 5' and 3' UTRs are shown as red filled boxes. Boxes with dashed stroke denote putative exonic sequence. Exons are denoted by uppercase characters IA, IB, IC and from II to IV. Introns are denoted by lowercase characters ia, ib and ic (denoting the first intron) and from ii to iv. Arrowheads indicate the position of clusters of transcription start sites identified in exon IA. Arrowheads with dashed lines indicate positions of putative transcription start sites within exon IB and exon IC. Thin lines indicate the regions that are spliced out from the primary transcript. Thin blue lines indicate the alternative, neural-specific splicing pattern. **(D)** Dashed lines indicate the region of NRSF/REST protein encoded by exons II, III, IV and exon N. A linear representation of the NRSF/REST protein is shown with functional domains indicated. Ovals denote zinc finger motifs.


(A)
B
X
B
B
B
B
B
X
1 Kb
Acc65I 10.0
PstI 5.3
HindIII 2.5
PCR 5.0
PstI 4.5
(B)
IA IB IC
II
III
N
IV
(C)
i
ii
iii
iv
ia
ib
ic
(D)
N-Terminal Repression Domain
Eight Zinc Finger DNA-binding Domain
C-Terminal Repression Domain

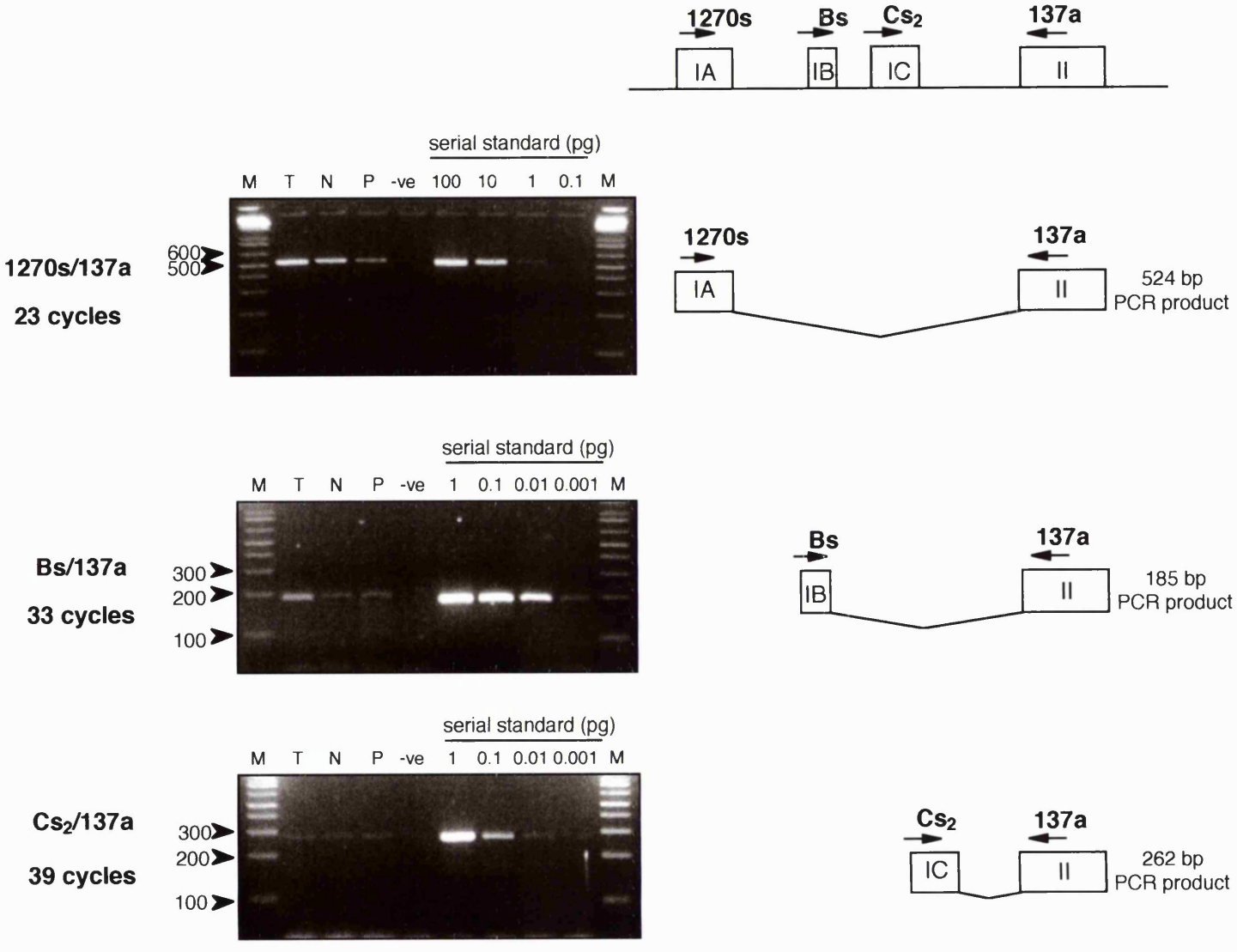
attributed to the high GC content contained within this region, therefore the size of this intron was possibly overestimated in the rat.

Having identified mouse sequences homologous to the rat alternative 5' UT exons, RT-PCR was performed to determine if these mouse 5' UT exons were transcribed. Primer pairs, consisting of an upstream primer derived from each 5' UT exon (1270s in exon IA (+161/+184); Bs in exon IB (+1125/+1149); and Cs_2 in exon IC (+1410/+1433)) in conjunction with a downstream primer derived from exon II, (137a), were used on cDNA derived from 3T3 fibroblast and Neuro2a neuroblastoma mouse cell lines and PC12 cells, a rat pheochromocytoma cell line that expresses low levels of NRSF/REST (Chong et al., 1995; Shimojo et al., 1999) (Figure 2). The lengths of PCR products obtained were consistent with all transcripts containing only exon IA, exon IB or exon IC spliced onto exon II and DNA sequencing confirmed this observation. No PCR products were obtained in absence of reverse transcriptase indicating that the observed PCR fragments were amplified from cDNA template. The presence of amplification products from 3T3 and Neuro2a cDNA template using primer pairs to each of the 5' UT exons indicates that like the rat gene, mouse NRSF/REST mRNA is transcribed from the start of three alternative 5' UT exons: exon IA, exon IB and exon IC, resulting in transcripts containing three different 5' UTRs which were designated type A, type B and type C, respectively.

mRNA containing type A, 5' UTR is the predominantly expressed NRSF/REST transcript in 3T3, Neuro2a and PC12 cell lines

The relative abundance of NRSF/REST mRNA containing different 5' UTRs was examined using semi-quantitative PCR. PCRs were performed on serial plasmid standards and cDNA derived from 3T3, Neuro2a and PC12 cells (Figure 2).

Figure 2 **Semi-quantitative RT-PCR analysis of NRSF/REST mRNA containing type A, type B and type C, 5'UTRs** RT-PCR analysis was performed on 0.8 μg DNased, total RNA extracted from mouse 3T3 (T) and Neuro2a (N) cells and rat PC12 (P) cells. PCRs were performed on serial plasmid standards to show signal strength was proportional to the amount of template. Decreasing amounts of plasmid DNA (given in picograms) were used as indicated above the respective lanes. RNA without reverse transcription is included as a negative control (-ve). PCRs were carried out in the presence of 5% (w/v) DMSO, using primer pairs consisting of an upstream primer to each of the 5' UT exons IA, IB and IC in conjunction with a downstream primer to exon II. PCR cycling conditions were 1 cycle 94°C/3 min; 23-39 cycles of 94°C/30 s, 55°C/30 s and 72°C/2 min. PCR products were electrophoresed through 2% MetaPhor® agarose; M, 1-Kb plus DNA marker (Life Technologies). The primer pair and number of cycles used in each PCR are indicated to the left of each gel photo. Arrowheads point to positions of DNA fragments in 1-Kb DNA ladder with sizes given in base pairs. A map of the 5' portion of the mouse NRSF/REST gene is shown on the right with positions of primers indicated by arrows and the predicted sizes of PCR products given.


1270s
Bs
Cs2
137a
IA
IB
IC
II
serial standard (pg)
M T N P -ve 100 10 1 0.1 M
1270s/137a
23 cycles
600
500
1270s
137a
IA
II
524 bp
PCR product
serial standard (pg)
M T N P -ve 1 0.1 0.01 0.001 M
Bs/137a
33 cycles
300
200
100
Bs
137a
IB
II
185 bp
PCR product
serial standard (pg)
M T N P -ve 1 0.1 0.01 0.001 M
Cs2/137a
39 cycles
300
200
100
Cs2
137a
IC
II
262 bp
PCR product

NRSF/REST mRNA containing type A, 5' UTR was detectable after only 23 cycles of amplification. In contrast, detection of mRNA containing type B and type C, 5' UTRs required 33 and 39 cycles of amplification, respectively. In 3T3 cells a strong signal was detected for type A, 5' UTR that was similar to, or stronger than, the signal seen in the plasmid standard containing 10 pg DNA and the bands corresponding to type B and type C, 5' UTR were of similar intensity to the band observed in plasmid standard containing 0.001 pg DNA. Thus in 3T3 cells the signal strength for mRNA containing type A, 5' UTR was approximately 1000-fold stronger than for mRNA containing type B and type C, 5' UTRs. Similar results were obtained using cDNA derived from Neuro2a and PC12 cells showing a much stronger signal strength for mRNA containing type A, 5' UTR compared to mRNA containing type B and type C, 5' UTR. Thus semi-quantitative PCR analyses indicate that NRSF/REST mRNA containing exon IA is, by far, the most abundant NRSF/REST transcript expressed in mouse 3T3 and Neuro2a cells and in rat PC12 cells.

The two repression domains of NRSF/REST protein are encoded by separate exons

Genomic DNA sequence from P1 clone restriction fragments was compared to mouse NRSF/REST cDNA sequence to determine the exon/intron structure of the mouse NRSF/REST gene (shown in Figure 1). Sequence of the entire open reading frame (ORF) was obtained by combining partial cDNA sequence on Genbank (accession number MMU13878) and sequence determined from a cDNA clone, pMT mNRSF (a gift from D.J. Anderson, Caltech). The mouse NRSF/REST ORF was found to be encoded by three exons (Figure 1). The first coding exon, exon II (0.9 kb), encodes part of the ORF extending from the translation initiation codon, through the N-terminal repression domain, to immediately beyond the end of the fourth zinc finger motif. Exon III (84 bp) encodes the fifth zinc finger motif, and exon IV separated from exon III by an

intron of approximately 5 kb, encodes the remainder of the ORF which includes the C-terminal repression domain. The mouse NRSF/REST gene structure, described here, compared with the partial descriptions of rat and human NRSF/REST gene structures that were recently reported by Palm, et al. (1998; 1999) reveal that the exon/intron organisation of the ORF is conserved across these species.

Northern blot analysis of human and rat NRSF/REST mRNA indicates the presence of a long 3' UTR (Chong et al., 1995; Palm et al., 1998). The 3' UTR of both rat and human NRSF/REST mRNA have been mapped to the exon encoding the C-terminal portion of NRSF/REST. The sequence of the 3' UTR of mouse NRSF/REST is unknown, however, alignment of 0.9 kb of mouse NRSF/REST genomic sequence, downstream of the stop codon, to Genbank showed similarity to the 5' portion of human and rat NRSF/REST 3' UTRs (accession number HSU22680, HSU22314 and AF037199). These data indicate that at least a portion of mouse NRSF/REST 3' UTR is present within exon IV which encodes the C-terminus of NRSF/REST (Figure 1). Since most aspects of organisation of the NRSF/REST gene are conserved between mouse, rat and human, the entire 3' UTR of mouse NRSF/REST mRNA may also be present in exon IV, although genomic DNA sequence through to the polyadenylation signal of mouse NRSF/REST was not obtained to provide confirmation.

Identification of a neural-specific exon that separates the two repression domains of NRSF/REST

During the course of performing RT-PCR analyses on cDNA derived from different cell lines to determine NRSF/REST gene expression, using primers to the DNA-binding domain of NRSF/REST, consistently generated two PCR products using cDNA derived from NB-OK-1 cells, a human neuroblastoma cell line (Figure 3). Both

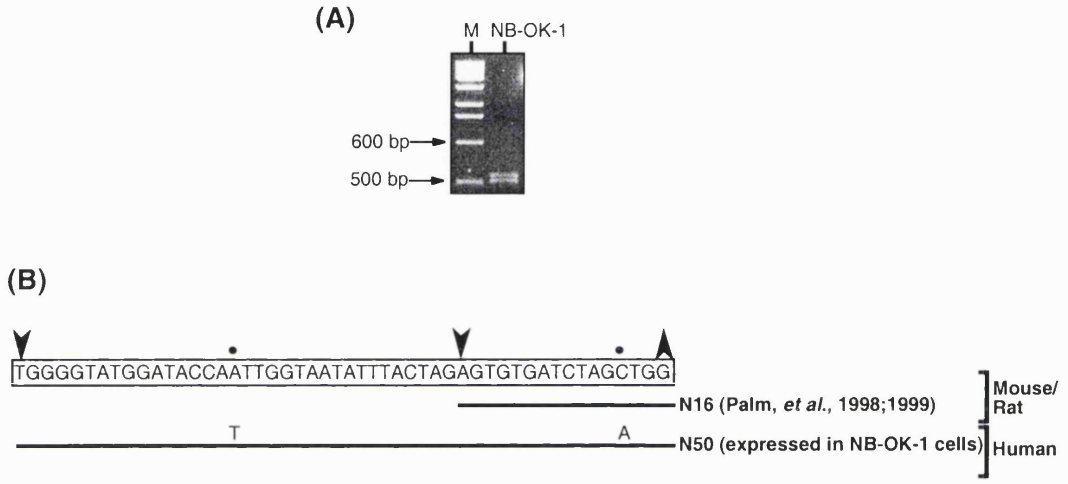


Figure 3 **Detection of NRSF/REST mRNA splice variants by RT-PCR analysis** RT-PCR analysis was performed on 0.4 μg of DNase I -treated, total RNA per reaction as detailed under Methods and Materials. (**A**) NB-OK-1 RNA was subjected to RT-PCR with primer pair 406s/934a which span the NRSF/REST DNA binding domain. M, 1-Kb DNA marker (Life Technologies). (**B**) Sequence of a 50 bp region of the mouse NRSF/REST gene is shown. Filled dots indicate nucleotide mismatches between the human and mouse sequence. Arrowheads pointing down denote splice acceptor sites and arrowheads pointing up denote splice donor sites. Lines represent the sequence spanned by exon N in different species.

PCR fragments were cloned and sequenced. A 529 bp PCR fragment was obtained that corresponded to NRSF/REST cDNA described earlier (Chong et al., 1995; Schoenherr and Anderson, 1995a). However, a 579 bp PCR fragment was also obtained and found to correspond to a novel NRSF/REST splice variant containing a 50 bp exon (exon N50) insertion between the fifth and sixth zinc fingers motifs. Sequence analysis revealed that human NRSF/REST mRNA, containing the alternative 50 bp exon, is predicted to encode an in-frame stop codon after the fifth zinc finger motif. Therefore the human NRSF/REST mRNA splice variant, expressed in NB-OK-1 cells, is predicted to encode a truncated NRSF/REST protein isoform, designated $NRSF_{SF}$, containing the N-terminal repression domain and five of the zinc fingers of the DNA binding domain. A recently identified human NRSF/REST splice variant, that is highly expressed in small cell lung cancer (Coulson et al., 2000), contains a 50 bp insertion between the fifth and sixth finger motifs and comparison of its sequence to the 50 bp insertion found in NRSF/REST expressed in NB-OK-1 cells shows the sequence is identical. Alignment of the sequence of exon N50 with mouse NRSF/REST genomic sequence revealed 96% homology with a 50 bp region within the intronic sequence between exon III and exon IV. During the course of this study the detection of mouse NRSF/REST mRNA splice variants, containing a 16 bp insert (exon N16) between the fifth and sixth zinc finger motifs, were reported to be expressed in neuroblastoma cell lines and brain tissue but not in non-neural tissues (Palm et al., 1999). Comparison of exon N16 sequence to mouse genomic sequence showed it is identical to the 3' portion of the 50 bp region in the mouse NRSF/REST gene that is homologous to exon N50 (Figure 3). This 16 bp sequence, positioned 4.2 kb downstream of exon III and 0.8 kb upstream of exon IV in the mouse gene, was designated exon N (for neural-specific exon) (Figure 1).

Exon/intron boundaries between mouse, rat and human NRSF/REST genes are conserved

The sequence at the predicted exon/intron junctions of the mouse, rat and human NRSF/REST genes are shown in Table 1. The sequences of mouse, rat and human splice donor sites were found to be 100% conserved while the sequences of mouse and rat splice acceptor sites were conserved but diverged from human splice acceptor sites. Nonetheless the positions of splice acceptor sites in mouse, rat and human genes have been conserved in most cases. The exception is the splice acceptor site which gives rise to exon N50 in human. The AG dinucleotide that is highly conserved in all splice acceptor sites is present as GG in mouse and rat and is therefore not used (Table 1). A splice acceptor site found a further 34 bp downstream is used instead, in mouse and rat, resulting in a shorter exon N. Although this splice acceptor site is also present in the human NRSF/REST gene no evidence demonstrating its use has been reported.

3.2.3 Detection of mouse NRSF/REST transcripts in P1 transfected HEK 293 cells

Sequence analysis of P1 restriction fragments showed that the entire mouse NRSF/REST gene including, at least, 10 kb of 5' flanking sequence and 9 kb of 3' flanking sequence is present in the P1 genomic clone. To determine whether sufficient NRSF/REST genomic sequence was present to direct expression of the mouse NRSF/REST gene the P1 genomic clone was transfected into HEK 293 cells, a human NRSF/REST expressing cell line and RT-PCR analysis was performed to determine if mouse transcripts could be detected. RNA extracted from wild-type HEK 293 cells and P1 transfected HEK 293 cells were subjected to RT-PCR using species specific primers to discriminate between the endogenous and transfected NRSF/REST genes. Primer, 1270s (containing sequence to exon IA) used in conjunction with primers 69a and 137a (containing sequence to exon II) yielded no PCR products from cDNA derived from

	5' splice site (splice donor)	3' splice site (splice acceptor)	
consensus	**MAG**IGTRAGT	Y_nNYAGI**G**	
exon IA	**GAGG**|GTAGGAC ····|xxx	TTTATCCAG|**AGTACAGTTA** ·····|-C·····	exon II
exon IB	**GAG**|GTAACT ···|·····	TTTATCCAG|**AGTACAGTTA** ·····|-C·····	exon II
exon IC	**GAG**|GTGCGT ···|·····	TTTATCCAG|**AGTACAGTTA** ·····|-C·····	exon II
exon II	**CAG**|GTAAG ···|····· ···|·····	ACTTTGCAG|**GAG** ·T·····|··· ·T·····|··	exon III
exon III	**CAG**|GTTGGT ···|····· ··|·····	TACCTTTAG|**GTG** ·····|··· ·····|···	exon IV
exon III	**CAG**|GTTGGT ···|····· ··|·····	TTGGACCGG|**TGG** (not used) ·····|··· (not used) ·····A·|··· ATTTACTAG|**AGT** ·····|··· ·····|··· (not used)	exon N50 exon N16
exon N16IN28 N50IN62	**TGG**|GTATGTATTCAG|GTA ···|·····|··· ···|·····|···	TACCTTTAG|**GTG** ·····|··· ·G···C··|···	exon IV

Table 1 **Comparison of sequence at exon/intron boundaries of mouse, rat and human NRSF/REST genes** Mouse sequences at the predicted exon/intron boundaries were derived by comparison of mouse NRSF/REST genomic sequence to cDNA sequence. Mouse sequences are shown in black; rat sequences are shown in blue (taken from Palm et al., 1998) and human sequences are shown in red (taken from Palm, et al., 1999). Exon sequences are shown in bold text and intron sequences in plain text. I denotes exon/intron boundary; x denotes sequences not determined.

untransfected HEK 293 cells (Figure 4, lanes 3 and 8). Primers to cyclophillin produced a robust PCR band (lanes 2 and 7) indicating that a lack of PCR product was not due to a lack of cDNA. In contrast, a PCR product was obtained using cDNA derived from transfected HEK 293 cells and the length of this PCR product was identical to the PCR product obtained from cDNA derived from mouse 3T3 cells (lanes 4, 5 and 9, 10). As no PCR products were obtained in absence of reverse transcriptase (lanes 6 and 11), and the primers were derived from separate exons, the observed PCR products indicate the presence of mouse NRSF/REST mRNA in transfected HEK 293 cells. The presence of correctly spliced mouse NRSF/REST transcripts in transfected HEK 293 cells demonstrates that the P1 genomic clone contains sufficient sequence to drive NRSF/REST gene expression. Moreover since the entire P1 clone was transfected *cis*-regulatory elements maybe exonic, intronic or within 3' flanking sequence. These latter possibilities have numerous precedents in genes encoding transcription factors, for example, *Mash1* (Verma-Kuvari et al., 1998) and *Csx/Nkx2.5* (a homeobox transcription factor) (Tanaka et al., 1998) which both require sequences that lie 3' of the coding sequence to generate appropriate gene expression patterns in transgenic mice.

