TRANSCRIPTIONAL REGULATORS OF NESTIN IN THE RAT CENTRAL NERVOUS SYSTEM

by

Richard Eric Josephson

B.A., University of Washington, 1988

Submitted to the Department of Biology in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

September 1996

© Massachusetts Institute of Technology 1996
All rights reserved

Signature of Author_______________________________.
Dept. of Biology
August 7, 1996

Certified by_______________________________.
Ronald D. G. McKay
Chief, Lab. of Molecular Biology, NINDS
Thesis Supervisor

Accepted by_______________________________.
Robert D. Rosenberg
Professor, Biology
Chair, Dept. Grad. Committee

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

AUG 14 1996 Science

LIBRARIES

TRANSCRIPTIONAL REGULATORS OF NESTIN IN THE RAT CENTRAL NERVOUS SYSTEM

by

Richard Eric Josephson

Submitted to the Department of Biology at the Massachusetts Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biology

Abstract:

Transcription factors are important control points in the patterning of the developing central nervous system. In order to understand how transcription factors function in fate determination, it is necessary to study transcriptional changes which are closely linked to changes in the determination state of the multipotential precursor cells. These cells can be identified by the expression of the intermediate filament nestin, and the nestin gene is sharply down-regulated upon the differentiation of precursors to neurons or glia. Expression of nestin in the central nervous system is dependent upon an enhancer located in the second intron of the gene. In this thesis, the nestin CNS-specific enhancer has been mapped to a 257 base pair region which is sufficient for expression throughout the embryonic neuroepithelium. This element also drives correct onset of expression in the embryonic neural tube and down-regulation in migrating post-mitotic neurons. Analysis of protein binding to the enhancer *in vitro* has revealed five sites which are recognized by nuclear proteins derived from embryonic CNS tissues. Each of these five sites interacts with multiple binding activities, with complex temporal and tissue distributions. Sequence comparison of the binding sites with known *cis*-element motifs suggests that the nestin enhancer contains two recognition sites for POU transcription factors, two sites for the AP-2 family, and one site possibly related to the cyclic AMP response element or the half-site for non-steroid nuclear hormone receptors. Antibody supershift assays demonstrate that the two POU sites are each bound by the CNS-specific transcription factors Brn-1 and Brn-2.

Supervisor: Dr. Ronald D.G. McKay
Title: Chief, Laboratory of Molecular Biology
National Institute of Neurological Disorders and Stroke

Biography: Richard Josephson was born in Seattle, Washington, in 1966 and attended the University of Washington. He obtained a B.A. with Honors in Biology in 1988. The last three years of this thesis work were completed at the National Institutes of Health in Bethesda, Maryland.

Acknowledgements

This thesis would not have been possible without the efforts of many previous investigators. Urban Lendahl and Lyle Zimmerman cloned the rat nestin gene, and Lyle localized the tissue-specific elements to the introns. The late Jonas Dahlstrand sequenced the human nestin gene at the Karolinska Institute and shared the second intron sequence with us. Dr. Keiko Ozato at NIH, Dr. Michael Rosenfeld at UCSD, and Dr. Trevor Williams at Yale generously provided antibodies. Tim Hayes taught me to do molecular biology through his cautionary koans. Shigeo Okabe made the transgenic mice and helped in the analysis, and Pat Turner made the transgenic mice which are still to be. Martha Marvin went boldly where no grad student had gone before, and made the way much easier for me to follow. Karin Forsberg-Nilsson, Jim Pickel, Heather Cameron, Montse Molne, and Ron McKay gave me many helpful comments on the manuscript. I also wish to thank Ron for creating a laboratory where a simple idea can be taken in any direction, and I am grateful to my many past and present colleagues in the McKay lab for providing guidance, friendship and support, and clean pipet tips late at night. I wish to thank my brother, Paul, for putting a roof over my head and cats in my lap, and my parents for encouraging their children to figure things out for themselves.

Table of Contents page

Chapter I: Why study the transcriptional regulation of nestin? 6

Chapter II: Localization of the neural tissue-specific element of the nestin gene 35

Figure 1. Map of nestin second intron fragments assayed in transgenic embryos 48

Figure 2. The 3' end of nestin's second intron directs reporter expression to the central nervous system 50

Figure 3. A middle portion of the second intron creates various patterns of ectopic expression 53

Figure 4. Deletion of a 70 bp EcoRI fragment affects the limits of the expression domain but not CNS-specificity 55

Figure 5. A 373 bp region highly conserved between the rat and human nestin second introns directs expression to the CNS 57

Figure 6. A 5' portion of the conserved region is insufficient for CNS-specific expression 59

Figure 7. The 3' portion of the conserved region drives reporter expression specifically to the central nervous system 61

Figure 8. The 257 bp CNS-specific element directs correct onset of expression in the neural tube 63

Figure 9. The 257 bp CNS-specific element directs expression to zones of neuronal precursor proliferation 65

Chapter III: Identification of *cis*-acting elements by *in vitro* binding assays 71

Figure 10. DNAseI footprinting assays on the 257 bp nestin CNS-specific element 85

Figure 11. Summary of protected regions detected by DNAse I footprinting 90

Figure 12. The footprint 1 region is bound in vitro by several proteins from various portions of the developing CNS 92

Figure 13. The footprint 5 region is bound by a small number of activities, one of which is specific to the cortex 94
Figure 14. The footprint 2 region is recognized by different sets of proteins in the cortex than in the rest of the CNS 96
Figure 15. The footprint 3 region is most strongly bound by extracts of cerebellum and spinal cord 98
Figure 16. Methylation interference assays show multiple complexes bind the same core recognition sequences 100
Figure 17. Summary of methylation interference patterns 106
Figure 18. Ftpt2 and Ftpt3 probes can compete with one another for binding of the same or similar proteins 108

Chapter IV: Identification of transcription factors which recognize the CNS-specific element *in vitro* **117**
Figure 19. Supershift assays show that Brn-1 and Brn-2 bind to the Ftpt1 probe *in vitro* 127
Figure 20. Brn-1 and Brn-2 form complexes on the Ftpt4 probe very similar to those on Ftpt1 129
Figure 21. Brn-1 and Brn-2 binding activities are present in telencephalon on embryonic day 13 131
Figure 22. Brn-1 and a single Brn-2 binding activity persist in cortex on postnatal day 5 133

Conclusions: A model of the nestin CNS-specific enhancer complex **142**
Figure 23. Model of the nestin CNS-specific enhancer 143
Table I. Summary of binding activities on the nestin CNS-specific enhancer 145

Chapter I: Why study the transcriptional regulation of nestin?

The development of the vertebrate central nervous system resembles the history of the Union Jack, the flag of the United Kingdom. The emblem of England has long been a simple upright red cross on a white field, the cross of St. George. At the beginning of the seventeenth century this flag was overlaid upon the diagonal white cross on blue of St. Andrew, to celebrate the union of England and Scotland. Two hundred years later the diagonal red cross on white of St. Patrick, patron saint of Ireland, was superimposed upon the other two crosses to form the complex pattern of the Union Jack, a symbol of unity among three nations.

The central nervous system is similarly fashioned by superimposing simple patterns to produce a complex design. In this case, we are talking about patterns of gene expression. The large number of differentiated cell types of the mammalian brain are produced in a cascade of determination steps in which the potential fates of totipotent embryonic cells are progressively restricted. Inductive events or intrinsic programs restrict cell fates by altering gene expression in a region of the embryo to set it apart from a more general region. These fate restrictions are implemented by transcription factors, which drive changes in overall gene expression at each determination step.

No transcription factor has yet been found which is active in only one of the many cell types of the central nervous system. Rather, it seems that a given cell type is established by the combined effects of numerous transcription factors over the course of development. However, it may be possible to dissect apart the overlapping patterns, as many Scottish and Irish people have suggested doing. By concentrating on one crucial stage of development, we can discover simple patterns underlying the whole.

For this reason, the intermediate filament nestin provides an opportunity to look for transcription factors involved in generation of the central nervous system (CNS). The expression of nestin in the CNS is largely limited to multipotential precursor cells, and is tightly

correlated to the proliferating cell state. While there is no evidence that intermediate filaments are necessary for the development of any cell type, it is likely that the upstream signals which activate or de-activate nestin are also involved in determination steps which affect neuroepithelial precursor cells.

Brief summary of CNS development, with overemphasis on transcription factors

The vertebrate central nervous system is formed by interaction between the mesodermal and ectodermal layers of the embryo. A region of dorsal mesoderm called the organizer or Hensen's node signals ectodermal cells to become neural, and also specifies the anteroposterior (A/P) axis of the embryo. Formation of the organizer in Xenopus embryos is dependent on physical rearrangements such as cortical rotation which specifies the dorsoventral polarity of the embryo [Black and Gerhart, 1985], and interactions between the animal and vegetal poles. A number of transcription factors are associated with the organizer. These include the homeobox gene *goosecoid*, which is expressed in mesodermal cells in the organizer region, and which can duplicate organizer activity when *goosecoid* mRNA is injected into Xenopus embryos [Cho *et al.*, 1991]. However, *goosecoid* is not required for gastrulation or neurulation in mice; knockout animals merely have craniofacial and rib abnormalities [Rivera-Perez *et al.*, 1995; Yamada *et al.*, 1995]. In addition, two Xenopus winged-helix factors, *Xenopus fork head* and *pintallavis*, are transcribed in the dorsal blastopore lip [Dirksen and Jamrich, 1992; Ruiz i Altaba and Jessell, 1992]. Additionally, a maternally transcribed LIM-domain factor, *Xlim-1*, is localized to dorsal mesoderm during gastrulation, and can promote neural or muscle differentiation if the inhibitory LIM domains are mutated [Taira *et al.*, 1994] .

The organizer is the source of neural inducing signals which travel both vertically from axial mesoderm to overlying ectoderm and horizontally through the ectoderm. Vertical signals from chordamesoderm induce neural tissue with the same A/P value [Mangold, 1933], but planar signals traveling through the neural

plate from posterior mesoderm can also induce a nearly complete A/P axis in exogastrulae or Keller explants [Doniach *et al.*, 1992; Papalopulu and Kintner, 1993; Ruiz i Altaba, 1992]. Cooperation of both kinds of signals is necessary for formation of certain structures, such as the eyes [Dixon and Kintner, 1989].

While planar signals alone can induce neural tissue, vertical signals are necessary for folding of the neural tube. For example, the notochord promotes bending of the neural plate by inducing medial cells to adopt a wedge-like shape [Smith and Schoenwolf, 1989; Van Straaten *et al.*, 1988]. However, intrinsic forces are also necessary for neurulation and appear to be generated by microfilaments and microtubules and to require internal calcium [reviewed by Jacobson, 1991]. Although the intermediate filament nestin is present during this process (see next subheading), it is not known to have a role in neurulation.

Both vertical and lateral signals are also involved in the mediolateral regionalization of the neural tube. The underlying notochord and prechordal mesoderm induce formation of the floorplate [Placzek *et al.*, 1990; Yamada *et al.*, 1991]. Notochord and floorplate can each induce ventral fates in the neural tube; these regions co-express the diffusible factor sonic hedgehog and the transcription factors HNF3α and HNF3β (Xenopus *Pintallavis*) [Ruiz i Altaba, 1994]. These signals also influence the expression patterns of the dorsally expressed transcription factors Pax-3 and Pax-7 [Goulding *et al.*, 1993]. Additional signals from the nonneural ectoderm signal the lateral neural plate cells to adopt dorsal fates; these signals include BMP-4 [Liem *et al.*, 1995].

The molecular nature of the neural inducing signals is just beginning to be understood. Two candidate neural inducers are the secreted polypeptides noggin and follistatin. Both are expressed in the organizer region and later in the notochord [Hemmati-Brivanlou *et al.*, 1994; Smith and Harland, 1992]. Each appears to act independently of the other, and can directly induce ectoderm to express both general and anterior neural markers but not caudal markers [Hemmati-Brivanlou *et al.*, 1994; Lamb *et al.*, 1993]. There is evidence that neural differentiation is actually the default pathway

for animal cap cells [Grunz and Tackle, 1989; Sato and Sargent, 1989]; the function of follistatin may be to bind and inhibit activin or a related TGFβ family member which inhibits the neural fate [Hemmati-Brivanlou and Melton, 1994].

The fact that noggin and follistatin cannot induce posterior neural tissue is in keeping with several earlier models of neural induction. It has been proposed that the A/P pattern of the CNS arises in at least two steps: the first step induces neural tissue with anterior potential, and a subsequent step transforms part of this to posterior neural tissue [Nieuwkoop, 1952]. Experiments in which anterior neural plate tissue was mixed in various ratios with posterior mesoderm suggest that intermediate neural structures are induced by a signal gradient from the mesoderm [Toivonen and Saxén, 1968]. However, the signalling molecule detected in this experiment is thought to be activin, which acts indirectly on neural tissue by inducing dorsal mesoderm [reviewed by Gilbert and Saxen, 1993]. More recent models propose that caudalizing signals diffuse through the plane of the ectoderm. One potential caudalizing signal is basic fibroblast growth factor (bFGF). bFGF is able to induce posterior neural markers in isolated animal caps [Cox and Hemmati-Brivanlou, 1995; Lamb and Harland, 1995] or anterior and posterior markers in a dose-dependent manner [Kengaku and Okamoto, 1995], but may be required to cooperate with anterior inducers in order to do so [reviewed by Doniach, 1995].

Retinoic acid (RA) is another potential caudalizing signal. Endogenous RA is detected in Xenopus gastrulae and neurulae [Durston *et al.*, 1989], and exhibits a ten-fold difference from posterior to anterior ends of neurulae [Chen *et al.*, 1994]. Exposure of embryos to exogenous retinoic acid truncates the A/P axis, and also respecifies the expression domains of region-specific markers [Durston *et al.*, 1989; Sive *et al.*, 1990]. RA affects both the A/P values signalled by the mesoderm and the ability of the ectoderm to respond [Sive and Cheng, 1991]. The responsive period of the embryo to RA is limited; embryonic tissues gradually lose their sensitivity to RA from the gastrula to early neurula stages [Sive *et al.*, 1990]. This suggests that the A/P pattern must be actively maintained by

signalling from gastrulation until committment occurs in the neurula. The receptors for retinoic acid are transcription factors and members of the nuclear receptor superfamily.

In the developing hindbrain, retinoic acid exerts its patterning effect in part by altering the expression of homeobox transcription factors [Marshall *et al.*, 1992]. The sixty amino acid homeobox DNA-binding domain is common to several protein families, but was first described in the Drosophila *antennapedia* and *bithorax* clusters which control segment identity. The mammalian homologues of these clusters are the Hox genes, whose role in the CNS is to specify the identity of several hindbrain segments [McGinnis and Krumlauf, 1992]. However, the Hox genes are not expressed anterior to the hindbrain. Patterning of the forebrain and midbrain may depend upon other homeobox genes: Otx2 appears very early in the anterior neural plate, and forms a nested set of expression patterns in the forebrain and midbrain with Otx1, Emx1, and Emx2 [Simeone *et al.*, 1992]. Otx2 expression is also repressed by retinoic acid [Simeone *et al.*, 1995]. Targeted disruption of Otx2 in mice produces embryos lacking the forebrain and midbrain regions [Acampora *et al.*, 1995].

Additionally, there is a great deal of evidence that specification or maintenance of the A/P axis is carried out by the neuroepithelial cells themselves. This has been observed in experiments in which portions of the neural plate were removed or rotated. Roach repeated experiments of Spemann, in which a large piece of the anterior neural plate is rotated 180 degrees with respect to the rostrocaudal axis; the grafted tissue develops according to its original orientation [Roach, 1945]. However, Slàdecek performed a similar rotation of neuroepithelium and found complete respecification of the A/P and mediolateral axes up to the early neurula stages [Slàdecek, 1955]. This difference may be because neural fold regions were not included in the latter experiment, or due to the smaller graft size; respecification of the axes is more common in smaller grafts [Jacobson, 1991]. However, recent studies on transplantation of tissue pieces or dissociated cells suggest that neuroepithelial cells are not committed to a regional fate; for example, dissociated cells from the postnatal cerebellum can integrate into the postnatal hippocampus

and differentiate in accord with their new position[Vicario-Abejón, 1995 #309; and references therein]. This suggests that individual precursor cells are not committed to a regional fate. Rather, active intercellular signalling maintains regional specifications long after the neurula stage.

Regionalization of the nervous system is quite well understood in the fruitfly. Studies have identified several distinct processes in Drosophila CNS formation, each of which is regulated by numerous interacting factors. The Drosophila central nervous system arises from a neuroectodermal region of the embryo whose dorsal boundary is determined by the secreted factors *decapentaplegic* and *short gastrulation*, while the ventral boundary depends upon the expression domains of the transcription factors *twist* and *snail* [reviewed by Doe and Skeath, 1996]. In each hemisegment, approximately 30 neuroblasts are singled out from this neuroectodermal region by the opposition of proneural and neurogenic genes. The proneural genes promote commitment to the neuroblast fate, while neurogenic genes mediate lateral inhibition of the neural fate in neighboring cells [Jan and Jan, 1993]. Proneural genes in the fly include basic-helix-loop-helix (bHLH) transcription factors of the *achaete-scute* complex (AS-C) and *atonal*. The neurogenic genes include the transmembrane proteins *Delta* and *Notch*, which transmit and receive the inhibitory signal, as well as the bHLH repressors of the *Enhancer-of-split* complex (E(spl)-C) and *groucho* which inhibit transcription of the AS-C genes [Heitzler *et al.*, 1996]. In addition, ligand binding to *Notch* causes an intracellular fragment to be cleaved from the receptor which can enter the nucleus and promote transcription of E(spl)-C genes in conjunction with another neurogenic factor, *Suppressor of Hairless* [Jarriault *et al.*, 1995; Kopan *et al.*, 1996; Lecourtois and Schweisguth, 1995]. After their formation, the neuroblasts are instructed by the proneural genes and the ubiquitously expressed bHLH protein*daughterless* to activate a set of 'neural precursor' genes including the bHLH factors *asense* and *deadpan* [Vaessin *et al.*, 1994].

Many of the genes which specify or inhibit neuroblast fate in drosophila have homologs in vertebrates. Several vertebrate

homologs of the achaete-scute genes are expressed exclusively in neural tissues. The Xenopus achaete-scute homolog *Xash-3* may have a role analogous to the Drosophila proneural genes in promoting neural fate: overexpression of *Xash-3* enlarges the neural tube at the expense of ectodermal cells [Turner and Weintraub, 1994]. Mash-1 is transiently expressed in some proliferating neural precursors, and is required for formation of olfactory and autonomic neurons [Guillemot *et al.*, 1993]. However, Mash-1 is thought to be necessary for differentiation of neuronal precursors and not for committment to the neuronal fate [Sommer *et al.*, 1995]. NeuroD is another bHLH gene expressed in postmitotic neurons in mouse and Xenopus, which also appears to promote neural differentiation [Lee *et al.*, 1995]. Mammals have homologs of other Drosophila proneural genes as well, such as the bHLH factors Math-1 and Math-2 which are related to *atonal* and are expressed in the nervous system [Akazawa *et al.*, 1995; Shimizu *et al.*, 1995].

Vertebrates also have numerous homologs of the Drosophila neurogenic genes, some of which are comparable in function. Multiple homologs of *Notch* have been found in mammals, zebrafish, and chick, as well as homologs of the Drosophila ligands *Delta* and *Serrate* [reviewed in Lewis, 1996]. The expression patterns of Notch-1 and Delta-1 in chick and frogs is consistent with a role in lateral inhibition, as Notch-1 is expressed throughout the proliferating regions of the CNS while Delta-1 is restricted to postmitotic cells which appear to be newborn neurons [Chitnis *et al.*, 1995; Myat *et al.*, 1996]. Moreover, in the chick retina Notch and Delta can be directly shown to regulate the production of neurons [Austin *et al.*, 1995]. This apparent conservation of function extends to another neurogenic gene: the intracellular domain of mammalian Notch cooperates with a homolog of *Suppressor of Hairless* to bind promoter regions [Jarriault *et al.*, 1995].

While many components of lateral inhibition are present in vertebrates, the role of these molecules in vertebrate development is not clear. One proposed role of vertebrate proneural and neurogenic genes is to control the size of the neural plate by regulating the competence of dorsal ectoderm to respond to neural induction. For

example, both *Xenopus Notch 1* (*XN-1*) and *Xash-3* are expressed in the dorsal ectoderm during gastrulation, and also in the neural plate and tube [Coffman *et al.*, 1990; Zimmerna *et al.*, 1993]. However, a presumably activating mutation of *XN-1* enlarges the neural tube rather than promoting ectodermal fates [Coffman *et al.*, 1993], while deletion of Mash-1 does not affect the size of the neural tube [Guillemot *et al.*, 1993]. Of course, many vertebrate homologs of the proneural and neurogenic genes exist, so unknown homologs may be involved in neural plate formation. On the other hand, a model has been proposed in which these genes act after neural plate formation to control the number of neural precursor cells [Chitnis and Kintner, 1995].

Other transcription factors in the neuroepithelium control the proliferation of neural precursor cells. The bHLH protein HES-1 is a homolog of Drosophila *hairy* and E(spl)-C genes and is a transcriptional repressor expressed only in the undifferentiated cells of the ventricular zone; disruption of HES-1 prematurely ends proliferation of neuronal precursors and causes severe neural tube defects [Ishibashi *et al.*, 1995]. The winged-helix factor BF-1 also promotes cell division, but is expressed only in the telencephalic neuroepithelium and eye structures. Targeted disruption of BF-1 reduces the proliferation of precursor cells and leads to their premature differentiation, which greatly reduces the size of the cerebral hemispheres. In addition, these mice may fail to specify ventral but not dorsal telencephalic cells [Xuan *et al.*, 1995].

During neurogenesis, it has been speculated that the precursor cells are asymmetrically dividing stem cells, creating one stem cell and one neuron at each division [Rakic, 1988]. Drosophila neuroblasts divide asymmetrically to produce a neuroblast and a ganglion mother cell. The homeodomain transcription factor *prospero* and the membrane molecule *numb* are inherited during mitosis by only the ganglion mother cell, and may functionally distinguish the mother from the uncommitted neuroblast [Knoblich *et al.*, 1995; Spana and Doe, 1995; Spana *et al.*, 1995]. The mouse gene Prox1 is a homologue of *prospero* expressed in newborn neurons in the subventricular zone, but it is not yet known to have a role in neural fate

determination [Oliver *et al.*, 1993]. In the mammalian cortex, unequal distribution of Notch may differentiate the postmitotic daughter of an asymmetrically dividing stem cell from the proliferating daughter [Chenn and McConnell, 1995].

Between their final division and maturity, newborn neurons must undergo many processes of migration and differentiation which are also dependent on correct transcription factor expression. First, the newly postmitotic neurons of the neural tube migrate outward from the ventricular zone along the processes of radial glia cells [Rakic, 1972]. A naturally arising mutation in the paired-box gene Pax-6 disrupts the radial migration of cortical neurons and the subsequent layering of the cortex [Schmahl *et al.*, 1993]. The same mutation, when homozygous, blocks eye formation; thus Pax-6 may be required for the determination of the entire eye structure [Hogan *et al.*, 1988].

Some transcription factors may have very general roles in differentiation. The *trans*-activating activity called AP-1 is elevated in all regions of the neonatal rat brain but declines to low basal levels in the adult. AP-1 activity is composed of heterodimeric proteins of the fos and jun families. These proto-oncogenes are basic region-leucine zipper (bZIP) transcription factors induced by adult neurons in response to activating stimuli such as light, odor, and pain [Pennypacker, 1995]. Fos and jun also have roles in neuronal differentiation. They are induced by growth factor signals; the first detectable event in the nerve growth factor (NGF) or epidermal growth factor (EGF)-induced differentiation of the PC12 cell line is the transcriptional activation of c-fos [Kruijer *et al.*, 1985]. In addition, the Drosophila jun homolog specifies the R7 photoreceptor [Bohmann *et al.*, 1994].

Another family of bZIP transcription factors also mediate a number of signals in CNS development. The cyclic AMP response element-binding proteins (CREBs) include at least three genes, called CREB, CREM, and ATF-1, which have numerous isoforms and can heterodimerize with one another or jun-related factors. Binding of CREB to the CRE is constitutive, but the protein has trans-activating function only when phosphorylated by cAMP-dependent protein

kinase (PKA), calcium/calmodulin protein kinase (CaMK), or CREB kinase (CREBK) [Vallejo, 1994]. Like AP-1, CREB transcriptional activity is induced in cell lines in response to differentiation signals such as NGF or transforming growth factor $\beta 1$ [Ginty *et al.*, 1994; Kramer *et al.*, 1991].

Other transcription factors are closely associated with the determination of a few specific cell types. The POU transcription factor Pit-1 is required for the proliferation and maintenance of three hormone-secreting cell types of the anterior pituitary, and also transactivates the hormone genes specific to each cell type [Li *et al.*, 1990]. Another POU transcription factor, brain-2 (Brn-2), is required for late steps in the maturation of just three neuronal cell types in the hypothalamus [Nakai *et al.*, 1995; Schonemann *et al.*, 1995]. However, Pit-1, Brn-2 and several other POU genes are also expressed widely in the early CNS, at the time of neural tube closure [Treacy and Rosenfeld, 1992]. The combinatorial pattern of POU genes may be involved in regionalization of the neural tube, assuming this early function is masked in mutant animals by the compensating effects of other factors [Alvarez-Bolado *et al.*, 1995] (see chapter IV for an in-depth discussion of POU factors).

Finally, a wide variety of transcription factors are expressed in the mature cell types of the adult CNS. Many show preferential expression in certain cell types, such as the POU gene Brn-5/Emb in layer IV of the cortex [Andersen *et al.*, 1993], or the lin-11/Isl-1/mec-3 (LIM) family member Isl-1 in motor neurons of the spinal cord and brainstem [Thor *et al.*, 1991]. However, no transcription factors have yet been found to be expressed only in a specific type of neuron. Therefore it is unlikely that mature neuronal phenotypes are specified by unique trans-acting factors; rather, combinations of transcription factors probably specify the identity of cells in the nervous system.

Nestin marks proliferating precursor cells in the CNS

It is apparent that a large number of transcription factors are expressed during development of the central nervous system. In order to identify a few factors out of this abundance which are

critical in the fate choices made by a given cell, it is necessary to study transcriptional changes which coincide with changes in the developmental state of that cell. The intermediate filament gene nestin provides such an opportunity to study the multipotential precursor cells of the mammalian CNS, as nestin expression is tightly linked to the transient precursor cell state.

Nestin was originally described in a search for markers of the neuroepithelial precursor cells. Monoclonal antibodies were made against the E15 rat neural tube, and one antibody, Rat 401, was found to stain radial cells spanning the neural tube from ventricle to pia, including radial glia [Hockfield and McKay, 1985]. The antigen, later identified as nestin [Lendahl *et al.*, 1990], was also found in some non-neuronal cells of the peripheral nervous system and in developing muscle of the somitic myotome. Nestin has since been shown to be transiently expressed in developing muscle and to colocalize with vimentin and desmin in myoblasts and myotubes [Kachinsky *et al.*, 1994; Sejersen and Lendahl, 1993]. Rat 401 staining was not found in the adult brain [Hockfield and McKay, 1985].

The nestin expressing cells of the neural tube are multipotential precursors to neurons and glia. The Rat 401 epitope first appears at E10, shortly after final neural tube closure. The nervous system at this stage is a simple columnar epithelium of rapidly dividing cells. At E11 to E12, approximately 98% of dissociated neural tube cells are Rat 401 positive [Frederiksen and McKay, 1988]; the abundance of nestin sharply declines during periods of neurogenesis in different regions of the brain. In spinal cord, the proportion of Rat 401 staining falls between E11 and E16; in cerebrum, this drop occurs from E16 to E21. This corresponds with the periods of final mitosis observed for neurons in each region [Berry *et al.*, 1964; Nornes and Das, 1974]. During the course of neurogenesis, nestin positive cells were shown to be actively proliferating [Frederiksen and McKay, 1988]. The very tight correlation between cell proliferation, neuronal birth, and Rat 401 staining indicates that expression of nestin is strongly linked to the precursor cell state.

Although Rat 401 staining is found in less than 0.1% of the cells of the CNS by the period of maximal gliogenesis after birth [Frederiksen and McKay, 1988], glial cells can also be derived from nestin positive precursors. Primary O-2A progenitors for oligodendrocytes and type II astrocytes express nestin, which is down-regulated upon differentiation to either fate [Gallo and Armstrong, 1995]. Cultured neural tube primary cells can transiently co-express Rat 401 antigen with neurofilaments or glial fibrillary acidic protein (GFAP), which specifically mark neurons and astrocytes, respectively [Antanitus *et al.*, 1975; Hirokawa, 1982; Schnapp and Reese, 1982]. Also, a nestin positive cerebellar cell line can be induced to express either neuronal or glial markers depending on culture conditions [Frederiksen *et al.*, 1988].

Thus the monoclonal antibody Rat 401 is believed to detect multipotential precursor cells in the developing brain. Other studies have also provided evidence for the existence of multipotential precursors in the nervous system by the immortalizion of precursor cells [Bartlett *et al.*, 1988; Evrard *et al.*, 1990; Ryder *et al.*, 1990], differentiating isolated primary cells in culture [Temple, 1989], or marking proliferating cells *in vivo* with dyes or lacZ-transducing retroviruses [Gray *et al.*, 1988; Turner and Cepko, 1987; Wetts and Fraser, 1988] (See also Kilpatrick *et al.*, 1995 for a review).

Recently, nestin expression has been found in the adult under several conditions. Nestin is strongly expressed by reactive astrocytes in response to brain injury or excitotoxic lesioning [Clarke *et al.*, 1994; Frisen *et al.*, 1995]. Several kinds of CNS tumors express nestin, including glioblastomas, primitive neuroectodermal tumors, and medulloblastomas [Dahlstrand *et al.*, 1992a; Valtz *et al.*, 1991]. Finally, nestin expression is detectable in the normal adult brain. A population of cells isolated from adult striatum can be induced by epidermal growth factor (EGF) to proliferate and express nestin *in vitro* [Reynolds and Weiss, 1992]. These cells were taken from the lateral subventricular zone of the adult brain, and are generally quiescent *in vivo* [Morshead *et al.*, 1994]. However, they can be induced to proliferate and migrate *in vivo* by infusion of EGF; more

than 95% of these EGF-responsive migrating precursor cells of the adult brain express nestin [Craig *et al.*, 1996].

Other CNS intermediate filaments are cell-type specific also

As mentioned previously, the antigen for the Rat 401 monoclonal antibody is called nestin, for neuroepithelial stem cell intermediate filament. The intermediate filament (IF) genes are grouped on the basis of sequence homology into six types, of which nestin alone forms type VI [Lendahl *et al.*, 1990]. The other five classes are: acidic keratins (type I); basic keratins (type II); desmin, vimentin, peripherin and GFAP (type III); neurofilaments and α-internexin (type IV); and nuclear lamins (type V). The conserved locations of the first two introns in the genes for nestin and the neurofilaments suggest that the two classes are derived from a common ancestral gene [Dahlstrand *et al.*, 1992b].