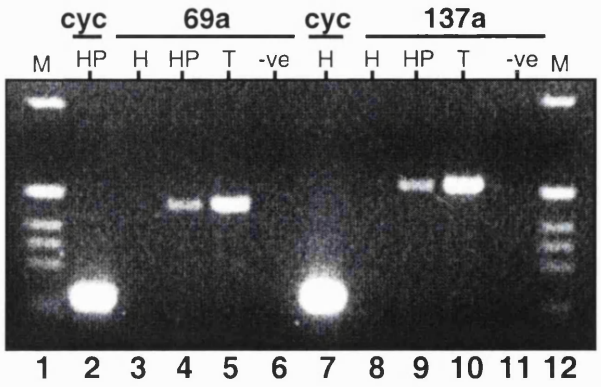


Figure 4 **RT-PCR analysis of P1 clone transfected HEK 293 cells** HEK 293 cells, grown in 600 mm dishes, were transiently co-transfected with 0.9 μg of P1 genomic clone and 0.1 μg pCMVβ DNA (Clontech) using TFx™-50 (Promega) as detailed under Methods and Materials. Subsequent LacZ histocytochemistry showed that >1% of cells were transfected. RT-PCR analysis was performed on 0.4 μg of DNase I-treated, total RNA extracted from 3T3 (T), HEK 293 (H) and P1 transfected HEK 293 (HP) cells. P1-transfected HEK RNA, without reverse transcription, was included as a negative control (-ve). PCRs were performed in the presence of 5% (w/v) DMSO using primer pairs to cyclophillin (lanes 2 and 7) and mouse NRSF/REST: 1270s/69a (lanes 3, 4, 5 and 6) and 1270s/137a (lanes 8, 9, 10 and 11). Cycling conditions used were: 1 cycle 94°C/3 min; 35 cycles 94°C/45 s, 63°C/30 s and 72°C/1 min; 1 cycle 72°C/5 min. PCR products were electrophoresed through 2% MetaPhor® agarose; M, 1-Kb DNA marker (Life Technologies).

3.2.4 Summary

A mouse genomic P1 clone was obtained, containing the entire *Rest* gene locus and at least 10 kb of 5' flanking sequence and 9 kb of 3' flanking sequence. This clone contains sufficient sequence to direct expression of mouse NRSF/REST transcripts in a human, NRSF/REST expressing cell line. Determination of the mouse *Rest* gene structure showed that the N-terminal repression domain and the first four zinc finger motifs forming the DNA binding domain; the fifth zinc finger motif; and the remainder of the coding region, including the remaining three zinc finger motifs of the DNA binding domain, the C-terminal repression domain, and the 3' UTR are encoded by exon II, exon III and exon IV, respectively. A human NRSF/REST splice variant containing a 50 bp exon between the fifth and sixth zinc finger motifs is expressed in NB-OK-1 human neuroblastoma cells and is predicted to encode a NRSF/REST isoform, designated $NRSF_{SF}$, consisting of the N-terminal repression domain and five of the zinc finger motifs forming the DNA binding domain. Similar splice variants have been identified in mouse, as well as rat, neuroblastoma cell lines and brain tissues (Palm et al., 1998; 1999) which contain a short exon corresponding to the distal portion of the 50 bp exon identified in human and these also encode $NRSF_{SF}$. Three 5' UT exons were identified, exon IA, exon IB and exon IC which are transcribed and alternativly spliced onto the first coding exon, exon II, resulting in mRNA containing three different 5' UTRs designated type A, type B and type C, respectively. The much higher abundance of mRNA containing type A, 5' UTR in mouse fibroblast and neuroblastoma cell lines suggests the promoter directing transcription of mRNA containing type A, 5' UTR is responsible for the vast majority of transcripts.

3.3 Transcription initiation site analysis

To identify the 5' flanking sequence of the mouse *Rest* gene, several complementary methods were used to determine the positions of the transcription initiation sites, consisting of mapping the 5' termini of mRNA and carrying out *in vitro* transcription assays using NRSF/REST genomic fragments.

3.3.1 5' RACE PCR analysis indicates that type A, 5' UTR is approximately 300 bp long.

As a means to mapping the 5' termini of mouse NRSF/REST mRNA transcripts, the length of the 5' UTR was determined by using 5' RACE PCR analysis as shown in Figure 5. The cDNA template, derived from 3T3 mouse cells, was tailed with dATP to form the 5' anchor site and subjected to two rounds of PCR amplification using, first, anchor primer RACE-1, in conjunction with NRSF/REST primer, 137a (the 5' end of which anneals 158 bp downstream of the 5' end of exon II) and, second, RACE-2 primer in conjunction with an internal NRSF/REST primer, 69a (the 5' end of which anneals 96 bp downstream of the 5' end of exon II). PCR products containing sequence corresponding to type A, 5' UTR, the most abundant mRNA, were detected by Southern hybridisation using radio-labelled oligodeoxynucleotide probes: As and Aa, corresponding to positions +273/+303 and +468/+442, respectively. A smeary signal that extends from approximately 280-480 bp was obtained, however, the strongest hybridisation signal corresponded to a PCR product of approximately 400 bp (Figure 5), and since 96 bp should correspond to constitutive exon II sequence, this PCR product was expected to contain approximately 300 bp of exon IA sequence extending from position +224 to position +528 corresponding to the 3' end of exon IA. The 5' RACE PCR products were cloned and screened by colony hybridisation using radio-labelled oligodeoxynucleotide probes, As and Aa, and six positive colonies were identified.

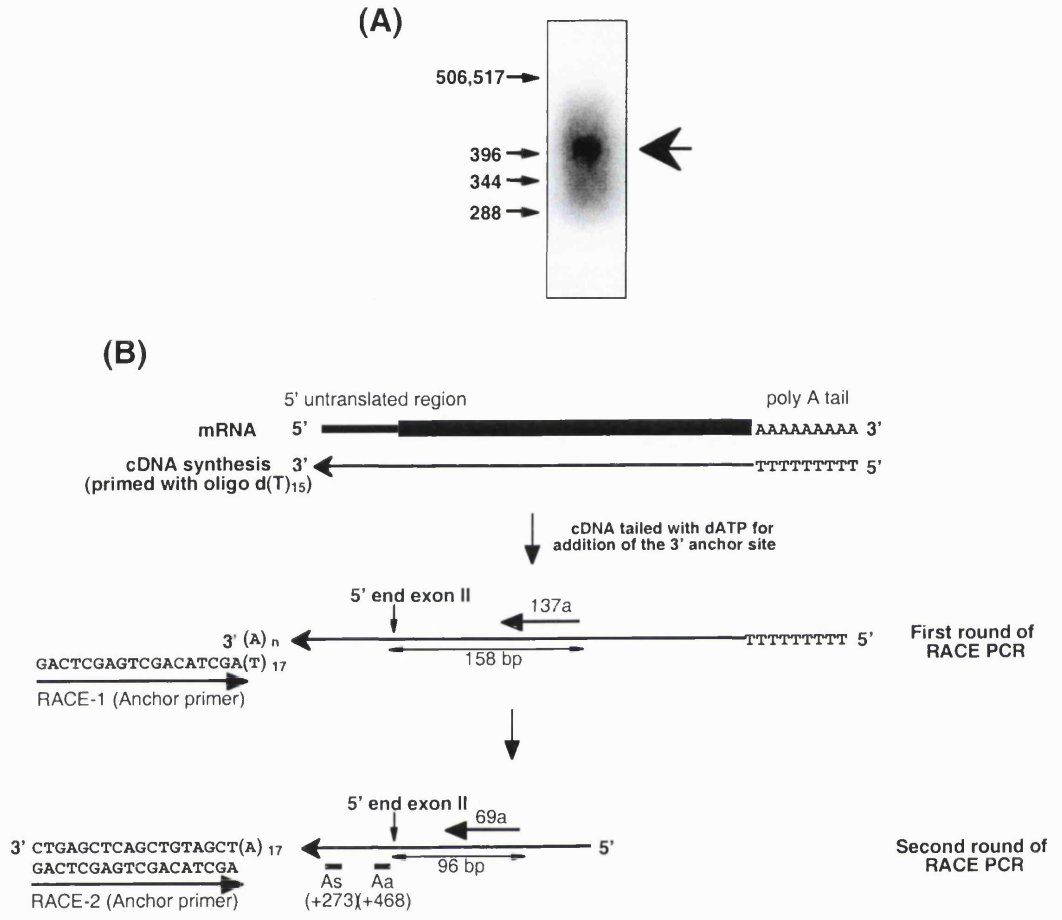


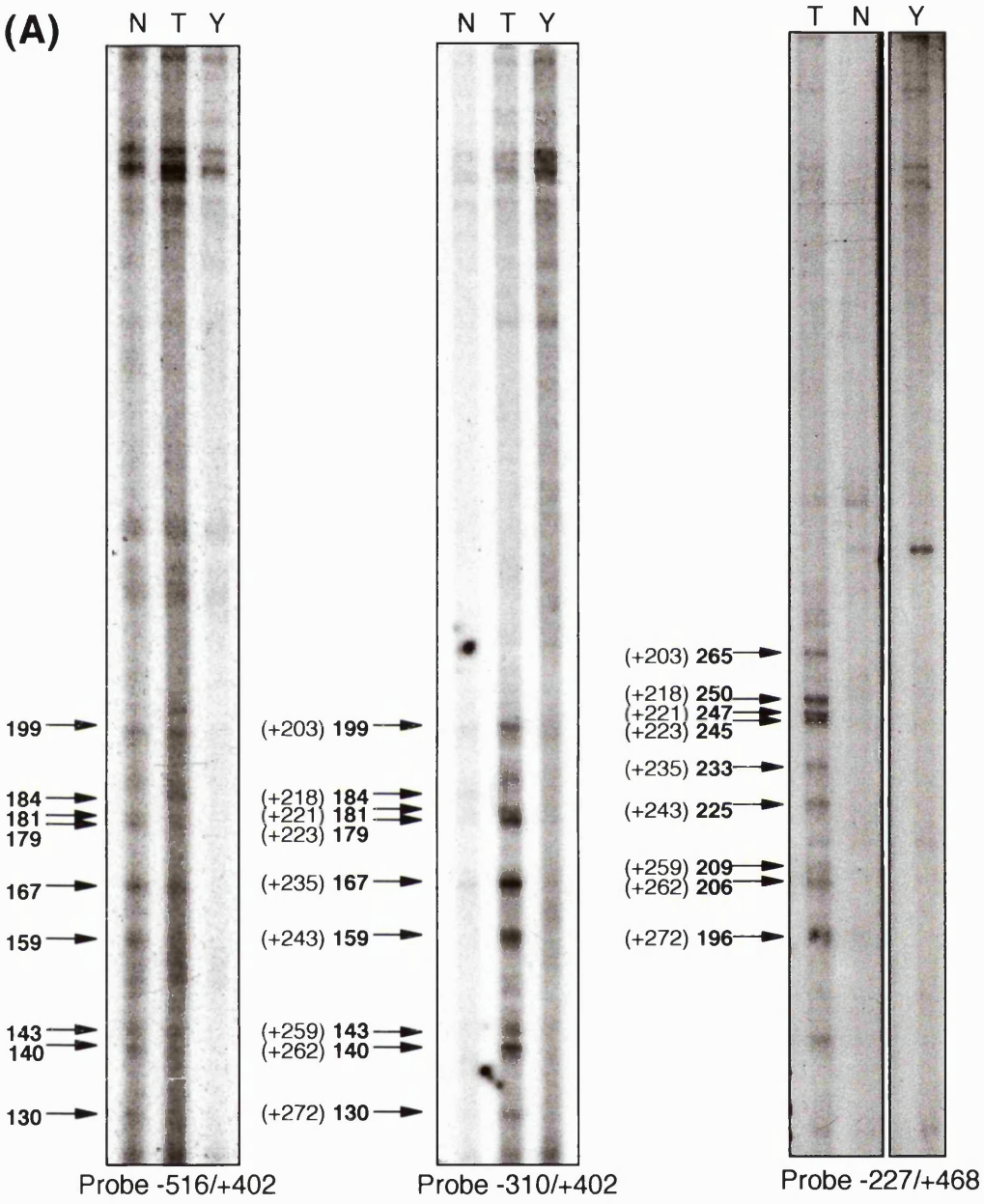
Figure 5 **5' RACE PCR analysis of mouse NRSF/REST mRNA (A)** 5' RACE PCR products from the second round of amplification were electrophoresed through 2% MetaPhor® agarose and subjected to Southern hybridisation using radio-labelled, oligodeoxynucleotide probes: As and Aa (the 5' ends of which correspond to positions +273/+303 and +468/+442 in exon IA). Arrows point to positions of DNA fragments in 1-Kb DNA ladder (Life Technologies) with sizes given in base pairs. The large arrow points to the strongly hybridising band. **(B)** Schematic showing procedure for 5' RACE PCR (adapted from Frohman et al., 1988).

Sequence analysis showed one clone contained sequence corresponding to exon IA sequence, but between positions +423 and +528 only, and the remaining clones were false positives (contained no insert). Although Southern hybridisation indicated that longer 5' RACE PCR products were obtained, a large number of 5' RACE clones would have been needed to be analysed to obtain 5' RACE clones containing full length type A, 5' UTR, therefore, alternative methods were used to map the position of the 5' termini of mouse NRSF/REST mRNA.

3.3.2 Identification of a cluster of transcription initiation sites within exon IA between positions +203 and +272

To map the exact position of the 5' termini of NRSF/REST mRNA containing type A, 5' UTR, RNase protection analyses were performed. Two internally radio-labelled cRNA probes, spanning the 5' end of exon IA and 5' flanking sequence, corresponding to positions -516 to +402 and positions -310 to +402 generated an identical series of protected bands that were specific to 3T3 and Neuro2a RNA and which were not generated using yeast RNA (Figure 6). A third cRNA probe spanning positions -227 to +468, generated a similar series of bands that migrated with an incremental increase in molecular weight that is consistent with the increase in length of the 3' end of the probe (Figure 6). Since RNA migrates slower than DNA in urea-acrylamide gels (Sambrook et al., 1989), specific protected fragments generated with cRNA probe -227/+468 must be 66 nucleotides longer than specific protected fragments produced with cRNA probes -516/+402 and -310/+402, however a difference of 67 bp indicated that the length of RNA molecules needed to be corrected by 1.5% (corrected lengths of protected fragments are given in Figure 6). The positions of the transcription initiation sites, corresponding to the 5' termini of NRSF/REST mRNA, were determined by subtracting the length of the protected fragment from the position corresponding to

Figure 6 **RNase protection analysis of the 5' end of mouse NRSF/REST mRNA (A)** Three cRNA probes, internally radio-labelled with [α-^{32}P] rCTP, were synthesised using T7 RNA polymerase using linearised plasmid templates. cRNA probes –516/+402 and –310/+402 were generated from pBSK Xho 1.0 cut with *Bam*H1 and *Spe*I, respectively. cRNA probe –227/+468 was generated from pGEMT-easy 1700s-Aa cut with *Sal*I. cRNA probes –516/+402 and –310/+402 were each annealed with 10 μg yeast RNA (Y), 10 μg 3T3 total RNA (T) and 50μg and 10μg Neuro2a total RNA (N), respectively. cRNA probe –227/+468 was annealed with 2 μg yeast RNA, 2 μg 3T3 poly A$^+$ RNA and 2 μg Neuro2a poly A$^+$ RNA. Hybridisation was performed in 20 μl of 80% formamide, 400 mM NaCl, 40 mM PIPES (pH 7.9) and 1 mM EDTA at 55°C overnight. Samples were subjected to RNase digestion as described under Methods and Materials and electrophoresed through a 3.5% urea-polyacrylamide gel. The gel was dried and exposed to X-ray film. A DNA sequence ladder was used as a size marker (not shown). Arrows indicate protected cRNA fragments specific to 3T3 and Neuro2a RNA. Apparent lengths of protected cRNA fragments (shown in bold) have been corrected for 1.5% slower migration of RNA compared to DNA on denaturing polyacrylamide gel. The corresponding positions of transcription initiation sites are given in parantheses **(B)** Map of the 5' end of the mouse NRSF/REST gene with the regions spanned by each probe indicated.


(A)
N T Y
199
184
181
179
167
159
143
140
130
Probe -516/+402
N T Y
(+203) 199
(+218) 184
(+221) 181
(+223) 179
(+235) 167
(+243) 159
(+259) 143
(+262) 140
(+272) 130
Probe -310/+402
T N Y
(+203) 265
(+218) 250
(+221) 247
(+223) 245
(+235) 233
(+243) 225
(+259) 209
(+262) 206
(+272) 196
Probe -227/+468
(B)
5'
Exon IA
3'
-516
+1
+528
-516
+402
-310
+402
-227
+468

the 3' end of the cRNA probe. A cluster of nine transcription initiation sites was mapped within a 69 bp segment of the mouse NRSF/REST gene between positions +203 and +272 (Figure 8). Interestingly a transcription start site at position +218 coincides with the position corresponding to the longest rat type A, 5' UTR containing cDNA clone isolated (Palm, et al., 1998). Moreover the transcription initiation sites identified by RNase protection analysis are predicted to generate type A, 5' UTR ranging between 265 bp and 334 bp long, and this is consistent with the 5' RACE PCR data which indicated that the length of the type A, 5' UTR is approximately 300 bp long. The cluster of nine transcription initiation sites are used in both 3T3 and Neuro2a cells, however, the intensity of the bands obtained using Neuro2a RNA were much lower than those obtained using 3T3 RNA (Figure 6), and this finding correlates with published data demonstrating that NRSF/REST mRNA levels are ~30-fold lower in Neuro2a cells compared to 3T3 cells (Lönnerberg et al., 1996).