Expression of the neurofilament genes, like nestin, is tightly correlated with the differentiation state of cells in the nervous system. In fact, throughout the development of neural lineages cells alter their expression of intermediate filaments with each change in the cell's developmental potential. This linkage of intermediate filaments and cell state is well known because it has allowed cell biologists to follow the process of specification through immunohistochemistry. Prior to neurulation, the embryonic epithelium expresses cytokeratins. These are absent from the neuroepithelium, which coexpresses vimentin and nestin [Bignami *et al.*, 1982; Hockfield and McKay, 1985]. Both proteins are detected by antibodies at E10, shortly after neural tube closure; in situ hybridization, however, reveals that nestin is expressed in the neural plate by E7.75 in the mouse, at the time the plate first becomes distinguishable [Dahlstrand *et al.*, 1995]. Later, nestin and vimentin disappear from newly postmitotic neurons and are absent from migrating young neurons in the intermediate zone [Bennett, 1987; Fliegner *et al.*, 1994].

These newly born neurons in the intermediate zone express α-internexin and often the light and medium weight neurofilaments (NF-L and NF-M) [Fliegner *et al.*, 1994]. Expression of the fourth type

IV IF, NF-H, follows the other two neurofilaments in many neurons but is absent from a few mature neuronal types. In general, NF-H is found in projection neurons with large cell bodies, whereas α-internexin is the only intermediate filament found in the small cerebellar granule cells [Fliegner *et al.*, 1994].

The other intermediate filaments of the nervous system are also highly selective of cell type. GFAP (type III) is found only in astrocytes and radial glia [Antanitus *et al.*, 1975], while peripherin (type III) is found in neurons of the peripheral nervous system (PNS) and some CNS neurons which project to the periphery [Portier *et al.*, 1983a; Portier *et al.*, 1983b].

Transcriptional regulation of intermediate filament genes

Because nestin and other intermediate filaments of the CNS are so closely identified with the differentiation state of the cell, the gene regulation of many intermediate filaments has been studied in hope of finding how this cell type specificity arises. In cell culture systems, several intermediate filament genes have been found to have *cis*-acting elements which confer expression inducible by known transcription factors or extracellular growth factors. Astrocyte-specific expression of the GFAP gene is dependent upon AP-1, nuclear factor I, and CRE motifs in the proximal promoter [Miura *et al.*, 1990]. Similarly, the proximal promoter of α-internexin contains AP-1, AP-2, AP-4, SP1, and octamer binding sites [Ching and Liem, 1991], while the upstream region of the NF-L gene is transactivated by fos/jun (AP-1) or Krox-24 [Pospelov *et al.*, 1994].

The upstream promoter elements of several IFs mediate the induction of these genes by extracellular growth factors. These inductions underscore the link between filaments and cell identity; several growth factors, including FGF and NGF, act upon neuroepithelial precursors to affect their survival, proliferation, or fate choice [reviewed in Kilpatrick *et al.*, 1995]. The upstream sequences of the NF-M and peripherin genes mediate their induction by NGF [Zopf *et al.*, 1990]. The human and chicken vimentin genes are serum-inducible, and this induction is mediated by AP-1 sites [Carey and Zehner, 1995; Rittling *et al.*, 1989]. The two upstream AP-

1 sites in the chicken vimentin enhancer flank a unique "antisilencer" element which is required for induction by FGF [Carey and Zehner, 1995].

The vimentin gene also illustrates that negative *cis*-acting elements are important for correct regulation of intermediate filament genes. The chicken vimentin gene contains three related silencer elements which bind a 95 kD protein; the abundance of this protein *in vivo* increases as vimentin expression drops in maturing mesoderm [Garzon and Zehner, 1994]. Negative elements have also been described in the mouse GFAP promoter and the human desmin (type III IF) promoter, while epidermal keratin genes are negatively regulated by retinoic acid and thyroid hormone receptors [reviewed by Oshima, 1992].

Clearly, a wealth of information about intermediate filament gene regulation has come from cell culture systems. However, the value of culture models is limited by the fact that several intermediate filaments are promiscuously expressed in cell lines derived from non-expressing tissues [Liem, 1993]. Desmin, α-internexin, and neurofilament light and medium weight genes are all expressed upon transfection into fibroblasts, although the endogenous genes are silent. Likewise, a 10 kilobase DNA fragment encompassing the human keratin 18 gene is correctly expressed in transgenic mice, but promiscuously expressed in various cell lines [Oshima, 1992]. Furthermore, the endogenous vimentin gene is expressed in many cell lines originating from tissues which do not express vimentin *in vivo* [Garzon and Zehner, 1994]. This suggests that native chromatin state plays an important role in silencing intermediate filament genes.

Reliable information about tissue-specific expression of intermediate filament genes therefore must be sought from *in vivo* assays such as the creation of transgenic mice. Fewer neuronal IF genes have been studied by this method. Genomic clones of each of the neurofilament triplet proteins have been shown to give neuron-specific expression in mice. Correct expression of the NF-L gene requires the presence of intragenic sequences; independent neuron-specific enhancers are present in both the first and second introns

[Beaudet *et al.*, 1992; Charron *et al.*, 1995a]. In addition, the 292 bp proximal promoter confers independence from insertion site effects. This phenomenon appears to be related to chromatin structure, as it requires the presence of a Matrix Attachment Region located in the 3 prime untranslated sequence of reporter transcripts [Charron *et al.*, 1995b]. Thus correct regulation of NF-L *in vivo* relies upon several distinct *cis*-acting regions.

Similarly, intragenic sequences are required for correct cell-type specific expression of peripherin in transgenic mice [Belecky-Adams *et al.*, 1993]. The presence of the introns also allows upregulation of peripherin following nerve injury. In contrast, several studies have determined that 2 kilobases of GFAP upstream sequences are sufficient for both astrocyte-specific expression in mice and upregulation of the gene following brain injury [Brenner *et al.*, 1994; Galou *et al.*, 1994; Johnson *et al.*, 1995].

Prior to the beginning of this thesis work, it had been established that the nestin gene also has multiple independent *cis*-acting elements. The upstream flanking sequence and basal promoter are sufficient for selective expression in nestin positive cell lines, but not in transgenic mice. Intragenic sequences are required for correct tissue-specific expression *in vivo*. The nestin first intron drives expression in developing muscle of the somites, while the second intron contains the elements necessary for expression in the neuroepithelium [Zimmerman *et al.*, 1994]. This neuroepithelial element within the second intron fits the classical definition of an enhancer, as it functions independent of position or orientation relative to a heterologous basal promoter.

In summary, studying the regulation of the nestin gene should provide insight into the developmental control of neuroepithelial stem cells in the adult as well as in the embryo. Expression of nestin is tightly correlated to the precursor cell state, and declines rapidly with the birth of postmitotic neurons or glia. The discovery of a tissue-specific element driving neuroepithelial expression of nestin has provided an opportunity to search for transcriptional control mechanisms which may directly influence the fate choices being made by CNS stem cells and their progeny.

This thesis will discuss:

Chapter II -Localization of the nestin neuroepithelial tissue-specific element to 257 base pairs

Chapter III -mapping of five potential *cis*-acting sequences through protein/DNA interactions *in vitro*

Chapter IV -identification of known transcription factors which bind these sequences

References

Acampora, D., S. Mazan, Y. Lallemand, V. Avantaggiato, M. Maury, A. Simeone and P. Brulet (1995) Forebrain and midbrain regions are deleted in Otx2-/- mutants due to a defective anterior neuroectoderm specification during gastrulation. Development **121**, 3279-3290.

Akazawa, C., M. Ishibashi, C. Shimizu, S. Nakanishi and R. Kageyama (1995) A mammalian helix-loop-helix factor structurally related to the product of Drosophila proneural gene atonal is a positive transcriptional regulator expressed in the developing nervous system. J. Biol. Chem. **270**, 8730-8.

Alvarez-Bolado, G., M.G. Rosenfeld and L.W. Swanson (1995) Model of forebrain regionalization based on spatiotemporal patterns of POU-III homeobox gene expression, birthdates, and morphological features. J. Comp. Neurol. **355**, 237-295.

Andersen, B., M.D. Schonemann, R.2. Pearse, K. Jenne, J. Sugarman and M.G. Rosenfeld (1993) Brn-5 is a divergent POU domain factor highly expressed in layer IV of the neocortex. J. Biol. Chem. **268**, 23390-8.

Antanitus, D.S., B.H. Choi and L.W. Lapham (1975) Immunofluorescence staining of astrocytes in vitro using antiserum to glial fibrillary acidic protein. Brain Res. **89**, 363-367.

Austin, C.P., D.E. Feldman, J.A. Ida and C.L. Cepko (1995) Vertebrate retinal ganglion cells are selected from competent progenitors by the action of Notch. Development **121**, 3637-3650.

Bartlett, P.F., H.H. Reid, K.A. Bailey and O. Bernard (1988) Immortalization of mouse neural precursor cells by the c-myc oncogene. Proc. Natl. Acad. Sci. **85**, 3255-3259.

Beaudet, L., G. Charron, D. Houle, I. Tretjakoff, A. Peterson and J.-P. Julien (1992) Intragenic regulatory elements contribute to transcriptional control of the neurofilament light gene. Gene **116**, 205-214.

Belecky-Adams, T., D.C. Wight, J.J. Kopchick and L.M. Parysek (1993) Intragenic sequences are required for cell type-specific and injury-induced expression of the rat peripherin gene. J. Neurosci. **13**, 5056-5065.

Bennett, G.S. (1987) Changes in intermediate filament composition during neurogenesis. Curr. Top. Dev. Biol. **21**, 151-184.

Berry, M., A.W. Rogers and J.T. Eayrs (1964) Pattern of cell migration during cortical histogenesis. Nature **203**, 591-593.

Bignami, A., T. Raju and D. Dahl (1982) Localization of vimentin, the nonspecific intermediate filament protein, in embryonal glia and in early differentiating neurons. Dev. Biol. **91**, 286-295.

Black, S.D. and J.C. Gerhart (1985) Experimental control of the site of embryonic axis formation in Xenopus laevis eggs centrifuged before first cleavage. Dev. Biol. **108**, 310-24.

Bohmann, D., M.C. Ellis, L.M. Staszewski and M. Mlodzik (1994) Drosophila Jun mediates Ras-dependent photoreceptor determination. Cell **78**, 973-86.

Brenner, M., W.C. Kisseberth, Y. Su, F. Besnard and A. Messing (1994) GFAP promoter directs astrocyte-specific expression in transgenic mice. J. Neurosci. **14**, 1030-7.

Carey, I. and Z.E. Zehner (1995) Regulation of chicken vimentin gene expression by serum, phorbol ester, and growth factors: identification of a novel fibroblast growth factor-inducible element. Cell Growth Diff. **6**, 899-908.

Charron, G., L.G. Guy, M. Bazinet and J.P. Julien (1995a) Multiple neuron-specific enhancers in the gene coding for the human neurofilament light chain. J. Biol. Chem. **270**, 30604-10.

Charron, G., J.P. Julien and V. Bibor-Hardy (1995b) Neuron specificity of the neurofilament light promoter in transgenic mice requires the presence of DNA unwinding elements. J. Biol. Chem. **270**, 25739-45.

Chen, Y.-P., L. Huang and M. Solursh (1994) A concentration gradient of retinoids in the early Xenopus laevis embryo. Dev. Biol. **161**, 70-76.

Chenn, A. and S.K. McConnell (1995) Cleavage orientation and the asymmetric inheritance of Notch1 immunoreactivity in mammalian neurogenesis. Cell **82**, 631-641.

Ching, G.Y. and R.K.H. Liem (1991) Structure of the gene for the neuronal intermediate filament protein α-internexin and functional analysis of its promoter. J. Biol. Chem. **266**, 19459-19468.

Chitnis, A., D. Henrique, J. Lewis, D. Ish-Horowicz and C. Kintner (1995) Primary neurogenesis in Xenopus embryos regulated by a homologue of the Drosophila neurogenic gene Delta [see comments]. Nature **375**, 761-6.

Chitnis, A. and C. Kintner (1995) Neural induction and neurogenesis in amphibian embryos. Persp. Dev. Neurobiol. **3**, 3-15.

Cho, K.W., B. Blumberg, H. Steinbeisser and E.M. De Robertis (1991) Molecular nature of Spemann's organizer: the role of the Xenopus homeobox gene goosecoid. Cell **67**, 1111-20.

Clarke, S.R., A.K. Shetty, J.L. Bradley and D.A. Turner (1994) Reactive astrocytes express the embryonic intermediate neurofilament nestin. Neuroreport **5**, 1885-8.

Coffman, C., W. Harris and C. Kintner (1990) Xotch, the Xenopus homolg of Drosophila notch. Science **249**, 1438-1441.

Coffman, C.R., P. Skoglund, W.A. Harris and C.R. Kintner (1993) Expression of an extracellular deletion of Xotch diverts cell fate in Xenopus embryos. Cell **73**, 659-671.

Cox, W.G. and A. Hemmati-Brivanlou (1995) Caudalization of neural fate by tissue recombination and bFGF. Development **121**, 4349-58.

Craig, C.G., V. Tropepe, C.M. Morshead, B.A. Reynolds, S. Weiss and D. van der Kooy (1996) In vivo growth factor expansion of endogenous subependymal neural precursor cell populations in the adult mouse brain. J. Neurosci. **16**, 2649-2658.

Dahlstrand, J., V.P. Collins and U. Lendahl (1992a) Expression of the class VI intermediate filament nestin in human central nervous system tumors. Cancer. Res. **52**, 5334-41.

Dahlstrand, J., M. Lardelli and U. Lendahl (1995) Nestin mRNA expression correlates with the central nervous system progenitor cell state in many, but not all, regions of developing central nervous system. Dev. Brain Res. **84**, 109-129.

Dahlstrand, J., L.B. Zimmerman, R.D. McKay and U. Lendahl (1992b) Characterization of the human nestin gene reveals a close evolutionary relationship to neurofilaments. J. Cell Sci. **103**, 589-97.

Dirksen, M.L. and M. Jamrich (1992) A novel, activin-inducible, blastopore lip-specific gene of Xenopus laevis contains a fork head DNA-binding domain. Genes Dev. **6**, 599-608.

Dixon, J.E. and C.R. Kintner (1989) Cellular contacts required for neural induction in Xenopus embryos: evidence for two signals. Development **106**, 749-757.

Doe, C.Q. and J.B. Skeath (1996) Neurogenesis in the insect central nervous system. Curr. Opin. Neuro. **6**, 18-24.

Doniach, T. (1995) Basic FGF as an inducer of anteroposterior neural pattern. Cell **83**, 1067-1070.

Doniach, T., C.R. Phillips and J.C. Gerhart (1992) Planar induction of anteroposterior pattern in the developing central nervous system of Xenopus laevis. Science **257**, 542-545.

Durston, A.J., J.P.M. Timmermans, W.J. Hage, H.F.J. Hendriks, N.J. de Vries, M. Heideveld and P.D. Nieuwkoop (1989) Retinoic acid causes an anteroposterior transformation in the developing central nervous system. Nature **340**, 140-144.

Evrard, C., I. Borde, P. Marin, E. Galiana, J. Premont, F. Gros and P. Rouget (1990) Immortalization of bipotential and plastic glio-neuronal precursor cells. Proc. Natl. Acad. Sci. **87**, 3062-3066.

Fliegner, K.H., M.P. Kaplan, T.L. Wood, J.E. Pintar and R.K.H. Liem (1994) Expression of the gene for the neuronal intermediate filament protein α-internexin coincides with the onset of neuronal differentiation in the developing rat nervous system. J. Comp. Neurol. **342**, 161-173.

Frederiksen, K., P.S. Jat, N. Valtz, D. Levy and R. McKay (1988) Immortalization of precursor cells from the mammalian CNS. Neuron **1**, 439-448.

Frederiksen, K. and R. McKay (1988) Proliferation and differentiation of rat neuroepithelial precursor cells in vivo. J. Neurosci. **8**, 1144-1151.

Frisen, J., C.B. Johansson, C. Torok, M. Risling and U. Lendahl (1995) Rapid, widespread, and longlasting induction of nestin contributes to the generation of glial scar tissue after CNS injury. J. Cell Biol. **131**, 453-64.

Gallo, V. and R.C. Armstrong (1995) Developmental and growth factor-induced regulation of nestin in oligodendrocyte lineage cells. J. Neurosci. **15**, 394-406.

Galou, M., S. Pournin, D. Ensergueix, J.L. Ridet, J.L. Tchelingerian, L. Lossouarn, A. Privat, C. Babinet and P. Dupouey (1994) Normal and pathological expression of GFAP promoter elements in transgenic mice. Glia **12**, 281-93.

Garzon, R.J. and Z.E. Zehner (1994) Multiple silencer elements are involved in regulating the chicken vimentin gene. Mol. Cell. Biol. **14**, 934-943.

Gilbert, S.F. and L. Saxen (1993) Spemann's organizer: models and molecules. Mech. Dev. **41**, 73-89.

Ginty, D.D., A. Bonni and M.E. Greenberg (1994) Nerve growth factor activates a Ras-dependent protein kinase that stimulates c-fos transcription via phosphorylation of CREB. Cell **77**, 713-725.

Goulding, M.D., A. Lumsden and P. Gruss (1993) Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord. Development **117**, 1001-1016.

Gray, G.E., J.C. Glover, J. Majors and J.R. Sanes (1988) Radial arrangement of clonally related cells in the chicken optic tectum: lineage analysis with a recombinant retrovirus. Proc. Natl. Acad. Sci. **85**, 7356-7360.

Grunz, H. and L. Tackle (1989) Neural differentiation of Xenopus laevis ectoderm takes place after disaggregation and delayed reaggregation without inducer. Cell Differ. Dev. **28**, 211-217.

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

Heitzler, P., M. Bourois, L. Ruel, C. Carteret and P. Simpson (1996) Genes of the Enhancer of split and achaete-scute complexes are required for a regulatory loop between Notch and Delta during lateral signalling in Drosophila. Development **122**, 161-171.

Hemmati-Brivanlou, A., O.G. Kelly and D.A. Melton (1994) Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. Cell **77**, 283-95.

Hemmati-Brivanlou, A. and D.A. Melton (1994) Inhibition of activin receptors signaling promotes neuralization in Xenopus. Cell **77**, 273-282.

Hirokawa, N. (1982) Cross-linker system between neurofilaments, microtubules, and membranus organelles in frog axons revealed by the quick-freeze, deep-etching method. J. Cell Biol. **94**, 129-142.

Hockfield, S. and R. McKay (1985) Identification of major cell classes in the developing mammalian nervous system. J. Neurosci. **5**, 3310-3328.

Hogan, B.L., E.M. Hirst, G. Horsburgh and C.M. Hetherington (1988) Small eye (Sey): a mouse model for the genetic analysis of craniofacial abnormalities. Development **103 Suppl**, 115-119.

Ishibashi, M., S.L. Ang, K. Shiota, S. Nakanishi, R. Kageyama and F. Guillemot (1995) Targeted disruption of mammalian hairy and Enhancer of split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. Genes Dev. **9**, 3136-48.

Jacobson, M. (1991) Developmental Neurobiology New York: Plenum.

Jan, Y. and L. Jan (1993) HLH proteins, fly neurogenesis, and vertebrate myogenesis. Cell **75**, 827-30.

Jarriault, S., C. Brou, F. Logeat, E.H. Schroeter, R. Kopan and A. Israel (1995) Signalling downstream of activated mammalian Notch [see comments]. Nature **377**, 355-8.

Johnson, W.B., M.D. Ruppe, E.M. Rockenstein, J. Price, V.P. Sarthy, L.C. Verderber and L. Mucke (1995) Indicator expression directed by regulatory sequences of the glial fibrillary acidic protein (GFAP) gene: in vivo comparison of distinct GFAP-lacZ transgenes. Glia **13**, 174-84.

Kachinsky, A.M., J.A. Dominov and J.B. Miller (1994) Myogenesis and the intermediate filament protein, nestin. Dev. Biol. **165**, 216-28.

Kengaku, M. and H. Okamoto (1995) bFGF as a possible morphogen for the anteroposterior axis of the central nervous system in Xenopus. **121**, 3121-30.

Kilpatrick, T.J., L.J. Richards and P.F. Bartlett (1995) The regulation of neural precursor cells within the mammalian brain. Mol. Cell. Neurosci. **6**, 2-15.

Knoblich, J.A., L.Y. Jan and Y.N. Jan (1995) Asymmetric segregation of Numb and Prospero during cell division. Nature **377**, 624-627.

Kopan, R., E.H. Schroeter, J.S. Nye and H. Weintraub (1996) Signal transduction by activated mNotch: Importance of proteolytic processing and its regulation by the extracellular domain. Proc. Natl. Acad. Sci. USA **93**, 1683-1688.

Kramer, I.M., I. Koornneef, S.W. de Laat and A.J.M. van den Eijnden-van Raaij (1991) TGF-β1 induces phosphorylation of the cyclic AMP responsive element binding protein in ML-CC164 cells. EMBO J. **10**, 1083-1089.

Kruijer, W., D. Schubert and I.M. Verma (1985) Induction of the proto-oncogene fos by nerve growth factor. Proc. Natl. Acad. Sci. USA **82**, 7330-7334.

Lamb, T.M. and R.M. Harland (1995) Fibroblast growth factor is a direct neural inducer, which combined with noggin generates anterior-posterior neural pattern. Development **121**, 3627-36.

Lamb, T.M., A.K. Knecht, W.C. Smith, S.E. Stachel, A.N. Economides, N. Stahl, G.D. Yancopolous and R.M. Harland (1993) Neural induction by the secreted polypeptide noggin. Science **262**, 713-718.

Lecourtois, M. and F. Schweisguth (1995) The neurogenic suppressor of hairless DNA-binding protein mediates the transcriptional activation of the enhancer of split complex genes triggered by Notch signaling. Genes Dev. **9**, 2598-608.

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

Lendahl, U., L.B. Zimmerman and R.D.G. McKay (1990) CNS stem cells express a new class of intermediate filament protein. Cell **60**, 585-595.

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

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

Liem, K.F., G. Tremml, H. Roelink and T.M. Jessell (1995) Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm. Cell **82**, 969-979.

Liem, R.K.H. (1993) Molecular biology of neuronal intermediate filaments. Curr. Op. Cell Biol. **5**, 12-16.

Mangold, O. (1933) Isolationsversuche zur analyse der entwicklung bestimmmter kopforgane. Naturwissen. **21**, 394-397.

Marshall, H., S. Nonchev, M.H. Sham, I. Muchamore, A. Lumsden and R. Krumlauf (1992) Retinoic acid alters hindbrain Hox code and induces transformation of rhombomeres 2/3 into a 4/5 identity. Nature **360**, 737-741.

McGinnis, W. and R. Krumlauf (1992) Homeobox genes and axial patterning. Cell **68**, 283-302.

Miura, M., T. Tamura and K. Mikoshiba (1990) Cell-specific expression of the mouse glial fibrillary acidic protein gene: identification of the cis- and trans-acting promoter elements for astrocyte-specific expression. J. Neurochem. **55**, 1180-8.

Morshead, C.M., B.A. Reynolds, C.G. Craig, M.W. McBurney, W.A. Staines, D. Morassutti, S. Weiss and D. van der Kooy (1994) Neural stem cells in the adult mammalian forebrain: a relatively quiescent subpopulation of subependymal cells. Neuron **13**, 1071-82.

Myat, A., D. Henrique, D. Ish-Horowicz and J. Lewis (1996) A chick homologue of Serrate, and its relationship with Notch and Delta homologues during central neurogenesis. Dev. Biol. **In Press**,

Nakai, S., H. Kawano, T. Yudate, M. Nishi, J. Kuno, A. Nagata, K. Jishage, H. Hamada, H. Fujii, K. Kawamura, K. Shiba and T. Noda (1995) The POU domain transcription factor Brn-2 is required for the determination of specific neuronal lineages in the hypothalamus of the mouse. Genes Dev. **9**, 3109-3121.

Nieuwkoop, P.D. (1952) Activation and organization of the amphibian central nervous system. J. Exp. Zool. **120**, 1-130.

Nornes, H.O. and G.D. Das (1974) Temporal pattern of neurogenesis in spinal cord of rat. I. An autoradiographic study-Time and sites of origin and migration and settling patterns of neuroblasts. Brain Res. **73**, 121-138.

Oliver, G., B. Sosa-Pineda, S. Geisendorf, E.P. Spana, C.Q. Doe and P. Gruss (1993) Prox 1, a prospero-related homeobox gene expressed during mouse development. Mech. Dev. **44**, 3-16.

Oshima, R.G. (1992) Intermediate filament molecular biology. Curr Opin Cell Biol **4**, 110-116.
Papalopulu, N. and C. Kintner (1993) Xenopus Distal-less related homeobox genes are expressed in the developing forebrain and are induced by planar signals. Development **117**, 961-75.
Pennypacker, K.R. (1995) AP-1 transcription factor complexes in CNS disorders and development. J. Florida M.A. **82**, 551-554.
Placzek, M., M. Tessier-Lavigne, T. Yamada, T. Jessell and J. Dodd (1990) Mesodermal control of neural cell identity: floor plate induction by the notochord. Science **250**, 985-8.
Portier, M.M., P. Brachet, B. Croizat and F. Gros (1983a) Regulation of peripherin in mouse neuroblastoma and rat PC 12 pheochromocytoma cell lines. Dev. Neurosci. **6**, 215-26.
Portier, M.M., B. de Nechaud and F. Gros (1983b) Peripherin, a new member of the intermediate filament protein family. Dev. Neurosci. **6**, 335-44.
Pospelov, V.A., T.V. Pospelova and J.P. Julien (1994) AP-1 and Krox-24 transcription factors activate the neurofilament light gene promoter in P19 embryonal carcinoma cells. Cell Growth. Differ. **5**, 187-96.
Rakic, P. (1972) Mode of cell migration to the superficial layers of fetal monkey neocortex. J. Comp. Neurol. **145**, 61-84.
Rakic, P. (1988) Specification of cerebral cortical areas. Science **241**, 170-176.
Reynolds, B.A. and S. Weiss (1992) Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. Science **255**, 1707-10.
Rittling, S.R., L. Coutinho, T. Amram and M. Kolbe (1989) AP-1/jun binding sites mediate serum inducibility of the human vimentin promoter. Nucleic Acids Res. **17**, 1619-1632.
Rivera-Perez, J.A., M. Mallo, M. Gendron-Maguire, T. Gridley and R.R. Behringer (1995) Goosecoid is not an essential component of the mouse gastrula organizer but is required for craniofacial and rib development. Development **121**, 3005-3012.
Roach, F.C. (1945) Differentiation of the central nervous system after axial reversals of the medullary plate of Amblystoma. J. Exp. Zool. **99**, 53-77.
Ruiz i Altaba, A. (1992) Planar and vertical signals in the induction and patterning of the Xenopus nervous system. Development **115**, 67-80.
Ruiz i Altaba, A. (1994) Pattern formation in the vertebrate neural plate. TINS **17**, 233-243.

Ruiz i Altaba, A. and T.M. Jessell (1992) Pintallavis, a gene expressed in the organizer and midline cells of frog embryos: involvement in the development of the neural axis. **116**, 81-93.

Ryder, E., E. Snyder and C.L. Cepko (1990) Establishment and characterization of multipotent neural cell lines using retrovirus vector-mediated oncogene transfer. J. Neurobiol. **21**, 356-375.

Sato, S.M. and T.D. Sargent (1989) Development of neural inducing capacity in dissociated Xenopus embryos. Dev. Biol. **134**, 263-266.

Schmahl, W., M. Knoedlseder, J. Favor and D. Davidson (1993) Defects of neuronal migration and the pathogenesis of cortical malformations are associated with Small eye (Sey) in the mouse, a point mutation at the *Pax-6* locus. Acta Neuropathol. **86**, 126-135.

Schnapp, B.J. and T.S. Reese (1982) Cytoplasmic structure in rapid-frozen axons. J. Cell Biol. **94**, 667-679.

Schonemann, M.D., A.K. Ryan, R.J. McEvilly, S.M. O'Connell, C.A. Arias, K.A. Kalla, P. Li, P.E. Sawchenko and M.G. Rosenfeld (1995) Development and survival of the endocrine hypothalamus and posterior pituitary gland requires the neuronal POU domain factor Brn-2. Genes Dev. **9**, 3122-3135.

Sejersen, T. and U. Lendahl (1993) Transient expression of the intermediate filament nestin during skeletal muscle development. J. Cell Sci. **106**, 1291-300.

Shimizu, C., C. Akazawa, S. Nakanishi and R. Kageyama (1995) MATH-2, a mammalian helix-loop-helix factor structurally related to the product of Drosophila proneural gene atonal, is specifically expressed in the nervous system. Eur. J. Biochem. **229**, 239-248.

Simeone, A., D. Acampora, M. Gulisano, A. Stornaiuolo and E. Boncinelli (1992) Nested expression domains of four homeobox genes in developing rostral brain. Nature **358**, 687-690.

Simeone, A., V. Avantaggiato, M.C. Moroni, F. Mavilio, C. Arra, F. Cotelli, V. Nigro and D. Acampora (1995) Retinoic acid induces stage-specific antero-posterior transformation of rostral central nervous system. Mech Dev **51**, 83-98.

Sive, H.L. and P.F. Cheng (1991) Retinoic acid perturbs the expression of Xhox.lab genes and alters mesodermal determination in Xenopus laevis. Genes Dev. **5**, 1321-1332.

Sive, H.L., B.W. Draper, R.M. Harland and H. Weintraub (1990) Identification of a retinoic acid-sensitive period during primary axis formation in Xenopus laevis. Genes Dev. **4**, 932-942.

Slàdecek, F. (1955) Regulative tendencies of the central nervous system during embryogenesis of the Axolotl (Amblystoma mexicanum cope). II. Regulation after simultaneous inversion of anteroposterior and medio-lateral axes of medullary plate. Acta Soc. Zool. Bohemoslav. **19**, 138-149.

Smith, J.L. and G.C. Schoenwolf (1989) Notochordal induction of cell wedging in the chick neural plate and its role in neural tube formation. J. Exp. Zool. **250**, 49-62.

Smith, W.C. and R.M. Harland (1992) Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in Xenopus embryos. Cell **70**, 829-840.

Sommer, L., N. Shah, M. Rao and D.J. Anderson (1995) The cellular function of Mash-1 in autonomic neurogenesis. Neuron **15**, 1245-1258.

Spana, E.P. and C.Q. Doe (1995) The prospero transcription factor is asymmetrically localized to the cell cortex during neuroblast mitosis in Drosophila. Development **121**, 3187-95.

Spana, E.P., C. Kopczynski, C.S. Goodman and C.Q. Doe (1995) Asymmetric localization of numb autonomously determines sibling neuron identity in the Drosophila CNS. Development **121**, 3489-94.

Taira, M., H. Otani, J.P. Saint-Jeannet and I.B. Dawid (1994) Role of the LIM class homeodomain protein Xlim-1 in neural and muscle induction by the Spemann organizer in Xenopus [published erratum appears in Nature 1995 Feb 2;373(6513):451]. Nature **372**, 677-9.