Intriguingly RT-PCR analysis of cDNA derived from 3T3 cells using a primer within exon IA (1270s) in conjunction with a primer within exon II (137a) demonstrated that sequence as far upstream as position +161 within exon IA is transcribed (Figure 4), thus RNase protection and 5' RACE PCR data still corresponded to mRNA transcripts initiating within known exon IA sequence. Failure to observe RNase protected fragments, that could correspond to mRNA transcripts initiated upstream of position +203, suggested that the abundance of such transcripts was below the limits of detection.

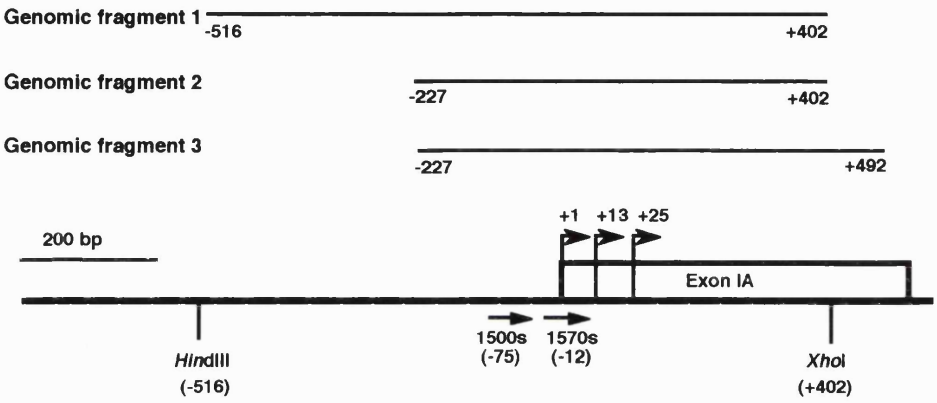
3.3.3 Identification of a second cluster of transcription initiation sites further upstream within exon IA

Since mRNA corresponding to transcripts initiating upstream of position +203 were detected by RT-PCR but not by RNase protection assays, an alternative method

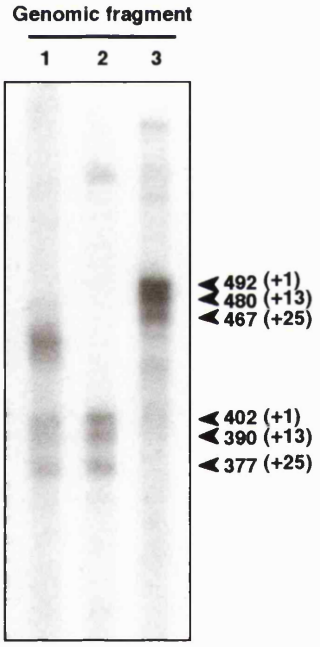
was used to identify the positions of transcription initiation sites that may exist further upstream within exon IA. *In vitro* transcription assays, as a method of determining the position of transcription start sites, have been used previously, for example, for the analysis of the human decorin (Santra et al., 1994) and N-methylpurine-DNA glycosylase (Tatsuka et al., 1995) genes. Mouse NRSF/REST genomic DNA fragments, containing exon IA and 5' flanking sequence, were assayed for the ability to initiate transcription using 'transcriptionally active' HeLa nuclear extracts, containing [α-^{32}P] rCTP, and the resulting run-off transcripts were analysed by autoradiography. HeLa cells express NRSF/REST (Chong et al., 1995) and therefore nuclear extracts derived from these cells should contain the necessary factors required to initiate transcription faithfully from NRSF/REST genomic templates. Mouse NRSF/REST genomic fragment 1 (corresponding to -516/+402) and genomic fragment 2 (corresponding to -227/+402) both generated identical sets of three run-off transcripts, corresponding to DNA sizes of 423, 410 and 397 nucleotides long (Figure 7). Genomic fragment 3, spanning positions –227 to +492, and thus extending 90 bp downstream compared to genomic fragments 1 and 2, produced run-off transcripts corresponding to DNA sizes of 518, 505 and 492 nucleotides long (Figure 7). The fact that specific run-off transcripts from genomic fragment 3, must be 90 nucleotides longer than run-off transcripts from genomic fragments 1 and 2, was used to determine that the length of RNA had to be corrected by 5% due to slower migration compared to DNA (Sambrook et al., 1989). Accordingly, the corrected sizes of run-off transcripts generated from genomic fragments 1 and 2 were found to correspond to 402, 390 and 377 nucleotides, whilst those generated from genomic fragment 3 correspond to 492, 480 and 467 nucleotides (Figure 7). The positions from which transcripts had been initiated were determined by subtracting the length of the transcript from the position of the 3' end of the genomic fragment. The length of these transcripts indicated the presence of transcription initiation sites, within

Figure 7 **Transcription initiation site analysis using an *in vitro* transcription assay and comparative PCR of genomic DNA and cDNA.** **(A)** Map of the 5' end of the mouse NRSF/REST gene showing exon IA and 5' flanking sequence. Thin lines mark the regions spanned by the genomic fragments 1, 2 and 3 that were used in the *in vitro* transcription assay. Arrowheads indicate the positions of transcription initiation sites identified using the *in vitro* transcription assay. Arrows indicate the position of primers used to carry out the comparative PCR analysis. **(B)** *In vitro* transcription assay performed as detailed under Methods and Materials (Chapter 2) using HeLa nuclear extract (Promega) and genomic fragments 1, 2 and 3. Samples were electrophoresed through a 4% urea-acrylamide gel, dried and exposed to X-ray film. Labelled arrowheads indicate transcript sizes (in nucleotides) and sizes have been corrected for 5% slower migration of RNA compared with DNA on this denaturing polyacrylamide gel. The number in parantheses indicates the positions of corresponding transcription initiation sites. **(C)** DNase I treated, total 3T3 RNA was reverse transcribed as described under Methods and Materials. PCRs were performed in the presence of 5% DMSO using NRSF/REST primer pairs to exon IA: 1500s/IAa and 1570s/IAa$_2$ on 3T3 cDNA (C), 3T3 cDNA without reverse transcription (N) and a subcloned mouse NRSF/REST gene restriction fragment *Pst*I 5.3 (G) (refer to Figure 1). The PCR cycling conditions used were: 1 cycle 94°C/3 min; 35 cycles 94°C/30 s, 63°C/30 s, 72°C/3 min; 1 cycle 72°C/5 min. PCR products were electrophoresed through 2% MetaPhor® agarose. M, 1-Kb ladder DNA marker.

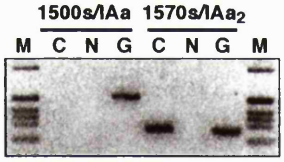
(A)


Genomic fragment 1
-516
+402
Genomic fragment 2
-227
+402
Genomic fragment 3
-227
+492
200 bp
+1 +13 +25
Exon IA
1500s
(-75)
1570s
(-12)
HindIII
(-516)
XhoI
(+402)

(B)


Genomic fragment
1 2 3
492 (+1)
480 (+13)
467 (+25)
402 (+1)
390 (+13)
377 (+25)

(C)


1500s/IAa
1570s/IAa2
M C N G C N G M

exon IA, upstream of the cluster of nine transcription initiation sites identified by RNase protection analysis. The position corresponding to the initiation site of the longest transcript was designated +1 and the two shorter transcripts accordingly as +13 and +25 (Figure 8). Interestingly, the intensity of the three run-off transcripts appeared to be similar suggesting that the three start sites are equally used, at least in *in vitro* transcription assays. The additional band generated from genomic fragment 1 may correspond to an initiation site positioned ~30 bases upstream of +1, but this band was not generated from genomic fragments 2 and 3 and is therefore likely to be non-specific.

3.3.4 The region spanning position +1 is transcribed within the endogenous NRSF/REST gene

To detect the presence of mRNA initiated at positions +1, +13 and +25 within exon IA, primer extension assays were performed using a primer derived from exon IA sequence (corresponding to positions +260/+240). An identical set of primer extension products were detected using run-off transcripts generated from HeLa cell extract and poly A^+ mRNA derived from 3T3 cells as template, however primer extension assays performed with a second primer to exon IA (corresponding to positions +247/+227) did not produce bands with the predicted incremental decrease in length of 13 bases (data not shown). These data indicate that primer extension products obtained with one or both of these primers were non-specific. Furthermore the length of primer extension products obtained with both of these primers did not correspond to any of the transcription start sites mapped by RNase protection or *in vitro* transcription assays. As an alternative method to detect endogenous mRNA transcripts initiated at position +1, comparative PCR analyses of mouse genomic DNA and 3T3 cDNA were carried out. Two different upstream primers were used: 1500s (position -74/-50) that contained sequence just upstream of the proposed transcription start site at +1, and 1570s (position

-12/+12) that contained sequence overlapping +1. These upstream primers were used in conjunction with two downstream primers: IAa and IAa_2, containing sequence derived from exon IA. The PCR products obtained from this analysis are shown in Figure 7. Primer pair 1570s/IAa_2 generated the same size PCR products from both genomic DNA and cDNA, indicating that this region is transcribed and that the primers span contiguous exonic sequence. However, primer pair 1500s/IAa generated PCR products only from genomic DNA, consistent with this sequence being upstream of the transcription initiation sites. In control reactions no PCR products were obtained using RNA without reverse transcription, demonstrating that amplified products obtained using cDNA could have not resulted from contaminating genomic DNA. The comparative PCR analysis produced data that is consistent with mRNA transcripts initiated at position +1 within the endogenous NRSF/REST gene, at least in 3T3 cells.

3.3.5 Transcription initiation sites within exon IB and exon IC

Semi-quantitative PCR analysis indicated that mRNA containing type B and type C, 5' UTRs are present at very low levels, relative to mRNA containing type A, 5' UTR in mouse NRSF/REST expressing cell lines (see Figure 2). In an attempt to map the positions from which mRNA which are initiated within exon IB and exon IC, a radio-labelled cRNA probe corresponding to positions +794 to +1533, and thus spanning both exon IB and exon IC was used in RNase protection assays. Protected fragments were observed using both 3T3 and Neuro2a RNA but protected fragments of an identical length were also observed using yeast RNA (data not shown), indicating that the protected bands observed were due to self-protection, possibly as a result of secondary structure. In fact the sequence between positions +794 and +1533 has a high GC content, thus predisposing this region to secondary structure. During the course of these studies Koenigsberger et al., (2000) reported the identification of several transcription initiation

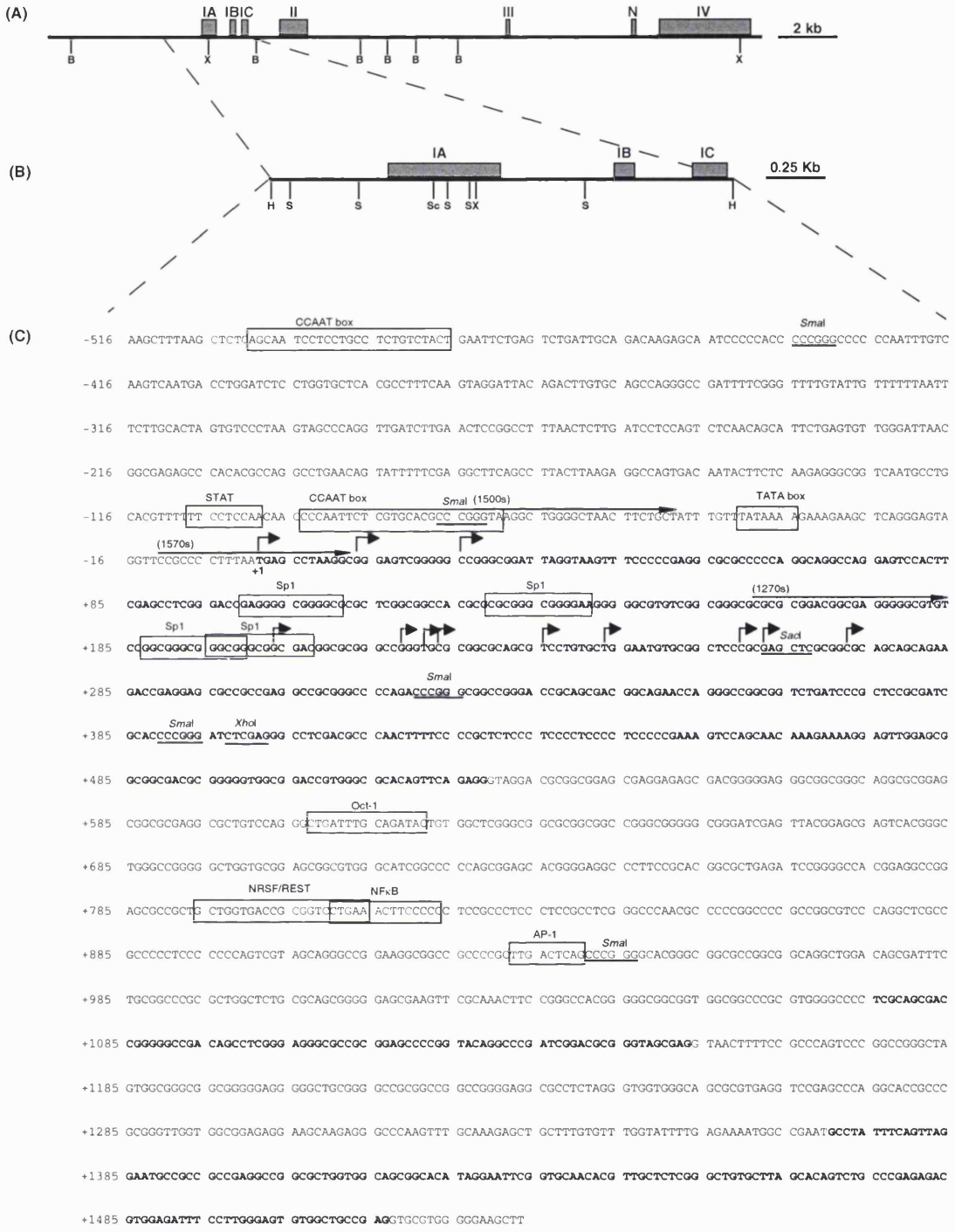
sites within exon IB and exon IC of the mouse NRSF/REST gene and mRNA initiated at these sites are predicted to generate mRNA containing type B and type C, 5' UTRs, respectively.

3.3.6 Sequence analysis of the 2.0 kb *Hin*dIII fragment containing the three 5' non-coding exons

To identify the DNA elements that direct and control transcription of mRNA initiated within exon IA, the DNA sequence extending 516 bp upstream and 1533 bp downstream of the start site at position +1 was determined (Figure 8). Inspection of the nucleotide sequence surrounding the cluster of transcription start sites around +1 within exon IA revealed a canonical TATA box at position -42 and two CCAAT boxes at positions -95 and -501 (Figure 8). Examination of the sequence surrounding the cluster of nine transcription initiation sites between +203 and +272 within exon IA revealed no homology with the TATA box or initiator element (Inr) $YYA^{+1}NWYY$ (Javahery et al., 1994) suggesting this is a TATA-less promoter. However the sequence spanning +203 to +272 is highly GC-rich, which is characteristic of many promoters that generate transcripts with multiple 5' termini. Four Sp1 sites were identified using MatInspector (Quandt et al., 1995), positioned 5 bp, 16 bp, 74 bp and 104 bp upstream of position +203, the most 5' start site within this cluster (Figure 8). Sp1 sites in close proximity to transcription start sites have been identified in many TATA-less promoters and have been shown to be essential for promoter activity. Examples include the insulin-like growth factor-binding protein-2 (Boisclair et al., 1993) and adenosine deaminase (Dusing and Wiginton, 1994) genes. Database analysis of the sequence between positions -516 and +947 using MatInspector Professional (Quandt et al., 1995) revealed 90 matches for potential transcription factor binding sites. A number of recognition sites were identified for AP-2, EGR-1 and EGR-3 (data not shown), but the sequence

extending 1.3 kb downstream of +1 is highly GC-rich (78%) which increases the likelihood of spuriously identifying these GC-rich transcription factor recognition sites. However the identification of sites for transcription factors belonging to the class of immediate early genes is interesting since these could potentially play a role in modulating NRSF/REST expression in neurones in response to extracellular stimuli. In addition, recognition sites for STAT, AP-1, Oct-1, NFκB and even NRSF/REST were amongst those identified (Figure 8).

Figure 8 **Sequence of a 2.0 kb *Hin*dIII fragment containing the NRSF/REST 5' UT exons and proximal regions.** (**A**) Restriction map of mouse NRSF/REST gene showing relative positions of exons (shown as boxes) and introns (shown as lines). (**B**) Restriction map of a 2.0 kb *Hind*III fragment that spans the 5' UT exons: IA, IB and IC and flanking sequences. B, *Bam*HI; X, *Xho*I; H, *Hin*dIII; S, *Sma*I; Sc, *Sac*I. (**C**) Nucleotide sequence of the 2.0 kb *Hin*dIII fragment. Nucleotides in bold indicate sequences for 5' UT exons: IA, IB and IC. Positions of *Xho*I, *Sma*I and *Sac*I restriction sites are underlined. Arrowheads indicate positions of transcription initiation sites within exon IA determined by RNase protection and *in vitro* transcription assays with the most 5' transcription start site assigned as +1. Primers used in comparative PCR analysis are identified by arrows lying above the sequence. Boxes indicate the positions of the TATA box and CAAT boxes, as well as binding sites for the transcription factors: Sp1, AP1, STAT, Oct-1, NFκB, and NRSF/REST which were all identified using MatInspector Professional (Quandt et al., 1995).