Temple, S. (1989) Division and differentiation of isolated CNS blast cells in microculture. Nature **340**, 471-473.

Thor, S., J. Ericson, T. Brannstrom and T. Edlund (1991) The homeodomain LIM protein Isl-1 is expressed in subsets of neurons and endocrine cells in the adult rat. Neuron **7**, 881-889.

Toivonen, S. and L. Saxén (1968) Morphogenetic interaction of presumptive neural and mesodermal cells mixed in different ratios. Science **159**, 539-540.

Treacy, M.N. and M.G. Rosenfeld (1992) Expression of a family of POU-domain protein regulatory genes during development of the central nervous system. Annu. Rev. Neurosci. **15**, 139-165.

Turner, D.L. and C.L. Cepko (1987) A common progenitor for neurons and glia persists in rat retina late in development. Nature **238**, 131-136.

Turner, D.L. and H. Weintraub (1994) Expression of achaete-scute homolog 3 in Xenopus embryos converts ectodermal cells to a neural fate. Genes Dev. **8**, 1434-1447.

Vaessin, H., M. Brand, L.Y. Jan and Y.N. Jan (1994) daughterless is essential for neuronal precursor differentiation but not for initiation of neuronal precursor formation in the Drosophila embryo. Development **120**, 935-945.

Vallejo, M. (1994) Transcriptional control of gene expression by cAMP-response element binding proteins. J. Neuroendo. **6**, 587-596.

Valtz, N.L., T.E. Hayes, T. Norregaard, S.M. Liu and R.D. McKay (1991) An embryonic origin for medulloblastoma. New Biol. **3**, 364-71.

Van Straaten, H.W.M., J.W.M. Hekking, E.J.L.M. Wiertz-Hoessels, F. Thors and J. Drukker (1988) Effect of the notochord on the differentiation of a floor plate area in the neural tube of the chick embryo. Anat. Embryol. **177**, 317-324.

Wetts, R. and S.E. Fraser (1988) Multipotent precursors can give rise to all major cell types of the frog retina. Science **239**, 1142-1145.

Xuan, S., C.A. Baptista, G. Balas, W. Tao, V.C. Soares and E. Lai (1995) Winged helix transcription factor BF-1 is essential for the development of the cerebral hemispheres. Neuron **14**, 1141-1152.

Yamada, G., A. Mansouri, M. Torres, E.T. Stuart, M. Blum, M. Schultz, E.M. De Robertis and P. Gruss (1995) Targeted mutation of the murine goosecoid gene results in craniofacial defects and neonatal death. Development **121**, 2917-22.

Yamada, T., M. Placzek, H. Tanaka, J. Dodd and T.M. Jessell (1991) Control of cell pattern in the developing nervous system: polarizing activity of the floor plate and notochord. Cell **64**, 635-47.

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

Zimmerna, K., J. Shih, J. Bars, A. Collazo and D.J. Anderson (1993) XASH3, a novel Xenopus achaete-scute homolog, provides an early marker of planar neural induction and position along the mediolateral axis of the neural plate. Development **119**, 221-232.

Zopf, D., B. Dineva, H. Betz and E.D. Gundelfinger (1990) Isolation of the chicken middle-molecular weight neurofilament (NF-M)

gene and characterization of its promoter. Nucl. Acids Res. **18**, 521-9.

Chapter II: Localization of the neural tissue-specific element of the nestin gene.

Introduction to the nestin neuroepithelial enhancer

The intermediate filament gene nestin has great value to cell biologists as a marker for multipotent precursor cells. Previous members of this laboratory recognized that the nestin gene could be very valuable to molecular biologists as well. They cloned the gene for nestin from the rat, and determined the approximate location of the *cis*-acting transcriptional control regions. Isolation of these control regions has allowed on-going experiments in which heterologous genes are mis-expressed in the neuroepithelial precursor cells for the purpose of marking or manipulating these cells. In the work presented in this thesis, we have characterized in detail the transcriptional enhancer which drives expression of nestin in the central nervous system. Our goal is to identify transcription factors which have regulatory roles in the fate determination of neural precursor cells.

Our predecessors began the work of studying the transcriptional control of nestin by cloning a long stretch of upstream flanking DNA to look for promoter elements. This 5.8 kilobasepairs (kb) upstream region was ligated to an E. coli beta galactosidase (lacZ) reporter gene, and transfected into immortal cerebellar cell lines to test whether the promoter could confer cell-type specificity. The cell lines ST15A and ST15I were immortalized at the same time, under identical conditions; ST15A expresses nestin while ST15I does not. The nestin upstream reporter plasmid (pNesPlacZ) drives lacZ expression selectively in ST15A, indicating that this upstream sequence confers cell-type specificity *in vitro* [Zimmerman, 1992].

At a later time, 5.3 kb of nestin intragenic sequences were added to the upstream region/lacZ construct. All three introns of the rat nestin gene, which are 964 basepairs (bp), 1654 bp, and 194 bp in length, are included in the intragenic fragment. This fragment was added downstream of the SV40 polyA site which follows the lacZ reporter gene, and the plasmid is called pNesPlacZ+3 introns. The introns have little additional effect in the cell culture assay; their

presence increased the reporter expression in the nestin-positive cell line by less than two-fold. It appears that the nestin upstream flanking region alone is sufficient *in vitro.*

The same reporter plasmid which was sufficient in cell lines was then inserted into transient transgenic mice, to test whether the 5.8 kb upstream region could drive expression in all nestin-positive cells of the CNS, peripheral ganglia, and somites. This assay is called transient because mouse embryos are analyzed within weeks of DNA injection, rather than being bred to make stable lines. The upstream plasmid pNesPlacZ was not correctly expressed in transgenic mice. Only a few of the transgenic animals turn on the reporter in portions of the CNS and somites, and the expression is weak, variable and incomplete [Zimmerman *et al.*, 1994].

Strong and complete tissue-specific expression in transgenic mice requires the three introns. The plasmid bearing 5.3 kb of nestin intragenic sequences (pNesPlacZ+3 introns) was strongly expressed in transient transgenic mice in a pattern closely matching endogenous nestin expression [Zimmerman *et al.*, 1994]. In embryos, lacZ staining is present throughout the CNS and in the somites and presomitic mesoderm. These are the areas which react most strongly to Rat 401 [Hockfield and McKay, 1985]. Furthermore, stable lines made with pNesPlacZ+3 introns demonstrate the expected onset of reporter gene expression in the ventral neural tube at E9.0. The lacZ staining at E10.5 is prominent in the cell bodies and pial endfeet of virtually all neuroepithelial cells, as is Rat 401. Finally, extinction of reporter expression takes place at the appropriate ages in various portions of the CNS. LacZ is gone from spinal cord at E16.5, after neurogenesis is complete [Nornes and Das, 1974] but is still strong on postnatal day 1 in cerebellum, where neuronal progenitors are still dividing after birth [Altman, 1972]. Therefore, the intragenic sequences of the nestin gene are required for correct tissue-specific and developmental regulation in transgenic mice.

These results demonstrated that the nestin gene contains multiple transcriptional regulatory elements. While the intragenic sequences are required for tissue-specific activity in mice, the upstream region alone is sufficient in a nestin-positive cell line. The

upstream flanking region seems to have another function *in vivo*. A 160 bp basal promoter element from the herpes simplex virus thymidine kinase (TK) gene can also be cis-activated by nestin's intragenic sequences. Mice with this construct (TKlacZ+3 introns) show the same expression pattern described above for pNesPlacZ+3 introns. About 60% of mice positive by PCR for the presence of the TKlacZ+3 introns construct express lacZ in the correct pattern, while the remainder are lacZ negative. This contrasts with the results for the native nestin upstream promoter region, which drives proper lacZ expression in 100% of transgenic mice [Zimmerman *et al.*, 1994].

The contrast between the effectiveness of the TK basal promoter and the native nestin promoter indicates that the nestin upstream regions have some sort of activity *in vivo*. Some possible interpretations are that the upstream sequence interacts better with the transcriptional machinery, or binds a factor which increases the likelihood of transcription. However, this would not adequately explain why some transgenic mice have very strong reporter expression overall while others are completely negative. Since every transient transgenic embryo integrates the injected DNA into a unique site, the nestin upstream sequences probably insulate the reporter gene from the influence of the insertion site. Meanwhile, expression from the TK basal promoter is subject to these position effects. Perhaps this is due to the greater length of the nestin sequences, or perhaps the nestin upstream region actively blocks positional effects on transcription.

Recently, it has begun to be understood how transcriptional events in the genome can be isolated from position effects. Sequences called Matrix Attachment Regions (MARs) are thought to organize the genome into "loop domains" which are insulated from epigenetic controls acting on the neighboring loops [Chung, 1993]. MARs are long A/T-rich sequences which bind nuclear matrix, and sometimes contain various motifs such as topoisomerase II sites, homeobox binding sites, and regions of kinked or single-stranded DNA. It is interesting to note that the furthest upstream portion of the cloned nestin flanking sequences is very A/T rich. The nestin upstream sequence has not been examined for the other features, nor the

function, of an MAR, but it is possible that an MAR allows the nestin upstream sequences to suppress insertion site effects.

Further analysis of the nestin intragenic sequences disclosed that the gene contains multiple independent regulatory elements [Zimmerman *et al.*, 1994]. A 1.3 kb fragment encompassing the first intron was tested in transgenic mice along with the upstream region (pNesPlacZ+1st intron). This construct drove expression of lacZ throughout the somites, similar to the staining of muscle precursors seen with Rat 401.

A similar construct was tested with a 1.8 kb fragment containing nestin's second intron downstream of the nestin promoter and lacZ (pNesPlacZ+2nd intron). This plasmid produced transgenic mice with lacZ staining throughout the embryonic central nervous system. The pattern and intensity of staining in the neuroepithelium were very similar to those obtained with pNesPlacZ+3 introns, but no lacZ was seen in somites. A third construct was made containing the third intron within a 0.5 kb fragment (pNesPlacZ+3rd intron). This produced only weak and variable staining identical to that observed with the nestin upstream alone (pNesPlacZ). Thus the nestin gene contains two independent tissue-specific regulatory elements, one for muscle precursors and another for neural precursors, in the first and second introns respectively.

The second intron element was also added to the basal promoter plasmid TKlacZ, both upstream and downstream of the promoter and in either orientation relative to the transcriptional unit. All of these constructs performed similarly, driving reporter expression to the neural tube. This neural tissue-specific element thus meets the classical definition of an enhancer:

> "Modular arrays of sequence elements that stimulate accurate initiation of transcription without strict requirements for distance or orientation with respect to the promoter [Hyman *et al.*, 1988]."

The goal of this thesis is to locate the nestin neuroepithelial enhancer and define it in terms of a "modular array of sequence

elements" bound by identifiable transcription factors. Chapter II deals with locating the enhancer within the 1800 base pair fragment previously defined. We used the transient transgenic mouse assay, analyzing all constructs for their effects at a single age, E13.5, to facilitate comparison to the prior data described above. The tissue-specific elements driving expression in the CNS were localized to a region of 257 base pairs. Transgenic embryos examined at multiple ages demonstrate that this 257 bp region reproduces the onset of nestin expression in the early neural tube, as well as the extinction of expression in postmitotic neurons.

Materials And Methods

All standard molecular biology techniques were performed as described in [Maniatis *et al.*, 1989].

Preparation of Transgenic Mice:

Cesium-chloride purified plasmid DNA was restriction digested to remove the pUC vector sequences. Enzymes used are listed for each construct. The digestion reactions were subsequently extracted once in phenol, once in 25:24:1 phenol/chloroform/isoamyl alcohol, and once in 24:1 chloroform/isoamyl alcohol. The DNA was concentrated by ethanol precipitation and layered atop a 10-40% sucrose gradient in 1 M NaCl, 10 mM Tris-HCl [pH 8.0], 1 mM EDTA. The gradient was centrifuged 16 hours at 35000 rpm in a Beckmann SW50.1 rotor.100-300 μL gradient fractions were checked on an agarose gel for separation of insert from vector. Insert-containing fractions were pooled and dialyzed in Spectrapor cellulose tubing (MWCO 12000-14000) versus injection buffer (10 mM Tris-HCl [pH 8.0], 0.1 mM EDTA) at 4°C overnight. The dialysed DNA was diluted in injection buffer to a concentration of 1-5 ng/μL and sterilized via a 0.22 μM syringe-tip filter (Millipore Millex-GV4).

Transgenic mice were produced as described by [Hogan *et al.*, 1986]. Embryo donors were C57Bl6 x C3H F1 female mice, stud males were C57Bl6 or C57Bl6 x C3H F1, and ICR mice were used as foster mothers. Incorporation of transgene was assayed (as previously described by [Zimmerman *et al.*, 1994]) by PCR on yolk sac DNA for 25 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, using the primers 5'-TACCACAGCGGATGGTTCGG-3' and 5'-GTGGTGGTTATGCCGATCGC-3'. In the presence of *lacZ* DNA, a 352 bp product results.

Whole-mount LacZ Procedure:

Embryos were fixed in 0.2% glutaraldehyde, 2 mM $MgCl_2$, 0.1% Triton X-100 in phosphate-buffered saline (PBS) at room temperature for 60 minutes. They were washed in two changes of PBS plus 2 mM $MgCl_2$, then soaked 20 minutes in PBS plus 2 mM $MgCl_2$, 0.1% Triton X-100, and stained in 1 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal), 5mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6$, 2 mM $MgCl_2$, 0.1% Triton X-100 in PBS at 37°C. All embryos were stained for 90

minutes to three hours, then washed with 70% ethanol to stop the reaction.

LacZ Staining of Frozen Sections:

Embryos were sacrificed at E16.5 and the brains were removed and soaked 20% sucrose until loss of buoyancy. The brains were frozen and cut by cryostat into 40 micron sections, which were collected on glass slides. The tissue was fixed 40 minutes in 0.2% glutaraldehyde and processed for X-Gal histochemistry as described above.

Plasmids:

RgIIBbsSpeTKlacZ: The nestin second intron was cut with SpeI and BbsI, which in this case gives a HindIII-compatible cohesive end. The 587 bp fragment was isolated using Geneclean II (Bio101) and ligated into pBluescript KS+ (Stratagene) digested by SpeI and HindIII. The resulting plasmid (gIIBSBt) was cut with SalI and NotI. The 615 bp fragment containing the nestin sequences was isolated and ligated into SalI and NotI digested pTKlacZ [Zimmerman *et al.*, 1994]. pUC sequences were removed by digestion with SmaI, NotI and ScaI prior to injection of embryos.

RgIISalStuTKlacZ: The plasmid RgIItklacZ [Zimmerman *et al.*, 1994] was cut with StuI and SalI. The overhangs were filled in with Klenow fragment and religated. pUC sequences were removed by digestion with SmaI and ScaI prior to injection of embryos.

NsiSalgIITKlacZ: The plasmid gIItklacZ [Zimmerman *et al.*, 1994] was cut with NsiI and SalI. The NsiI 3' overhang was blunted with bacteriophage T4 polymerase, and the SalI 5' overhang was filled in with Klenow fragment. These blunted ends were religated. pUC sequences were removed by digestion with SmaI and ScaI prior to injection of embryos.

NsiSalgIIBt: The plasmid NSgIItklacZ was cut with HindIII and BstYI. An 857 bp fragment containing all the nestin sequence of the plasmid above was isolated following agarose gel electrophoresis, and ligated into HindIII and BamHI-digested pBluescript KS+.

SpeSalgIITKlacZ: The plasmid NSgIIBt was cut with SpeI. A 770 bp fragment was isolated and ligated into pTKlacZ digested with SpeI and dephosphorylated with calf intestinal alkaline phosphatase. pUC

sequences were removed by digestion with NarI and XmnI prior to injection of embryos.

ΔRI-RINSgIItklacZ: NSgIIbt was digested by EcoRI at two sites 70 bp apart in the nestin second intron and religated. The remaining 777 bp of nestin sequences were cut out with NotI and SalI, and ligated into Not I and Sal I digested pTKlacZ. pUC sequences were removed by digestion with SmaI, NotI and ScaI prior to injection of embryos.

gIIConservedTKlacZ, gIICons5'TKlacZ, 257gIITKlacZ: portions of the second intron were PCR amplified from a rat nestin genomic clone. Amplification was performed using Vent polymerase (New England Biolabs) for 30 cycles as follows: 94°C 30"/55°C 30"/72°C 60", plus a final 8 minute extension at 72°C. PCR products were isolated using QIAEX silicon beads (Qiagen) and cloned into the vector pCRII using the TA cloning kit (Invitrogen). The resulting plasmids were digested with SpeI and XhoI, and the inserts ligated into pTKlacZ cut with SpeI and SalI. This group of plasmids was prepared for injection into mouse embryos using SmaI, SpeI and XmnI to remove the pUC sequences.

Oligonucleotides used for the above amplifications:

gIIConsTKlacZ:	5'-CGGACTAGTACGTTAAGACCTCTGCCCTGAAG-3'
	5'-CCGCTCGAGACTCCTCAGATCAGTCTCCGCC-3'
	392 bp product
gIICons5'TKlacZ:	5'-CACGACGTTGTAAAACGACGG-3'
	5'-TAAGAAGGGAGGGCTCCTCAGG-3'
	348 bp product
257gIITKlacZ:	5'-GGTCTGAAAAGGATTTGGAGAAGG-3'
	5'-CTCACTAAAGGGAACAAAAGCTGG-3'
	307 bp product

Results

The nestin neuroepithelial enhancer is located near the 3' end of the second intron

The neuroepithelial enhancer was located by subdividing the nestin second intron. Each of three overlapping restriction fragments was placed upstream of the herpesvirus thymidine kinase (TK) promoter driving E. coli lacZ and tested in transient transgenic mice (see figure 1.) All of the transgenic embryos were sacrificed at E13.5 or 14.5, ages at which we can expect to see nestin expressed in all regions of the central nervous system. An 847 bp fragment containing the 3' end of the second intron and the entire third exon (NsiSal gIITKlacZ, construct D in figure 1) drove reporter expression throughout the central nervous systems of 3 out of 5 embryos identified as transgenic by PCR on yolk sac DNA (see figure 2a). This ratio of lacZ-expressing embryos to total transgenic mice is comparable to that seen with the entire 1.8 kb second intron region (10/18, construct A in figure 1.)

In contrast, there appears to be no tissue-specific enhancer activity in the 5' end or middle of the second intron. The 5' construct, RgIISalStuTKlacZ (B in figure 1), gave no beta galactosidase expression in any of 3 transgenic embryos (data not shown). The middle fragment, RgIIBbsSpeTKlacZ (C in figure 1), drove reporter expression in 3 out of 5 transgenic embryos. However, the strongest staining in all embryos was outside the CNS, and the pattern of expression was different in each embryo. Only one embryo had staining in the CNS, which was weaker than that in the face and nose of the same animal (see figure 3a). Another embryo has staining on the upper snout and at the proximal end of the limbs, but none in the CNS (figure 3b). Since the staining pattern with construct C varies so greatly from mouse to mouse, we conclude that it does not contain any CNS tissue-specific activity.

The 847 bp enhancer fragment, which drives CNS-specific expression in construct D, could not be easily subdivided with restriction enzymes. Instead, we deleted relatively short restriction fragments from it. Removal of 83 basepairs from the upstream end

of this 847 bp region by digestion with SpeI does not affect the activity of the tissue-specific element: the construct SpeSal gIITKlacZ (E in figure 1) drives CNS lacZ expression in 3 of 4 transgenic embryos (figure 2b). The pattern of staining is not distinguishable from construct D.

On the other hand, deletion of a 70 bp EcoRI fragment from the middle of this 847 bp piece (deltaRIRI-NsiSal gIITKlacZ, F in figure 1) affects the activity of the enhancer. Six of ten transgenic embyos had staining specifically in the CNS, but only one stained the full length of the CNS. The others had abnormal rostral or caudal limits of expression (see arrowheads in figure 4). 4 embryos have no expression posterior to the hindbrain, and one is lacking expression in most of the forebrain. Notably, the tissue specificity of the enhancer is not altered; lacZ is not found outside the CNS in any of these embryos. It may also be noteworthy that two of these embryos have strong expression in the forebrain and midbrain, but none in the spinal cord (4b and 4c).

The position of the neuroepithelial element is predicted by sequence conservation between the rat and human nestin genes

The deletion of the 70 bp EcoRI fragment suggested that sequences roughly in the middle of the 847 bp enhancer region are important for activity. As discussed previously, the sequence of the human nestin second intron was made available to us by Dr. Urban Lendahl. While the rat and human second introns are 52% identical overall, the area of highest homology is also located in the middle of the 847 bp fragment. A stretch of approximately 375 bp was identified in the rat gene with 78% identity to the human intron.

This highly conserved region was isolated by PCR and placed upstream of the TK promoter-lacZ construct. Six of Ten transgenic mice produced with this plasmid (gIIConsTKlacZ, G in figure 1) show lacZ staining throughout the CNS (see figure 5). However, only three of these six show continuous expression in the various regions of the brain and along the length of the spinal cord. The other three express lacZ in discrete patches in the brain and lateral stripes in the spinal

cord; it seems that small groups of cells are lacZ positive while neighboring groups are unstained (figure 5b,c). This patchy expression pattern was observed in only one mouse bearing a larger portion of the nestin enhancer, the 847 bp of construct D.

The conserved 375 bp region was further divided into two overlapping fragments of approximately 250 bp. These were generated by PCR and tested in transgenics. The 5' portion (gIICons5' TKlacZ, H in figure 1) gave various ectopic expression patterns in 4 of 10 transgenic embryos (see figure 6). Portions of the CNS are stained in two of these animals: the spinal cord in figure 6a and the dorsal midbrain and a portion of the spinal cord in figure 6b. However, the second of these two embryos has much stronger staining in the face and limbs than in the CNS. The expression pattern is very different in each of the four embryos, so it appears that the 5' end of the conserved region is not sufficient for tissue-specific activity.

The 3' 257 bp portion of the conserved region (257gIITKlacZ, I in figure 1) drove expression of lacZ in the CNS of 6 out of 8 transgenic embryos. Again, discrete patches or stripes of beta galactosidase were seen in half of the six embryos, while the others showed continuous expression (see figure 7). The embryo shown in figure 7c has relatively little expression in the CNS, so the discrete stripes of lacZ in the spinal cord are easily seen. From the rear the brain appears negative, but the telencephalon is dotted with points of staining. Overall, the expression pattern obtained with 257gIITKlacZ is very similar to that given by the entire second intron. The embyos shown in figures 7a and 7b appear to be lacking expression in the area between the cerebral cortices and the midbrain, and staining of this area is relatively weak even with the complete second intron. This area may stain poorly because the diencephalon is further from the surface of a whole-mounted embryo than are the cortex, midbrain, or spinal cord. Alternatively, additional elements may be necessary to direct expression to these areas.

The 257 bp CNS tissue-specific element recapitulates the neural expression pattern of pNesPlacZ+3 introns

Additional transient transgenic mice injected with the plasmid 257gIITKlacZ were sacrificed at E9.5 or E16.5 to examine the 257 bp element's function at early and late periods of nestin expression. These embryos are compared here to the patterns of CNS staining seen in pNesPlacZ+3 introns stable lines, which express lacZ in both neural and muscle precursors.

At E9.5, expression of 257gIITKlacZ is mostly limited to the ventral neural tube. LacZ staining is seen in ventral diencephalon, mesencephalon, metencephalon and spinal cord. The dorsal neural tube is stained only in the metencephalon and rostral spinal cord (figure 8). The anterior portion of the hindbrain, surrounding the otic vesicle, is virtually lacZ negative in these embryos. Just posterior to the otic vesicle there is a sharp boundary where lacZ expression begins, and this staining extends down the spinal cord beyond the forelimb bud. The staining pattern in the CNS of E9.5 pNesPlacZ+3 introns embryos is very similar. Reporter expression is mainly confined to the ventral neural tube posterior to the optic vesicle, with dorsal expression first appearing in the roof plate of the metencephalon and myelencephalon, then spreading to the dorsal spinal cord (figure 8b, reproduced from [Zimmerman *et al.*, 1994].) Thus, with the possible exception of the hindbrain, 257gIITKlacZ reproduces the early expression pattern seen in a nestin reporter construct with 5.8 kb of upstream sequences and 5.3 kb of intragenic sequences.

The 257gIITKlacZ construct also appears to allow extinction of reporter expression during neurogenesis. Figure 9 shows that transgenic embryos at E16.5 have strong lacZ staining in the subventricular zones surrounding the lateral ventricles of the telencephalon (see figure 9a). It has been previously observed that nestin expression at this age is largely limited to the ventricular and subventricular zones of the telencephalon [Dahlstrand *et al.*, 1995], as is the lacZ reporter here. At higher magnification (figure 9b), strong reporter expression is seen in the SVZ, with blue radial processes extending across the intermediate zone (arrowhead) and a blue layer

at the pial surface (arrow), which is probably the pial endfeet of the radial cells [Zimmerman *et al.*, 1994]. Very little lacZ staining is seen in the outer layers of the cortex where the maturing neurons are located. It appears that differentiating neurons are able to down-regulate expression via the 257 bp element.

Figure 1. Map of nestin second intron fragments assayed for CNS-specific activity in transient transgenic embryos

Black bars on this map show the regions of a rat nestin genomic clone (A) which were inserted into the plasmid pTKlacZ and tested in transient transgenic embryos for the capacity to drive lacZ transcription in the neural tube. Restriction fragments or PCR fragments of the rat nestin second intron were inserted upstream of the herpesvirus thymidine kinase basal promoter and the E. coli lacZ gene in the plasmid pTKlacZ. The resulting constructs were injected into fertilized mouse eggs and removed from foster mothers on embryonic day 13.5 or 14.5. Embryos were assayed for central nervous system-specific βgalactosidase activity by X-Gal histochemistry, and for the presence of the transgene by PCR on yolk sac DNA.

A. Second intron region of a rat nestin genomic clone. This 1832 bp fragment produced strong CNS-specific expression in 10 out of 18 transient transgenic embryos [Zimmerman *et al.*, 1994]. The CNS-specific activity of this fragment is independent of its orientation or order relative to the reporter gene.

B. RgIISalStuTKlacZ A 585 bp BstYI-StuI fragment of nestin genomic DNA was inserted upstream of TKlacZ in the reversed orientation, i.e. coding strand orientation relative to lacZ is opposite to that found in the genome. No βgalactosidase expression was found in any of three transgenic embryos.

C. RgIIBbsSpeTKlacZ A 587 bp BbsI-SpeI fragment from the middle of the second intron was inserted in a reversed orientation upstream of TKlacZ. Various patterns of ectopic staining were seen in three of five transient transgenic embryos.

D. NsiSalgIITKlacZ An 847 bp NsiI-BstYI fragment including the 3' end of the second intron and the entire third exon was inserted in the forward orientation upstream of TKlacZ. Strong expression was seen through the full length of the CNS in three of five transgenic embryos.

E. SpeSalgIITKlacZ Identical to construct D, except an 83 bp NsiI-SpeI fragment was removed from the 5' end of the genomic DNA before insertion into pTKlacZ. Expression pattern was identical to that of construct D in three of four transgenic embryos.

F. ΔRI-RI NSgIITKlacZ Identical to construct D, except a 70 bp EcoRI fragment was removed from the middle of the nestin genomic DNA and religated. Six of ten transgenic embryos displayed expression in the brain, but only two expressed in the spinal cord and only one had strong expression through the length of the neuraxis.

G. gIIConsTKlacZ A 373 bp fragment of genomic DNA highly conserved between the rat and human nestin second introns was isolated by PCR and inserted into pTKlacZ. Six of ten transgenic embryos display CNS-specific βgalactosidase expression.

H. gIICons5'TKlacZ A 253 bp fragment corresponding to the 5' portion of the region tested in construct G was isolated by PCR and inserted into pTKlacZ. Various ectopic expression patterns, including some CNS activity, were seen in four of ten transgenic embryos. None of the embryos had expression selectively in the CNS.

I. 257gIITKlacZ A 257 bp fragment corresponding to the 3' portion of the region tested in construct G, and overlapping by 137 bp the region used in construct H, was isolated by PCR. This construct produced CNS-specific βgalactosidase expression in six of eight transgenic embryos.

The Nestin Neuroepithelial Enhancer is Located Within 257 Basepairs

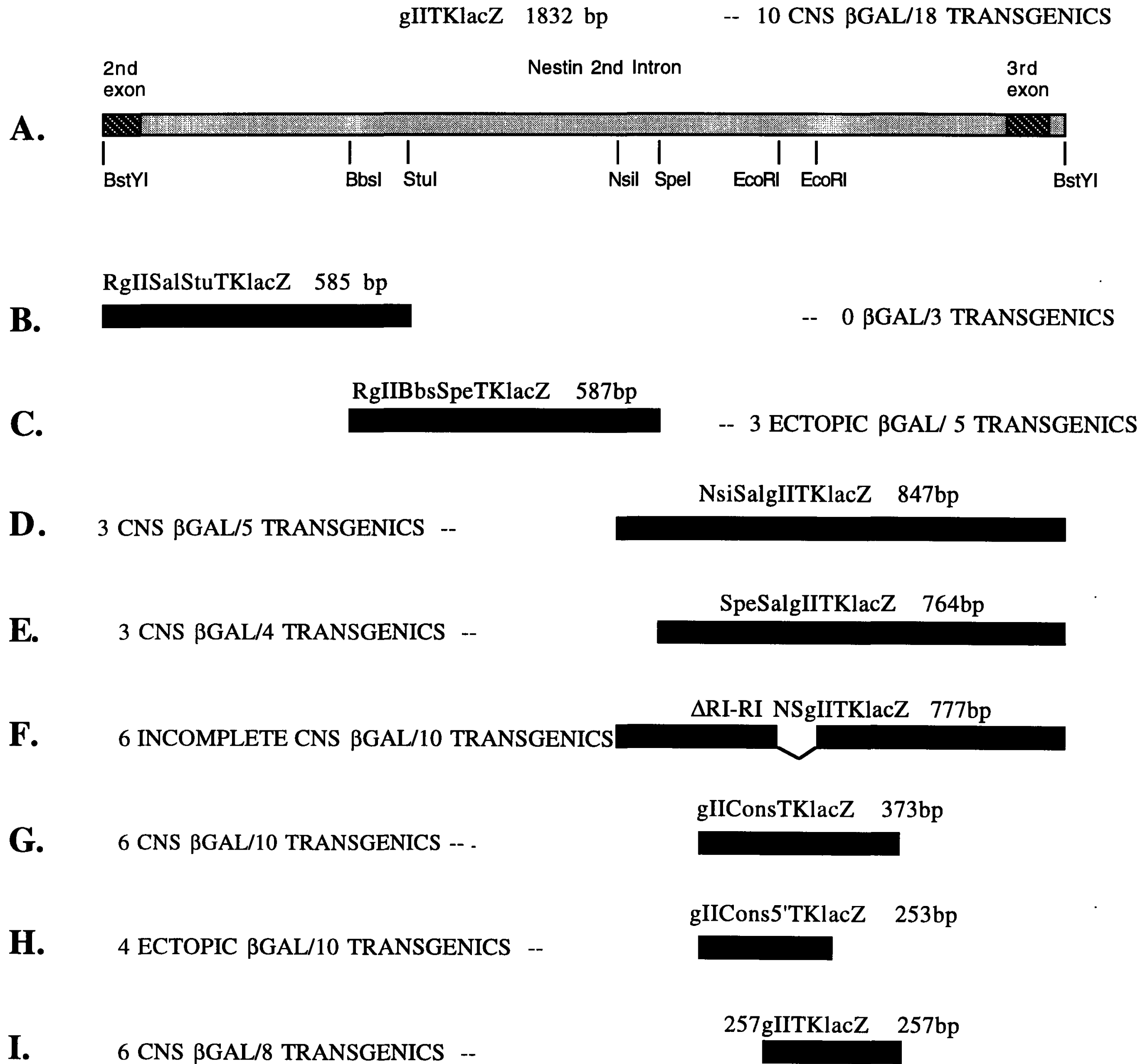
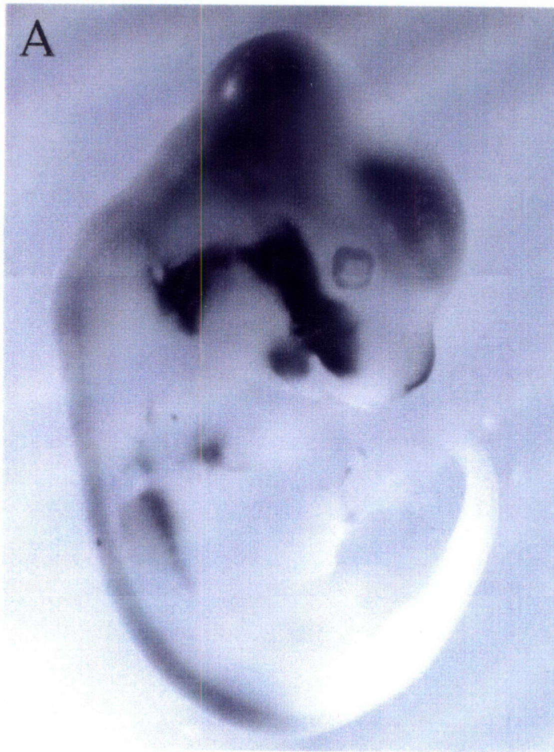


Figure 2. The 3' end of nestin's second intron directs reporter expression to the central nervous system

(A) Whole-mount E13.5 transient transgenic embryo produced with the plasmid NsiSalgIITKlacZ (see figure 1D). The 847 bp fragment of nestin genomic DNA includes the downstream end of the second intron plus the entire third exon. Strong CNS-specific staining is seen along the entire length of the neuraxis.
(B) Whole-mount E13.5 transient transgenic mouse made with the plasmid SpeSalgIITKlacZ (see figure 1E). Strong CNS-specific staining is still obtained with this slightly smaller portion of the second intron.

Figure 3. A middle portion of the second intron creates various patterns of ectopic expression

Two whole-mount E13.5 transient transgenic embryos made with the plasmid RgIIBbsSpeTKlacZ (see figure 1C). Although weak CNS staining is present in one embryo (A), the other two transgenic embryos have no CNS staining (B) and data not shown.


A

B

Figure 4. Deletion of a 70bp EcoRI fragment affects the limits of the expression domain but not CNS-specificity

Six whole-mount E13.5 transgenic mice bearing the plasmid ΔRI-RI NSgIITKlacZ (see figure 1F). Expression is specific for the CNS in all six, but strong and complete in only one embryo (A). Other embryos have more restricted domains of expression; the observed anterior and posterior limits of βgalactosidase activity are marked with arrowheads. Two embryos (B, C) have strong expression in the forebrain and midbrain but lack all expression in the spinal cord. The other three are weakly expressed in the brain and lacking in some portions of the CNS.

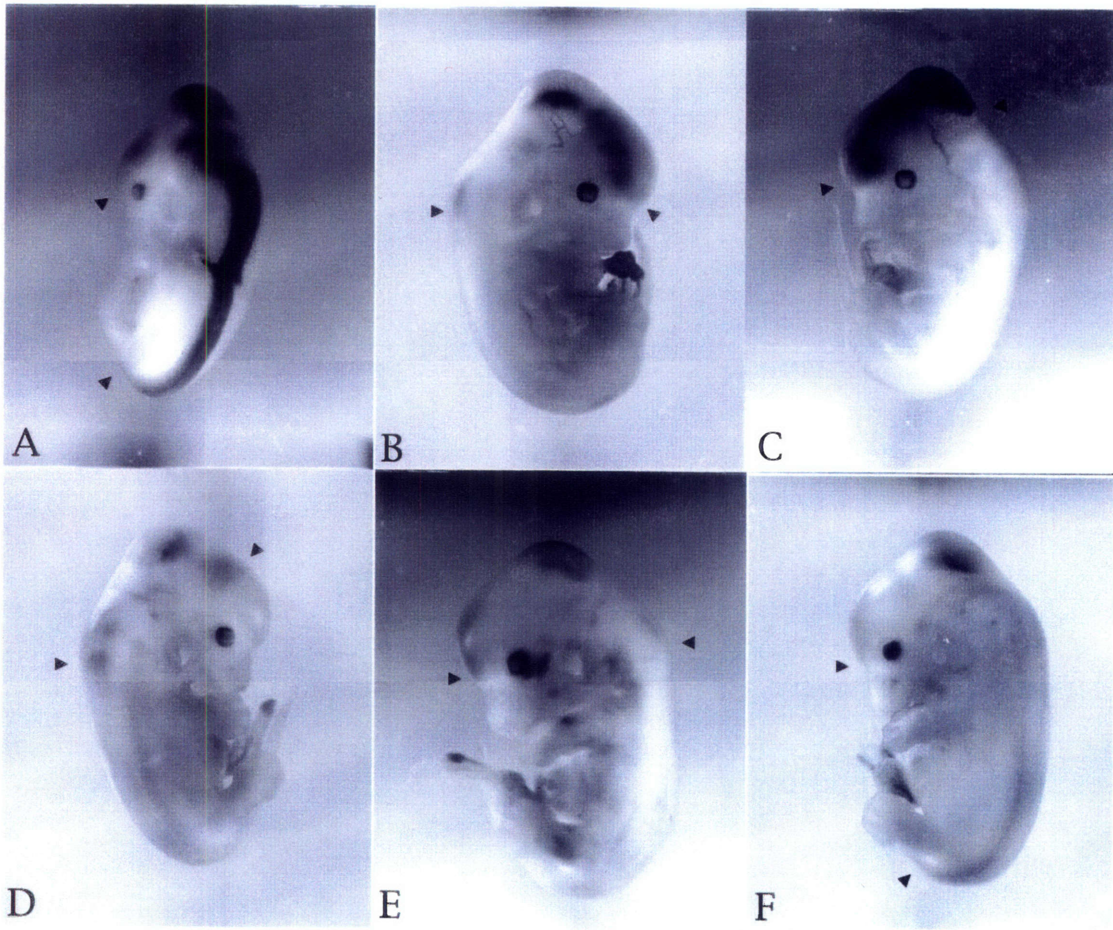

A
B
C
D
E
F

Figure 5. A 373 bp region highly conserved between the rat and human nestin second introns directs expression to the CNS

Three whole-mount E13.5 embryos made with the plasmid gIIConsTKlacZ (G in figure 1). βgalactosidase staining is seen only in the CNS of each embryo. About half of the stained embryos display transverse stripes of staining in the spinal cord (B) or amorphous patches of staining in the cortical hemispheres (C).

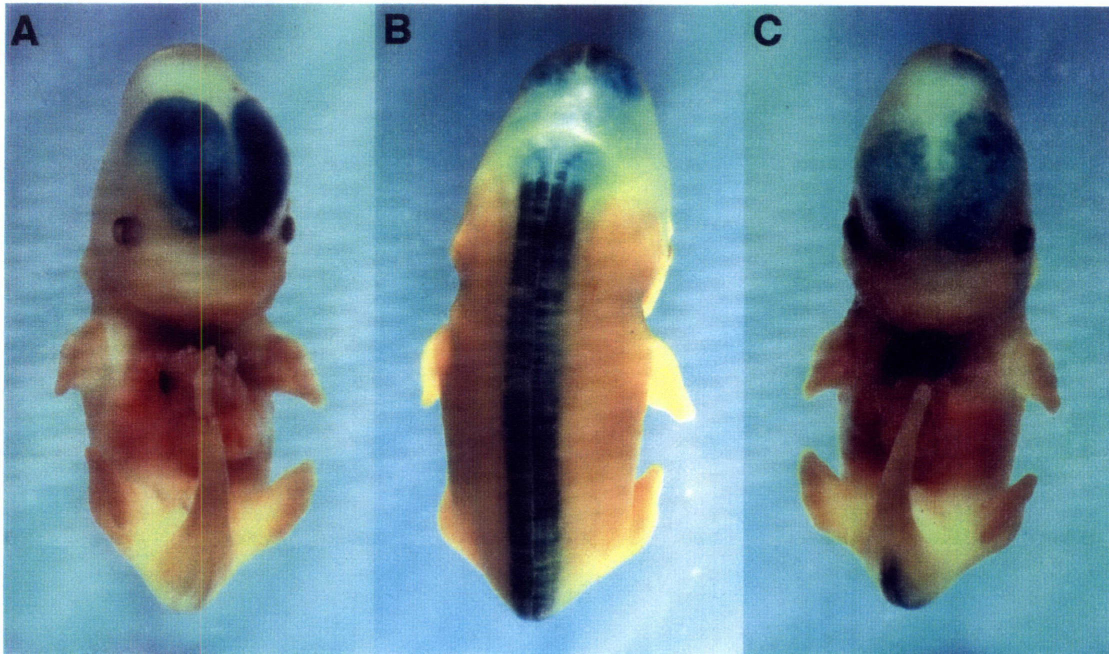
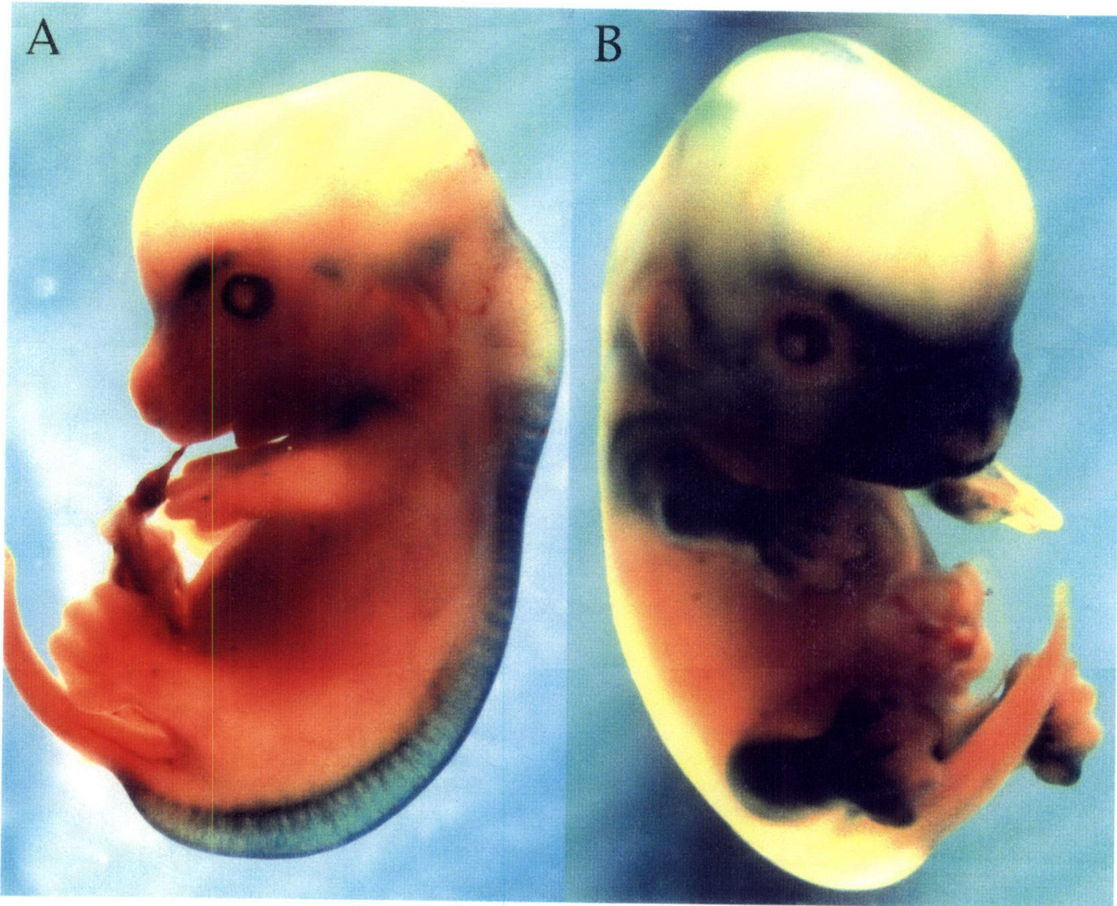

A
B
C

Figure 6. A 5' portion of the conserved region is insufficient for CNS-specific expression

Two whole-mount E14.5 embryos produced with the plasmid gIICons5'TKlacZ, containing the upstream 253 bp of the conserved region (H in figure 1). Staining of portions of the CNS is sometimes seen, as in the two embryos shown here. The embryo in (A) has strong reporter expression in the spinal cord and hindbrain, but none in the brain. The embryo in (B) has faint expression in the midbrain and spinal cord, along with very strong expression in the face and limbs. However, each of four βGal expressing transgenic embryos has a different pattern of ectopic expression.



A
B

Figure 7. The 3' portion of the conserved region drives reporter expression specifically to the central nervous system

Whole-mount E13.5 embryos bearing the plasmid 257gIITKlacZ (construct I in figure 1). Staining is seen only in the CNS of each mouse. In one embryo, expression is limited to a number of transverse stripes in the spinal cord (C) and small patches in the cerebral hemispheres (not shown).

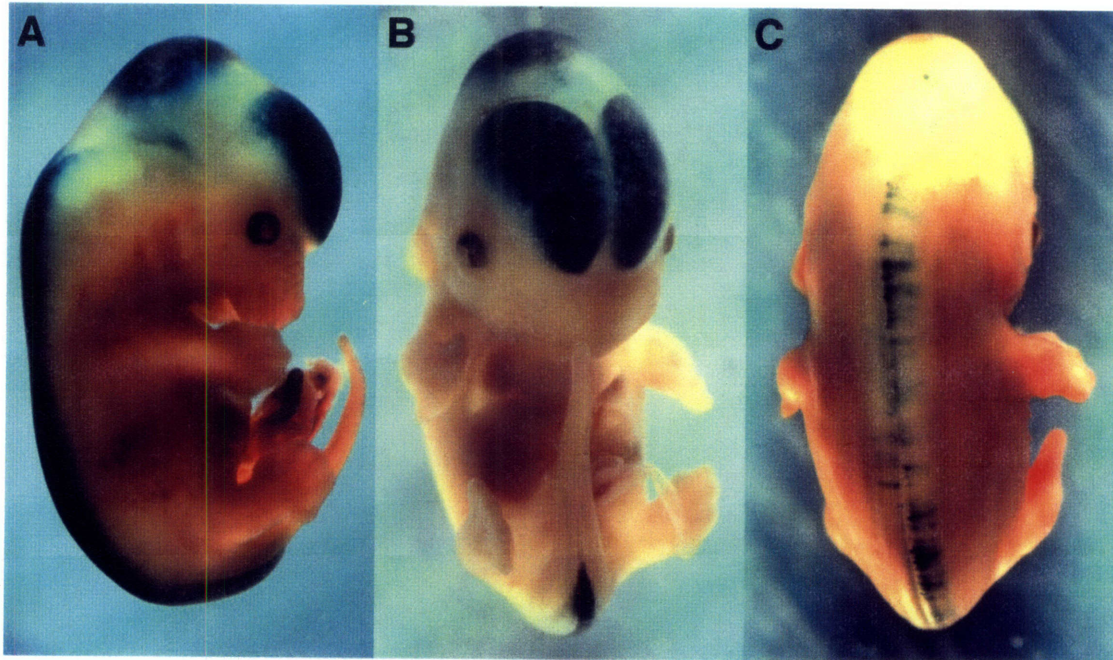
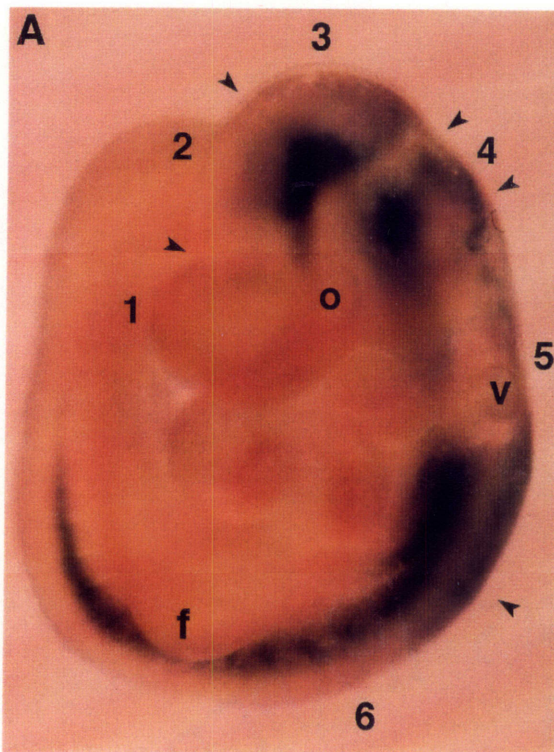

A
B
C

Figure 8. The 257 bp CNS-specific element directs correct onset of expression in the neural tube

Whole-mount E9.5 transgenic embryos made with 257gIITKlacZ (A) or pNesPlacZ+3 introns (B). Expression in the early CNS is similar whether driven by 257 bp of the second intron (A) or 11.1 kb of nestin genomic sequences containing all three introns (B). The embryo in (B) also has staining in the somites, due to the presence of the first intron muscle-specific element. CNS staining is seen only from the optic vesicle to just posterior of the forelimb bud in either embryo. In both cases, CNS expression appears first in the ventral neural tube and discrete dorsal regions of the hindbrain. Expression from the 257 bp element is noticeably weaker in the ventral myelencephalon and telencephalon than in other regions.

f, forelimb bud; o, optic vesicle; v, otic vesicle; 1, telencephalon; 2, diencephalon; 3, mesencephalon; 4, metencephalon; 5, myelencephalon.
(B) reproduced by permission from [Zimmerman *et al.*, 1994].

257gIITKlacZ



pNesPlacZ+3 introns

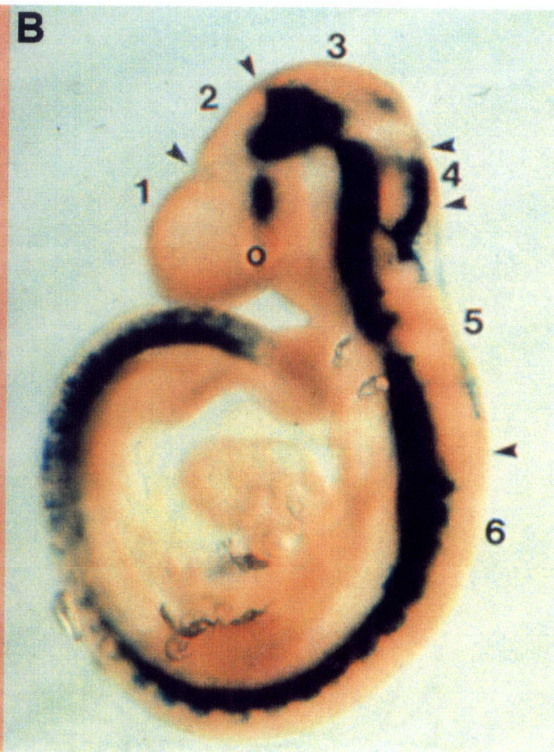
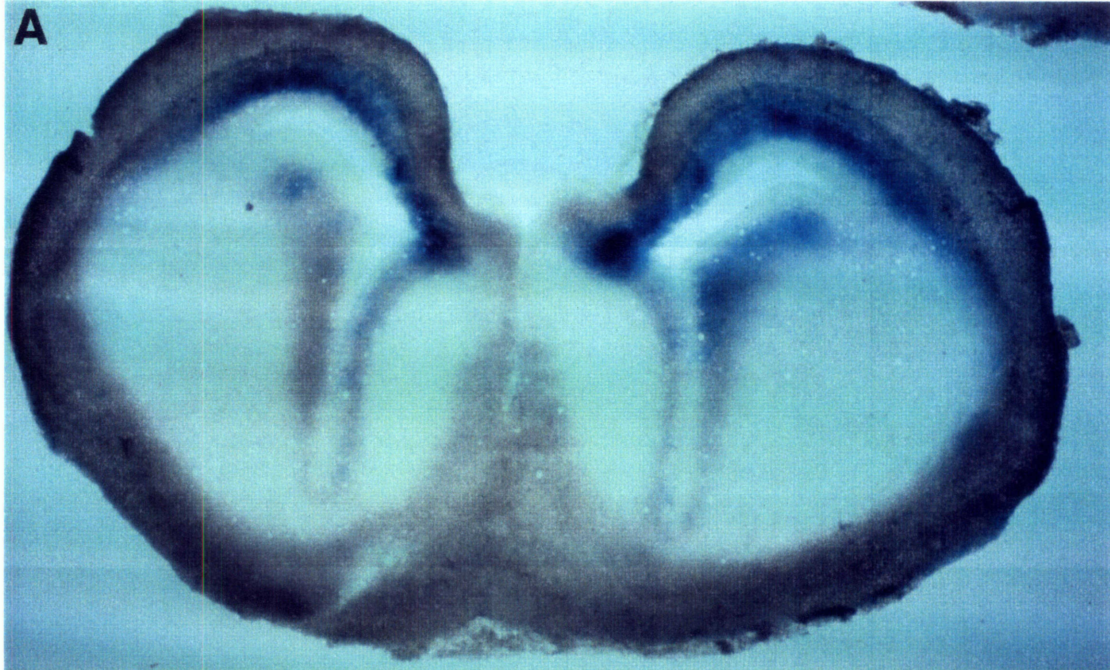
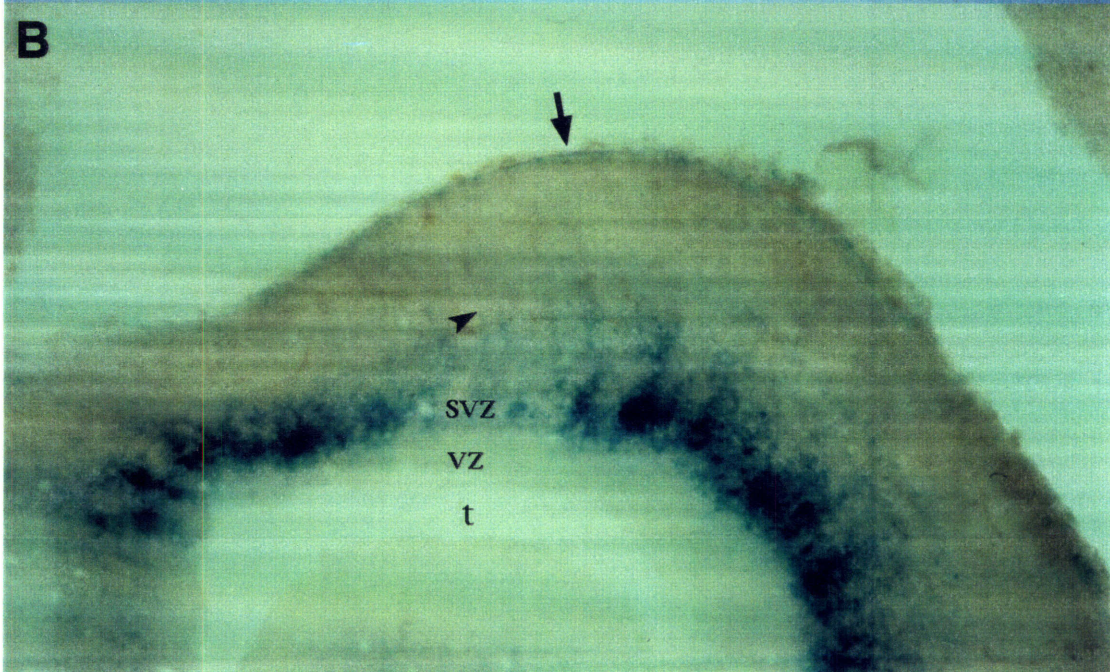


Figure 9. The 257 bp CNS-specific element directs expression to zones of neuronal precursor proliferation

Coronal sections through forebrain of an E16.5 transient transgenic embryo made with the plasmid 257gIITKlacZ. Strong lacZ staining is seen surrounding the telencephalic ventricles in (A). An adjacent section at higher magnification shows that the strongest expression is in the subventricular zone, the proliferating region of cortex at this age. The arrowhead points to fine radial processes which stain for βgalactosidase, and the arrow indicates a band of staining along the pial surface. These are possibly the elongated cell bodies and radial endfeet, respectively, of radial glia.

t, telencephalic vesicle; vz, ventricular zone; svz, subventricular zone


A


B
svz
vz
t

Discussion

The transgenic mouse assays presented here demonstrate that a 257 bp fragment of the nestin second intron contains the elements necessary for CNS-specific expression of the lacZ reporter gene. This expression is strong in most embryos and stains all regions of the brain and spinal cord. Analysis of transgenic mice at E9.5, during the period of neural tube closure, shows that the reporter gene follows the previously observed sequence of nestin activation. LacZ appears mainly in the ventral neural tube at E9.5, and dorsally only in the metencephalon and rostral spinal cord. Likewise, sections of older transgenic embryos at E16.5 show that the reporter gene is extinguished in newly postmitotic neurons, as has been shown for the nestin gene [Dahlstrand *et al.*, 1995]. Thus this 257 bp region appears to contain the appropriate *cis*-acting sequences for the onset and extinction of nestin expression as well as the tissue-specific elements.

Although there is no loss of tissue-specific or developmental regulation in narrowing down these elements from the 1832 bp second intron region to 257 bp, some possible changes in activity are seen. Only half of the lacZ-positive embryos generated with the 257 bp or 373 bp constructs showed continuous reporter expression through the CNS. The other half showed discrete patches of staining in the brain or transverse stripes of staining in the spinal cord. These patches and stripes are possibly related to the chimerism of the transient transgenic mouse embryos; DNA injected into the pronucleus of a one cell embryo is thought to integrate into the genome only after several cell divisions. Even in those animals where lacZ expression seems continuous, some cells of the CNS are probably derived from non-transgenic blastomeres. The diffusion of the blue lacZ reaction product usually hides this fact. In the mice with patchy staining patterns shown here, it may be that most of the embryo is derived from non-transgenic cells, so the patchiness is caused by a lower rate of transgene insertion into blastomeres. On the other hand, the patchiness could also be caused by the failure of many transgenic cells to express the reporter, which would suggest that the

smaller fragments of the nestin enhancer are less able to drive expression.

An interesting aspect of these patches and stripes is that they may mark clones of cells. It has been shown in the chick spinal cord that there is little lateral movement of cells in the early ventricular zone and that early neuronal clones form tight transverse bands [Leber and Sanes, 1995]. This could be why expression in the spinal cord of some transgenic mice shows transverse stripes.

It should be noted that the 257 bp tissue-specific enhancer region was located with the help of sequence conservation. The rat and human nestin second introns share an approximately 375 bp region of 78% identity. Using PCR to isolate this fragment, we simultaneously isolated the CNS-specific control region. This shortcut may prove valuable to others seeking to locate transcriptional regulatory elements and who have cloned genes available from more than one species.

Finally, according to the definition quoted earlier, enhancers are "modular arrays of elements" and the removal of one or a few elements is not usually catastrophic to the function of the enhancer [Hyman *et al.*, 1988]. That seems to be the case in the deletion of the 70 bp EcoRI fragment from the 847 bp NsiSalgIITKlacZ vector. This fragment is part of the highly conserved 257 bp region, but its deletion does not change the tissue specificity of expression. Rather, in most animals this deletion prevents lacZ expression from filling the entire CNS and in some cases limits it to the brain. Any regulatory elements contained within the 70 base pair deleted region are not required for CNS specificity, but have effects on the strength or stochastic likelihood of expression, especially in the posterior CNS.

The same 70 base pair region is part of the 257 bp fragment which is sufficient for expression throughout the early CNS. However, it is also part of the construct gIICons5'tklacZ (H in figure 1), in which the upstream end of the highly conserved region was shown to produce various patterns of ectopic expression (figure 6). Half of the lacZ-expressing mice made with this construct do have expression in the spinal cord, or midbrain and spinal cord, but all have different strong areas of expression outside the CNS. Clearly, the region of

overlap between this construct and the 257 bp enhancer, which includes the 70 bp region, is not sufficient for CNS-specific expression. However, it may contain *cis*-elements capable of increasing transcription in the posterior CNS. On the other hand, that portion of the 257 bp enhancer which lies outside the 70 bp deletion may be sufficient to drive CNS-specific expression.

In summary, the nestin neuroepithelial regulatory elements have been located within a 257 bp fragment of the second intron. This piece is small enough for *in vitro* analysis of protein-DNA interactions. In the next chapter we will use DNAse I footprinting to locate potential *cis*-acting sequences within this 257 bp. Then, gel shift and methylation interference assays will be performed to examine protein binding sites in finer detail, hoping that we will be able to recognize known transcription factor binding motifs that will point us toward the *trans*-acting factors which drive nestin expression in the CNS.

References

Altman, J. (1972) Postnatal development of the cerebellar cortex in the rat. III. Maturation of the components of the granular layer. J. Comp. Neurol. **145**, 465-514.

Chung, J.H. (1993) A 5' element of the chicken β-globin domain serves as an insulator in human ertyhroid cells and protects against position effect in drosophila. Cell **74**, 505-514.

Dahlstrand, J., M. Lardelli and U. Lendahl (1995) Nestin mRNA expression correlates with the central nervous system progenitor cell state in many, but not all, regions of developing central nervous system. Dev. Brain Res. **84**, 109-129.

Hockfield, S. and R. McKay (1985) Identification of major cell classes in the developing mammalian nervous system. J. Neurosci. **5**, 3310-3328.

Hogan, B., F. Constantini and E. Lacey (1986) Manipulating the Mouse Embryo: A Laboratory Manual Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Hyman, S.E., M. Comb, Y.S. Lin, J. Pearlberg, M.R. Green and H.M. Goodman (1988) A common trans-acting factor is involved in transcriptional regulation of neurotransmitter genes by cyclic AMP. Mol. Cell. Biol. **8**, 4225-33.

Leber, S.M. and J.R. Sanes (1995) Migratory paths of neurons and glia in the embryonic chick spinal cord. J. Neurosci. **15**, 1236-1248.

Maniatis, T., E.F. Fritsch and J. Sambrook (1989) Molecular Cloning: a Laboratory Manual Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Nornes, H.O. and G.D. Das (1974) Temporal pattern of neurogenesis in spinal cord of rat. I. An autoradiographic study-Time and sites of origin and migration and settling patterns of neuroblasts. Brain Res. **73**, 121-138.