(A)
IA IBIC II III N IV
2 kb
B X B B B B B X
(B)
IA IB IC
0.25 Kb
H S S Sc S SX S H
(C)
CCAAT box
SmaI
-516 AAGCTTTAAG CTCTCAGCAA TCCTCCTGCC TCTGTCTACT GAATTCTGAG TCTGATTGCA GACAAGAGCA ATCCCCCACC CCCGGGCCCC CCAATTTGTC
-416 AAGTCAATGA CCTGGATCTC CTGGTGCTCA CGCCTTTCAA GTAGGATTAC AGACTTGTGC AGCCAGGGCC GATTTTCGGG TTTTGTATTG TTTTTTAATT
-316 TCTTGCACTA GTGTCCCTAA GTAGCCCAGG TTGATCTTGA ACTCCGGCCT TTAACTCTTG ATCCTCCAGT CTCAACAGCA TTCTGAGTGT TGGGATTAAC
-216 GGCGAGAGCC CACACGCCAG GCCTGAACAG TATTTTTCGA GGCTTCAGCC TTACTTAAGA GGCCAGTGAC AATACTTCTC AAGAGGGCGG TCAATGCCTG
STAT
CCAAT box
SmaI (1500s)
TATA box
-116 CACGTTTTTT CCTCCAACAA CCCCAATTCT CGTGCACGCC CGGGTAAGGC TGGGGCTAAC TTCTGCTATT TGTTTATAAA AGAAAGAAGC TCAGGGAGTA
(1570s)
-16 GGTTCCGCCC CTTTAATGAG CCTAAGGCGG GAGTCGGGGG CCGGGCGGAT TAGGTAAGTT TCCCCCGAGG CGCGCCCCCA GGCAGGCCAG GAGTCCACTT
+1
Sp1
Sp1
(1270s)
+85 CGAGCCTCGG GACCGAGGGG CGGGGCCGCC TCGGCGGCCA CGCCGCGGGG CGGGGAAGGG GGCGTGTCGG CGGGCGCGCG CGGACGGCGA GGGGGCGTGT
Sp1
Sp1
SacI
+185 CCGGCGGGCG GGCGGGCGGC GACGGCGCGG GCCGGGTGCG CGGCGCAGCG TCCTGTGCTG GAATGTGCGG CTCCCGCGAG CTCGCGGCGC AGCAGCAGAA
SmaI
+285 GACCGAGGAG CGCCGCCGAG GCCGCGGGCC CCAGACCCGG GCGGCCGGGA CCGCAGCGAC GGCAGAACCA GGGCCGGCGG TCTGATCCCG CTCCGCGATC
SmaI
XhoI
+385 GCACCCCGGG ATCTCGAGGG CCTCGACGCC CAACTTTTCC CCGCTCTCCC TCCCCTCCCC TCCCCCGAAA GTCCAGCAAC AAAGAAAAGG AGTTGGAGCG
+485 GCGGCGACGC GGGGGTGGCG GACCGTGGGC GCACAGTTCA GAGGGTAGGA CGCGGCGGAG CGAGGAGAGC GACGGGGGAG GGCGGCGGGC AGGCGCGGAG
Oct-1
+585 CGGCGCGAGG CGCTGTCCAG GCCTGATTTG CAGATACTGT GGCTCGGGCG GCGCGGCGGC CGGGCGGGGG CGGGATCGAG TTACGGAGCG AGTCACGGGC
+685 TGGGCCGGGG GCTGGTGCGG AGCGGCGTGG GCATCGGCCC CCAGCGGAGC ACGGGGAGGC CCTTCCGCAC GGCGCTGAGA TCCGGGGCCA CGGAGGCCGG
NRSF/REST
NFκB
+785 AGCGCCGCTG CTGGTGACCG CGGTGCTGAA ACTTCCCCGC TCCGCCCTCC CTCCGCCTCG GGCCCAACGC CCCCGGCCCC GCCGGCGTCC CAGGCTCGCC
AP-1
SmaI
+885 GCCCCCTCCC CCCCAGTCGT AGCAGGGCCG GAAGGCGGCC GCCCCGCTTG ACTCAGCCCG GGGCACGGGC GGCGCCGGCG GCAGGCTGGA CAGCGATTTC
+985 TGCGGCCCGC GCTGGCTCTG CGCAGCGGGG GAGCGAAGTT CGCAAACTTC CGGGCCACGG GGGCGGCGGT GGCGGCCCGC GTGGGGCCCC TCGCAGCGAC
+1085 CGGGGGCCGA CAGCCTCGGG AGGGCGCCGC GGAGCCCCGG TACAGGCCCG ATCGGACGCG GGTAGCGAGG TAACTTTTCC GCCCAGTCCC GGCCGGGCTA
+1185 GTGGCGGGCG GCGGGGGAGG GGGCTGCGGG GCCGCGGCCG GCCGGGGAGG CGCCTCTAGG GTGGTGGGCA GCGCGTGAGG TCCGAGCCCA GGCACCGCCC
+1285 GCGGGTTGGT GGCGGAGAGG AAGCAAGAGG GCCCAAGTTT GCAAAGAGCT GCTTTGTGTT TGGTATTTTG AGAAAATGGC CGAATGCCTA TTTCAGTTAG
+1385 GAATGCCGCC GCCGAGGCCG GCGCTGGTGG CAGCGGCACA TAGGAATTCG GTGCAACACG TTGCTCTCGG GCTGTGCTTA GCACAGTCTG CCCGAGAGAC
+1485 GTGGAGATTT CCTTGGGAGT GTGGCTGCCG AGGTGCGTGG GGGAAGCTT

3.3.7 Summary

5' RACE PCR analysis demonstrated that NRSF/REST mRNA containing type A, 5' UTR extends approximately 300 bp upstream of the initiation codon. Using RNase protection assays, a cluster of nine transcription initiation sites were mapped within a 69 bp segment within exon IA (between positions +203 and +272) that are used in both 3T3 and Neuro2a cells, however, as expected the abundance of NRSF/REST mRNA in Neuro2a cells was much lower. Thus the 5' RACE PCR and RNase protection analysis produced consistent results, indicating the presence of a cluster of transcription start sites within exon IA that generate type A, 5' UTR of a length ranging between 265 bp and 334 bp. However, RT-PCR analysis demonstrated that exon IA sequence upstream of position +203 is transcribed in 3T3 cells, although these mRNA transcripts are likely to be present at very low levels since they were not detected using RNase protection assays. Using an *in vitro* transcription assay, genomic fragments containing exon IA and the 5' flanking region generated run-off transcripts from a triplet of transcription initiation sites (the most upstream site was assigned +1) in a 25 bp segment that is positioned a further 178 bp upstream of the cluster of nine transcription start sites. Furthermore RT-PCR analysis of 3T3 RNA using a primer upstream of +1 generated a PCR product from genomic DNA but not cDNA, while using a primer spanning +1 generated PCR products from both cDNA and genomic DNA, demonstrating that the upstream start sites are utilised in the endogenous NRSF/REST gene, at least in 3T3 cells. Sequence analysis of the 5' proximal sequence revealed that the two clusters of transcription initiation sites that are positioned 178 bp apart in exon IA are associated with distinct promoter elements.

3.4 Identification and characterisation of *cis*-acting regulatory domains within the mouse NRSF/REST gene

To localise the genomic regions necessary and sufficient for promoter activity, fusion constructs were prepared containing different NRSF/REST genomic fragments, directing expression of the luciferase reporter gene. The reporter constructs were transiently transfected into non-neural and neuronal cell lines and primary cultures that display different levels of endogenous NRSF/REST gene expression.

3.4.1 Identification of a region of the NRSF/REST gene able to direct cell-specific reporter gene expression

Transient transfections were performed in 3T3 cells, a fibroblast cell line that expresses high levels of NRSF/REST mRNA (Lönnerberg et al., 1996); Neuro2a cells, a neuroblastoma cell line that expresses moderate levels of NRSF/REST mRNA (Lönnerberg et al., 1996) and PC12 cells, an adrenal pheochromocytoma cell line that expresses very low levels of NRSF/REST mRNA (Chong et al., 1995; Shimojo et al., 1999). In accord with these published data, RNase protection assays performed with cRNA probes directed towards the 5' termini of mouse NRSF/REST mRNA demonstrated that 3T3 cells express much higher levels of NRSF/REST mRNA than Neuro2a cells (Figure 6). To test for promoter activity an 11 kb mouse NRSF/REST genomic fragment, starting at approximately position -8500 and extending downstream to the translation initiation codon in exon II, was cloned into a modified luciferase vector, pGL3 splice (see chapter 2, section 2.3.3). This construct, pGL3 -8.5, contains exon IA, exon IB, exon IC and the splice acceptor site of exon II, thus the prediction is that reporter transcripts containing three different 5' UTRs are generated, the same as in the endogenous gene. The relative luciferase activities obtained by transient transfection of this construct into cell lines are shown in Figure 9. The highest level of reporter gene

expression was detected in 3T3 cells which produced 13±2.9 –fold (n = 4) activity, relative to promoterless pGL3 splice. A moderate level of activity, of 3.1±0.79 –fold (n = 3), was measured in Neuro2a cells, and no activity above background level was detected in PC12 cells, which produced 0.93±0.57 –fold (n = 7) activity. The luciferase activity produced by pGL3 -8.5 in 3T3, Neuro2a and PC12 cells correlates with the high, moderate and low levels, respectively, of endogenous NRSF/REST gene expression that have been reported for these cell lines (Chong et al., 1995; Lönnerberg et al., 1996; Shimojo et al., 1999). Thus pGL3 -8.5 appears to contain sufficient genomic sequence to direct cell-specific reporter gene expression in transient transfections. As shown in Figure 9, removal of 4 kb of 5' flanking sequence and a part of the first intron (between nucleotides +2000 and +2400), to generate pGL3 -4.5 Δ2.0/2.4, containing exon IA, exon IB, exon IC and the exon II splice acceptor (the latter allowing intronic sequence to be spliced out of reporter transcripts), produced similar high levels of luciferase activity in 3T3 cells (15±1.2 –fold (n = 8)), moderate levels of activity in Neuro2a cells (4.6±0.69 –fold (n = 6)) and only slight activity in PC12 cells (1.8±0.39 –fold (n = 8)). Compared with pGL3 -8.5 this construct, pGL3 -4.5 Δ2.0/2.4, produced a slight but not significant increase in luciferase activity in both Neuro2a and PC12 cells ($P > 0.1$, unpaired student's t test). Thus the 6.5 kb NRSF/REST genomic fragment between nucleotides -4500 and +2000 still contained sufficient sequence to direct cell-specific reporter gene expression in 3T3, Neuro2a and PC12 cells.

The lower reporter gene expression produced by both pGL3 -8.5 and pGL3 -4.5 Δ2.0/2.4 in Neuro2a and PC12 cells, compared to 3T3 cells, could be due to repressor elements that are selectively active in Neuro2a and PC12 cells and/or enhancer elements that are selectively active in 3T3 cells. To test these possibilities, a series deletions were made from the 5' and 3' end of the 6.5 kb genomic fragment in pGL3 -4.5 Δ2.0/2.4.

These were assayed in 3T3, Neuro2a and PC12 cell lines and the results are summarised in Figure 9.

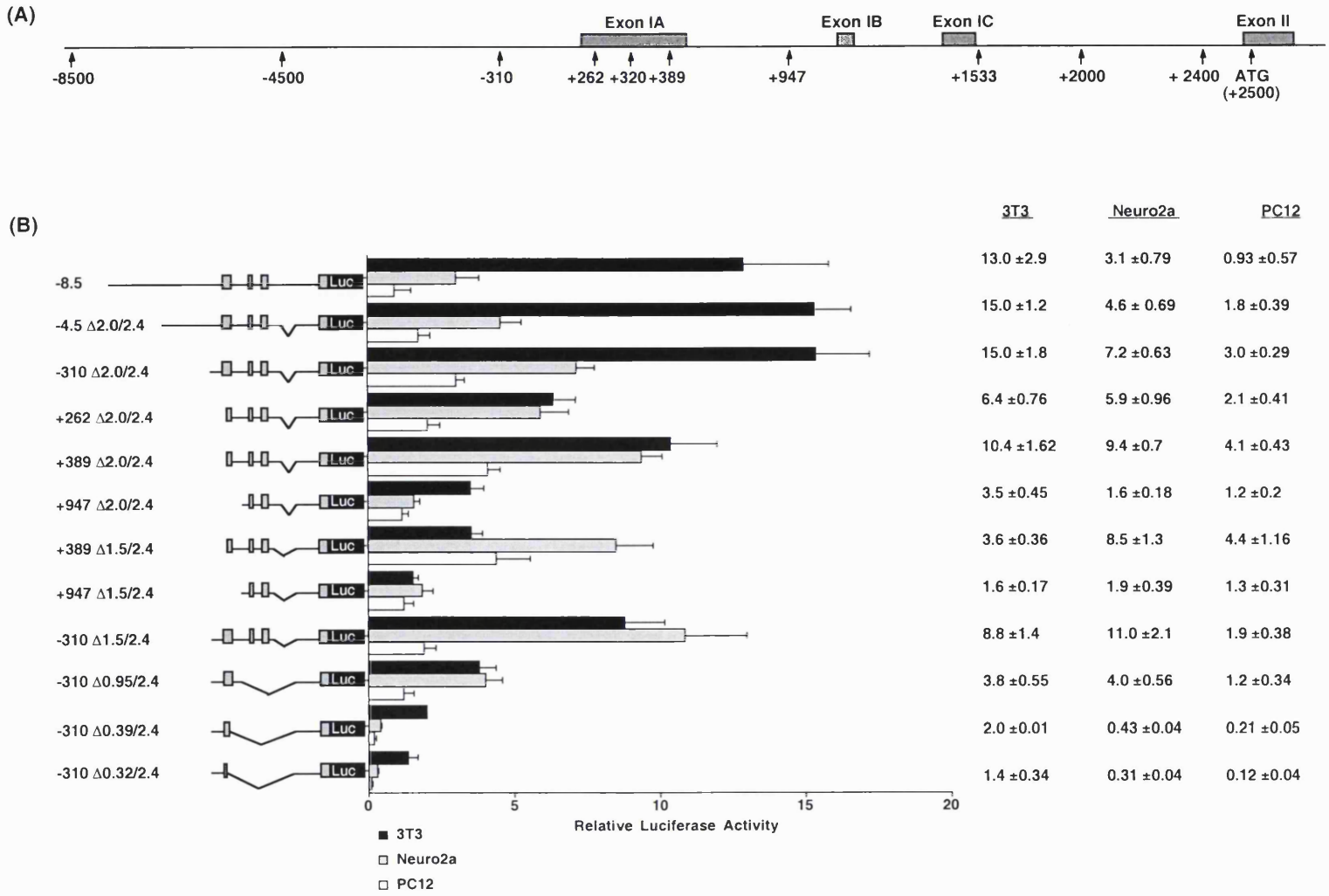
3.4.2 5' deletion analysis

Lower NRSF/REST gene expression in Neuro2a and PC12 cells is achieved by cell selective repressor elements

Deletion of 4.2 kb of 5' flanking sequence between nucleotides -4500 and -310, to generate pGL3 -310 Δ2.0/2.4, increased luciferase activity in Neuro2a and PC12 cells by 1.6 and 1.7 –fold, respectively. However no change in reporter gene activity was detected in 3T3 cells. The luciferase activity displayed by pGL3 -310 Δ2.0/2.4 in Neuro2a and PC12 cells, compared to pGL3 -4.5 Δ2.0/2.4, was a slight but significant increase ($P < 0.05$, unpaired student's t test). Thus the region between -4500 and -310 may harbour a weak repressor element that is selectively active in Neuro2a and PC12 cell lines.

NRSF/REST mRNA containing type A, 5' UTR is far more abundant than mRNA containing type B and type C, 5' UTRs in 3T3, Neuro2a and PC12 cell lines (see Figure 2). In accord with these data, deletion of the region between -310 and +262 which contains both clusters of transcription initiation sites identified within exon IA, to generate pGL3 +262 Δ2.0/2.4, resulted in a significant reduction in luciferase activity of 2.4 –fold in 3T3 cells. Strikingly, removal of nucleotides -310 to +262 had no significant effect on reporter gene expression in Neuro2a and PC12 cells. In pGL3 +262 Δ2.0/2.4 transcription is potentially initiated within exon IB and exon IC, since low levels of mRNA transcripts that are initiated within exon IB and exon IC were detected by RT-PCR in all three cell lines (see Figure 2), and several transcription start sites have been identified within exon IB and exon IC (Koenigsberger et al. 2000). Thus reporter gene

Figure 9 **Transient transfection analysis of a deletion series of mouse NRSF/REST gene reporter constructs.** **(A)** Map of the 5' end of mouse NRSF/REST gene and flanking region. Arrows indicate positions where deletions were introduced and these are numbered with respect to the 5' most transcription start site defined as +1. **(B)** For Neuro2a, 3T3 and PC12 cells, respectively 0.25 μg, 0.5 μg and 1 μg of DNA consisting of equimolar amounts of each reporter plasmid, and Bluescript® II (Stratagene) DNA to make up the total amount of DNA, were transfected using Tfx™-50 (Promega) into each well of a 24-well plate. Cells were co-transfected with pRL-CMV (Promega) which contains *Renilla* cDNA driven by the CMV promoter, at a ratio of 1:250 of pRL CMV to reporter plasmid. Luciferase assays were carried out using 5 μl (Neuro2a) and 20 μl (3T3 and PC12) of cell extract and reagents of the Dual-Luciferase™ assay system (Promega). Relative luciferase activity is defined as firefly luminescence normalised to *Renilla* luminescence expressed relative to normalised luminescence driven from the promoterless pGL3 splice. A representation of each reporter plasmid is indicated on the left of the graph and values of relative luciferase activity are given on the right. Data represent the mean ±S.E.M of at least three independent experiments each performed in triplicate.



	3T3	Neuro2a	PC12
-8.5	13.0 ±2.9	3.1 ±0.79	0.93 ±0.57
-4.5 Δ2.0/2.4	15.0 ±1.2	4.6 ± 0.69	1.8 ±0.39
-310 Δ2.0/2.4	15.0 ±1.8	7.2 ±0.63	3.0 ±0.29
+262 Δ2.0/2.4	6.4 ±0.76	5.9 ±0.96	2.1 ±0.41
+389 Δ2.0/2.4	10.4 ±1.62	9.4 ±0.7	4.1 ±0.43
+947 Δ2.0/2.4	3.5 ±0.45	1.6 ±0.18	1.2 ±0.2
+389 Δ1.5/2.4	3.6 ±0.36	8.5 ±1.3	4.4 ±1.16
+947 Δ1.5/2.4	1.6 ±0.17	1.9 ±0.39	1.3 ±0.31
-310 Δ1.5/2.4	8.8 ±1.4	11.0 ±2.1	1.9 ±0.38
-310 Δ0.95/2.4	3.8 ±0.55	4.0 ±0.56	1.2 ±0.34
-310 Δ0.39/2.4	2.0 ±0.01	0.43 ±0.04	0.21 ±0.05
-310 Δ0.32/2.4	1.4 ±0.34	0.31 ±0.04	0.12 ±0.04

expression directed by the promoters upstream of exon IB and exon IC appear to have increased. One rationalisation for this is that the segment between positions -310 to +262 contains a repressor element(s) that suppresses transcription initiated within all three 5' UT exons in Neuro2a and PC12 cells but not in 3T3 cells. Removal of the transcription start sites within -310/+262 would be mitigated by the concomitant removal of the repressor element(s), hence resulting in little change in reporter gene activity. Also the level of luciferase activity in Neuro2a and PC12 cells would be expected to be the same as the level of activity in 3T3 cells. The transfection data suggest that additional *cis*-elements that maintain lower expression in PC12 cells are present in pGL3 +262 Δ2.0/2.4. A further deletion from positions +262 to +389, to generate pGL3 +389 Δ2.0/2.4, produced a further small increase in luciferase activity ranging between 1.6 and 2.0 –fold in all three cell lines indicating that this region contains a weak repressor element(s). In fact the reporter construct pGL3 +389 Δ2.0/2.4 produced the maximal luciferase activity observed in Neuro2a and PC12 cell lines. In summary the 5' deletion data suggest that the lower levels of reporter gene expression observed in Neuro2a and PC12 cells are due to repressor elements that are present between positions -4500 and +389, however the repressor element(s) present between nucleotides +262 and +389 also suppresses activity in 3T3 cells.

Identification of a proximal enhancer region spanning the 3' portion of exon IA and part of the first intron

Further 5' deletion from positions +389 to +947, resulting in removal of the remaining 3' end of exon IA and part of the first intron, to generate pGL3 +947 Δ2.0/2.4 produced a dramatic reduction in luciferase activity ranging between 3 and 6 –fold in all three cell lines. Since, the reporter construct pGL3 +947 Δ2.0/2.4 still retains the transcription start sites within exon IB and exon IC that have been mapped by

Koenigsberger et al. (2000), the region between +389 and +947 appears to contain enhancer element(s) that promote transcription from these exons. However, to demonstrate that reduced luciferase activity was not due to deletion of initiation sites within exon IB, the region between +389 and +947 was cloned, in both orientations, upstream of a heterologous promoter in pGL3 Inr (Wood et al., 1999) containing the initiator sequence from the adenovirus major late promoter. As can be seen in Figure 11 this fragment enhances transcription from pGL3 Inr between 5.1±0.55 and 6.6±1.06 –fold when cloned in the forward orientation, and to a similar degree ranging between 6.1±0.68 and 8.0±0.89 –fold when cloned in the reverse orientation, in all three cell lines. These data showing that the region between +389 and +947 activates reporter gene expression from a heterologous promoter in an orientation-independent manner indicates that the reduction in reporter gene expression seen upon deletion of this fragment in pGL3 +389 Δ2.0/2.4 is due to the removal of an enhancer element(s) that is present between +389 and +947. The genomic segment corresponding to the sequence between positions +389 and +947 was designated the proximal enhancer region.