Zimmerman, L. (1992) Structure and Regulation of the Nestin Gene Doctor of Philosophy Thesis, Massachusetts Institute of Technology.

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

Chapter III: Identification of *cis*-acting elements by *in vitro* binding assays

Introduction

In the previous chapter, the region of the nestin gene driving transcription in the CNS was pared down to 257 basepairs. It would have been possible to continue testing fragments of this region in transgenic mice, and then look for factors which bind those fragments whose deletion causes a loss of enhancer function. Instead, we turned entirely to *in vitro* assays at this point, partly to avoid the effort and expense of making so many transgenic mice. This chapter presents the results of *in vitro* DNA binding studies, designed to determine how many proteins can bind to the 257 bp element, and what the core binding motifs are. Comparison of the bound sequences to consensus binding sites for known transcription factor families will allow us to guess the identities of the proteins which regulate the nestin gene.

The first step was to roughly locate the binding sites and see how many there are. This was done by footprinting *in vitro* with deoxyribonuclease I (DNAse I). DNAse I footprinting involves mixing an end-labeled DNA probe, in this case encompassing the 257 bp region, with proteins from a nuclear extract. We made nuclear extracts from various parts of the rat brain at several embryonic and early postnatal ages. Regions of the DNA probe which are bound by proteins *in vitro* are protected from subsequent random digestion with DNAse I. This region of protection is visualized on a denaturing gel as a set of bands which are weaker in the presence of protein extract. This assay allows the binding sites of several proteins to be detected at once over a large region of DNA.

The next step was to break up the 257 bp region into fragments containing only a single binding site. We therefore synthesized short double-stranded probes containing only one DNAse I footprint. These were again mixed *in vitro* with nuclear extract protein, but instead of being digested with DNAse, the probe/protein mixture was analyzed on a non denaturing polyacrylamide gel. Negatively charged DNA usually moves quickly toward the positive

pole at the bottom of the gel, but those probe molecules which are bound by protein are slowed by the protein's bulk and different charge density and migrate more slowly through a gel. This is called an electrophoretic mobility shift assay (EMSA), or "gel shift" in the vernacular. The EMSA is a more sensitive assay for DNA/protein interaction than the DNAse I footprinting, because in the former the bound protein must nearly saturate the probe in order to be detected. This is not easy in extracts from tissues containing many cell types. The gel shift assay can also show that multiple proteins are binding a probe. However, this assay can also detect proteins which bind DNA in a non-sequence-dependent manner, so it is always necessary to test the specificity of the interactions detected.

The third *in vitro* assay in these studies uses a variation of the gel shift. Methylation interference assays are a means of determining which bases in a binding site most closely interact with the binding protein. This is done by methylating the binding site to sterically inhibit the interaction. A short end-labeled probe is first methylated randomly at one or a few guanine or adenine bases per probe molecule, then mixed with protein and gel shifted as above. The retarded band (bound probe) and normally migrating band (free probe) are cut out of the EMSA gel, and cleaved at the methylated bases with piperidine. Comparison of the cleavage patterns of bound and free probes by denaturing PAGE shows that cleavage at some bases never occurs in the bound fraction. This is because methylation of these bases inhibits protein binding. We conclude that methylation blocks the especially close interaction of protein with these bases, which are part of the core binding motif.

By these means we have located five DNAse I footprints in the 257 bp nestin tissue-specific element. Each of these five protein binding sites has been further analyzed by gel shifts and/or methylation interference to identify a potential core motif. It is our intention to use these core sequence motifs to identify the transcription factors which bind this region and regulate the transcription of the nestin gene in the CNS.

Materials and Methods

Nuclear Extracts:

Nuclear extracts were prepared by the method of [Dignam *et al.*, 1983]. Timed pregnant Sprague-Dawley rats (Taconic) were sacrificed by CO_2 asphyxiation. Embryos were dissected in Hank's Balanced Salt Solution (Gibco BRL) with 3.7 g/L $NaHCO_3$ and 3.9 g/L HEPES [pH 7.2]. Collected tissues were pelleted at 1000 rpm in a Sorvall RT-6000 centrifuge and resuspended in five volumes buffer A (10 mM HEPES [pH 7.9], 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride [PMSF], 10 µg/mL leupeptin, 10 µg/mL aprotinin, and 2 µg/mL pepstatin.) This was incubated on ice 10 minutes during hypotonic lysis, then triturated a few times through a 200 µL pipet tip and transferred to a 1.5 mL eppendorf tube. This was spun 3 minutes at 2000 rpm in a microcentrifuge, and the pellet was resuspended in 2 volumes buffer A. The nuclei were released by 16 strokes with a teflon microfuge tube pestle (BelArt), then pelleted for 20 minutes at 25000 g in a Sorvall F-20 Micro rotor. The pellet was resuspended in 2 volumes buffer C (20 mM HEPES [pH 7.9], 420 mM NaCl, 25% glycerol [vol/vol], 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 10 µg/mL leupeptin, 10 µg/mL aprotinin, and 2 µg/mL pepstatin) and homogenized with 16 strokes of a pestle as above. The contents were mixed 30 minutes at 4°C on a rocker platform, then spun 30 minutes at 25000g in the Sorvall F-20 Micro rotor. The supernatant was transferred to Spectrapor dialysis tubing (MWCO 12000-14000) and dialyzed 5 hrs at 4°C versus 100 mL of buffer D (20 mM HEPES [pH 7.9], 100 mM KCl, 20% glycerol [vol/vol], 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF.) After dialysis, the material was spun 20 minutes at 25000g as above, then aliquoted and frozen in liquid nitrogen before storage at -80°C. Protein concentration was determined using a Pierce BCA protein assay kit and BSA standards.

DNAse I footprinting:

The plasmid pCRII-5669, containing 256 bp of the nestin second intron, was digested with HindIII or EcoRV and labeled with [α ^{32}P]-dCTP and Klenow fragment. Unincorporated label was removed via a sephadex G-50 spin column (5 Prime→3 Prime.) A second digestion with EcoRV or HindIII released the single end-labeled probe, which was separated from the vector on low-melting agarose and purified

on an Elutip column (Schleicher & Schuell) before being ethanol precipitated and resuspended in TE [pH 8.0].
10000 cpm of probe were mixed with 14 μgs of tissue nuclear extract or 15 μgs of bovine serum albumin (BSA) plus 9 μgs of poly (dI-dC) in footprint buffer (12 mM Tris-HCl [pH 8.0], 50 mM KCl, 1mM MgCl2, 1 mM CaCl2, 5 mM NaCl, 1 mM dithiothreitol [DTT], 5% glycerol [vol/vol].) This reaction was incubated 30 minutes at room temperature. One volume (50 μL) of footprint buffer was added just before the addition of DNAse I (RQ1 DNAse, Promega). BSA controls received 0-0.5 μgs DNAse I, while nuclear extract samples received 0.4-0.8 μgs. Digestion proceeded at room temperature for exactly two minutes before being stopped with an equal volume of stop buffer (50mM EDTA, 0.2% SDS, 100μg/mL yeast tRNA.) The reaction was extracted once in phenol and ethanol precipitated. The DNA was resuspended in formamide loading dyes and heated 5 minutes at 90° C before analysis on a 6% acrylamide, 7 M urea sequencing gel.

Double stranded oligonucleotide probes:

Synthetic oligonucleotides were denatured and separated by polyacrylamide gel electrophoresis (PAGE) on 8% polyacrylamide/ 7 M urea gels. Full length oligonucleotides were eluted from crushed gel slices in O.1% sodium dodecyl sulfate (SDS), O.5 M ammonium acetate, 10 mM magnesium acetate, and purified via a Sep-Pak C18 reversed-phase column (Waters). Complementary oligonucleotides were annealed at 10 picomoles/μL in 100 mM NaCl, 10 mM Tris-HCl [pH 8.0], 1 mM EDTA while slowly cooling from 80° C. The double stranded probes were labelled at a single end with [α ^{32}P]-dCTP and Klenow fragment, adding 200mM dATP and dGTP to fill in the HindIII overhanging end or 200mM dTTP to label the SalI overhang. Unincorporated nucleotides were removed by a G-25 sephadex spin column (Select-D, 5 Prime→3 Prime.) All probes were subsequently checked by denaturing PAGE to ensure that a single labeled species was present.

Electrophoretic Mobility Shift Assay:

Protein binding reactions were performed by combining10,000 cpm of double-stranded oligonucleotide probe and 4 μgs of tissue nuclear extract with 2.2 μgs poly(dI-dC) and 6.0 μgs BSA in binding buffer (12% glycerol [vol/vol], 12 mM N-2-hydroxyethylpiperazine-N'-2-

ethanesulfonic acid [HEPES] [pH 7.9], 4 mM Tris-HCl [pH 7.9], 60 mM KCl, 1 mM EDTA, 1mM dithiothreitol [DTT].) In competition assays, a 100- or 500-fold molar excess of unlabeled probe or unlabeled nonspecific competitor was added. The binding reaction occurred at room temperature (23° C) for 30 minutes or 1 hour. Binding reactions were electrophoresed at 200 volts on a 4% 80:1 acrylamide/bis-acrylamide/2.5% glycerol gel. The gels were dried 1 hour at 80° C and visualized on a Molecular Dynamics PhosphorImager.

Methylation Interference:

1-2$\times 10^6$ cpm of probe was methylated with 1 µL of 99+% dimethyl sulfate (DMS) in 200 µL of 50 mM sodium cacodylate [pH 8.0] , 1 mM EDTA. After 5 minutes at room temperature the reaction was stopped with one-fifth volume of 1.5 M sodium acetate, 1.0 M β-mercaptoethanol. Magnesium chloride was added to 10 mM, along with 20 µg of yeast tRNA before the probe was ethanol precipitated three times. The methylated probe was resuspended in 10 mM Tris-HCl [pH 8.0], 1 mM EDTA (TE), at about 10000 cpm/µL.
Protein binding reactions were performed by combining100,000 cpm of methylated probe and 40 µgs of tissue nuclear extract in binding buffer (12% glycerol [vol/vol], 12 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid [HEPES] [pH 7.9], 4 mM Tris-HCl [pH 7.9], 60 mM KCl, 1 mM EDTA, 1mM dithiothreitol [DTT]) for 1 hour at room temperature (23° C). Binding reactions were electrophoresed at 200 volts on a 4% 80:1 acrylamide/bis-acrylamide gel. The wet gel was exposed to a phosphor screen (Molecular Dynamics) to locate and cut out the bound and free probe fractions. These gel slices were covered in 1% agarose, 0.5x Tris-Borate-EDTA buffer (TBE), and the DNA electrophoresed into NA45 paper (Schleicher and Schuell.) Probe was eluted from the paper at 65° C in 50 mM Tris-HCl [pH 8.0], 1.0 M NaCl, 10 mM EDTA. The elution was extracted once with phenol and ethanol precipitated. The DNA was resuspended in 1.0 M piperidine and cleaved at 90° C for 30 minutes. The piperidine was removed by three cycles of lyophilization, then the probe was resuspended in formamide loading dyes (98% formamide, 10 mM EDTA, .025% xylene cyanol, .025% bromophenol blue) at about 200 cpm/µL. Equal activities of bound and free probes were analyzed by denaturing PAGE on a 12% acrylamide, 7 M urea sequencing gel. Labelled probe which had been subjected to the Maxam and Gilbert G+A reaction was run alongside as a positional marker. The dried gel was

visualized on X-OMAT AR film (Kodak) or a Molecular Dynamics PhosphorImager screen.

Oligonucleotides used in Methylation Interference and Electrophoretic Mobility Shift Assays:

Oligonucleotide probes were designed as complementary pairs with 4 bp 5' overhangs at each end. The forward sequence of each pair is listed below; the reverse sequence is complementary to this. In addition, a HindIII overhang (AGCT) was placed at the 5' end of each forward oligo, and a Sal I overhang (TCGA) at the 5' end of each reverse oligo.

Ftpt 1: 5'-GGAGAAGGGGAGCTGAATTCATTTGCTTTTGTCTGTTACCAG-3'
Ftpt 2: 5'-GCAGAGAGAGAGCCATCCCCTGGGAACAGCCTGAGAATTC-3'
Ftpt 3: 5'-CAGCCTGAGAATTCCCACTTCCCCTGAGGAGCCCTCCCTTCT-3'
Ftpt 4: 5'-GTGTGGACAAAAGGCAATAATTAGCATGAGAATCGGCCTC-3'
Ftpt 5: 5'-CCCTCCCAGAGGATGAGGTCATCGGCCTTGGCCTTGGGTGGG-3'

TREp is an artificial palindromic thyroid hormone response element from [Marks *et al.*, 1992].
TREp forward: 5'-GCCAGATTCAGGTCATGACCTGAGGAGA-3'
TREp reverse: 5'-CTCTCCTCAGGTCATGACCTGAATCT-3'

Results

DNAse I footprinting reveals five sites of protein interaction with the 257 bp nestin neural enhancer region

Figure 10 shows the results of DNAse I footprinting assays on a double-stranded DNA probe containing the 257 bp neural element defined in the previous chapter. Nuclear extracts [Dignam *et al.*, 1983] in these experiments were prepared from midbrain, cerebellum, spinal cord, and cerebral cortex or telencephalon of rats at embryonic days 13, 16, 17 and 20 or postnatal day 5 (P5). Footprinting with brain tissue extracts is challenging because of the variety of cell types present in each tissue. Complete protection of binding sites from DNAse I digestion requires saturation of the probe with binding protein, which can be difficult to achieve with purified protein. However, the protections obtained with brain extracts are strong enough to map footprinted regions and roughly estimate the tissue and age distribution of binding activities.

Labeling the forward strand of the DNA probe leads to the footprints shown in figures 10a and b. A total of five regions are protected from DNAseI digestion in various extracts, but all five protections are not seen in any single extract. These five regions are diagrammed at the right of the figures with their approximate endpoints marked. The base number listed for each endpoint refers to its position in a nestin genomic clone, not its position relative to the transcriptional start site (at 2473 in the genomic clone). Positions of enhanced DNAseI digestion in the presence of protein, called hypersensitive sites, are found adjacent to the protected regions (footprints) numbers 2 and 5.

The same five footprints are found again when the DNA probe is labeled on the reverse strand (figure 10c and d). The endpoints of the protected regions are generally not the same as on the forward strand, but the protected regions roughly correspond on the two strands (see figure 11). Each protected region on the forward strand overlaps one on the reverse strand, so the five pairs of protections on opposite strands have been named footprints 1-5.

None of the five footprints is detectable in every age and tissue from which nuclear extracts were prepared. In these and other experiments not shown, footprints 1 and 4 were most generally occupied. These two sites were protected by most extracts from the four brain regions taken at ages from E13 to P5. Footprint number 5 was almost as widespread, but was missing from some midbrain and spinal cord extracts which protected footprints 1 and 4. By contrast, footprints 2 and 3 are detectable in only a few tissues. Footprint 2 is found only in the midbrain, spinal cord, and midbrain at E16, and never in cortex. It is also found in these tissues, but not cortex, at P5 (data not shown.) Footprint 3 seems to be the most restricted of the five. It is detectable only in cerebellum and perhaps spinal cord at E16 (figure 10a and c) and in P5 cerebellum (data not shown.) It appears that the transcriptional control element driving nestin expression in all regions of the CNS may interact with some transcription factors which are active only in limited areas of it.

Electrophoretic Mobility Shift Assays show the numbers and tissue distributions of factors which bind the nestin neural enhancer

DNAse I footprinting indicated five areas of the 257 bp CNS-specific element which are bound by nuclear proteins, but suggested that these factors may not be active in all regions of the brain. The EMSA or gel shift assay was used to test this idea. Gel shifts are a much more sensitive measure of DNA binding activity than footprinting assays, because the latter requires near-saturation of the DNA binding site by protein. The gel shift assay, in contrast, can detect protein binding on just a few molecules of labeled probe. In addition, proteins in the EMSA may retard migration of the DNA probe to varying degrees depending on their bulk and charge density. This generally produces separate bands for each binding protein and each multiprotein complex, allowing an estimate to be made of the number of proteins capable of binding a given probe.

Gel shifts on a double-stranded oligonucleotide containing 42 bp of nestin sequence surrounding footprinted region 1 (probe Ftpt1, see figure 17 for probe sequences) confirms that binding

activity for this probe is widespread in the developing CNS (figure 12). Two prominent bands are seen in every CNS tissue extract tested, whether from embryonic (E13, E16) or postnatal (P5) rats. Three weaker and faster migrating bands are present in E13 telencephalon, and in a cortical extract from E17, but are absent at P5. Interestingly, the a similar EMSA with the footprint 4 region (Ftpt4) yields almost identical results (data not shown). The similar binding patterns for these two probes may be explained by sequence homology (see figure 17). Both DNAse I protected regions include 8 basepairs matching (at 7 of 8 bases) the "octamer" consensus binding site for the Pit1-Oct-Unc86 (POU) family of transcription factors. In chapter IV it will be determined whether POU proteins can bind to footprinted regions 1 and 4 of the nestin second intron.

The last three lanes of figure 12 demonstrate that the DNA binding activities detected in the Ftpt1 EMSA recognize this probe in a sequence-specific manner. The five strongly shifted bands produced on this probe by E17 cortical extract are displaced by a 100-fold excess of the same unlabeled probe. However, a 100-fold excess of a probe unrelated in sequence does not compete for binding of these proteins. Therefore these binding activities do not recognize all DNA, but require some particular sequence present in the Ftpt1 probe.

Gel shift assays on the Ftpt5 probe show that binding activity specific to footprint 5 is found in most extracts. In figure 13, two prominent bands are present in the cortex at E17 or P5 (arrowheads), and to a much smaller degree in the E16 cortical extract. Only the faster-migrating of these two complexes is present in E16 or P5 midbrain, cerebellum and spinal cord. A much weaker band of slightly higher mobility is present in all but the cortical extracts. Both bands of the E17 cortical doublet bind specifically to the footprint 5 probe and are not competed away by heterologous DNA. It therefore appears that footprint 5 is strongly bound by one protein in most regions of the brain, but an additional binding activity is found only in the cerebral cortex.

In contrast, binding activity for the footprint 2 and 3 probes is much stronger in CNS regions other than the cortex. Figure 14 shows

gel shifts on the Ftpt2 probe, which bears 40 bp of nestin sequence including footprint number 2. In E16 midbrain, cerebellum, or spinal cord two strong bands are seen, the stronger of which is not present in E16 cortex. E16 cortex has instead a weak band which migrates slightly slower (grey arrowhead in figure 14). The quickly-migrating band detected in all extracts at E16 is absent from all but the cerebellum at P5; the slower-migrating band is strongest in cerebellum and absent from P5 cortex. In summary, the Ftpt2 probe is bound by: an activity specific to the caudal CNS, that is midbrain, cerebellum and spinal cord; another which is present in all four extracts at E16; and a third which may be specific to the E16 cortex.

Similarly, a probe containing the sequence of footprint 3 is much more strongly bound by extracts from cerebellum or spinal cord than by more rostral tissue extracts (see figure 15). Two strong complexes are seen in most extracts (arrowheads). The faster-running complex is present in all four brain regions at E16 but restricted to cerebellum at P5. The slower-running complex is again strongest in cerebellum and spinal cord, but does appear to be present in cortex at both ages. In between these two complexes are two sets of bands which appear to be cortex-specific. Therefore the distribution of footprint 3 binding activity is not strictly limited to cerebellum and spinal cord, as was suggested in DNAseI footprinting experiments. However, both footprint 2 and footprint 3 regions are more strongly bound in cerebellum and spinal cord.

Methylation Interference assays identify the core binding sequences of the proteins interacting with the nestin enhancer

The DNAse I footprinting and gel shift assays described above show that several proteins can bind *in vitro* to each of five footprinted regions of the nestin neuroepithelial enhancer. Efforts to identify these proteins begin with more precise description of the DNA sequences to which they bind. Comparison of these sequences with known *cis*-acting sequences may point to a few transcription factor families as possible regulators of nestin.

Sites on DNA which closely interact with proteins can be mapped by the methylation interference assay. This assay was performed on the same double-stranded oligonucleotide probes used in gel shift assays, following their modification with dimethylsulfate (DMS). Figure 16 shows the results of methylation interference assays on probes corresponding to each of the five footprinted regions of the nestin CNS enhancer. For reference, a typical EMSA performed with the same probe and nuclear extract is shown alongside the methylation interference results. Following a large-scale gel shift of methylated DNA, the probes are isolated and cleaved with piperidine prior to PAGE analysis. Comparison of the probes bound by protein (lanes marked Bound or Bd) with uncomplexed probe (lanes marked Free) indicates at which guanine or adenine residues methylation can sterically inhibit protein binding.

Protein interactions with the bottom strand of the footprint 1 region are shown in figure 16a. Nuclear proteins from the E17 rat cortex produce several complexes with this probe, as shown in a typical gel shift (inset). The lanes of the methylation interference gel correspond to the bands marked Bd1, Bd2, Bd3 and Free in the gel shift assay. Comparison of piperidine cleavage patterns between the three bound lanes and the free probe show that six adenines are relatively uncleaved in bound DNA (arrowheads, fig. 16a). Methylation of these same six adenines inhibits formation of all three bound complexes to approximately the same degree, even though the Bd3 complex can be seen in the gel shift lane to actually be composed of three discretely shifted bands. It appears that the binding site defined by methylation interference of the Ftpt1 probe is involved in formation of at least four different protein-DNA complexes. These complexes may represent different binding proteins, or a variety of multi-protein combinations, or could be produced by subtle differences in charge of a single binding protein.

The location of the methylation interference sites of these complexes relative to the footprinted area is summarized in figure 17a. The adenine/thymine-rich sequence which includes the six blocked adenines overlaps a seven of eight base-pair homology to

the well known octamer sequence 5'-ATTTGCAT-3', the consensus binding site of the Pit-1/Oct/Unc-86 (POU) transcription factor family [Wirth *et al.*, 1987] (see figure 17a).

Methylation interference on an oligonucleotide probe corresponding to footprint 4 tells a similar story (figures 16d and 17d). Methylation of any of three adenines on the bottom strand inhibits formation of four isolated protein-DNA complexes. As with footprint 1, the fastest-migrating group is actually three different shifted bands too close to be separated in the methylation interference assay. The three blocked adenines on the Ftpt4 probe are again part of an A/T-rich sequence overlapping a seven-of-eight base pair homology to the octamer sequence. Methylation interference was not performed on the top strand of this probe. The two octamer-like sequences differ somewhat: 5'-ATTTGCTT-3' in Ftpt1 versus 5'-ATTAGCAT-3' in Ftpt4. The similarity of the complexes seen in EMSAs and the methylation interference sites of these two probes suggest that the two sites are bound by the same protein or group of proteins. The presence of octamer-like motifs adds the possibility that footprints 1 and 4 are occupied by POU family members.

Methylation interference demonstrates strong similarities between footprints 2 and 3. In EMSAs, each oligonucleotide probe shows a major complex most abundant in spinal cord and cerebellum, and relatively lacking in cortex and midbrain. Methylation interference of this major complex with E16 spinal cord nuclear extract points out the similarity in the binding sites. Binding of the Ftpt2 probe is blocked by methylation of two adjacent guanines on the top strand, and three of four nearby guanines on the bottom strand (figures 16b and 17b). The sequence encompassing the five blocked guanines reads (along the top strand) 5'-TCCCCTGGGAA-3'. This is a close match to a sequence occurring within footprint 3, reading along the top strand: 5'-TCCCCTGAGGA-3'. Methylation interference was performed on the bottom strand of the Ftpt3 oligonucleotide probe, and revealed that this sequence is part of the protein binding site (figures 16c and 17c). Methylation at any of the four guanines complementary to the shared top strand sequence

inhibits protein binding to the Ftpt3 probe. In addition, three guanines further 3' along the bottom strand are also required for protein binding. The similar methylation interference sites on the two probes suggest that footprints 2 and 3 are recognized by the same or related proteins.

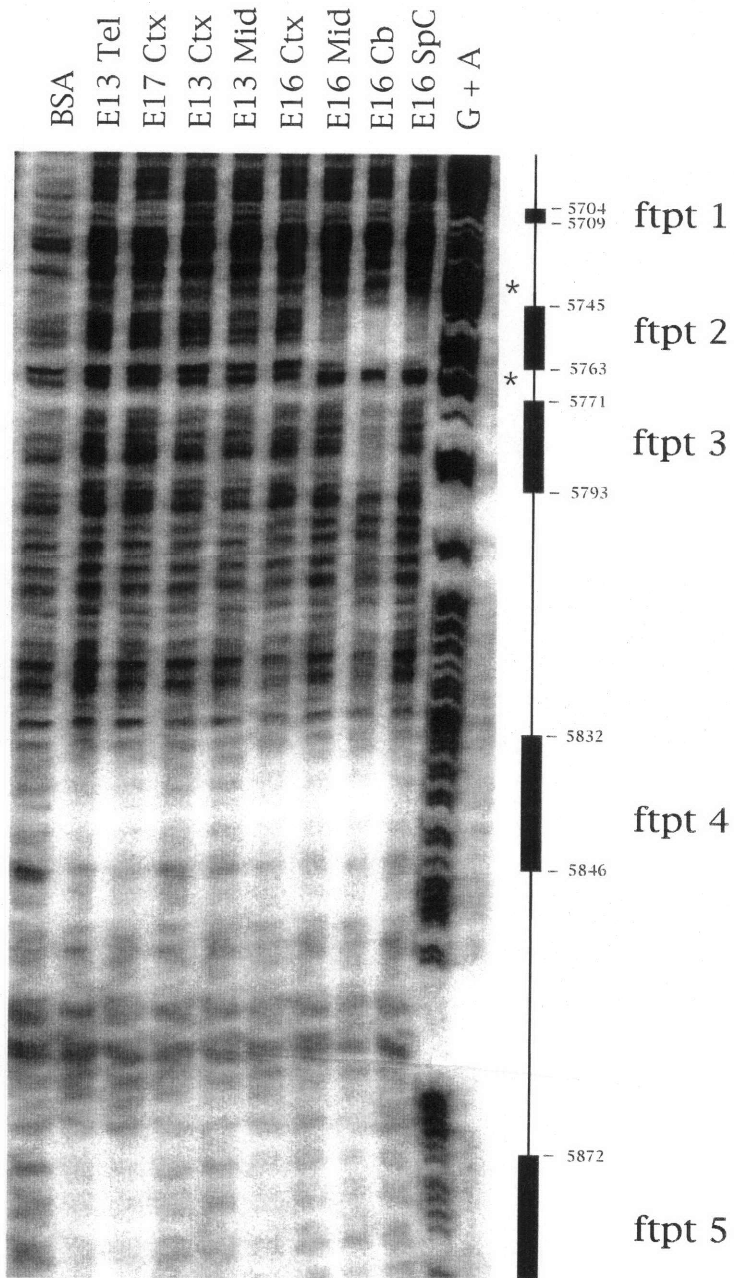
The conclusion that related proteins bind footprints 2 and 3 *in vitro* is further supported by gel shift assays using double-stranded oligonucleotide competitors. Figure 18 shows that binding of nuclear proteins from E16 rat spinal cord to a labeled Ftpt3 probe is completely inhibited by a five hundred-fold molar excess of unlabeled probe for either Ftpt3 or Ftpt2. A similar quantity of an unrelated probe, in this case an artificial palindromic thyroid hormone response element (TREp), does not compete away the binding protein. The same cross-competition is observed with a labeled Ftpt2 probe and excess unlabeled Ftpt3 probe. This competition assay strongly suggests that the two probes bind the same protein. It must be noted that the two oligonucleotide probes used here have fourteen base pairs of genomic sequence in common, in addition to homology between them described above. However, methylation interference is observed at only the last base of these fourteen (see figure 17c). It is more likely that the binding site of the related proteins is the region of sequence homology in the middle of each probe which is blocked by methylation. Notably, this sequence homology is similar to the consensus motif reported by [Imagawa *et al.*, 1987]as the binding site of the transcription factor AP-2.

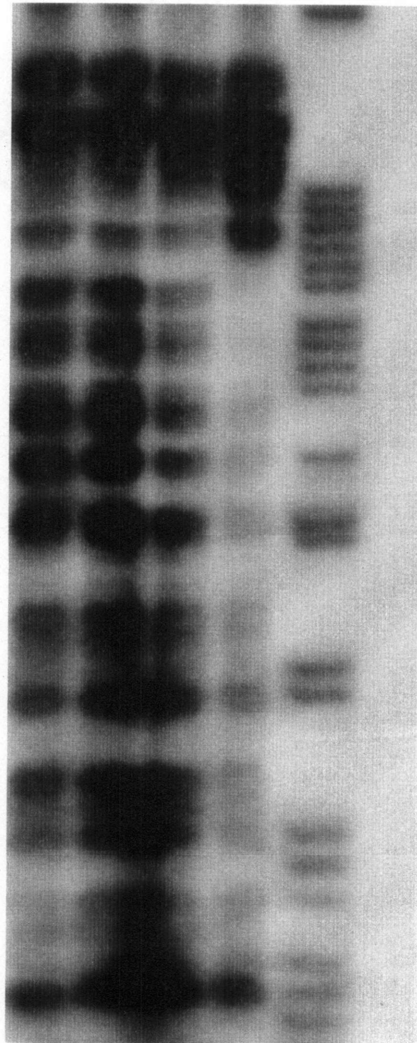
Finally, methylation interference of the Ftpt5 probe reveals a third type of binding site (figures 16e and 17e). In the gel shift assay, two strong protein-DNA complexes migrate very close together, and a third slower-migrating complex (Bd1) is sometimes detected. On the top strand, all three complexes are inhibited by methylation at either of two adjacent guanines. A nearby guanine is the only detectable site of interaction on the opposite strand with the two more abundant proteins (Bd2 and Bd3). The sequence comprising these three methylation interference sites on Ftpt5 has homologies to a number of transcription factor binding sites. The eight base pair sequence 5'-TGAGGTCA-3' is identical to a site in the

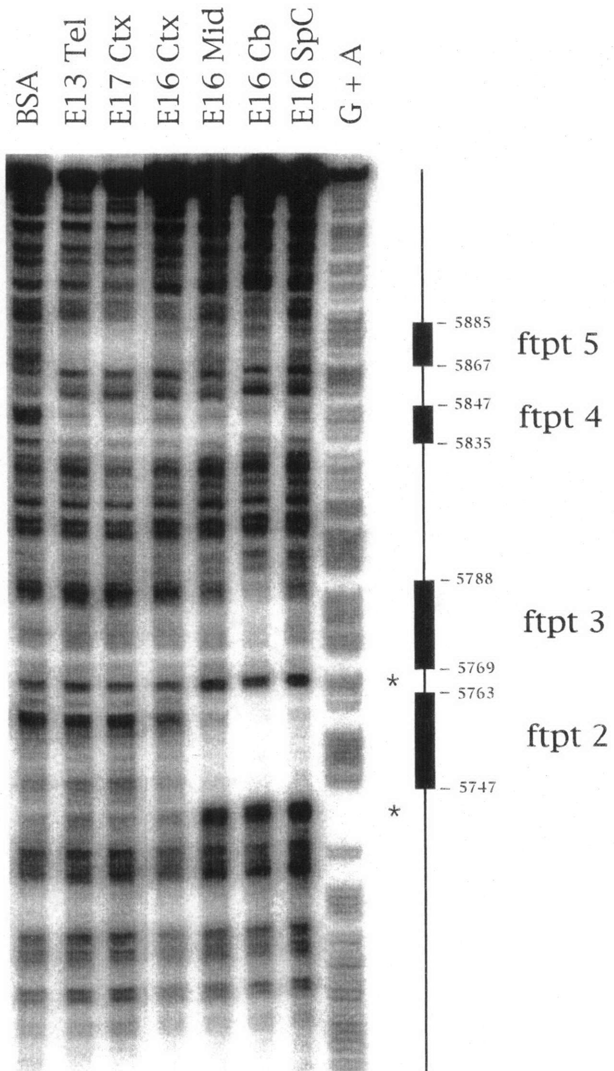
MHC class I promoter which binds a tri-iodothyronine receptor/9-*cis*-retinoic acid receptor beta heterodimer (T3R/RXRβ) [Marks *et al.*, 1992]. However, this site is an atypical one, as most T3R/RXRβ binding sites are palindromes or direct repeats spaced by 4 nucleotides [Umesono *et al.*, 1991]. The same eight bases are also similar to the consensus cAMP response element (CRE), the binding site of CRE-binding protein (CREB). Since CREB binds as a dimer to inverted repeat sequences, it may be relevant that the methylation interference sites of footprint 5 are at the center of a five base pair inverted repeat: 5'-GATGAGGTCATC-3'. Another trans-acting activity which binds similar inverted repeats is AP-1, a dimer formed by fos, jun, or related proteins. All of these proteins may be candidates for binding to this region of the nestin CNS enhancer.

Figure 10. DNAseI footprinting assays on the 257 bp nestin CNS-specific element

DNAse I footprinting on a HindIII-XbaI fragment of pCRII-257 which contains the entire 257 bp CNS-specific element of the rat nestin second intron. In each lane, end-labeled DNA was mixed with 14 μgs of rat brain nuclear extracts, or 14 μgs of bovine serum albumin prior to digestion with DNAse I for exactly 2 minutes. Tissues were collected on embryonic days E13, E16, or E17 or postnatal day P5.
(A) Footprinting on the forward strand of the DNA probe. Five regions of protection are marked by black boxes at the right of the figure. The five footprints are not all visible in any one extract. In this experiment, footprint 2 is apparent only in E16 midbrain, cerebellum, and spinal cord; footprint 3 is visible only in E16 cerebellum. Approximate endpoints of footprints are indicated with the corresponding base number in a nestin genomic clone. The transcriptional start site in this genomic clone is at 2473. Sites of greater digestion in the presence of some tissue extracts (hypersensitive sites) are marked with an asterisk.
(B) A portion of another footprinting experiment on the forward strand, showing footprint 5 more clearly.
(C) Footprinting on the reverse strand of the same DNA fragment. Four footprinted regions are detected and numbered 2-5 to correspond to the footprints on the forward strand. Footprints 2 and 3 are seen in the same tissues as on the forward strand.
(D) A portion of another footprinting experiment on the reversed probe shows that footprint 1 can also be detected on this strand.
BSA, bovine serum albumin; Tel, telencephalon nuclear extract; Ctx, cortex; Mid, midbrain; Cb, cerebellum; SpC, spinal cord; G+A, Maxam-Gilbert guanine+adenine sequencing reaction performed on the same end-labeled probe as the footprinting assays.


BSA
E13 Tel
E17 Ctx
E13 Ctx
E13 Mid
E16 Ctx
E16 Mid
E16 Cb
E16 SpC
G + A
5704
5709
ftpt 1
*
5745
ftpt 2
5763
*
5771
ftpt 3
5793
5832
ftpt 4
5846
5872
ftpt 5


BSA
BSA
E20 Ctx
E17 Ctx
G + A
* * *
5872
ftpt 5
5898


BSA
E13 Tel
E17 Ctx
E16 Ctx
E16 Mid
E16 Cb
E16 SpC
G + A
5885
5867
ftpt 5
5847
5835
ftpt 4
5788
5769
ftpt 3
5763
5747
ftpt 2

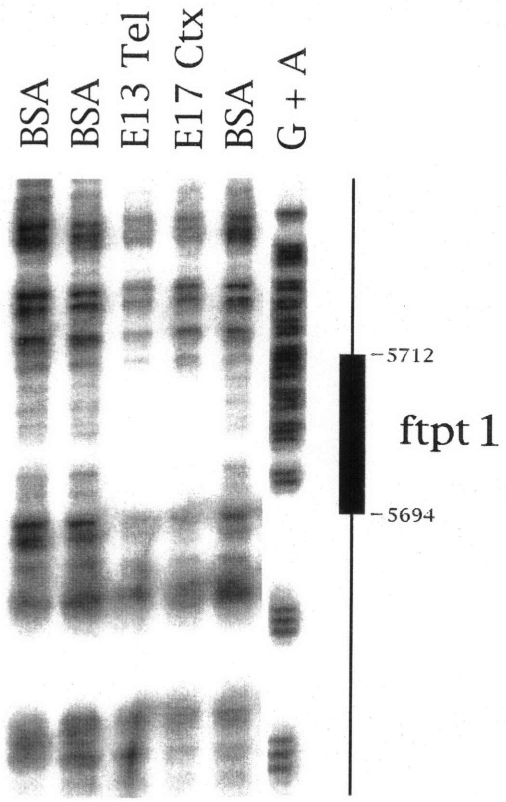

BSA
BSA
E13 Tel
E17 Ctx
BSA
G + A
-5712
ftpt 1
-5694

Figure 11. Summary of protected regions detected by DNAse I footprinting

Sequence of both strands of the 257 base pair CNS-specific element of the rat nestin second intron. Footprinted regions are marked by stippled boxes, and hypersensitive sites by asterisks. The protected regions detected on the forward strand roughly line up with those on the reverse strand, so aligned footprints on opposite strands are given a single number from one to five. The bases are numbered relative to the genomic clone, as in figure 11.

DNAseI Footprinting Reveals Five Protein-Binding Sites in the Nestin Enhancer

```
                                                  Footprint 1
5669
GGTCTGAAAA GGATTTGGAG AAGGGGAGCT GAATTCATTT GCTTTTGTCT
CCAGACTTTT CCTAAACCTC TTCCCCTCGA CTTAAGTAAA CGAAAACAGA

                                   *         Footprint 2
5719
GTTACCAGCT CTGGGGGCAG AGAGAGAGCC ATCCCCTGGG AACAGCCTGA
CAATGGTCGA GACCCCCGTC TCTCTCTCGG TAGGGGACCC TTGTCGGACT
                                   *                       *

               Footprint 3
5769
GAATTCCCAC TTCCCCTGAG GAGCCCTCCC TTCTTAGGCC CTCCAGATGG
CTTAAGGGTG AAGGGGACTC CTCGGGAGGG AAGAATCCGG GAGGTCTACC

                       Footprint 4                 *          **
5819
TAGTGTGGAC AAAAGGCAAT AATTAGCATG AGAATCGGCC TCCCTCCCAG
ATCACACCTG TTTTCCGTTA TTAATCGTAC TCTTAGCCGG AGGGAGGGTC

5869
*                 Footprint 5
AGGATGAGGT CATCGGCCTT GGCCTTGGGT GGGGAGGCGG AGACTGATCT
TCCTACTCCA GTAGCCGGAA CCGGAACCCA CCCCTCCGCC TCTGACTAGA

5919
GAGGAGT
CTCCTCA
```

* = DNAseI hypersensitive site

Figure 12. The footprint 1 region is bound in vitro by several proteins from various portions of the developing CNS

Electrophoretic mobility shift assays (EMSA) performed with a double-stranded oligonucleotide probe containing the footprint 1 region of the rat nestin second intron (see figure 17 for sequence of probe Ftpt1). The labeled probe was mixed with four μgs of nuclear extract from each tissue tested. DNA probes which are bound by proteins are retarded in mobility and are thus closer to the top of the gel (left of figure) than unbound probes, which in lanes 1-10 have been run just off the bottom of the gel (at the right of the figure.) Sequence specificity of binding is demonstrated in lanes 11-13. The complexes formed by E17 cortical extract on the Ftpt 1 probe are displaced by a 100-fold molar excess of the unlabeled Ftpt1 probe (lane 12), but not by the unrelated probe TREp (lane 13) which contains an artificial palindromic thyroid hormone response element.

Tel, telencephalon nuclear extract; Ctx, cortex; Mid, midbrain; Cb, cerebellum; SpC, spinal cord

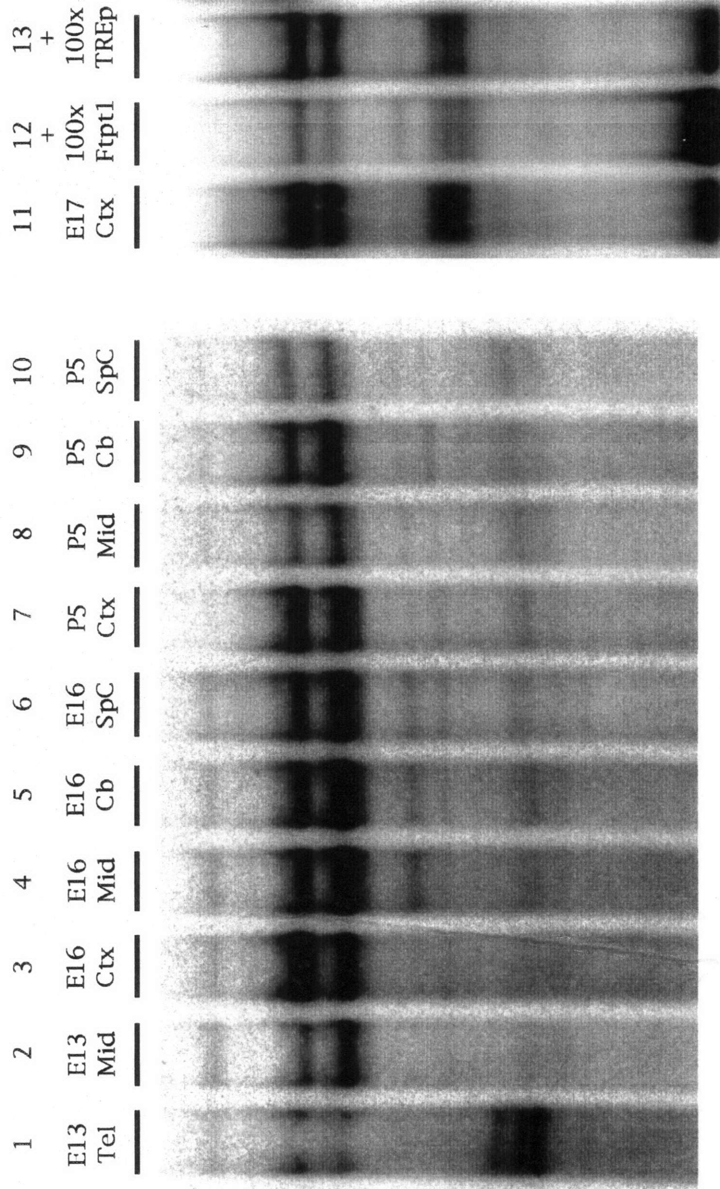

Ftpt_1
1 E13 Tel
2 E13 Mid
3 E16 Ctx
4 E16 Mid
5 E16 Cb
6 E16 SpC
7 P5 Ctx
8 P5 Mid
9 P5 Cb
10 P5 SpC
11 E17 Ctx
12 + 100x Ftpt1
13 + 100x TREp

Figure 13. The footprint 5 region is bound by a small number of activities, one of which is specific to the cortex

EMSA performed with the Ftpt5 probe, containing 42 bp of nestin second intron sequence encompassing footprint number 5. The free probe has been run just off the bottom of the gel in order to better separate the shifted bands. Two strong bands (arrowheads) are seen in P5 cortex or E17 cortex (lanes 5 and 9). Only one of these is present, along with a weaker band of greater mobility, in midbrain, cerebellum and spinal cord. Lanes 9-11 show that the two strong bands produced by E17 cortical extract are sequence-specific, as they are displaced by 100-fold excess of unlabeled Ftpt5 probe but not by a similar quantity of cold Ftpt4.

Tel, telencephalon nuclear extract; Ctx, cortex; Mid, midbrain; Cb, cerebellum; SpC, spinal cord

Ftpt 5

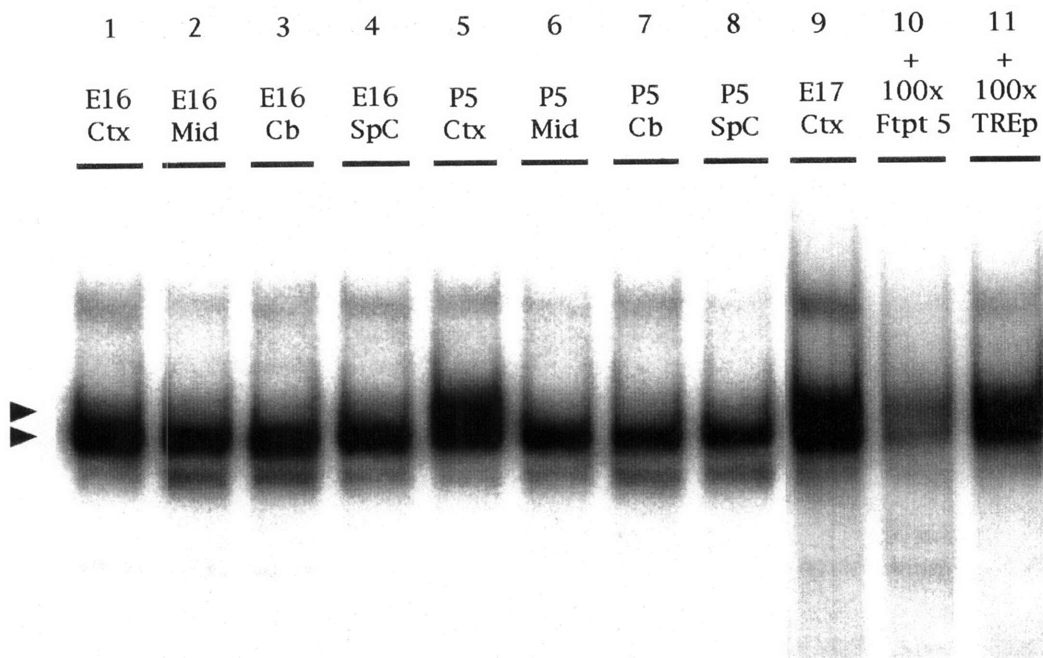


Figure 14. The footprint 2 region is recognized by different sets of proteins in the cortex than in the rest of the CNS

EMSA performed with the Ftpt2 probe (see figure 17). The E16 midbrain, cerebellum, and spinal cord extracts show two prominent binding activities, the slower-migrating of which (black arrowhead) is not present in cortex either at E16 or P5. The faster-migrating band is virtually absent from all but cerebellum at P5. E16 cortex has a second weak and slow-migrating band (grey arrowhead) which may not be present in the rest of the CNS extracts. Lanes 9-11 demonstrate the sequence-specificity of the two binding activities present in E16 spinal cord, by competition with a 500-fold molar excess of the cold probe (Ftpt2) or nonspecific oligo (TREp).

Tel, telencephalon nuclear extract; Ctx, cortex; Mid, midbrain; Cb, cerebellum; SpC, spinal cord; Free, unbound probe

Ftpt2

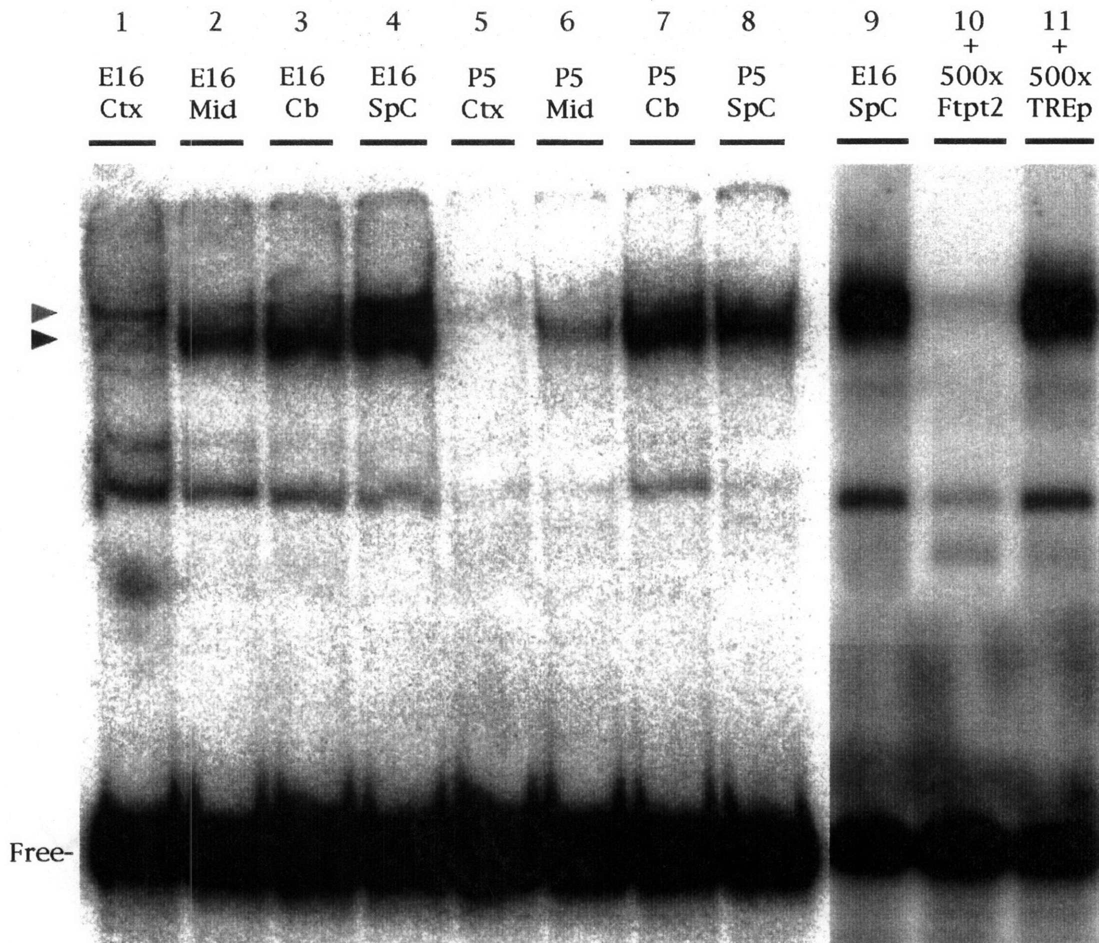


Figure 15. The footprint 3 region is most strongly bound by extracts of cerebellum and spinal cord

EMSA on the Ftpt3 probe shows two complexes (arrowheads) which are present in all extracts tested. The lower band of this pair is much stronger at E16 than at P5, but occurs in all four brain regions at both ages. The upper band is very strong only in E16 cerebellum and spinal cord, and in P5 cerebellum. It is weak but present in cortex and midbrain at both ages. The sequence-specificity of these bands in established in figure 18.

Tel, telencephalon nuclear extract; Ctx, cortex; Mid, midbrain; Cb, cerebellum; SpC, spinal cord; Free, unbound probe

Ftpt 3

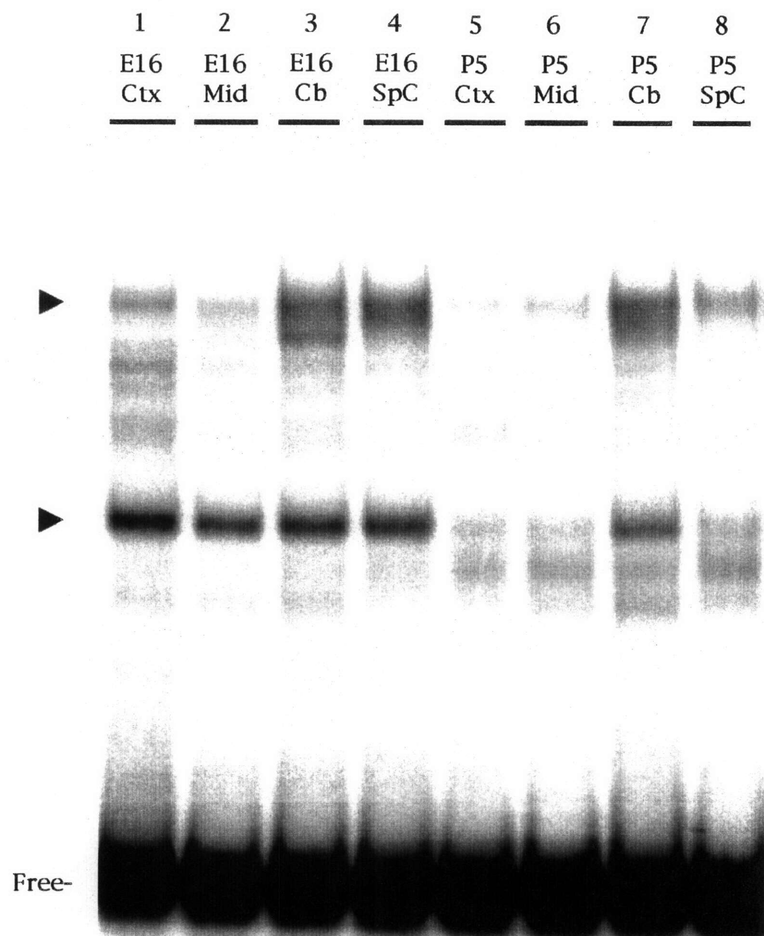


Figure 16. Methylation interference assays show multiple complexes bind the same core recognition sequences

The five double-stranded oligonucleotide probes (Ftpt1-Ftpt5, see figure 17) originally used in EMSA assays were partially methylated with dimethylsulfate prior to being gel shifted with 40 μgs of nuclear extract. Bound protein-DNA complexes and free probe were isolated from a nondenaturing gel, then cleaved by piperidine at methylated bases. Comparison of cleavage sites between bound and free DNA indicates bases at which methylation sterically inhibits protein binding.

At the left of each figure is a typical EMSA performed with 4 μgs of nuclear extract. This illustrates the relative migration rates of the complexes isolated in methylation interference assays. Complexes which run too closely to be individually isolated are labeled together. Bd, protein-DNA complex; Free, unbound DNA; G+A, Maxam-Gilbert sequencing reaction for purines. The sequence of the probe appears next to this lane.

Arrowheads indicate positions at which methylation interferes with protein binding. Results are summarized in figure 17.

(A) Methylation interference on the reverse strand of the probe **Ftpt1**, using E17 cortical extract. Methylation interference is observed at six adenines.

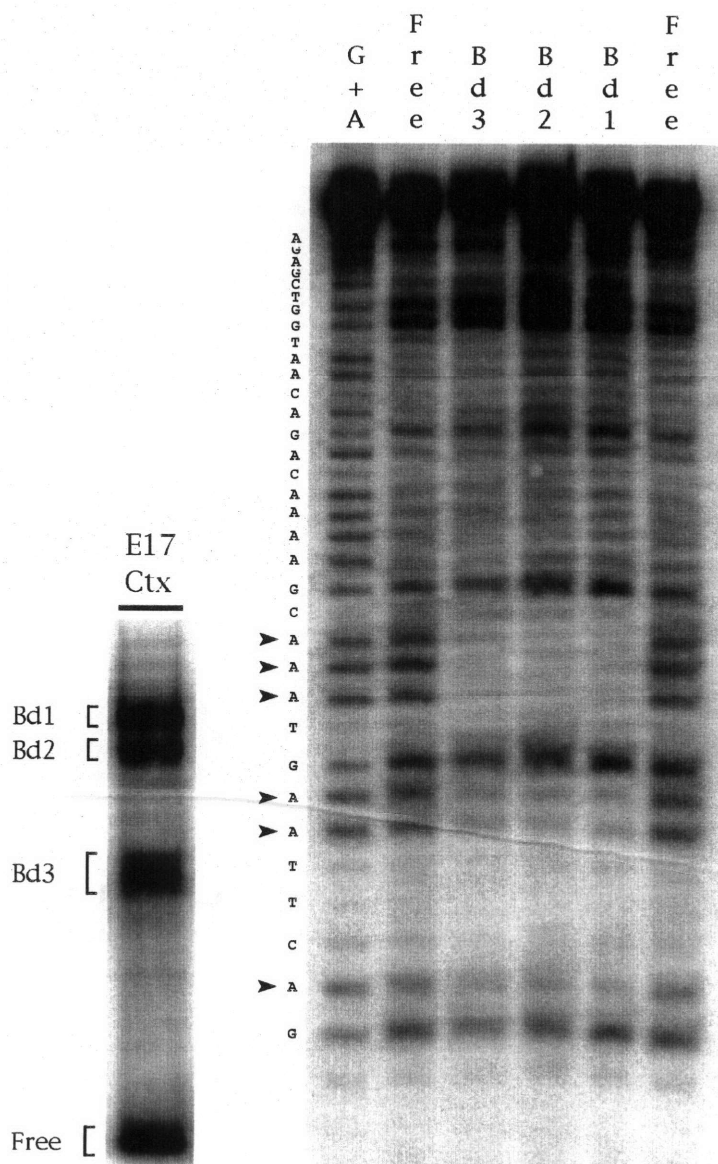
(B) Methylation interference on both strands of **Ftpt2**, with E16 spinal cord extract. Interference occurs at two guanines on each strand.

(C) Methylation interference on the bottom strand of **Ftpt3**, using E16 spinal cord extract. The faster-running bound complex seen in the EMSA was not present in the scaled-up reaction for methylation interference. In the slower-migrating complex, interference occurs at seven guanines.

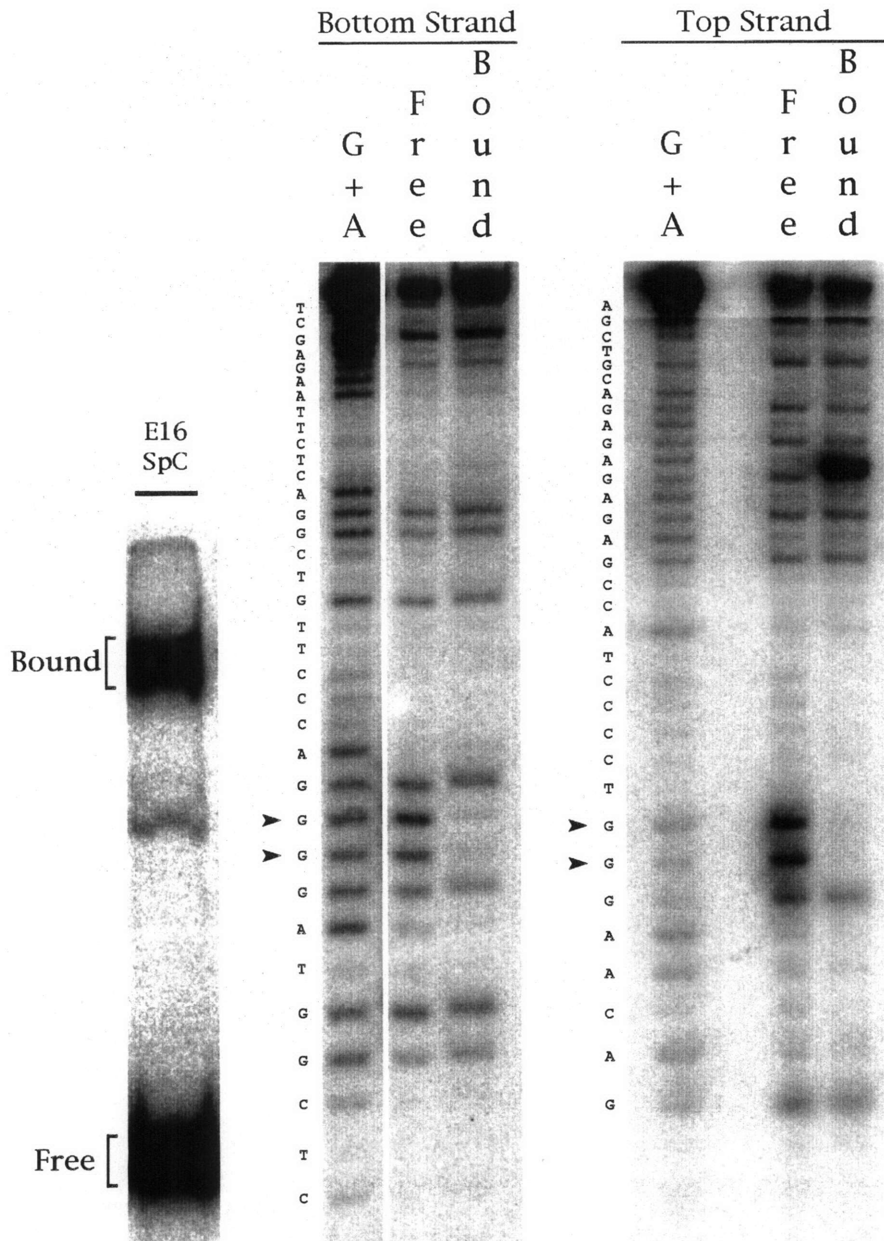
(D) Methylation interference on the bottom strand of **Ftpt4**, performed with E17 cortical extract, blocks protein binding to three adenines. Note that the same three adenines block at least four different complexes, as "Bd4" is actually three separate bands.

(E) Methylation interference on both strands of **Ftpt5** with E17 cortical extract. Interference with binding occurs at a single guanine on the reverse strand and at two guanines on the forward strand. This pattern is the same for each isolated complex. The apparent interference in the Bd1 lane at two guanines nearer the bottom of the figure has not been reproducible.