3.4.3 3' deletion analysis

The fact that NRSF/REST mRNA containing type A, 5' UTR is the most abundant NRSF/REST transcript in 3T3, Neuro2a and PC12 cells, motivated the search for transcriptional control elements that direct transcription of mRNA with type A, 5' UTR. Therefore a series of deletions were generated from the 3' end of the genomic insert within pGL3 -310 Δ2.0/2.4 to remove exon IC and exon IB, however the splice acceptor site between +2400 and +2500 was retained.

Identification of a distal enhancer region in the first intron which is required for the higher luciferase activity in 3T3 cells

Deletion of nucleotides +2000 to +1533, to generate pGL3 -310 Δ1.5/2.4 which retains exon IA, exon IB and exon IC, almost halved the luciferase activity produced in 3T3 cells, reduced activity by 1.6 –fold in PC12 cells and increased activity by 1.5 –fold in Neuro2a cells. Deletion of the same fragment to generate pGL3 +389 Δ1.5/2.4 and pGL3 +947 Δ1.5/2.4, both of which contain exon IB and exon IC but not exon IA, led to a 2.9 and 2.2 –fold reduction in luciferase activity in 3T3 cells, respectively. In contrast no significant change to reporter gene activity was observed in Neuro2a and PC12 cells. These data suggest that the segment between +1533 and +2000 increases reporter gene expression in 3T3 cells, and perhaps to some extent in PC12 cells. In fact Koenigsberger et al., (2000) also report the detecting enhancer activity in a fragment corresponding to +1518/+2010 (+1517/+2000 as defined here) which enhanced reporter gene expression directed by a heterologous promoter. The genomic region between nucleotides +1533 and +2000 was designated the distal enhancer.

Further deletion of the sequence between +1533 and +947, which contains the transcription start sites mapped within exon IC and most of start sites mapped within exon IB (Koenigsberger, et al. 2000), to give pGL3 -310 Δ0.95/2.4, decreased luciferase activity by between 1.6 and 2.3 –fold in all three cell lines. Decreased reporter gene activity may reflect a loss of low levels of reporter transcripts initiated from exon IB and exon IC however the possibility that this region contains an enhancer is not ruled out.

The proximal enhancer region is essential for exon IA promoter activity

In the 3' deletion series, the most dramatic change in luciferase activity was produced by a deletion of 558 bp, corresponding to the proximal enhancer region

between positions +947 and +389, to generate pGL3 -310 Δ0.39/2.4. This reporter construct still retained the 5' portion of exon IA, the two clusters of transcription start sites within exon IA, and 310 bp of 5' flanking sequence, however, reporter gene expression was decreased by 9.4 –fold and 5.9 –fold in Neuro2a and PC12 cells respectively, and 1.9 –fold in 3T3 cells. Luciferase activity produced by pGL3 -310 Δ0.39/2.4 in Neuro2a and PC12 cells was lower than background levels. However approximately 2.0±0.01 –fold activity was detected in 3T3 cells. These data are consistent with the 5' deletion analysis which suggested that the region between -310 and +262 contains a repressor element that is selectively active in Neuro2a and PC12 cells. Moreover the promoter activity associated with exon IA is dependent on the proximal enhancer region. A further deletion to generate pGL3 -310 Δ0.32/2.4 resulted in a further modest decrease in luciferase activity ranging from between 1.4 and 1.8 –fold in all three cell lines.

3.4.4 The proximal enhancer is essential for exon IA promoter activity in rat primary cultures of neuroepithelial cells

In situ hybridisation analysis of mouse and chick embryos has shown that low levels of NRSF/REST mRNA are present in the developing nervous system but expression is restricted mainly to the ventricular zone (Chen et al., 1998; Chong et al., 1995; Schoenherr and Anderson, 1995a). Loss-of-function and gain-of-function studies demonstrate that the expression of NRSF/REST in the ventricular zone is crucial for the proper spatial and temporal expression of NRSE/RE1-bearing genes (Chen et al., 1998; Paquette et al., 2000). Moreover the precise control of NRSF/REST gene expression in the ventricular zone is essential for the proper development of neurones (Paquette et al., 2000). Primary cultures of neuroepithelial (NE) cells derived from the cortical ventricular zone of rat embryos can differentiate into neurones, astrocytes and

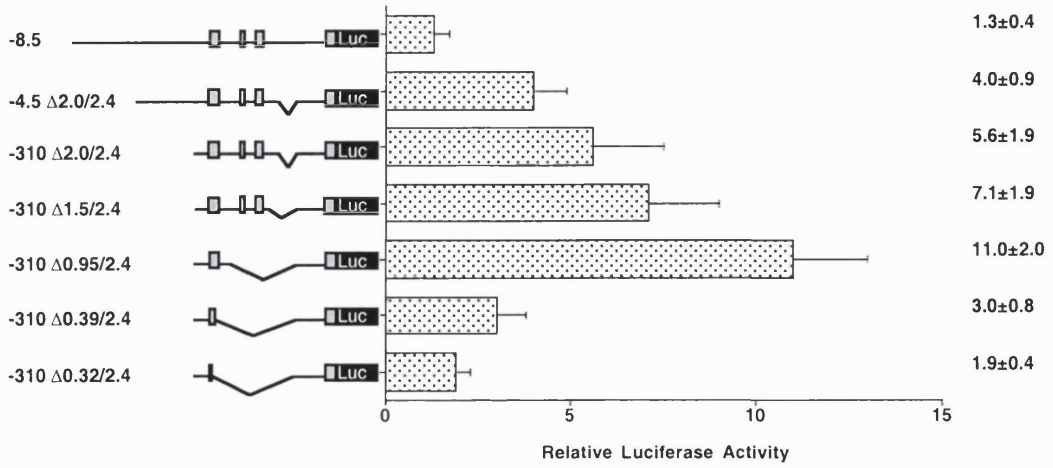


Figure 10 **Transient transfection analysis of the 3' deletion series of NRSF/REST gene reporter constructs in neuroepithelial cells**. For neuroepithelial cells prepared from rat E14 cortex 0.2 μg DNA containing equimolar amounts of each reporter plasmid, and Bluescript® II (Stratagene) DNA to make up the total amount of DNA, was transfected using Effectene (Qiagen) in each well of a 24 well plate. Neuroepithelial cells were co-transfected with pRL-CMV (Promega) which contains *Renilla* cDNA driven by the CMV promoter, at a ratio of 1:40 of pRL CMV to reporter plasmid. Luciferase assays were carried out using 20 μl cell extract and Dual-Luciferase™ assay system (Promega). Relative luciferase activity is defined as firefly luminescence normalised to *Renilla* luminescence expressed relative to normalised luminescence driven from the promoterless pGL3 splice. A representation of each reporter plasmid is indicated on the left of the graph and values of relative luciferase activity are given on the right. Data represent the mean ±S.E.M of at least three independent experiments each performed in triplicate.

oligodendrocytes (Williams and Price, 1995). RT-PCR analysis of primary NE cell cultures showed that NRSF/REST mRNA is present (N. J. Buckley, unpublished observation). Therefore, the NRSF/REST reporter constructs pGL3 -8.5 and pGL3 –4.5Δ2.0/2.4, that are sufficient to support cell specific reporter gene expression in fibroblast, neuroblastoma and pheochromocytoma cell lines, were tested for activity in primary cultures of NE cells. In all tissues examined, including embryonic and adult rat brain, NRSF/REST mRNA containing type A, 5' UTR is the most abundant NRSF/REST transcript (Palm et al., 1998). Accordingly the 3' deletion series of reporter constructs were tested to identify the transcriptional control elements required for exon IA promoter activity in NE cells. The data obtained from transient transfection assays of reporter constructs performed in NE cells are summarised in Figure 10. The luciferase activity produced by pGL3 -8.5 in NE cells was only slightly above background levels. In contrast reporter gene expression produced by pGL3 -4.5 Δ2.0/2.4 was 4.0±0.9 -fold above background. Removal of -4500 to -310 to generate pGL3 -310 Δ2.0-2.4, which in Neuro2a and PC12 cells produced a slight increase in reporter gene activity, produced no significant change in activity in NE cells. Deletion of the segments from positions +2000 to +1533, to generate pGL3 -310 Δ1.5/2.4, and from positions +1533 to +947, to generate pGL3 -310 Δ0.95/2.4, each produced an incremental increase in luciferase activity of 1.3 –fold. Although NRSF/REST mRNA containing type B and type C, 5' UTRs are present at very low levels in embryonic rat brain (Palm et al., 1998) and in rat primary neural progenitor cell cultures (Koenigsberger et al., 2000), the transfection data in this study suggest that the promoter regions upstream of exon IB and exon IC have very little activity in NE cells, and the most prominent promoter activity in these cells is upstream of exon IA. In accord with these data deletion of the proximal enhancer region, to generate pGL3 -310 Δ0.39/2.4, which is essential for transcription initiated within exon IA in Neuro2a and PC12 cells, and to a lesser extent in 3T3 cells, led to a marked

3.7 –fold reduction in activity. Finally, removal of nucleotides between +389 and +320, to generate pGL3 -310 Δ0.32/2.4 produced a further small decrease in luciferase activity.

3.4.5 The 5' and 3' portion of the proximal enhancer activates reporter gene expression driven from a heterologous promoter

Transient transfection analysis of the mouse NRSF/REST gene reporter constructs defined the proximal enhancer region, spanning the 3' portion of exon IA and part of the first intron between +389 and +947, as important for promoter activity in Neuro2a, PC12 and NE neuronal cells and to a lesser extent in 3T3 fibroblast cells. In order to define the boundaries of the proximal enhancer region a series of constructs were made by insertion of *Sau*3AI fragments spanning this region upstream of the heterologous promoter in pGL3 Inr, resulting in pGL3 Inr +394/+659, pGL3 Inr +659/+763 and pGL3 Inr +947/+763 and these were tested in 3T3, Neuro2a and PC12 cells. Since the proximal enhancer region was equally effective in both the reverse and forward orientations, the fact that clones containing the fragment between positions +763 and +947 were obtained in the reverse orientation only should not effect the ability of this region to enhance expression. As shown in Figure 11, both the 5' fragment, +394/+659, and the 3' fragment, +947/+763, each retained approximately 50% of the enhancer activity of the entire proximal enhancer region. The centre fragment +659/+763 did not appear to have any enhancer activity. Thus, the 5' and 3' portions of the proximal enhancer were able to activate reporter gene expression from a heterologous promoter, although neither fragment alone was as effective as the whole fragment.

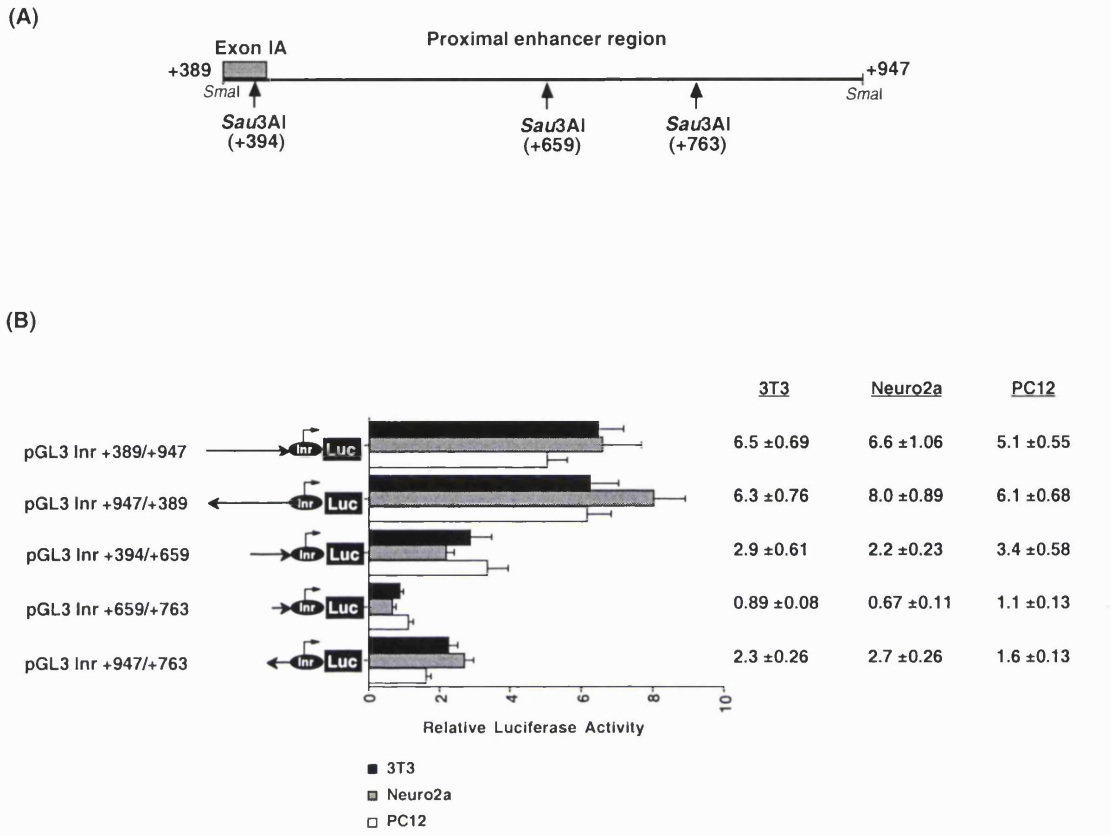


Figure 11 **Analysis of the NRSF/REST proximal enhancer region (A)** Map of the mouse NRSF/REST +389/+947 proximal enhancer region. Labelled arrows indicate positions of *Sau*3AI sites **(B)** For Neuro2a, 3T3 and PC12 cells, respectively 0.25 µg, 0.5 µg and 1 µg of reporter plasmid DNA were transfected using Tfx™-50 (Promega) into each well of a 24-well plate. Cells were co-transfected with pRL-CMV (Promega) which contains *Renilla* cDNA driven by the CMV promoter, at a ratio of 1:250 of pRL CMV to reporter plasmid. Luciferase assays were carried out using 5 µl (Neuro2a) and 20 µl (3T3 and PC12) of cell extract and reagents of the Dual-Luciferase™ assay system (Promega). Indicated on the left of the graph are schematic representations of pGL3 Inr containing fragments of the proximal enhancer cloned upstream of the adenovirus major late promoter initiator sequence. Values for relative luciferase activity are given on the right. Relative luciferase activity is defined as firefly luminescence normalised to *Renilla* luminescence expressed relative to normalised luminescence driven from the empty pGL3 Inr. Data represent the mean ±S.E.M of at least three independent experiments each performed in triplicate.

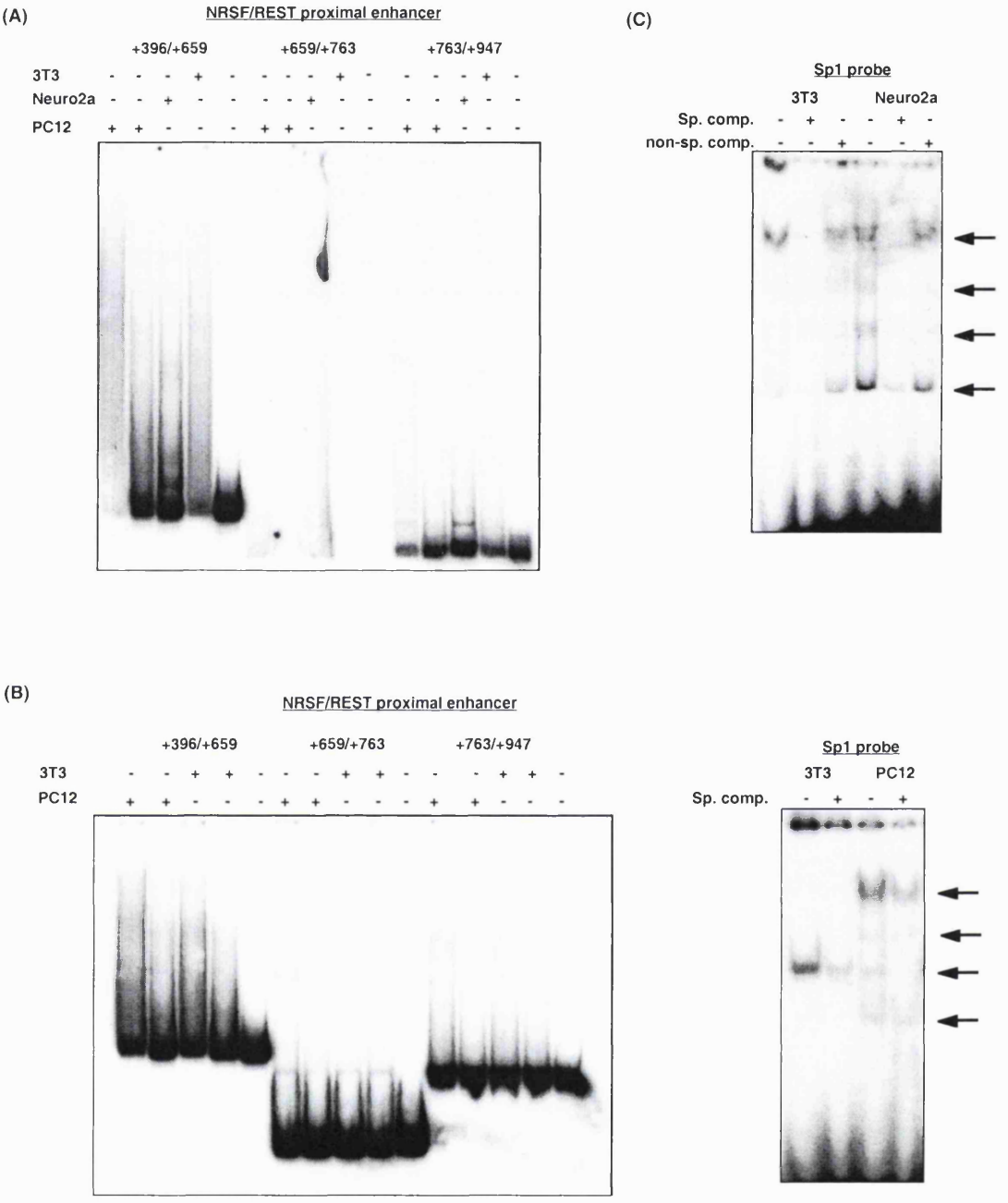
3.4.6 Electromobility shift assays performed with DNA fragments spanning the proximal enhancer region

To identify *trans*-acting factors that bind the proximal enhancer region, electromobility shift assays were performed using end-labelled *Sau*3AI fragments as probes corresponding to nucleotides +394/+659, +659/+763 and +763/+947. However, despite using different cell extract preparations from 3T3, Neuro2a and PC12 cell lines, no DNA-protein complexes were detected using any of the three probes under the conditions employed in this assay (Figure 12). The quality of the protein extracts was tested by performing EMSAs using a radio-labelled Sp1 oligodeoxynucleotide probe (Promega), and multiple DNA-protein complexes were detected (Figure 12). The detection of two or more specific complexes in EMSAs performed with a Sp1 oligodeoxynucleotide probe has also been reported by others (Myers et al., 1998; Wood et al., 1996). The specificity of the DNA-protein complexes was tested by competing the shift by inclusion of a 70 –fold excess of unlabeled Sp1 oligodeoxynucleotide (Figure 12).