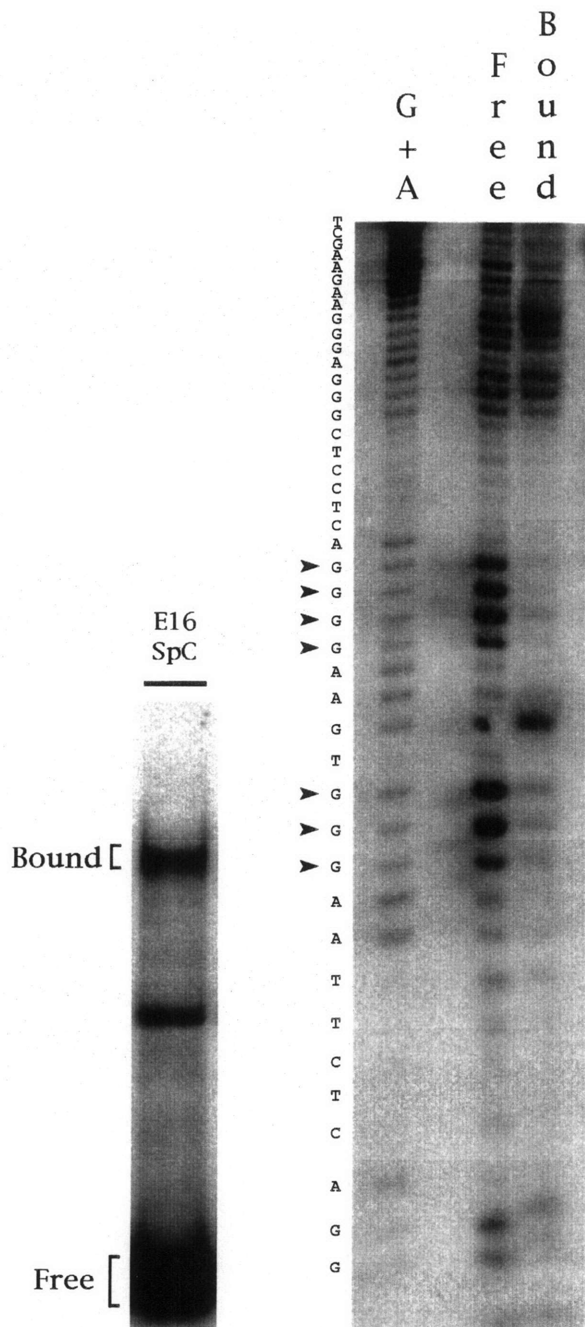
A. Methylation Interference on Ftpt1



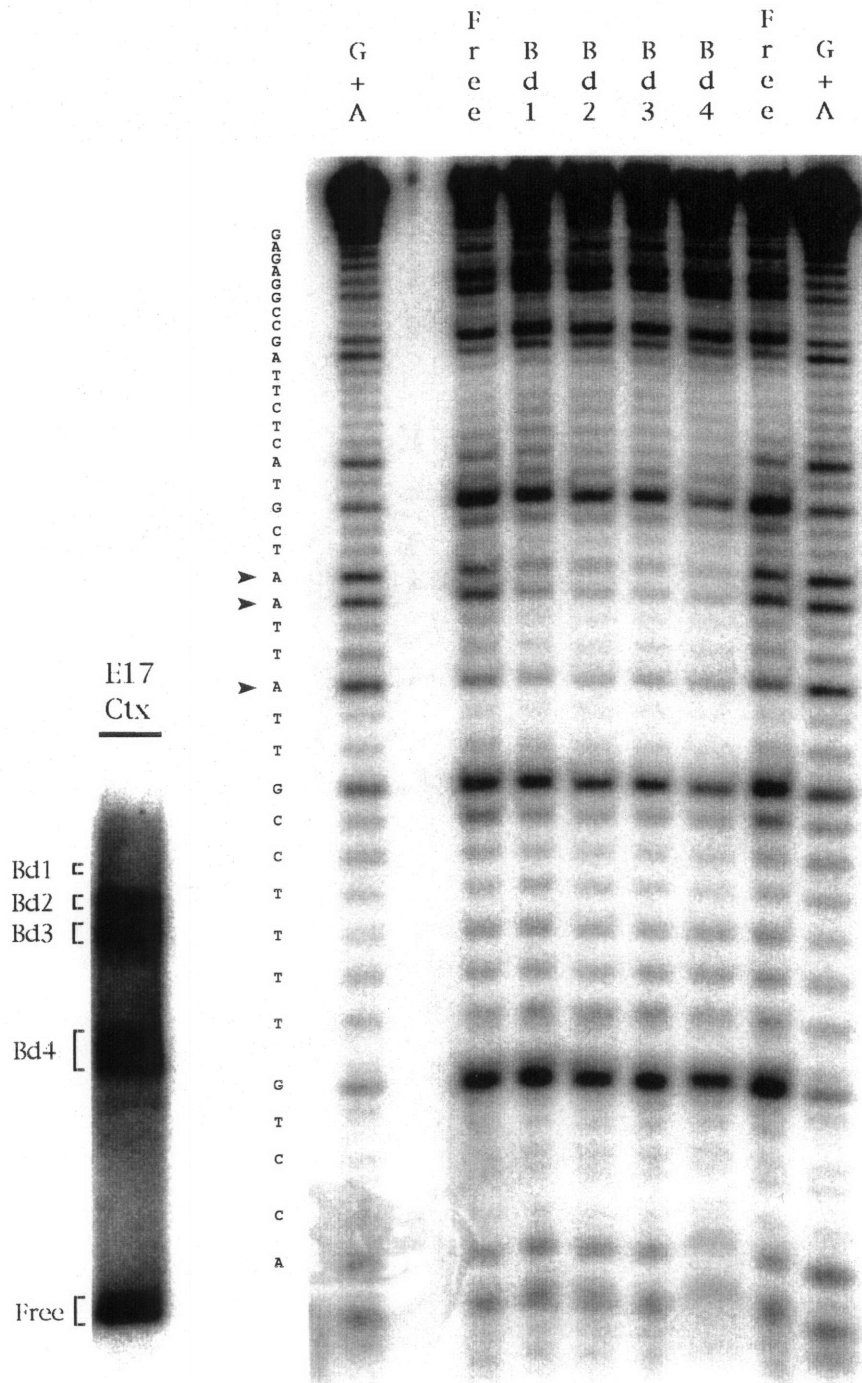
B. Methylation Interference on Ftpt2



C. Methylation Interference on Ftpt3



D. Methylation Interference on Ftpt4



E. Methylation Interference on Ftpt5

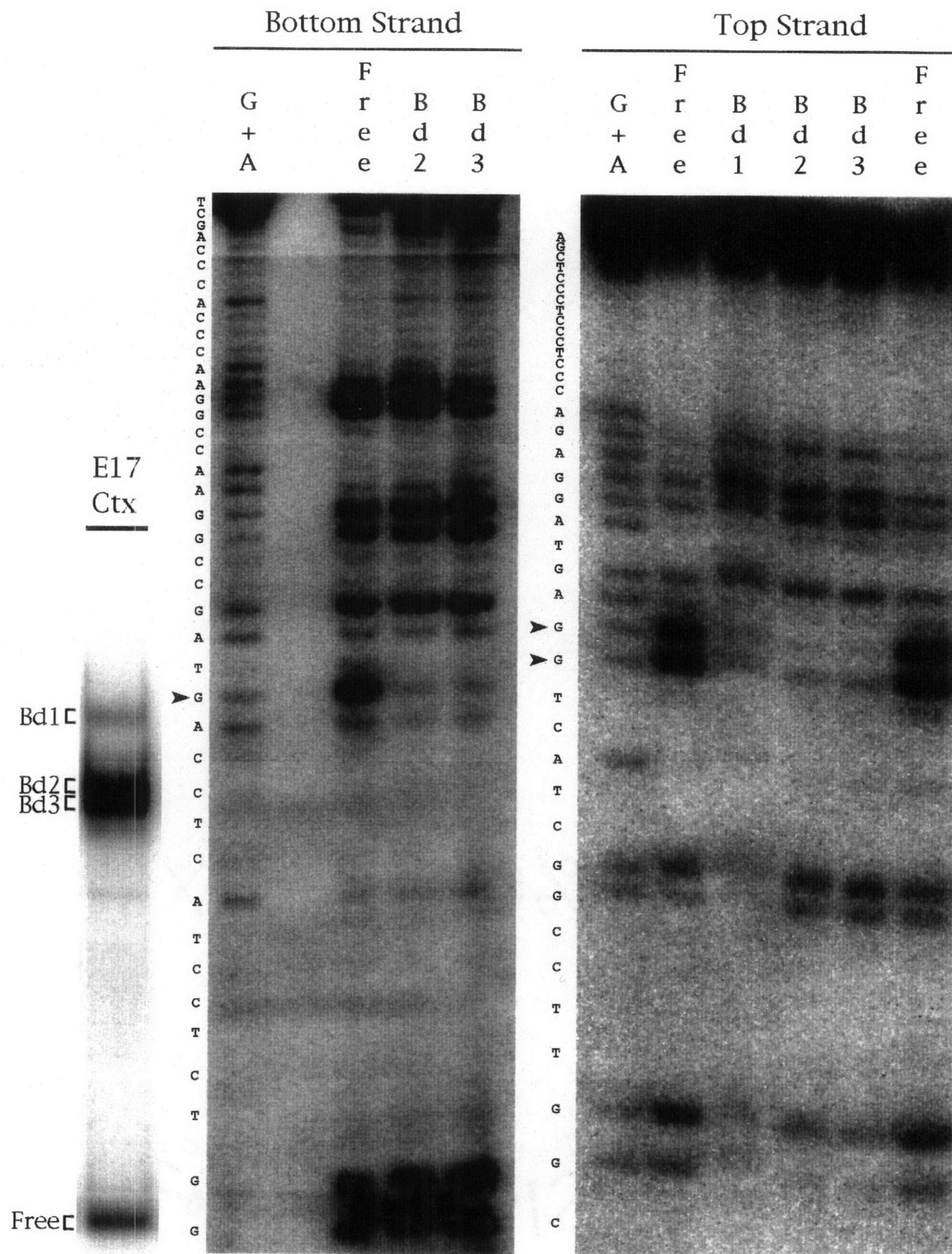


Figure 17. Summary of methylation interference patterns.

Sequences of the double-stranded oligonucleotides used for EMSA and methylation interference assays are shown. Each probe contains 40-42 bp of nestin second intron sequence, with added HindIII and SalI 5' overhangs to allow end-labeling with Klenow.
Over- or under- lining shows bases protected from DNAseI in footprinting assays on the 257 bp enhancer fragment (see figure 10).
v or ^ indicates site of methylation interference (see figure 16).
"ATTTGCaT" or similar short sequence under the probe sequence indicates a region homologous to the recognition site of a known transcription factor, or points out the homology between footprint 2 and footprint 3. Capital letters indicate identity, small letters are a mismatch.
(A) **Ftpt1 probe.** Three complexes bind a site overlapping with a seven-of-eight base pair match to the octamer recognition site of POU proteins.
(B) **Ftpt2 probe.** One complex binds a site very similar between Ftpt2 and Ftpt3. This sequence is also homologous to the recognition motif for AP-2 as reported by [Imagawa *et al.*, 1987].
(C) **Ftpt3 probe.** One complex binds an extended site, part of which is similar to Ftpt2.
(D) **Ftpt4 probe.** At least four complexes bind three adenines which overlap another seven-of-eight base pair match to the octamer motif.
(E) **Ftpt5 probe.** Three complexes bind a site with a number of interesting homologies. First, eight bases are identical to a binding site for thyroid hormone receptor/retinoid X receptor β found in the class I Major Histocompatibility Complex promoter [Marks *et al.*, 1992]. Second, seven of these eight bases are identical to the consensus cyclic AMP response element, and six of eight are identical to the consensus AP-1 site [Meyer and Habener, 1993].

Summary: Methylation Interference Patterns and Potential Transcription Factor Binding Sites

a. Ftpt 1:

```
                                ______
AGCTGGAGAAGGGGAGCTGAATTCATTTGCTTTTGTCTGTTACCAG
    CCTCTTCCCCTCGACTTAAGTAAACGAAAACAGACAATGGTCAGCT
                  ^   ^^  ^^^

                        ATTTGCaT            --octamer
```

b. Ftpt 2:

```
             vv
AGCTGCAGAGAGAGAGCCATCCCCTGGGAACAGCCTGAGAATTC
    CGTCTCTCTCTCGGTAGGGGACCCTTGTCGGACTCTTAAGAGCT
                      ^^

                   TCCCCTGaGgA              --Ftpt 3
                   TCCCCaN Gcg/cGc          --AP-2
```

c. Ftpt 3:

```
AGCTCAGCCTGAGAATTCCCACTTCCCCTGAGGAGCCCTCCCTTCT
    GTCGGACTCTTAAGGGTGAAGGGGACTCCTCGGGAGGGAAGAAGCT
                 ^^^    ^^^^

                       TCCCCTGaGgA          --Ftpt 2
                       TCCCCaN gcG/cGc      --AP-2
```

d. Ftpt 4:

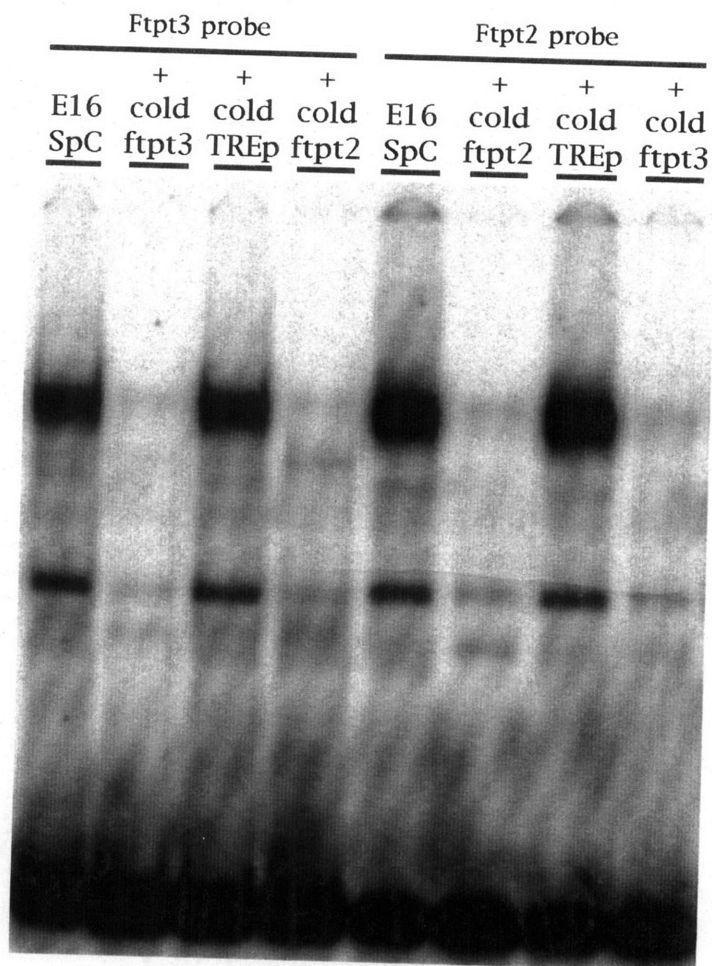
```
AGCTGTGTGGACAAAAGGCAATAATTAGCATGAGAATCGGCCTC
    CACACCTGTTTTCCGTTATTAATCGTACTCTTAGCCGGAGAGCT
                      ^  ^^

                       ATTtGCAT             --octamer
```

e. Ftpt 5:

```
               vv
AGCTCCCTCCCAGAGGATGAGGTCATCGGCCTTGGCCTTGGGTGGG
    GGGAGGGTCTCCTACTCCAGTAGCCGGAACCGGAACCCACCCAGCT
                        ^
                TGAGGTCA                    --T3R/RXRβ
                TGAcGTCA                    --CRE
                TGAc_TCA                    --AP-1
```

Figure 18. Ftpt2 and Ftpt3 probes can compete with one another for binding of the same or similar proteins.

EMSA with cross-competition between Ftpt2 and Ftpt3 using E16 spinal cord nuclear extract. Two complexes are formed by E16 SpC extract on both probes, and the mobilities and strengths of these complexes are very similar. Each complex binds in a sequence-specific manner, as excess unlabeled probe displaces it but not excess nonspecific oligonucleotide (TREp). In addition, each complex is also disrupted by excess probe for the neighboring footprinted region. All unlabeled competitors are used at 500-fold molar excess over labeled probe.
This experiment indicates that the two probes share a sequence recognized by the two proteins. As noted in figure 17, these two probes are homologous to one another at a sequence which is the core binding site as determined by methylation interference. However, footprints 2 and 3 are so closely spaced that the sequences of these two probes also overlap. A 14 bp stretch is part of both probes. So while these two probes definitely bind the same protein or different proteins with the same binding site, the identity of that sequence remains ambiguous.


Ftpt3 probe
Ftpt2 probe
E16 SpC
+ cold ftpt3
+ cold TREp
+ cold ftpt2
E16 SpC
+ cold ftpt2
+ cold TREp
+ cold ftpt3

Discussion

The experiments described in this chapter have shown that the 257 base pair fragment of the nestin second intron which functions *in vivo* as a CNS-specific transcriptional regulatory sequence is specifically bound by nuclear proteins *in vitro*. Protein binding assays performed *in vitro* are thus a convenient method to identify transcription factors which could be responsible for the *in vivo* function of the CNS regulatory element.

The nestin CNS regulatory element was first analyzed in vitro by DNAse I footprinting. This method revealed 5 areas of nuclear protein binding within the 257 base pair fragment, although some of these areas were not detectably bound in all tissues or all ages tested. The proteins which bind these footprints cannot be said on this basis to be absent from some tissues, since the footprinting assay requires near-saturation of the available DNA by bound proteins. This may be difficult to achieve when tissue extracts are used, as in the experiments reported here, due to the presence of multiple cell types. Generally speaking, the five footprinted areas fall into two categories. Footprints 1, 4, and 5 are present in the cortex as well as the other CNS regions tested (midbrain, cerebellum, and spinal cord) at almost all of the developmental ages examined. By contrast, protection of footprints 2 and 3 was not seen in cortex, and was strongest in spinal cord and cerebellum. Since terminal differentiation of neurons in the spinal cord precedes that in cortex, which precedes that in cerebellum, there is no correlation of footprint 2 and 3 binding activities with developmental progression. Rather, the protection of these two areas is stronger in more caudal regions of the CNS.

When each of the five footprinted areas is isolated on a short double-stranded probe, named Ftpt1-5, specific protein binding to each area is detectable by the gel shift assay. Gel shift assays show that multiple protein activities interact with each of these five binding sites, and the tissue distributions of these are much more complex than indicated by DNAse I footprinting. One important point about the tissue distributions is that nuclear extracts which fail to protect a given binding site from DNAse I are usually not completely

lacking in binding activity for that site. For example, factors which bind footprints 2 and 3 are present in embryonic cortex. However, the low abundance of these factors in cortex relative to spinal cord as measured by the gel shift assay probably explains why these sites are not footprinted by cortical extracts.

The presence of multiple binding activities in gel shift assays on any of the five probes also complicates the picture of nestin transcriptional regulation. As discussed previously, the multiple bands could be different proteins, or multiple bands could be generated by a single protein having a range of phosphorylation states or associated molecules. In addition, in nuclear extracts prepared from whole tissues, different binding activities may be contributed by different cell types. Another possible explanation for the wealth of binding activities detected by gel shift assays may be the lack of sequence context normally provided by the flanking regions and other binding sites. This could allow the formation of complexes which would ordinarily be incompatible with the complex of factors binding the enhancer.

To demonstrate that the multiple complexes detected on a given gel shift probe interact with the same DNA site, methylation interference assays were performed. It was determined that any methylated base found to block formation of one protein-DNA complex also blocked the formation of every other complex which could be isolated with that probe. For example, at least four separable complexes with the Ftpt4 probe interact with the same set of three adenine bases on one DNA strand (figure 16d). In addition, methylation interference studies suggest that the complexes detected on short probes are the same as those detected by DNAse I footprinting on the 257 base pair enhancer, since all binding sites blocked by methylation fall within the limits of a previously detected footprint (see figure 17).

Finally, the methylation interference patterns of the various binding proteins provide clues to the identity of these factors. Comparison of the mapped sequences with a database of transcription factor binding sites has identified homologies to several well-known DNA motifs (figure 17). The two probes Ftpt1 and Ftpt4

produce similar gel shift patterns with CNS nuclear extracts, and both contain homologies to the canonical octamer binding site of POU transcription factors. The POU genes form a diverse superfamily of proteins whose roles in neural development will be discussed in detail in the next chapter.

Two other probes, Ftpt2 and Ftpt3, also show remarkably congruent gel shift patterns and each probe competes with the other for protein binding. Methylation interference studies point to a region of sequence homology between the two probes, which is strongly reminiscent of the consensus binding site for the dimeric transcription factor AP-2. Unfortunately, this does not limit the candidate binding factors for footprints 2 and 3 to a single protein. The AP-2 family contains at least three genes designated AP-2α, AP-2β, and AP-2.2 [Chazaud *et al.*, 1996; Moser *et al.*, 1995; Williams *et al.*, 1988], and the first described member is alternatively spliced into four mRNAs with different tissue distributions [Meier *et al.*, 1995].

Binding sequence homologies also suggest a number of proteins which may produce footprint 5. Methylation interference shows that only a very small area of this long footprint is in close contact with the binding factor. The sequence of this small area has homology to previously identified binding motifs for nuclear hormone receptors, CRE-binding proteins, or AP-1 factors. Furthermore, each of these three groups is composed of a number of molecules. The cAMP response element binding proteins have at least three members: CREB, CREM, and ATF-1 [reviewed by Vallejo, 1994]. The binding activity called AP-1 can bind to the same inverted repeat sequence as the CREBs, and AP-1 is a dimer composed of c-fos, c-jun, or any of their numerous relatives. Finally, the nuclear hormone receptors are a large superfamily. The binding sites of these dimeric factors are typically direct or inverted repeats of the half-site 5'-AGGTCA-3'. However, some orphan receptors within this family, such as NGFI-B/nur77 or FTZ-F1 have been found to bind as monomers to a single half-site [Ueda *et al.*, 1992; Wilson *et al.*, 1992].

So while methylation interference studies have focused our attention on a few short sequences within nestin's CNS-specific

element, the homology of these sequences to known *cis*-acting elements does not point to a small number of *trans*-acting factors. In fact, we may question whether the discovery of these consensus motifs identifies the possible binding proteins at all. This is because the similarity of a sequence to the consensus binding site for a given transcription factor does not necessarily predict how well that factor will bind the sequence. Most transcription factors will bind a range of sequences in various genes which can deviate substantially from the consensus binding site.

The POU proteins, some of which may interact with the nestin CNS element, provide a good example of the range of permissible binding sequences for transcription factors. The ubiquitously expressed POU protein Oct-1 binds strongly to the consensus binding site for this family, the octamer motif 5'-ATGCAAAT-3' [Singh *et al.*, 1986]. It also binds with lower affinity to the the sequence 5'-TAATGARAT-3', where R is any purine [Baumruker *et al.*, 1988]. The octamer sequence can be viewed as separate binding sites for the two DNA-binding domains of Oct-1: The 3' AAAT sequence is bound by the homeodomain, while the POU-specific domain interacts with the 5' ATGC sequence [Kristie and Sharp, 1990; Verrijzer *et al.*, 1990].

Several brain-specific POU proteins also appear to have binding sites composed of two parts. The consensus sites for these proteins resemble a canonical octamer in which the half which binds the POU-specific domain is reversed, and the spacing between the two halves may vary. The consensus binding site for Brn-2 is 5'-CATn(T/A)AAT-3', where n represents zero, two, or three bases of any sequence [Li *et al.*, 1993]. A distant relative, Brn-3.0, binds the same consensus site as Brn-2 but the spacing (n) must be 3 bases [Li *et al.*, 1993]. Brn-3.0 binds poorly to the canonical octamer sequence. A third distant relation, Brn-5/Emb, binds similar sites with the consensus sequence 5'-GCATnTAAT-3' where n is 2 or 3 bases; it also binds strongly to the octamer-like sequence 5'-ATATGATAAT-3' [Andersen *et al.*, 1993].

This degenerate set of possible binding sites for POU proteins makes it harder to argue that members of this family must bind the nestin second intron element, but does not rule it out. The consensus

binding sites reported here for brain-specific POUs are A/T-rich sequences different from the octamer-like sequences found in footprints 1 and 4. However, the methylation interference patterns found with these two probes point to A/T-rich sequences which overlap the 5' ends of the two octamer-like sequences. So these binding sites may be described as related to the octamer, and related to the A/T-rich consensus sites for several brain-specific POUs. For this reason, the work in the next chapter will address the question of whether brain-specific POU proteins bind in vitro to footprints 1 and 4.

References

Andersen, B., M.D. Schonemann, R.d. Pearse, K. Jenne, J. Sugarman and M.G. Rosenfeld (1993) Brn-5 is a divergent POU domain factor highly expressed in layer IV of the neocortex. J. Biol. Chem. **268**, 23390-8.

Baumruker, T., R. Sturm and W. Herr (1988) OBP100 binds remarkably degenerate octamer motifs through specific interactions with flanking sequences. Genes Dev. **2**, 1400-1413.

Chazaud, C., M. Oulad-Abdelghani, P. Bouillet, D. Decimo, P. Chambon and P. Dolle (1996) AP-2.2, a novel gene related to AP-2, is expressed in the forebrain, limbs and face during mouse embryogenesis. Mech. Dev. **54**, 83-94.

Dignam, J.D., R.M. Lebovitz and R.G. Roeder (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucl. Acids Res. **11**, 1475-1489.

Imagawa, M., R. Chiu and M. Karin (1987) Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell **51**, 251-260.

Kristie, T. and P.A. Sharp (1990) Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV α-*trans*-activator protein. Genes Dev. **4**, 2383-2396.

Li, P., X. He, M.R. Gerrero, M. Mok, A. Aggarwal and M.G. Rosenfeld (1993) Spacing and orientation of bipartite DNA-binding motifs as potential functional determinants for POU domain factors. Genes Dev. **7**, 2483-96.

Marks, M.S., P.L. Hallenbeck, T. Nagata, J.H. Segars, E. Appella, V.M. Nikodem and K. Ozato (1992) H-2RIIBP (RXRβ)

heterodimerization provides a mechanism for combinatorial diversity in the regulation of retinoic acid and thyroid hormone responsive genes. EMBO J. **11**, 1419-1435.

Meier, P., M. Koedood, J. Philipp, A. Fontana and P.J. Mitchell (1995) Alternative mRNAs encode multiple isoforms of transcription factor AP-2 during murine embryogenesis. Dev. Biol. **169**, 1-14.

Meyer, T.E. and J.F. Habener (1993) Cyclic adenosine 3',5'-monophosphate response element binding protein (CREB) and related transcription-activating deoxyribonucleic acid-binding proteins. Endocr. Revs. **14**, 269-290.

Moser, M., A. Imhof, A. Pscherer, R. Bauer, W. Amselgruber, F. Sinowatz, F. Hofstader, R. Schule and R. Buettner (1995) Cloning and characterization of a second AP-2 transcription factor: AP-2β. Development **121**, 2779-2788.

Singh, H., R. Sen, D. Baltimore and P.A. Sharp (1986) A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature **319**, 154-158.

Ueda, H., G.-C. Sun, T. Murata and S. Hirose (1992) A novel DNA-binding motif abuts the zinc finger domain of insect nuclear hormone receptor FTZ-F1 and mouse embryonal long terminal repeat-binding protein. Mol. Cell. Biol. **12**, 5667-5672.

Umesono, K., K.K. Murakami, C.C. Thompson and R.M. Evans (1991) Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D_3 receptors. Cell **65**, 1255-1266.

Vallejo, M. (1994) Transcriptional control of gene expression by cAMP-response element binding proteins. J. Neuroendo. **6**, 587-596.

Verrijzer, C.P., A.J. Kal and P.C. van der Vliet (1990) The Oct-1 homeo domain contacts only part of the octamer sequence and full Oct-1 DNA-binding activity requires the POU-specific domain. Genes Dev. **4**, 1964-1974.

Williams, T., A. Admon, B. Luscher and R. Tjian (1988) Cloning and expression of AP-2, a cell-type-specific transcription factor that activates inducible enhancer elements. Genes Dev. **2**, 1557-1569.

Wilson, T.E., R.E. Paulsen, K.A. Padgett and J. Milbrandt (1992) Participation of non-zinc finger residues in DNA binding by two nuclear orphan receptors. Science **256**, 107-110.

Wirth, T., L. Staudt and D. Baltimore (1987) An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. Nature **329**, 174-178.

Chapter IV: Identification of transcription factors which recognize the CNS-specific element *in vitro*

Introduction

The experiments of the previous chapter identified five sequences in the 257 bp nestin neuroepithelial enhancer which are bound by proteins *in vitro*. Comparison of these sequences to the consensus recognition motifs of known transcription factor families suggests that the enhancer-binding proteins may belong to three groups: POU proteins, AP-2-like proteins, and a yet unknown group which could be the nuclear hormone receptor family, the CREB family, or the AP-1 family. Members of each of these groups have roles in adult neuronal function, but a few of the POU proteins are especially interesting for their actions in neuronal development. This chapter explores whether the expression of nestin is regulated by the class III POU proteins, which are present in the early neural tube and some of which have been shown to have roles in neural fate determination.

The nestin neuroepithelial enhancer contains two sites related to the octamer binding motif (5'-ATGCAAAT-3'), which is the canonical binding site for most members of the POU transcription factor family. This large group of homeobox-containing genes was initially identified by the discovery of the factors Pit-1 [Bodner *et al.*, 1988; Ingraham *et al.*, 1988], Oct-1 [Sturm *et al.*, 1988], Oct-2 [Clerc *et al.*, 1988; Ko *et al.*, 1988; Muller *et al.*, 1988; Scheidereit *et al.*, 1988], and the *Caenorhabditis elegans* gene unc-86 [Finney *et al.*, 1988]. These genes are now grouped into six classes (POU I-POU VI) based on amino acid sequence homologies within the highly conserved POU-specific (POUs) and POU-homeodomain (POUhd) regions [Wegner *et al.*, 1993].

While some POU genes, such as Oct-1, are expressed in virtually all cells [Sturm *et al.*, 1988], most members of this family are tissue-specific. The tissue-specific POU proteins are often critical at late stages in the formation of the cells in which they are found. For example, Pit-1 regulates proliferation and/or survival of three anterior pituitary cell types [reviewed in Andersen and Rosenfeld,

1995]. Oct-2 is necessary for terminal differentiation of B-cells [Corcoran *et al.*, 1993] and unc-86 is required to specify sensory neurons in *C. elegans* [Finney and Ruvkun, 1990].

A number of these tissue-specific POU proteins are expressed only in the brain. This distinguishes the POU family from the classical homeobox genes; no Hox gene transcripts are found in the CNS anterior to the hindbrain. Brain-specific POU genes include the entire class III family, comprising Brain-1 (Brn-1), Brn-2, Brn-4, and Tst-1/Oct-6/SCIP [Hara *et al.*, 1992; He *et al.*, 1989; Le Moine and Young, 1992; Mathis *et al.*, 1992; Monuki *et al.*, 1990; Suzuki *et al.*, 1990]. In addition, three class IV genes, Brn-3a, Brn-3b and Brn-3c, are predominantly expressed in the nervous system [Gerrero *et al.*, 1993; He *et al.*, 1989; Lillycrop *et al.*, 1992], as is the sole class VI gene, Emb/Brn-5 [Andersen *et al.*, 1993; Okamoto *et al.*, 1993].

Transcripts of most of these genes are found in the developing neural tube, which makes them candidates for *trans*-acting regulators of the nestin gene. Three POU-III members (Brn-1, Brn-2, and Brn-4) are widely expressed in the developing neural tube from about 10.5 days post-coitum (E10.5) in the rat [Alvarez-Bolado *et al.*, 1995]. Translation of Brn-2 begins prior to E8.5, when the protein is found in limited areas of the neural folds [Schonemann *et al.*, 1995]. The early onset and wide expression of Brn-1, Brn-2, and Brn-4 in the neural tube is reminiscent of nestin, which is first detected at E7.75 in limited areas of the neural plate [Dahlstrand *et al.*, 1995]. The fourth POU-III member, Tst-1, is much more restricted early in development. It expands from a single tiny patch ventral to the optic stalk at E10 [Alvarez-Bolado *et al.*, 1995] to weakly mark several discrete areas of the telencephalon, mesencephalon, and spinal cord at E13, and is limited to the forebrain and brainstem at E16 [Treacy and Rosenfeld, 1992]. This contrasts with the very widespread expression of Tst-1 in the adult brain [Alvarez-Bolado *et al.*, 1995].