3.4.7 Identification of an NRSE/RE1-like sequence within the proximal enhancer region

Sequence analysis using MatInspector Professional (Quandt et al., 1995) had revealed an NRSE/RE1-like sequence between positions +794 and +814 in the proximal enhancer region (Figure 8), raising the possibility that NRSF/REST may regulate its own expression. However EMSAs performed with a segment of the mouse NRSF/REST gene containing the NRSE/RE1-like sequence as a probe (+763/+947) did not produce any DNA-protein complexes (Figure 12) with 3T3 cell extracts which have previously been demonstrated to contain NRSF/REST binding activity (Wood et al., 1996). In addition the transfection data had not indicated that the proximal enhancer contains a repressor

Figure 12 **Electromobility shift assays performed with DNA probes spanning the NRSF/REST proximal enhancer region between positions +389 and +947 (A)** Electromobility shift assays were performed using ^{32}P-end-labelled DNA fragments corresponding to regions +394/+659, +659/+763 and +763/+947, of the mouse NRSF/REST gene, and whole cell extracts prepared from 3T3, Neuro2a and PC12 cell lines. **(B)** Repetition of electromobility shift assays using different 3T3 and PC12 whole cell extracts **(C)** Electromobility shift assays performed using a ^{32}P-labelled double-stranded consensus Sp1 probe (Promega) and 3T3, Neuro2a and PC12 cell extracts pre-incubated with a 70-fold molar excess of Sp1 oligodeoxynucleotide (sp. comp.), or 100-fold molar excess of oligodeoxynucleotide containing unrelated sequence (non-sp. comp.).


(A)
NRSF/REST proximal enhancer
+396/+659
+659/+763
+763/+947
3T3
Neuro2a
PC12
(C)
Sp1 probe
3T3
Neuro2a
Sp. comp.
non-sp. comp.
(B)
NRSF/REST proximal enhancer
+396/+659
+659/+763
+763/+947
3T3
PC12
Sp1 probe
3T3
PC12
Sp. comp.

element since deletion of this region had led to a marked reduction in reporter gene expression. However the presence of a repressor element may be masked by positive-acting elements within the same region. Table 2 shows the NRSE/RE1-like sequence within the NRSF/REST gene aligned to a published NRSE/RE1 consensus sequence (Schoenherr et al., 1996) and NRSE/RE1 sequences that have been functionally defined in seven other genes. The consensus NRSE/RE1 sequence has been derived by comparison of 19 different NRSE/RE1 sequences that can bind NRSF/REST (Schoenherr et al., 1996). The 19 sequences were chosen on the basis of their high degree of similarity and a few were selected for their divergence to a composite NRSE/RE1 sequence to define the limits of NRSF/REST binding. The consensus NRSE/RE1 was derived by assigning the most commonly found nucleotide to each position. Thirteen of these positions were either invariant or differed from the consensus only once among 19 sequences, while eight positions were variable. In some instances NRSF/REST was shown to tolerate a large number of differences from the consensus at these variable positions, for example the GABA-A receptor δ subunit and human cytochrome P450-11β have four and five mismatches, respectively, yet they bind NRSF/REST (Schoenherr et al., 1996). However, all sequences with six or more deviations from the consensus were found not to bind NRSF/REST. Shown in Table 2, the NRSE/RE1-like sequence in the mouse NRSF/REST gene diverges from the consensus sequence at six nucleotide positions which suggested that this sequence may not bind NRSF/REST. However, this interpretation was tempered by the observation that nucleotides seen to be critical in the context of one NRSE/RE1 sequence may be redundant in the context of another NRSE/RE1 sequence (Schoenherr et al., 1996). For example, mutation of the two adjacent G residues at position 12 and 13 to T residues has been shown to drastically reduce binding and silencing by NRSF/REST (Mori et al., 1992). However, the rat glycine receptor has an A residue at position 13 (as well as two

other deviations) and can bind NRSF/REST much better than the double mutant (Schoenherr et al., 1996). Since the NRSE/RE1 sequence in the mouse NRSF/REST gene may bind with a lower affinity than previously described NRSE/RE1 sequences, competition EMSAs were performed which would allow the relative concentration of NRSE/RE1-like sequence to be greatly increased in the binding reaction. Cell extracts were incubated with a radio-labelled probe generated from the promoter region of the rat Na type II gene containing an NRSE/RE1 that has previously been shown to bind NRSF/REST (Chong et al., 1995). Prior to addition of the probe, cell extracts prepared from PC12 cells, stably transfected with an NRSF/REST expression plasmid (described in Wood et al., 2001), were incubated with double-stranded oligodeoxynucleotide competitors corresponding to the NRSE/RE1 sequence from the mouse NRSF/REST gene or the rat Na type II gene. A 100-fold molar excess of Na type II NRSE/RE1 was sufficient to partially compete the observed shift although a 500-fold molar excess was required to fully compete (Figure 13). However even a 10,000-fold molar excess of the NRSE/RE1-like sequence from the NRSF/REST gene was unable to compete for NRSF/REST binding to the Na type II promoter (Figure 13). Thus these results suggest that the NRSE/RE1-like sequence between positions +794 and +814 within the mouse NRSF/REST gene does not bind NRSF/REST.

Consensus	TTCAGCACCACGGACAGCGCC
Mouse NRSF/REST	TTCAG**G**ACC**G**CGG**T**CA**C**C**AG**C
Rat SCG10	TTCAGCACCACGGA**G**AGTGCC
Rat Na type II channel	TTCAG**A**ACCACGGACAGCACC
Rat M_4 muscarinic receptor	TTCAGCACC**T**CGGACAGC**T**CC
Rat GluR2 receptor	TT**T**AGCACCG**A**GGACAGCGC**T**
Mouse nAChR, ß2 subunit	TTCAGCACCACGGACAGCGC**T**
Rat BDNF	TTCAGCACCT**T**GGACAGAGCC
Rat GABA-A receptor, δ subunit	TTCAGCACC**GA**GGAC**G**GCG**GA**

Table 2 **Alignment of mouse NRSF/REST gene NRSE/RE1-like sequence with functional NRSE/RE1 sequences** Nucleotides in bold indicate mismatches with the published consensus NRSE/RE1 sequence (Schoenherr et al., 1996). Bold, underlined nucleotides indicate mismatches that are invariant in other functional NRSE/RE1 sequences. Rat SCG10 (Mori et al., 1992); rat Na type II channel (Kraner, et al., 1992); rat M_4 muscarinic receptor (Wood, 1995, et al.; Mieda, et al., 1996); rat GluR2 receptor (Myers et al., 1998); mouse nAChR ß2 subunit (Bessis et al., 1997); rat BDNF (Timmusk et al., 1993) and rat GABA-A receptor, δ subunit (Schoenherr et al., 1996) NRSE/RE1 sequences shown.

How big is Bg171 / Atlas 7H

fragments void.

How big are compatibles

Why not label NAME NOSE side

for Emma?

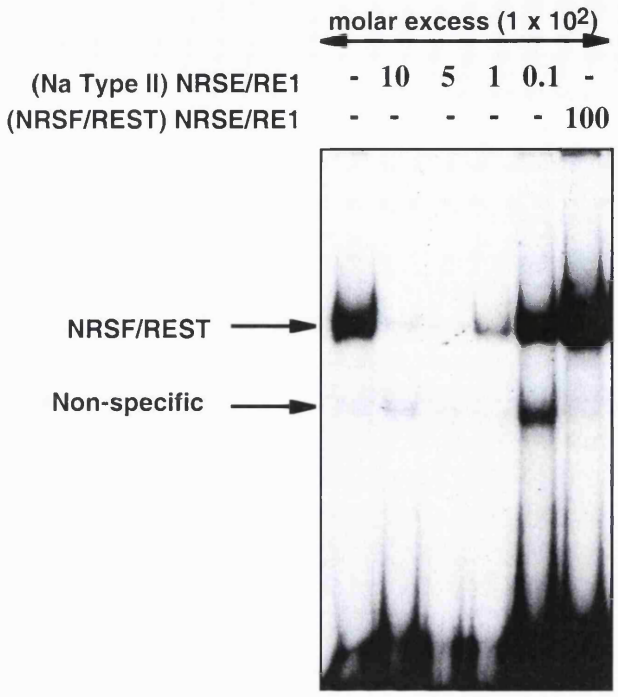


Figure 13 **Competition electromobility shift assays performed with the sequence corresponding to the NRSE/RE1-like sequence within the mouse NRSF-REST gene.** Electromobility shift assays were performed using a radio-labelled NRSE/RE1 containing *Bgl*II/*Hin*dIII fragment of the promoter of the type II sodium channel gene. PC12 C1 (Wood et al., 2001) whole cell extracts, expressing myc-tagged human NRSF/REST (pMT HZ4, gift from D.A. Anderson, Caltech), were pre-incubated with different amounts of double stranded oligodeoxynucleotide corresponding to the NRSE/RE1 sequence from either the type II sodium channel or mouse NRSF/REST genes. Labelled arrows show position of complexes.

3.4.8 Summary

A NRSF/REST genomic fragment, containing 11 kb of sequence upstream of the initiation codon, and a 6.5 kb genomic fragment internal to this larger fragment were both sufficient to support reporter gene expression in 3T3, Neuro2a, PC12 and NE cells. Furthermore the relative luciferase activities produced by pGL3 -8.5 and pGL3 -4.5 Δ2.0/2.4 in 3T3, Neuro2a and PC12 cell lines were consistent with the relative levels of endogenous NRSF/REST gene expression that have been reported in these cells. A series of 5' and 3' deletions within pGL3 -4.5 Δ2.0/2.4 were generated and their activity tested in 3T3, Neuro2a, PC12 and NE cells leading to the identification of several negative *cis*-acting regions that are required to suppress activity in Neuro2a and PC12 cells and a positive *cis*-acting region that is required to enhance expression in 3T3 cells. Moreover the proximal enhancer region, spanning the 3' portion of exon IA and part of the downstream intron, was found to be crucial for the promoter activity directing transcription initiated within exon IA, particularly in Neuro2a and PC12 neuronal cell lines and NE cell primary cultures. When placed upstream of a heterologous promoter, the proximal enhancer region activated transcription in an orientation and position-independent manner. Delineation of the proximal enhancer region, by generating heterologous promoter constructs containing smaller fragments spanning this region, showed that the 5' and 3' portions were each partially responsible for enhancer activity. However electromobility shift assays performed by using the 5' and 3' portions of the proximal enhancer region as probes failed to produce any detectable DNA-protein complexes. Finally competition EMSAs, performed with an oligonucleotide corresponding to an NRSE/RE1–like sequence present in the proximal enhancer region demonstrated that NRSF/REST is unlikely to interact with this sequence.

Chapter 4

4.0 DISCUSSION

4.1 Introduction

During neurogenesis, progenitor cells must select between neuronal versus non-neuronal cell fates; and in the former case, must select among a large repertoire of neuronal subtype fates. These developmental decisions require the action of transcriptional regulatory proteins, as has been indicated by genetic analyses first in invertebrates (reviewed in Jan and Jan, 1994) and more recently in vertebrates (Joyner and Guillemot, 1994). Neuronal cell fate specification usually involves a cascade of regulatory events in which transcriptional regulators are controlled by other transcriptional regulators acting in a hierarchical manner. Furthermore, these hierarchies of transcription factors are themselves controlled by signalling pathways that respond to cell-cell interactions and environmental signals. Thus transcriptional regulation is important in the genesis, as well as function, of neurones. A great deal of work has been devoted to defining the *cis*-acting DNA elements and *trans*-acting protein factors involved in the transcriptional regulation of neurone-specific genes (reviewed in Mandel and McKinnon, 1993) and most of these studies have identified positive-acting factors (reviewed in He and Rosenfeld, 1991). However recent studies have revealed that negative-acting transcription factors also play a significant role in directing neurone-specific gene expression (reviewed by Schoenherr and Anderson, 1995b). NRSF/REST represents one of the first transcriptional repressors to be identified that directs neural-specific expression of multiple, terminal-differentiation genes. NRSF/REST is a multi-zinc finger DNA-binding protein that is expressed at high levels in non-neural tissue during development and adulthood, at lower levels in undifferentiated neural progenitor cells and at very low levels in mature neurones. This transcription factor has been implicated in silencing, and more recently in both repressing and activating, the expression of multiple neurone-specific, terminal-differentiation genes (Schoenherr et

al., 1996) in non-neural and neuronal cells, respectively. Loss-of-function studies indicate that NRSF/REST expression is essential for normal growth and morphogenesis in early development (Chen et al., 1998), and gain-of-function studies indicate that downregulation of NRSF/REST expression in the developing nervous system during neuronal differentiation is essential for specific aspects of neuronal development (Paquette et al., 2000). In order to address how NRSF/REST expression is controlled at the transcriptional level, a study was initiated to investigate the structure of the NRSF/REST gene and its promoter to try and identify the *cis*-acting DNA elements and *trans*-acting proteins that govern cell-specific expression of the NRSF/REST gene.

4.2 Discussion of results

4.2.1 Alternative splicing at the 5' end of mouse NRSF/REST primary transcripts results in mRNA containing three alternative 5' UTRs

Human NRSF/REST cDNA, has been independently cloned in three laboratories and these all contain 5' UTR of different lengths, but the same sequence (Chong et al., 1995; Schoenherr and Anderson, 1995a; Scholl et al., 1996). Rat NRSF/REST cDNA clones have been isolated containing three different 5' UTRs (designated type A, type B and type C) (Palm et al., 1998), although expression studies have shown that rat NRSF/REST mRNA containing type A, 5' UTR (which is homologous to human NRSF/REST 5' UTR) is the most abundant (~80% of all NRSF/REST transcripts) in all tissues analysed, including brain (Palm et al., 1998). So far no reports of tissue-specific or developmental stage-specific expression of any of the 5' UTRs has been detected. One report claims that primary cultures of cortical neurones, prepared from E17 rat embryos, express mRNA containing type B, 5' UTR more abundantly than mRNA containing type A and type C, 5' UTRs (Koenigsberger et al., 2000), however the RT-PCR data in Koenigsberger et al., (2000) does not support their claim. The mouse

NRSF/REST gene contains three alternative 5' UT exons, designated exon IA, exon IB and exon IC. RT-PCR analysis of mouse 3T3 fibroblast and Neuro2a neuroblastoma cell lines led to the detection of splice variants in which each of the 5' UT exons is spliced onto the first coding exon, exon II, resulting in mRNA containing type A, type B and type C, 5' UTRs. Furthermore expression of the latter two splice variants was estimated to be approximately 1000-fold less abundant in all three cell lines. The fact that these alternative 5' UT exons are transcribed indicates that exon IA, exon IB and exon IC each has a promoter. However the exclusive or predominant expression of mRNA containing type A, 5' UTR in both murine and human NRSF/REST expressing cells suggests that the promoter directing transcription initiated within exon IA has the strongest activity.

4.2.2 NRSF/REST mRNA containing type A, 5' UTR are initiated from multiple start sites within exon IA that are arranged in two clusters

To define the promoter region upstream of exon IA, the positions of the transcription initiation sites were determined. Southern hybridisation analysis of 5' RACE PCR products showed type A, 5' UTR is predominantly about 300 bp long. In accord with these data, mapping the 5' termini of mRNA by RNase protection assays, using three probes spanning the 5' end of exon IA and the upstream sequence, indicated the presence of a cluster of nine transcription initiation sites within a 69 bp segment and the use of which are predicted to generate mRNA containing 5' UTR ranging between 265 bp and 334 bp long. The 5' proximal sequence to the transcription start sites between +203 and +272 is GC-rich and does not contain any TATA box or initiator sequences, however four Sp1 elements are present that may be required to direct transcription from these positions. However the fact that RT-PCR indicated that the sequence upstream of the most 5' transcription start site at position +203 is transcribed prompted the use of an alternative method to define the start sites. Using *in vitro*

transcription a second cluster of three transcription initiation sites was mapped within a 25 bp segment positioned 178 bp upstream and the most 5' start site within this cluster was assigned +1. Although mRNA transcripts that are initiated from the upstream cluster of start sites were not detected by RNase protection assays, the sequence was detected by RT-PCR suggesting the mRNA transcripts are present albeit at very low levels. Hence the utilisation of these upstream start sites appears quite low, at least in Neuro2a and 3T3 cells. The 5' proximal sequence contains a canonical TATA box positioned 42 bp upstream of +1, and two CAAT boxes positioned 95 bp and 501 bp upstream, and these may be involved in initiating transcription from these start sites. However in higher eukaryotes the TATA box normally lies 25–30 bp upstream of the transcription start site and usually is associated with a single initiation site. Thus the role of the TATA box in initiating correct transcription within the NRSF/REST gene can only be unequivocally determined by mutational analysis.

During the course of this study, a report was published identifying multiple transcription initiation sites within three alternative 5' UT exons in the mouse NRSF/REST gene (Koenigsberger et al., 2000). However very few of the transcription start sites were identified using both primer extension and *in vitro* transcription assays. Within exon A (corresponding to exon IA in this study) a cluster of seven sites were mapped between positions +142 and +285 (positions +128 and +271 as defined here) and a second cluster of sites were mapped between positions -65 and +56 (positions -78 and +42) (Koenigsberger et al., 2000). Comparison of the start sites identified by Koenigsberger, et al., (2000) and those identified in this study revealed that while some of the sites correspond to approximately the same positions, only two of the start sites at positions +39 and +232 (+25 and +218 as defined in this study), correspond to exactly the same nucleotide. Furthermore of the transcription start sites mapped by

Koenigsberger et al., (2000), thirteen are positioned upstream of the nucleotide defined as +203 in this study, and RNase protection assays indicated that the abundance of NRSF/REST mRNA transcripts initiated upstream of +203 is below the levels of detection. In addition, seven of the transcription initiation sites Koenigsberger et al., (2000) identified are positioned between -65 and +11 (corresponding to positions -78 to -4 as defined here) and hence lie upstream of the nucleotide defined as +1 in this study. Since RT-PCR performed with a primer corresponding to positions -75 to -50 did not generate any amplified products the abundance of mRNA initiated from the majority of transcription start sites mapped by Koenigsberger et al., (2000) must be present at extremely low levels.

In this study, exon IA was found to contain two clusters of transcription start sites separated by a region of 178 bp. Transcripts initiated from the upstream cluster of start sites within exon IA are predicted to result in NRSF/REST mRNA containing a GC-rich, 5' UT leader sequence ranging between 537 bp and 512 bp long. In contrast, transcripts initiated from the downstream cluster of start sites, distributed between positions +203 and +272, are predicted to result in mRNA leader sequence ranging between 265 bp and 334 bp long. NRSF/REST is believed to contain one translation initiation codon (Chong et al., 1995; Schoenherr and Anderson, 1995a) that resides in exon II, however, sequence analysis has revealed that a short ORF of 40 amino acids is present between positions +247 and +369 within exon IA (data not shown). Also an out-of-frame AUG codon is present in exon IC (data not shown). Thus mRNA transcripts initiated from either the upstream or downstream cluster of transcription start sites within exon IA, except for those mRNA initiated from positions +259, +262 and +272 in the downstream cluster, will contain the 40 bp ORF. The presence of a short open reading frame or one or more out-of-frame initiation codons upstream of the main open

reading frame can inhibit cap-dependent translation (Kozak, 1991). In addition, GC-rich 5' leader sequences are not as efficiently translated (Kozak, 1991). For example the rat AMPA receptor subunit GluR2 gene contains multiple transcription initiation sites some of which result in GluR2 mRNAs with long GC-rich 5' UTRs and several out of frame AUG codons which impair translation (Myers et al., 1997). Therefore the choice of start sites within exon IA from which transcripts are initiated potentially affects the translational efficiency of NRSF/REST mRNA. The downstream cluster of transcription start sites within exon IA produce mRNA that appears abundant in both neuroblastoma and fibroblast cell lines. However, whether there is stage- or tissue-specific use of the clusters transcription start sites within exon IA is currently unknown.

4.2.3 The promoter region upstream of exon IA displays the strongest activity and is dependent on the proximal enhancer region

A reporter construct containing 310 bp of 5' flanking sequence and the 5' portion of exon IA including the two clusters of transcription start sites, and 116 bp of exonic sequence downstream of the most distal transcription start site (pGL3 -310 Δ0.39/2.4) produced 2 –fold activity above background levels in a fibroblast cell line although the activity was lower than background in both a neuroblastoma and pheochromocytoma cell line. However, addition of the proximal enhancer region increased promoter activity by 9.4, 5.7 and 3.6 –fold in Neuro2a, PC12 and NE cells, respectively, and by 1.9 –fold in 3T3 cells. Thus activated transcription within exon IA is dependent on the presence of the proximal enhancer region, corresponding to the segment from positions +389 to +947.

Reporter constructs containing the exon IA promoter region (pGL3 -310 Δ0.95/2.4) generated approximately 2 –fold higher reporter gene activity in 3T3 and

Neuro2a cells, than reporter constructs containing both exon IB and exon IC and their respective 5' flanking sequences (pGL3 +947 Δ1.5/2.4). Koenigsburger et al. (2000) report similar results, showing that the 5' proximal sequence of exon IA supports the strongest reporter gene expression followed by the 5' proximal sequence of exon IB and then exon IC. However in this study the abundance of mRNA containing type A, 5' UTR was found to be 1000 –fold higher than the abundance of mRNA containing type B and type C, 5' UTRs. Thus, the relative promoter activities upstream of exon IA, exon IB and exon IC in transient transfection assays do not reflect these larger differences in the abundance of NRSF/REST mRNA containing different 5' UTRs. These data suggest that *in vivo* differences in mRNA stability may result in a higher abundance of mRNA containing type A, 5' UTR. Alternatively the reporter constructs may not contain all the necessary *cis*-regulatory elements to produce levels of endogenous gene transcription from exon IA, exon IB and exon IC.

The proximal enhancer region, which is able to activate transcription in an orientation-and position-independent manner when placed upstream of a heterologous promoter, is essential for transription initiated within exon IA. Enhancers that are positioned downstream of a transcription start site and increase transcription have been widely reported in higher eukaryotic genes such as the M_1 muscarinic receptor gene (Wood et al., 1999), utropin gene (Galvagni and Oliviero, 2000), β-enolase gene (Feo et al., 1995) and sonic hedgehog gene (Müller et al., 1999). The proximal enhancer region had a more marked effect on reporter gene expression in Neuro2a, PC12 and primary cultured neuroepithelial cells than in 3T3 cells. The apparent cell selectivity of the proximal enhancer region was lost when placed upstream of a heterologous promoter, which may reflect differences in the transcriptional platforms recruited by the different promoters. Koenigsburger et al. (2000) reported that a region, corresponding to

+521/+948 (+508/+947 as defined here), in the mouse NRSF/REST gene acts as an enhancer for transcription initiated within exon IB, but no data was presented on its action on transcription initiated within exon IA. Interestingly this region enhanced transcription initiated within exon IB as effectively in non-neural cells as in neuronal cells supporting the notion that apparent enhancer selectivity is promoter dependent. Delineation of the proximal enhancer region showed that enhancer activity is located within the proximal 250 bp and the distal 185 bp but both are required for full enhancer activity. EMSAs performed with probes corresponding to the 5' and 3' portions of the proximal enhancer region did not produce any detectable DNA-protein complexes, suggesting that any *trans*-acting factors are either present at low levels or exhibit weak binding under the conditions of the assay. On the other hand, the configuration of linear duplex DNA corresponding to the proximal enhancer region in this assay may not allow DNA binding proteins to bind their cognate sites. For example proteins that bind *cis*-regulatory elements in the c-myc gene interact with torsionally strained but not linear duplex DNA (Michelotti et al., 1996). Whether this is the reason for not detecting DNA:protein interactions could be tested by performing further EMSAs using single-stranded DNA fragments spanning the proximal enhancer region. A NRSE/RE1-like sequence was identified within the distal 185 bp of the proximal enhancer region. Although, more usually associated with repression, NRSE/RE1 sequences that are located downstream of the transcription initiation sites have nonetheless been shown to mediate transcriptional activation (Bessis et al., 1997; Kallunki, et al., 1998). However competition EMSAs performed with the NRSE/RE1 derived from the mouse NRSF/REST gene demonstrated that this sequence does not interact with recombinant NRSF/REST protein.

4.2.4 Cell-specific NRSF/REST gene expression is controlled by an array of positive and negative *cis* -acting regulatory domains

The *cis*-acting regulatory elements that control transcription of eukaryotic genes are often found upstream of the transcription start sites, however many genes have been found that contain *cis*-regulatory elements within exons, introns and even 3' flanking sequence. The *cis*-regulatory elements sufficient to direct cell-specific expression of the NRSF/REST gene, at least in transient transfection assays, are present within 11 kb of genomic sequence upstream of the initiation codon. An internal 6.5 kb genomic fragment, containing 4.5 kb of 5' flanking sequence, the three 5' UT exons and part of the first intron is also sufficient to produce cell-specific expression. The 3' end of this 6.5 kb fragment terminates in the first intron and retention of the exon II splice acceptor site ensures intronic sequence is removed from the 5' UTR of reporter transcripts. The 5' UTR of reporter transcripts, containing intronic sequence, could lower translational their efficiency and hence effect the luciferase activity. This is the most likely explanation for not being able to detect luciferase activity when a 2.0 kb *Hin*dIII fragment (nucleotides –516 to +1533), containing exon IA, exon IB, exon IC and terminating in the first intron, was cloned upstream of the luciferase gene in pGL3 Basic (Promega). Addition of a *Hin*dIII/*Nco*I fragment, containing a splice acceptor site, immediately downstream of the 2.0 kb *Hin*dIII fragment resulted in the production of 10-fold luciferase activity over empty vector. Subsequently, all reporter constructs that were used in this study were designed to contain the splice acceptor region. Several of the reporter constructs that were tested in transient transfection assays by Koenigsberger, et al. (2000) contained mouse NRSF/REST genomic fragments terminating in an intron and did not contain a splice acceptor site upstream of the reporter gene. These reporter constructs are predicted to generate reporter transcripts harbouring intronic sequence within their 5' UTR and examination of their data shows that reporter constructs containing a genomic

fragment spanning all three 5' UT exons and terminating in the first intron, produced considerably lower luciferase activity than reporter constructs containing genomic fragments that terminate in either exon IA, exon IB or exon IC.

Although the proximal enhancer region has an essential role in directing expression of the mouse NRSF/REST gene, other *cis*-regulatory domains that reside elsewhere in the gene may have important functions in directing cell-specific expression. Sequential deletions made from the 5' and 3' end of the 6.5 kb NRSF/REST genomic fragment, that produces cell-specific expression in 3T3, Neuro2a and PC12 cell lines and primary neuroepithelia cells, led to the identification of several *cis*-acting enhancer and repressor regions. Three putative negative regulatory regions were identified. Two of the regions located from positions -4500 to -310 and positions -310 to +262 appear to inhibit expression selectively in Neuro2a and PC12 cells. A third fragment between positions +262 and +389 suppresses expression in 3T3, Neuro2a and PC12 cell lines. Since NRSF/REST mRNA is only present at low levels in the ventricular zone and at extremely low levels in differentiated neurones (Chong et al., 1995; Schoenherr and Anderson, 1995) cell-selective repressor activities between -4500 and +262 may be responsible for the lower NRSF/REST expression in these cell types. However, removal of the *cis*-acting repressor region between -4500 and -310 produced no change in reporter gene activity in NE cells implying that this region is not responsible for the lower expression of NRSF/REST observed in these cells. A positive regulatory domain identified between positions +2000 and +2400, which was designated the distal enhancer region, increases expression in 3T3 cells. Similarly the distal enhancer region may also enhance expression in other non-neural cells which express high levels of NRSF/REST.

4.2.5 The organisation of the NRSF/REST gene allows generation of mRNA splice variants encoding truncated NRSF/REST protein isoforms

Determination of the mouse NRSF/REST gene structure and comparison to the recently described rat and human gene structures (Palm et al., 1998; 1999) shows that the exon/intron organisation is conserved. An alternatively spliced exon, exon N, which is expressed in brain tissue results in NRSF/REST mRNA splice variants containing either a 16 bp (REST4) or 28 bp (REST5) insertion after the fifth zinc finger motif (Palm et al., 1998; Palm et al., 1999). These neural specific mRNA splice variants are predicted to encode an NRSF/REST isoform ($NRSF_{SF}$) containing the N-terminal repression domain and five of the eight zinc finger motifs of the DNA binding domain. As NRSF/REST is the only protein that has been found to bind the NRSE/RE1 sequence, post-transcriptional regulation resulting in alternative splicing is an important mechanism for increasing the repertoire of NRSE/RE1 binding proteins.

4.3 Concluding remarks

Since the identification of the NRSE/RE1 *cis*-regulatory element in multiple neurone-specific genes over the last decade much has been learnt about NRSE/RE1 mediated transcriptional regulation, and more recently about NRSF/REST and its isoforms: the only NRSE/RE1 binding protein identified thus far. The expression pattern of different NRSF/REST mRNA splice variants has been studied, the mechanisms by which NRSF/REST represses transcription have begun to be determined, and loss-of-function and gain-of-function studies have provided clues about the role of NRSF/REST during development. However little is known about the mechanisms that govern cell-specific expression of the *Rest* gene during development and adulthood. This study, which has mapped the transcription start sites of the predominant NRSF/REST mRNA transcript and identified several *cis*-regulatory domains within the mouse NRSF/REST

gene, has opened the doorway to study the mechanisms underlying transcriptional regulation of the NRSF/REST gene.

4.4 Future Experiments

RNase protection assays have shown that NRSF/REST mRNA containing type A, 5' UTR is the most abundant in all rat tissues analysed (Palm et al., 1998). In accord with these data, semi-quantitative PCR showed that NRSF/REST-expressing cell lines contain 1000 –fold more mRNA containing type A, 5' UTR than mRNA containing type B and C, 5' UTRs. Two clusters of transcription start sites were identified within mouse exon IA. As discussed earlier (see chapter 4, section 4.2.2) the use of alternative transcription start sites within exon IA may result in NRSF/REST mRNA transcripts with different half-lives or translational efficiencies. To determine if this is the case experiments can be performed using expression plasmids containing different lengths of type A 5' UTR, corresponding to the different start sites, fused to luciferase coding region. The translational efficiency of transcripts can then be compared by assaying luciferase activity following *in vitro* translation or transient transfection of these expression constructs. Furthermore the half-life of mRNA containing different NRSF/REST 5' UTRs can be compared by treating transiently transfected cells with actinomycin D, which halts mRNA synthesis, and then assaying luciferase activity. To determine whether the two clusters of transcription initiation sites within exon IA are used in a developmental stage- or tissue-specific manner, *in situ* hybridisation can be performed, using a probe to the common region of type A, 5' UTR and another probe to 5' UT sequence present in mRNA initiated from the start sites at positions +1, +13 and +25 only.

In situ hybridisation analysis of mouse and chick embryos has shown that low levels of NRSF/REST mRNA are present in the developing nervous system and expression is restricted mainly to the ventricular zone (Chen et al., 1998; Chong et al., 1995; Schoenherr and Anderson, 1995). Furthermore, loss-of-function and gain-of-function studies demonstrate that the precise control of NRSF/REST gene expression in the ventricular zone is essential for the proper development of neurones (Chen et al., 1998; Paquette et al., 2000). Primary cultures of NE cells derived from the cortical ventricular zone of rat embryos can differentiate into neurones, astrocytes and oligodendrocytes (Williams and Price, 1995). However, NE primary cultures, derived from cortex of E14 rat embryos, that have been exposed to platelet-derived growth factor (PDGF) differentiate into neurones only (Williams et al., 1997). If the levels of NRSF/REST mRNA in PDGF-treated NE cells show a reduction compared to untreated cells, then these cells represent a convenient model system to assay NRSF/REST reporter constructs and identify the *cis*-regulatory elements required for downregulation of NRSF/REST gene expression during neuronal differentiation.

The proximal enhancer region is important for reporter gene expression directed by the NRSF/REST promoter. Heterologous promoter constructs show that the proximal and distal portions of this *cis*-regulatory domain contain enhancer activity, however the precise *cis*-acting regulatory elements have not been mapped and the *trans*-acting factors that bind have not been identified because DNA-protein interactions were not detected using EMSAs. As an alternative approach to try and identify *trans*-acting factors that bind the proximal enhancer region, a yeast one-hybrid screen of a cDNA expression library can be performed. This technique has the advantage of allowing larger DNA fragments, than can be used in EMSAs, to be tested. Thus the proximal enhancer region would not need to be tested as three separate DNA fragments, but the whole region can

be used to identify cDNA clones encoding DNA binding proteins. The resultant cDNA clones, encoding transcription factors can be tested functionally by co-transfection with NRSF/REST reporter contructs containing the proximal enhancer region to determine the effects on reporter gene expression.

Finally, it would appear from the transient transfection assays in the present study, that the NRSF/REST reporter constructs, pGL3 -8.5 and pGL3 -4.5 Δ2.0/2.4, contain the *cis*-acting elements required to direct cell-specific expression, however rigorous testing of this hypothesis can only be carried out using a transgenic animal model. Once a genomic fragment is identified that is sufficient to direct reporter gene expression in transgenic mice that recapitulates expression of the endogenous NRSF/REST gene, then deletion and mutation analyses of *cis*-regulatory domains can be performed. Comparison of the expression of wild-type and mutated transgenes in development and adulthood could provide useful insights into the role of specific *cis*-regulatory elements in directing NRSF/REST expression *in vivo*. Furthermore, several mouse lines have been generated with targeted mutations of genes encoding transcription factors that play a role in neuronal determination and differentiation (Joyner and Guillemot, 1994). Generation of transgenic mice, expressing NRSF/REST reporter genes, in genetically mutated backgrounds can enable the genetic pathways that control NRSF/REST expression in development to be determined.

5.0 REFERENCES

Aasland, R., Stewart, A. F., and Gibson, T. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. Trends Biochem Sci ***21***, 87-88.

Akazawa, C., Sasai, Y., Nakanishi, S., and Kageyama, R. (1992). Molecular characterisation of a rat negative regulator with a basic helix-loop-helix structure predominantly expressed in the developing nervous system. J Biol Chem ***267***, 21879-21885.

Anderson, D. J., Groves, A., Lo, L., Ma, Q., Rao, M., Shah, N. M., and Sommer, L. (1997). Cell lineage determination and the control of neuronal identity in the neural crest. Cold Spring Harb Symp Quant. Biol ***62***, 493-504.

Anderson, D.J., Jan, Y. N. (1997) Molecular and Cellular Approaches to Neural Development, The determination of neuronal phenotype (Oxford Univ. Press, New York).

Andrés, M. E., Burger, C., Peral-Rubio, M. J., Battaglioli, E., Anderson, M. E., Grimes, J., Dallman, J., Ballas, N., and Mandel, G. (1999). CoREST: A functional corepressor required for regulation of neural specific gene expression. Proc Natl Acad Sci USA ***96***, 9873-9878.

Atouf, F., Czernichow, P., and Scharfmann, R. (1997). Expression of neuronal traits in pancreatic beta cells. J Biol Chem ***272***, 1929-1934.

Bai, G., Norton, D. D., Prenger, M. S., and Kusiak, J. W. (1998). Single-stranded DNA-binding proteins and neuron-restrictive silencer factor participate in cell-specific transcriptional control of the *NMDAR1* gene. J Biol Chem ***273***, 1086-1091.

Bailey, A. M., and Posakony, J. W. (1995). Suppressor of Hairless directly activates transcription of *Enhancer of split* Complex genes in response to Notch receptor activity. Genes Dev ***9***, 2609-2622.

Battaglioli, E., Andres, M. E., Gnadt, J., and Mandel, G. (2000). Unraveling the REST repressor complex. Soc Neurosci Abstr ***26***, 1107.

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.

Bessis, A., Salmon, A.-M., Zoli, M., Le Novère, N., Picciotto, M., and Changeux, J.-P. (1995). Promoter elements conferring neuron-specific expression of the β2-subunit of the neuronal nicotinic acetylcholine receptor studied *in vitro* and in transgenic mice. Neuroscience ***69***, 807-819.

Bessis, A., Champtiaux, N., Chatelin, L., and Changeux, J. P. (1997). The neuron-restrictive silencer element: a dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain. Proc Natl Acad Sci U S A ***94***, 5906-5911.

Boisclair, Y. R., Brown, A. L., Casola, S., Rechler, M. M. (1993). Three clustered Sp1 sites are required for efficient transcription of the TATA-less promoter of the gene for insulin-like growth factor binding protein-2 from the rat. J Biol Chem ***268***, 24892-24901.

Calkhoven, C. F., and Ab, G. (1996). Multiple steps in the regulation of transcription-factor level and activity. Biochem J ***317***, 329-342.

Castella, P., Wagner, J. A., and Caudy, M. (1999). Regulation of hippocampal neuronal differentiation by the basic helix-loop-helix transcription factors HES-1 and MASH-1. J Neurosci Res ***56***, 229-240.

Chen, H., Thiagalingam, A., Chopra, H., Borges, M., Feder, J. N., Nelkin, B. D., Baylin, S. B., and Ball, D. W. (1997). Conservation of the *Drosophila* lateral inhibition pathway in human lung cancer: a hairy-related protein (HES1) directly represses *achaete-scute* homolog-1 expression. Proc Natl Acad Sci USA ***94***, 5355-5360.

Chen, Z.-F., Paquette, A. J., and Anderson, D. J. (1998). NRSF/REST is required *in vivo* for repression of multiple neuronal target genes during embryogenesis. Nature Genet ***20***, 136-142.

Chitnis, A., Henrique, D., Lewis, J., Ish-Horowicz, D., and Kintner, C. (1995). Primary neurogenesis in *Xenopus* embryos regulated by a homologue of the *Drosophila* neurogenic gene *Delta*. Nature ***375***, 761-766.

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 gene expression to neurons. Cell ***80***, 949-957.

Coulson, J. M., Fiskerstrand, C. E., Woll, P. J., and Quinn, J. P. (1999). Arginine vasopressin promoter regulation is mediated by a neuron-restrictive silencer element in small cell lung cancer. Cancer Res ***59***, 5123-5127.

Coulson, J. M., Edgson, J. L., Woll, P. J., and Quinn, J. P. (2000). A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. Cancer Res ***60***, 1840-1844.

de la Pompa, J. L., Wakeham, A., Correia, K. M., Samper, E., Brown, S., Aguilera, R. J., Nakano, T., Honjo, T., Mak, T. W., Rossant, J., and Conlon, R. A. (1997). Conservation of the Notch signalling pathway in mammalian neurogenesis. Development ***124***, 1139-1148.

Duncan, M., DiCicco-Bloom, E. M., Xiang, X., Benezra, R., and Chada, K. (1992). The gene for the helix-loop-helix protein, Id. is expressed in neural precursors. Dev Biol ***154***, 1-10.

Dusing, M. R., and Wiginton D. A. (1994). Sp1 is essential for both enhancer-mediated and basal activation of the TATA-less human adenosine deaminase promoter Nucleic Acids Res ***22***, 669-677.

Falvey, E., and Schibler, U. (1991). How are regulators regulated? FASEB J ***5***, 309-314.

Feo, S., Antona, V., Barbieri, G., Passantino, R., Calì, L., and Giallongo, A. (1995). Transcription of the human β enolase gene (*ENO-3*) is regulated by an intronic muscle-specific enhancer that binds myocyte-specific enhancer factor 2 proteins and ubiquitous G-rich box binding factors. Mol Cell Biol ***15***, 5991-6002.

Frohman, M. A., Dush, M. K., Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA ***85***, 8998-9002.

Galvagni, F., and Oliviero, S. (2000). Utropin transcription is activated by an intronic enhancer. J Biol Chem ***275***, 3168-3172.

Goodman, R. H., and Mandel, G. (1998). Activation and repression in the nervous system. Curr Opin Neurobiol ***8***, 413-417.

Green, J. B. A. (1994). Roads to neuralness: embryonic neural induction as derepression of a default state. Cell ***77***, 317-320.

Grimes, J. A., Nielson, S. J., Battaglioli, E., Miska, E. A., Speh, J. C., Berry, D. L., Atouf, F., Holdener, B. C., Mandel, G., and Kouzarides, T. (2000). The co-repressor mSin3A is a functional component of the REST-CoREST repressor complex. J Biol Chem ***275***, 9461-9467.

Groves, A. K., and Anderson, D. J. (1996). Role of environmental signals and transcriptional regulators in neural crest development. Dev Gen ***18***, 64-72.

Guillemot, F., and Joyner, A. L. (1993). Dynamic expression of the murine *Achaete-Scute* homolog *Mash1* in the developing nervous system. Mech Dev ***42***, 171-185.

Guillemot, F., Lo, 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.

Guillemot, F. (1999). Vertebrate bHLH genes and the determination of neuronal fates. Exp Cell Res ***253***, 357-364.

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

Hemmati-Brivanlou, A. (1998). Should the master regulator *Rest* in peace? Nature Genet ***20***, 109-110.

Hirata, H., Ohtsuka, T., Bessho, Y., and Kageyama, R. (2000). Generation of structurally and functionally distinct factors from the basic helix-loop-helix gene *Hes3* by alternative first exons. J Biol Chem ***275***, 19083-19089.

Horlein, A. J., Näär, A.M., Henzel, T., Torchin, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Soderström, M., Glass, C. K., Rosenfeld, M.G . (1995). Ligand-

dependent repression by the thyroid hormone receptor mediated by a nuclear co-repressor. Nature ***377***, 387-388.

Huang, Y., Myers, S. J., and Dingledine, R. (1999). Transcriptional repression by REST:recruitment of Sin3A and histone deacetylase to neuronal genes. Nature Neurosci ***2***, 867-872.

Ishibashi, M., Sasai, Y., Nakanishi, S., and Kageyama, R. (1993). Molecular characterisation of HES2, a mammalian helix-loop-helix factor structurally related to *Drosophila hairy* and *Enhancer of split*. Eur J Biochem ***215***, 645-652.

Ishibashi, M., Moriyoshi, K., Sasai, Y., Shiota, K., Nakanishi, S., and Kageyama, R. (1994). Persistent expression of helix-loop-helix factor HES1 prevents mammalian neural differentiation in the central nervous system. EMBO J ***13***, 1799-1805.

Ishibashi, M., Ang, S.-L., Shiota, K., Nakanishi, S., Kageyama, R., and Guillemot, F. (1995). Targeted disruption of mammalian *hairy* and *Enhancer of split* homolog-1 (*HES1*) leads to upregulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. Genes Dev ***9***, 3136-3148.

Ishiguro, H., Kim, K. T., Joh, T. H., and Kim, K.-S. (1993). Neuron-specific expression of the human dopamine β-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 Opin Neurobiol ***4***, 8-13.

Jarriault, S., Brou, C., Logeat, F., Schroeter, E. H., Kopan, R., and Israel, A. (1995). Signalling downstream of activated mammalian Notch. Nature ***377***, 355-358.

Jarriault, S., Le Bail, O., Hirsinger, E., Pourquie, O., Logeat, F., Strong, C. F., Brou, C., Seidah, N. G., and Israel, A. (1998). Delta-1 activation of Notch-1 signalling results in *HES-1* transctivation. Mol Cell Biol ***18***, 7423-7431.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., and Smale, S. T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell Biol ***14***, 116-27.

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

Jones, F.S., Meech, R. (1999). Knockout of REST/NRSF shows that the protein is a potent repressor of neuronally expressed genes in non-neural tissues. BioEssays ***21***, 372-376.

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

Kallunki, P., Jenkinson, S., Edelman, G. M., and Jones, F. S. (1995). Silencer elements modulate the expression of the gene for the neuron-glia cell adhesion molecule, Ng-CAM. J Biol Chem ***270***, 21291-21298.

Kallunki, P., Edelman, G. M., and Jones, F. S. (1998). The neural-restrictive silencer element can act as both a repressor and enhancer of L1 cell adhesion molecule gene expression during postnatal development. Proc Natl Acad Sci USA ***95***, 3233-3238.

Kallunki, P., Edelman, G.M., and Jones, F.S. (1997). Tissue-specific expression of the L1 cell adhesion molecule is modulated by the neural-restrictive silencer element. J Cell Biol ***138***, 1343-1354.

Koenigsberger, C., Chicca II, J. J., Amoureux, M.-C., Edelman, G. M., and Jones, F. S. (2000). Differential regulation by multiple promoters of the gene encoding the neuron-restrictive silencer factor. Proc Natl Acad Sci USA ***97***, 2291-2296.

Kozak, M. (1991). An analysis of vertebrate mRNA sequences: intimations of translational control. J Cell Biol ***115***, 887-903.

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.

Kroon, E., and MacDonald, R. J. (1997). Higher yields of P1-phage recombinant plasmid DNA. Elsevier Trends Journals Technical Tips Online.

Discuss Lautiger paper?

Lawinger, P., Venugopal, R., Guo, Z.-S., Immaneni, A., Sengupta, D., Lu, W., Ratelli, L., Carniero, A. M. D., Levin, V., Fuller, G. N., Echelard, Y., and Majumder, S. (2000). The neuronal repressor REST/NRSF is an essential regulator in medulloblastoma cells. Nature Med ***6***, 826-831.

LeCourtois, M., and Schweisguth, F. (1995). The neurogenic Suppressor of Hairless DNA-binding protein mediates the transcriptional activation of the *Enhancer of split* complex genes triggered by Notch signalling. Gens Dev ***9***, 2598-2608.

Lee, J. H., Shimojo, M., Chai, Y. G., and Hersh, L. B. (2000). Studies on the interaction of REST4 with the cholinergic repressor element 1/neuron-restrictive silencer factor. Brain Res Mol Brain Res ***80***, 88-98.

Leichter, M., and Thiel, G. (1999). Transcriptional repression by the zinc finger protein REST is mediated by titratable nuclear factors. Eur J Neurosci ***11***, 1937-1946.

Lewis, J. (1996). Neurogenic genes and vertebrate neurogenesis. Curr Opin Neurobiol ***6***, 3-10.

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 USA ***90***, 1460-1464.

Lietz, M., Cicchetti, P., and Thiel, G. (1998). Inverse expression pattern of REST and synapsin I in human neuroblastoma cells. Biol Chem ***379***, 1301-1304.

Lo, S., Johnson, J. E., Wuenschell, C. W., Saito, T., and Anderson, D. J. (1991). Mammalian *achaete-scute* homolog-1 is transiently expressed by spatially-restricted subsets of early neuroepithelial and neural crest cells. Genes Dev ***5***, 1524-1537.

Lo, L., Sommer, L., and Anderson, D. J. (1997). MASH1 maintains competence for BMP2-induced neuronal differentiation in post-migratory neural crest cells. Curr Biol ***7***, 440-450.

Lo, L., Tiveron, M.-C., and Anderson, D. J. (1998). MASH1 activates expression of the paired homeodomain transcription factor Phox2a, and couples pan-neuronal and subtype-specific components of autonomic neuronal identity. Development ***125***, 609-620.

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

Lönnerberg, P., Schoenherr, C. J., Anderson, D. J., and Ibanez, C. F. (1996). Cell type-specific regulation of choline acetyltransferase gene expression. J Biol Chem ***271***, 33358-33365.

Mandel, G., and McKinnon, D. (1993). Molecular basis of neural-specific gene expression. Annu Rev Neurosci ***16***, 323-345.

Maniatis, T., Goodbourn, S., and Fischer, J. (1987). Regulation of inducible and tissue-specific gene expression. Science ***236***, 1237-1244.

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

Michelotti, G. A., Michelotti, E. F., Pullner, A., Duncan, R., Eick, D., and Levens, D. (1996). Multiple single-stranded *cis* elements are associated with activated chromatin of the human c-*myc* gene in vivo. Mol Cell Biol ***16***, 2656-2669.

Mieda, M., Haga, T., and Saffen, T. W. (1996). Promoter region of the rat m4 muscarinic receptor gene contains a cell type-specific silencer element. J Biol Chem ***271***, 5177-5182.

Mieda, M., Haga, T., and Saffen, D. W. (1997). Expression of the rat m4 muscarinic acetylcholine receptor gene is regulated by the neuron-restrictive silencer element/repressor element 1. J Biol Chem ***272***, 5854-5860.

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

Mori, N., Stein, D., Sigmund, D., 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.

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 non-neuronal cells but not in neuronal cells. Neuron ***9***, 45-54.

Müller, F., Chang, B.-E., Albert, S., Fischer, N., Tora, L., and Strähle, U. (1999). Intronic enhancers control expression of zebrafish *sonic hedgehog* in floor plate and notochord. Development ***126***, 2103-2116.

Myers, S., Peters, J., Huang, Y., Comer, M., Barthel, F., and Dingledine, R. (1998). Transcriptional regulation of the GluR2 gene: neural-specific expression, multiple promoters, and regulatory elements. J Neurosci ***18***, 6723-6739.

Myers, S. J., Peters, J. P., and Dingledine, R. (1997). A translational inhibitory motif resides in the GluR2 5'-untranslated leader. Soc Neurosci Abstr ***23***, 923.

Naruse, Y., Aoki, T., Kojima, T., and Mori, N. (1999). Neural restrictive silencer factor recruits mSin3 and histone deacetylase complex to repress neuron specific target genes. Proc Natl Acad Sci USA ***96***, 13691-13696.

Nishimura, E., Sasaki, K., Maruyama, K., Tsukada, T., and Yamaguchi, K. (1996). Decrease in neuron-restrictive silencer factor (NRSF) mRNA levels during differntiation of cultured neuroblastoma cells. Neuroscience Letters ***211***, 101-104.

Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F., and Kageyama, R. (1995). Hes1 and Hes5 as Notch effectors in mammalian neuronal differentiation. EMBO J ***18***, 2196-2207.

Okamoto, S.-I., Sherman, K., and Lipton, S. A. (1999). Absence of binding activity of neuron-restrictive silencer factor is necessary, but not sufficient for transcription of NMDA receptor subunit type 1 in neuronal cells. Mol Brain Res ***74***, 44-54.

Palm, K., Belluardo, N., Metsis, M., and Timmusk, T. (1998). Neuronal expression of zinc finger transcripton factor REST/NRSF/XBR gene. J Neurosci ***18***, 1280-1296.

Palm, K., Metsis, M., and Timmusk, T. (1999). Neuron-specific splicing of Zinc finger transcription factor REST/NRSF/XBR is frequent in neuroblastomas and conserved in human, mouse and rat. Mol Brain Res ***72***, 30-39.

Paquette, A. J., Perez, S. E., and Anderson, D. J. (2000). Constitutive expression of the neuron-restrictive silencer factor (NRSF)/REST in differentiating neurons disrupts neuronal gene expression and causes axon pathfinding errors *in vivo*. Proc Natl Acad Sci USA ***97***, 12318-12323.

Park, E. J., Schroen, D. J., Yang, M., Li, H., Li, L., and Chen, J. D. (1999). SMRTe, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear receptor corepressor. Proc Natl Acad Sci USA ***96***, 3519-3524.

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. Nucleic Acids Res ***22***, 4748-4755.

Pazin, M. J., and Kadonaga, J. T. (1997). What's up and down with histone deacetylation and transcription? Cell ***89***, 325-328.

Quandt, K., Frech, K., Karas, H., Wingeder, E., and Werner, T. (1995). MatInd and MatInspector-New fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res ***23***, 4878-84.

Quinn, J. P. (1996). Neuronal-specific gene expression-the interactions of both positive and negative transcriptional regulators. Prog Neurobiol ***50***, 363-379.

Roopra, A., Sharling, L., Wood, I. C., Briggs, T., Bachfischer, U., Paquette, A. J., Anderson, D. J., and Buckley, N. J. (2000). Transcriptional repression by the neuron-restrictive silencer factor (REST/NRSF) is mediated via the Sin3-Histone deacetylase complex. Mol Cell Biol ***20***, 2147-2157

Ryan, A. K., and Rosenfeld, M. G. (1997). POU domain family values, flexibility, partnerships and developmental codes. Genes Dev ***11***, 1207-1225.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbour Laboratory Press).

Santra, M., Danielson, K., and Iozzo, R. (1994). Structural and functional characterization of the human decorin gene promoter. A homopurine-

homopyrimidine S1nuclease-sensitive region is involved in transcriptional control. J Biol Chem ***269***, 579-87.

Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoo, 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.

Sasai, Y., Lu, B., Steinbeisser, H., and De Robertis, E. M. (1995). Regulation of neural induction by Chd and Bmp-4 antagonistic patterning signals in *Xenopus*. Nature ***376***, 333-336.

Schoch, S., Cibelli, G., and Thiel, G. (1996). Neuron-specific gene expression of synapsin I. J Biol Chem ***271***, 3317-3323.

Schoenherr, C. J., and Anderson, D. J. (1995a). The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science ***267***, 1360-1363.

Schoenherr, C. J., and Anderson, D. J. (1995b). Silencing is golden: negative regulation in the control of neuronal gene transcription. Curr Opin Neurobiol ***5***, 566-571.

Schoenherr, C. J., Paquette, A. J., and Anderson, D. J. (1996). Identification of potential target genes for the neuron-restrictive silencer factor. Proc Natl Acad Sci USA, 9881-9886.

Scholl, T., Stevens, M. B., Mahanta, S., and Strominger, J. L. (1996). A zinc finger protein that represses transcription of the human MHC classII gene, DPA. J Immunology ***156***, 1448-1457.

Shimojo, M., Paquette, A. J., Anderson, D. J., and Hersh, L. B. (1999). Protein Kinase A regulates cholinergic gene expression in PC12 cells: REST4 silences the silencing activity of neuron restrictive silencer factor/REST. Mol Cell Biol ***19***, 6788-6795.

Simpson, P. (1995). Positive and negative regulators of neural fate. Neuron ***15***, 739-742.

Sommer, L., Shah, N., Rao, M., and Anderson, D. J. (1995). The cellular function of MASH1 in autonomic neurogenesis. Neuron ***15***, 1245-1258.

Sternberg, N. (1992). Cloning high molecular weight DNA fragments by the bacteriophage P1 system. Trends Genet ***8,*** 11-16

Sternberg, N., Smoller, D., and Braden, T. (1994). Three new developments in P1 cloning. Genet Anal Tech Appl ***11***, 171-180.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. Genes Dev ***12***, 599-606.

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

Tanabe, Y., and Jessell, T. M. (1996). Diversity and pattern in the devloping spinal cord. Science ***15***, 1115-1123.

Tapia-Ramírez, J., Eggen, B., Peral-Rubio, M., Toledo-Aral, J., and Mandel, G. (1997). A single zinc finger motif in the silencing factor REST represses the neural-specific type II sodium channel promoter. Proc Natl Acad Sci USA ***94***, 1177-1182.

Tanaka, M., Wechler, S.B., Lee, I. W., Yamasaki, N., Lawitts, J. A., and Izumo, S. (1999) Complex modular *cis*-acting elements regulate exression of the cardiac specifying homeobox gene Csx/Nkx2.5. Development ***126***, 1439-1450.

Tatsuka, M., Ibeanu, G. C., Izumi, T., Narayan, S., Ramana, C. V., Kim, N. K., Kang, W., Roy, G., and Mitra, S. (1995). Structural organisation of the mouse repair gene, N-methylpurine-DNA glycosylase. DNA Cell Biol ***14***, 37-45.

Thiel, G., Lietz, M., and Cramer, M. (1998). Biological activity and modular structure of RE1-silencing transcription factor (REST), a repressor of neuronal genes. J Biol Chem ***273***, 26891-26899.

Thiel, G., Lietz, M., and Leichter, M. (1999). Regulation of neuronal gene expression. Naturwissenschaften ***86***, 1-7.

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.

Timmusk, T., Palm, K., Lendahl, U., and Metsis, M. (1999). Brain-derived neurotrophic factor expression *in vivo* is under the control of neuron-restrictive silencer element. J Biol Chem ***274***, 1078-84.

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

Verma-Kuvari, S., Savage, T., Smith, D., and Johnson, J. E. (1998). Multiple elements regulate *Mash1* expression in the developing CNS. Dev Biol ***197***, 106-116.

Wang, W., Chi, T., Xue, Y., Zhou, S., Kuo, A., and Crabtree, G. R. (1998). Architectural DNA binding by a high-mobility-group/kinesin-like subunit in mammalian SWI/SNF-related complexes. Proc Natl Acad Sci USA ***95***, 492-498.

Weber, J. R., and Skene, J. H. P. (1997). Identification of a novel repressive element that contributes to neuron-specific gene expression. J Neurosci ***17***, 7583-7593.

Williams, B. P., Park, J. K., Alberta, J. A., Muhlebach, S. G., Hwang, G. Y., Roberts, T. M., and Stiles, C. D. (1997). A PDGF-regulated immediate early gene response initiates neuronal differentiation in ventricular zone progenitor cells. Neuron ***18***, 553-562.

Williams, B. P., and Price, J. (1995). Evidence for multiple precursor types in the embryonic rat cerebral cortex. Neuron ***14***, 1181-1188.

Wilson, P. A., and Hemmati-Brivanlou, A. (1995). Inhibition of epidermis and inhibition of neural fate by Bmp-4. Nature ***376***, 331-333.

Wood, I. C., Roopra.A., Harrington, C., and Buckley, N. J. (1995). Structure of the m4 cholinergic muscarinic receptor gene and its promoter. J Biol Chem ***270***, 30933-33940.

Wood, I. C., Roopra, A., and Buckley, N. J. (1996). Neural specific expression of the m4 muscarinic acetylcholine receptor gene is mediated by a RE1/NRSE-type silencing element. J Biol Chem ***271***, 14221-14225.

Wood, I. C., Garriga-Canut, M., Palmer, C. L., Pepitoni, S., and Buckley, N. J. (1999). Neuronal expression of the rat M_1 muscarinic acetylcholine receptor gene is regulated by elements in the first exon. Biochem J ***340***, 475-483.

Wood, I. C., Jones, C. J., Mistry, M., Roopra, A., and Buckley, N. J. (2001). The neuron-restrictive silencer factor remodels chromatin and represses transcription in neurons. submitted.

Wuenchell, 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.