Brn-3a is also strongly expressed in the developing CNS. In contrast to Brn-1 and Brn-2, its expression is discontinuous, localized to discrete regions of the mesencephalon, rhombencephalon, and spinal cord, as well as outside the CNS in the trigeminal and dorsal root ganglia [Treacy and Rosenfeld, 1992]. The class VI gene Emb is

detectable in the CNS as early as E10 in the mouse, and is found diffusely through the neural tube and stronger in the spinal cord at E13 [Okamoto *et al.*, 1993].

Interestingly, two of the founding POU family members which are restricted to other tissues in the adult are transiently expressed in the neural tube. Pit-1 first appears in the neural plate and tube from E10 to E13, then the signal is lost until it reappears at E16 in the anterior pituitary [Treacy and Rosenfeld, 1992]. While Pit-1 has been shown to be necessary for pituitary development, it may have unknown functions for a very brief period in the neural tube. Similarly, the Oct-2 gene is closely identified with differentiation of B lymphocytes [Ko *et al.*, 1988; Muller *et al.*, 1988; Scheidereit *et al.*, 1988]. However, Oct-2 transcripts are found through the length of the neural tube from at E13 and E16, and can also be found in a few regions of the adult brain, including the hypothalamic suprachiasmatic nucleus (SCN) and medial mammillary nucleus (MM) [Treacy and Rosenfeld, 1992].

Thus the developing nervous system is the site of expression for many POU family genes. It must come as no surprise that roles have been found for several of these genes in the generation of the mature central nervous system. The class III POU genes have been shown to be involved in neural differentiation both *in vitro* and *in vivo*. In P19 embryonal carcinoma cells, antisense inhibition of Brn-2 expression prevents differentiation by retinoic acid to the neuronal phenotype, but does not stop expression of muscle markers [Fujii and Hamada, 1993]. P19 cells can also be differentiated by fusion with L cell fibroblasts, to generate a neuroepithelial-like cell which co-expresses Brn-2 and nestin [Shimazaki *et al.*, 1993]. This differentiation seems to require the down-regulation of a class V POU gene, Oct-3, which is expressed only in pluripotent early embryonic cells and some germ cells [Okamoto *et al.*, 1990; Rosner *et al.*, 1990; Schöler *et al.*, 1990]. Other POU genes such as Brn-1, Tst-1, Brn-4, and Emb are not expressed in P19 either before or after retinoic acid differentiation [Fujii and Hamada, 1993] Their unimportance in P19 is counterbalanced by the observation that most neurons derived from P19 are cholinergic [Jones-Villeneuve *et al.*, 1983];

differentiation of other neuronal types may require other transcription factors. Brn-4 is important in *Xenopus laevis* , where its homolog XlPOU2 is activated by the neural inducer *noggin* and has direct neuralizing activity in isolated ectoderm [Witta *et al.*, 1995].

Finally, Brn-2 has been shown to be required for the terminal differentiation and survival of specific neuronal lineages. The disruption of this gene by homologous recombination results in the loss of the paraventricular nucleus (PVH) and supraoptic nucleus (SO) of the hypothalamus as well as the posterior pituitary [Nakai *et al.*, 1995; Schonemann *et al.*, 1995]. The neuroendocrine cells of these two nuclei die either during migration [Nakai *et al.*, 1995], or following migration and the failure of terminal differentiation [Schonemann *et al.*, 1995]. No other brain regions are grossly abnormal, and no disruption of development is seen prior to E12.5. However, Brn-2 is expressed in the forebrain as early as E8.5. Its function in other brain regions, and in hypothalamus prior to E12.5, is likely to be complemented in knockout animals by redundant genes. Since the primordium of the hypothalamus adjacent to the third ventricle co-expresses Brn-1, Brn-2, and Brn-4, it seems possible that POU-III proteins have overlapping functions to a certain extent [Schonemann *et al.*, 1995].

Among the functions of the POU-III proteins may be the transcriptional regulation of the nestin gene. We have discussed a good deal of circumstantial evidence for this role, from the widespread expression of Brn-1, Brn-2 and Brn-4 in the early neural tube to the co-expression of Brn-2 and nestin in the neural differentiation of P19 cells. The work presented in this chapter tests the hypothesis that the nestin CNS-specific element is bound in vitro by POU-III family proteins.

This test is performed by modifying the familiar gel shift assays of the previous chapter. Proteins which bind to labeled DNA probes slow their migration through a nondenaturing gel; when antibodies to the binding protein are added, the size of the complex may be increased. The antibody-containing complex migrates more slowly and is called a supershift. Alternatively, the antibody may block interaction of the protein with DNA and cause the loss of the

shifted band. Supershift assays are an excellent method to identify proteins which bind a sequence of interest *in vitro*, as the high specificity of antibodies for their epitopes allows accurate identification of the target proteins.

Materials and Methods

Antibody supershift assays:

Supershift assays were identical to the EMSAs reported in the previous chapter, except that antibody was added to the protein-DNA binding reactions after 30 minutes at room temperature, and incubated a further 30 minutes before gel loading.
Protein binding reactions were performed by combining10,000 cpm of double-stranded oligonucleotide probe and 4 µgs of tissue nuclear extract with 2.2 µgs poly(dI-dC) and 6.0 µgs BSA in binding buffer (12% glycerol [vol/vol], 12 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid [HEPES] [pH 7.9], 4 mM Tris-HCl [pH 7.9], 60 mM KCl, 1 mM EDTA, 1mM dithiothreitol [DTT].) The binding reaction proceeded at room temperature (23° C) for 30 minutes before antibody addition, then was incubated another 30 minutes before gel loading. Binding reactions were electrophoresed at 200 volts on a 4% 80:1 acrylamide/bis-acrylamide/2.5% glycerol gel.

A second supershifting method was also attempted for antibodies which produced no supershifts by the method above. 1 µL of antibody was mixed with 4 µgs of tissue nuclear extract and TBST (10mM Tris-HCl [pH 7.9], 150mM NaCl, 0.05% Tween-20) to10 µLs. This was incubated 60 minutes on ice prior to the addition of 2.2 µgs poly(dI-dC) and 6.0 µgs BSA in 10 µLs 2x binding buffer. This mixture was incubated 10 minutes on ice before addition of 10,000 cpm of labeled probe, and kept on ice 15 minutes prior to gel loading. Binding reactions were electrophoresed at 200 volts on a 4% 80:1 acrylamide/bis-acrylamide/2.5% glycerol gel at 4° C. Gels were dried 1 hour at 80° C and examined on a Molecular Dynamics PhosphorImager.

Antibodies used for supershift assays:

1G2: guinea pig anti-Brain-1
2GB: guinea pig anti-Brain-2
Tst1: rabbit anti-Tst-1
4R3: rabbit anti-Brain-4
Pit-1: rabbit anti-Pit-1
These antibodies were the generous gift of Dr. Michael G. Rosenfeld. Each was diluted 1 to 10 in 10 mM Tris-HCl [pH 8.0], 1 mM EDTA before use.

RXRβ: rabbit anti-9-*cis*-Retinoic Acid Receptor beta (RXRβ)
This antibody was used undiluted, and was the kind gift of Dr. Keiko Ozato.

3B5: mouse anti-AP-2
This antibody was used undiluted. It was a magnanimous gift from Dr. Trevor Williams.

Anti-CREB and anti-phosphoCREB antibodies were purchased from New England Biolabs, a second anti-phosphoCREB antibody from Upstate Biotechnical, and anti-c-fos from Oncogene Science. Each was used as a 1:10 dilution in 10 mM Tris-HCl [pH 8.0], 1 mM EDTA.

Results

Brn-1 and Brn-2 bind to both octamer sites of the nestin enhancer

Supershift assays were performed to determine whether any of the complexes on the two octamer-containing probes are formed by class III POU proteins. Antibodies to the four POU-III proteins (Brn-1, Brn-2, Tst-1 and Brn-4) were kindly provided by Dr. M.G. Rosenfeld. The double-stranded oligonucleotide probes used here are the same ones described in the previous chapter as "footprint 1" and "footprint 4". These probes were mixed with cortical extracts prior to addition of antibodies.

Figure 19 shows a supershift assay in which antibodies were added to identical binding reactions composed of labeled footprint 1 probe and E17 rat cortex nuclear extract. Lane 1 shows that in the absence of any antibody, E17 cortical tissue produces five strong complexes on the footprint 1 probe. Lane two shows that addition of antibody to Brn-1 blocks the formation of the uppermost complex seen in lane 1, and produces a supershifted complex which has migrated only a short distance from the top of the gel (arrowhead). It appears that the Brn-1 antibody has bound to just one protein-DNA complex and greatly slowed its rate of migration. Therefore, this complex is labeled Brn-1 at the left of the figure.

Similarly, addition of Brn-2 antibody in lane 3 blocks the formation of two complexes seen with cortical extract alone. However, the supershifted complex in lane 3 contains less labeled DNA than that in lane 2. This indicates that the Brn-2 antibody may be interfering with the binding of DNA by its target protein. The two complexes which contain Brn-2 are marked at the left. The difference between the two complexes may arise from phosphorylation, binding of a co-factor, or alternative forms of Brn-2 itself.

Antibodies to Tst-1 and Brn-4 neither supershift nor interfere with the formation of any of the complexes seen here. Neither does an antibody to the 9-*cis*-retinoic acid β receptor (RXRβ), provided by Dr. K. Ozato and used here as a negative control. The last two lanes of the figure are a competion experiment which shows that all five

protein-DNA complexes demonstrate sequence-specific binding to the footprint 1 probe.

The complexes formed on the other octamer-containing probe, Ftpt4, are virtually identical with those on Ftpt1. E17 cortical extract forms five complexes of very similar mobilities on each of the two probes (compare figures 19 and 20). Addition of antibodies show that once again, Brn-1 forms one of these complexes and Brn-2 forms another two. Binding of Tst-1, Brn-4, and RXRβ are not detected. Although their octamer-like sequences differ, the probes footprint 1 and footprint 2 bind the same proteins present in the E17 rat cortex.

Binding of Brn-1 and Brn-2 persists through cortical development while other octamer-binding activities abate

Figure 21 shows a series of supershift assays performed on the footprint 4 probe using cortical extract from E13 rats. The protein-DNA complexes formed by E13 cortex are very similar to those detected at E17, except an for an additional slow-migrating complex. Supershift assays show that one Brn-1 complex and two Brn-2 complexes are again formed at E13. Antibodies to Tst-1 and Brn-4 do not block formation of any of the six complexes, leaving three unidentified.

By postnatal day 5, only two strong complexes remain (see figure 22). Antibody supershifts reveal that these two bands are formed by Brn-1 and Brn-2. A third complex is detected migrating just faster than Brn-2, and is most visible when the Brn-2 band is removed by supershifting (lane 3). This activity may be supershifted by addition of antibody to Tst-1, but confirmation of this may require a double supershift with antibodies to Brn-2 and Tst-1.

Supershift assays were also attempted on the footprint 1 and 4 probes using an antibody to the class I POU protein Pit-1, which is transiently expressed in the developing CNS. However, Pit-1 antibody did not recognize any complexes formed by E13 or E17 cortical extracts on either probe (data not shown).

Activities which recognize other binding sites of the nestin enhancer are not identified by supershift assays

The oligonucleotide probes Ftpt2 and Ftpt3 contain sequences similar to the recognition motif for AP-2 proteins (see chapter III). An antibody raised against the AP-2α protein was provided by Dr. T. Williams and used in supershift assays. Both probes are strongly bound by proteins present in E16 rat spinal cord which are not recognized by the AP-2α antibody (data not shown.)

The probe footprint 5 contains a binding site similar to the cAMP response element or the half-site for non-steroid nuclear receptors (see chapter III). E17 rat cortex nuclear extract produces two strong complexes on this probe. Antibodies against RXRβ, c-fos, Brn-1, CREB, and two antibodies against serine133-phosphorylated CREB were used in efforts to supershift the footprint 5 probe, without success (data not shown).

Figure 19. Supershift assays show that Brn-1 and Brn-2 bind to the Ftpt1 probe *in vitro*

Supershift assays were performed by adding antibodies versus class III POU proteins or RXRβ to EMSAs. In lanes 2-6, antibodies were added to identical EMSAs halfway through the incubation of protein with DNA.

Lane 1: Gel shift with E17 rat cortex nuclear extract on the labeled Ftpt1 probe produces five abundant protein-DNA complexes.

Lane 2: antibody versus Brn-1 protein inhibits formation of the uppermost complex (marked Brn-1 at left), and produces instead a much slower-migrating complex at the top of the gel (arrowhead at left of figure).

Lane 3: antibody versus Brn-2 protein inhibits formation of two complexes, both marked Brn-2 at the left of the figure. A diffuse set of supershifted bands appears near the top of the gel.

Lanes 4-6: antibodies versus Tst-1, Brn-4, or RXRβ have no effect on the typical EMSA pattern.

Lanes 7 and 8 demonstrate that all five complexes bind in a sequence specific manner to the Ftpt1 probe. The unlabeled competitor is Ftpt1 in lane 7, and TREp in lane 8. Both are added at 500-fold molar excess over labeled probe.

Brn-1 and Brn-2 bind to Ftpt1

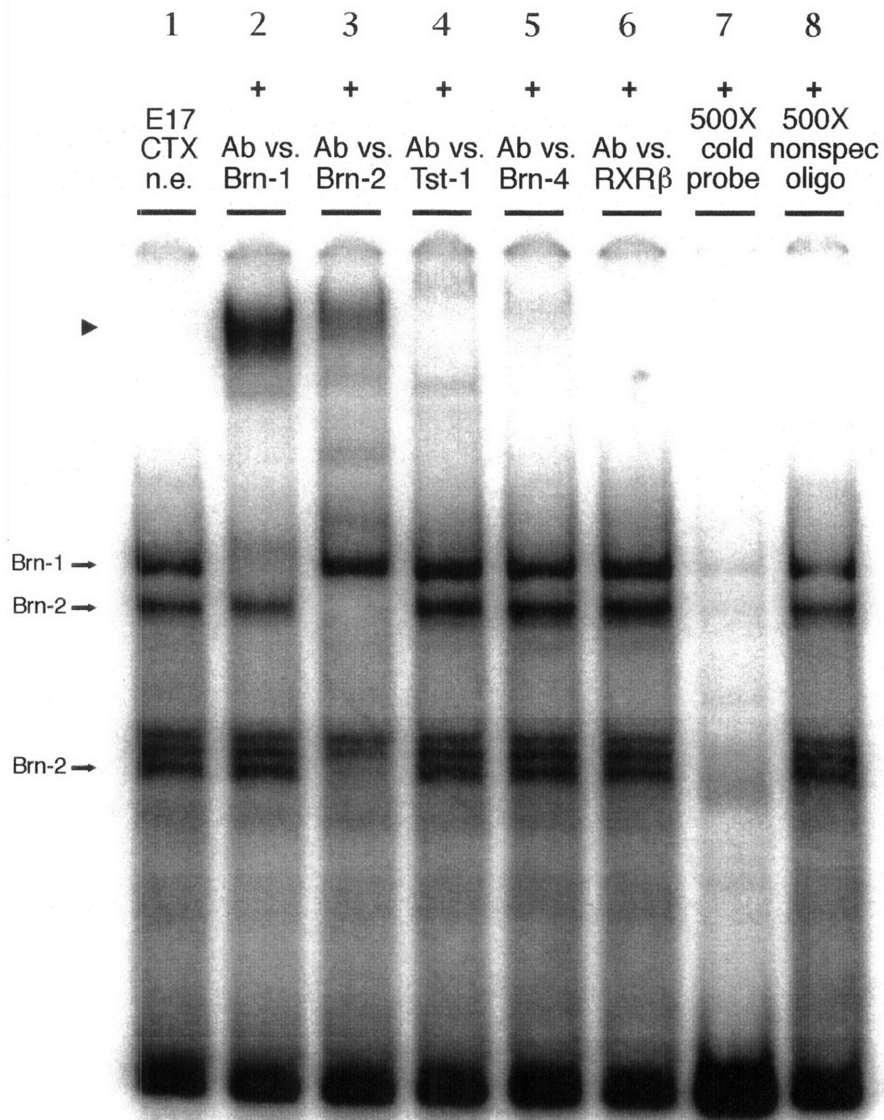


Figure 20. Brn-1 and Brn-2 form complexes on the Ftpt4 probe very similar to those on Ftpt1

Supershift assays with antibodies to class III POU proteins and RXRβ on the Ftpt4 probe.
Lane 1: in the absence of antibody, E17 rat cortex nuclear extract forms five complexes with the labeled Ftpt4 probe. These are similar in mobility and relative abundance to the complexes detected with the Ftpt1 probe (figure 19).
Lane 2: antibody versus Brn-1 inhibits formation of the slowest-migrating complex and produces a supershift (arrowhead at left of figure).
Lane 3: antibody versus Brn-2 blocks formation of two complexes and produces a diffuse set of supershifted bands.
Lanes 4-6: antibodies against Tst-1, Brn-4, or RXRβ have no detectable effect on any of the five protein-DNA complexes.
Lanes 7 and 8 demonstrate that all five complexes bind in a sequence specific manner to the Ftpt4 probe. The unlabeled competitors are Ftpt4 in lane 7, TREp in lane 8. Both are added at 500-fold molar excess over labeled probe.

Brn-1 and Brn-2 bind to Ftpt4

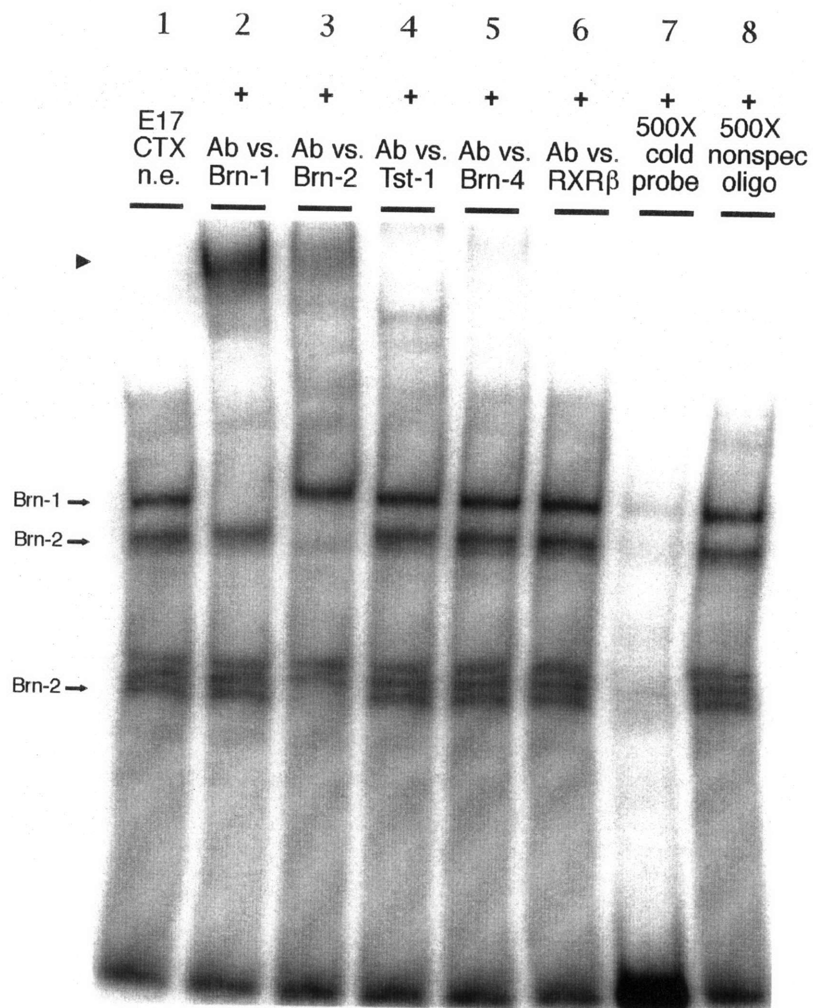


Figure 21. Brn-1 and Brn-2 binding activities are present in telencephalon on embryonic day 13

Antibodies to class III POU proteins were added to incubations of E13 rat telencephalon nuclear extract with labeled Ftpt4 probe.
Lane 1: E13 telencephalon proteins produce six complexes with Ftpt4. Five of these correspond to bands produced by E17 cortex (figure 20), while the slowest-migrating band is only detected at E13.
Lane 2: antibody versus Brn-1 supershifts one band, corresponding to the Brn-1 complex present at E17.
Lane 3: antibody versus Brn-2 supershifts two bands corresponding to the two Brn-2 complexes detected at E17.
Lanes 3 and 4: antibodies to Tst-1 and Brn-4 do not affect any of the six complexes produced by E13 telencephalon extract.

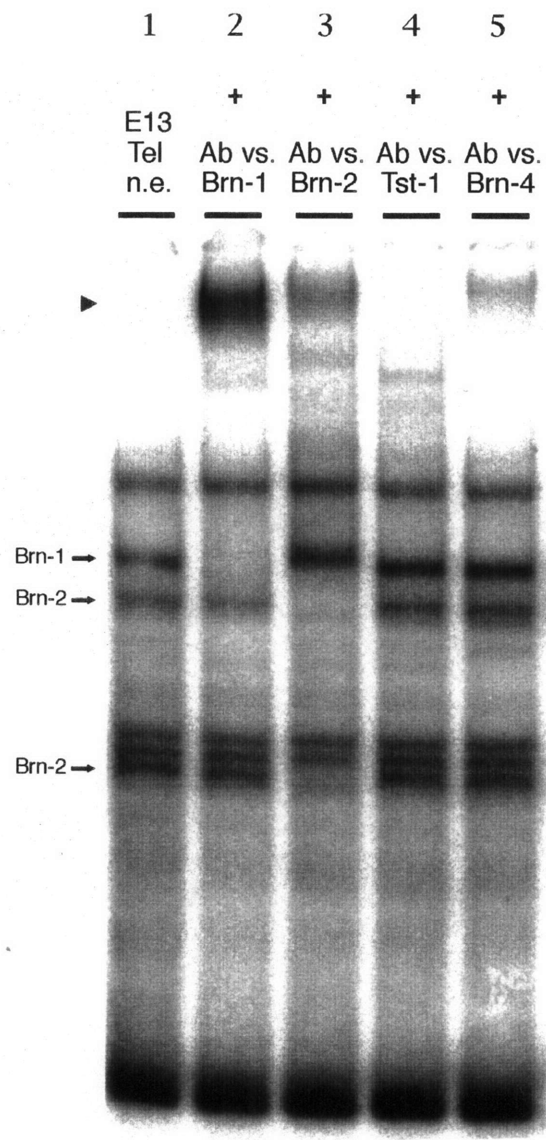
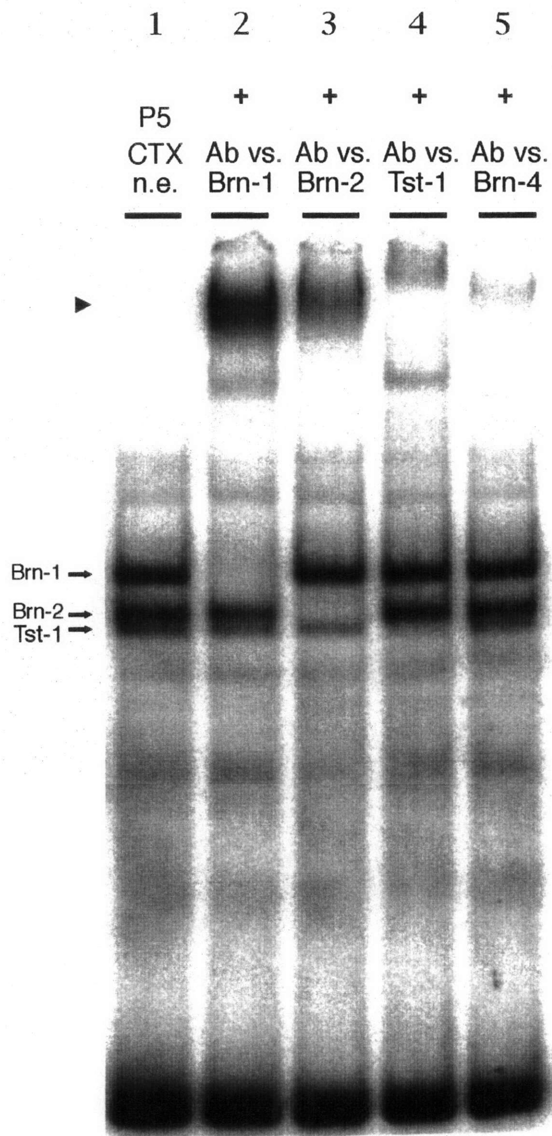

1
2
3
4
5
+
+
+
+
E13
Tel
n.e.
Ab vs.
Brn-1
Ab vs.
Brn-2
Ab vs.
Tst-1
Ab vs.
Brn-4
Brn-1
Brn-2
Brn-2

Figure 22. Brn-1 and a single Brn-2 binding activity persist in cortex on postnatal day 5

Antibodies to class III POU proteins were added to Ftpt4 binding reactions with P5 rat cortex nuclear extract.
Lane 1: only two strong binding activities for Ftpt4 are detected in P5 cortex extract.
Lane 2: antibody versus Brn-1 completely displaces the slower-migrating complex and produces a prominent supershift (arrowhead).
Lane 3: antibody versus Brn-2 inhibits the faster-migrating complex and produces a supershift. When the Brn-2 complex is displaced, a fainter band is visible which runs only slightly ahead of the Brn-2 band in lane 1.
Lane 4: antibody versus Tst-1 appears to displace the weak complex which runs just ahead of Brn-2 (compare doublet in lane 1 with lanes 3 and 4).
Lane 5: antibody versus Brn-4 has no detectable effect on any of the bands produced by P5 cortex.


1
2
3
4
5
+
+
+
+
P5
CTX
n.e.
Ab vs.
Brn-1
Ab vs.
Brn-2
Ab vs.
Tst-1
Ab vs.
Brn-4
Brn-1
Brn-2
Tst-1

Discussion

Antibody supershift assays are a very powerful method for identifying DNA-binding proteins *in vitro*. Assuming that antibodies to candidate transcription factors are available from colleagues, supershift assays are simpler and faster than using an oligonucleotide probe to purify a binding protein or screen an expression library. Also, the precision of the identification made by a supershift is as good as the specificity of the antibody itself. Another simple means of identifying a *cis*-acting element is to assay it for binding by a purified transcription factor; however such an identification is tenuous because multiple factors within a given family can usually bind to the same site, and members of different families sometimes cooperate or compete for binding to a single site. Supershift assays are therefore the best method for discriminating among several members of a known transcription factor family.

Numerous members of the POU transcription factor family are good candidates for binding to the two octamer-like sites of the nestin neural enhancer. Supershift assays have identified two of these, Brn-1 and Brn-2, among the activities which recognize both octamer sites in rat CNS extracts. Both proteins are detected by supershifts through the period of neurogenesis in the cortex: in E13 telencephalon, E17 cortex, and P5 cortex. Furthermore, the two slow-migrating bands identified as Brn-1 and Brn-2 by antibody are present in gel shifts using extracts from all levels of the neural tube (see figure 12).

Brn-2 is also detected in the E13 and E17 telencephalon as a faster-migrating complex of lower abundance (figures 19-21). This form of Brn-2 activity, along with two unidentified bands of similar mobilities, appears to be specific to the forebrain (figure 12). This second activity is not detected at P5, an age by which neurogenesis in the cortex has ended [Berry *et al.*, 1964]. Perhaps the difference between the transient and persistent forms of Brn-2 could be detected in immunoprecipitation assays to look for phosphorylation or associated proteins.

Another significant feature of Brn-1 and Brn-2 binding revealed by supershifting is that the two proteins are not present in

a single complex with DNA. This suggests either that these two proteins compete for binding to the Ftpt1 and Ftpt4 sites or that they are not present in the same cells. Extensive co-expression of Brn-1 and Brn-2 has been shown by *in situ* hybridization [Alvarez-Bolado *et al.*, 1995], so it seems likely that they compete.

Brn-1 and Brn-2 also appear to compete for binding to the two octamers with several unidentified proteins. Supershift assays for the other POU-III proteins revealed only a minor Tst-1 band at P5 (figure 22), and no Brn-4 binding was detected. Methylation interference on the two octamer sites showed that all of the isolatable complexes recognized the same bases (figure 16d), which suggests that all contain POU proteins. Many of the known POU genes are expressed in the brain at some period. This includes the ubiquitous Oct-1, the transient CNS expression of Pit-1 and Oct-2 [Treacy and Rosenfeld, 1992], the three Brn-3 genes [Gerrero *et al.*, 1993; He *et al.*, 1989; Lillycrop *et al.*, 1992] and Brn-5/Emb [Andersen *et al.*, 1993] as well as the class III POUs assayed here. In these experiments one binding activity was strong only at E13 (figure 21), so it was hypothesized to be Pit-1; however, antibody to Pit-1 protein did not recognize any of the complexes on either octamer probe.

Similarly, we have proposed several known transcription factors as candidate binding activities for the other three footprinted regions. Footprint 5 is suggested to be a cyclic AMP response element or a non-steroid nuclear hormone receptor binding site, based on sequence homologies and its inverted repeat structure. However, no supershifts have yet been produced on this site using antibodies against CREB, phosphorylated CREB, c-fos, or RXRβ. Also, footprints 2 and 3 are proposed to be AP-2 binding sites, based on sequence homology with the AP-2 consensus motif, and the predominance of binding activity in the posterior CNS. Yet this binding activity is not recognized by an antibody to AP-2.

These negative results with various antibodies do not absolutely rule out the tentative identifications of transcription factors which we have made. First, not all antibodies are capable of supershifting their target protein. This may be because some

antibodies are generated against denatured proteins or peptide fragments for use in western blots, and cannot recognize their epitope under non-denaturing conditions. Alternatively, the binding conditions for gel shifts may not be compatible in other ways with the antibody's function. Second, consensus binding motifs are usually recognized by a number of related transcription factors, while the antibodies are specific to one or a few factors.

For example, one of the CREB antibodies used here binds an epitope common to CREB, CREM, and ATF-1 [Ginty *et al.*, 1993]. The antibody recognizes all three proteins in western blots, but its range in supershifts is unknown. However, there are additional proteins in this family which also recognize the same DNA sequence [reviewed by Meyer and Habener, 1993]. This site also has homology to the non-steroid hormone response elements. Although no supershift was produced with an antibody to RXRβ, this is only one of three RXRs which are promiscuous heterodimerization partners for the retinoic acid receptors(RARs), thyroid hormone receptors (T3Rs), and vitamin D3 receptors (VDRs)[reviewed by Mangelsdorf and Evans, 1995]. In addition, such sites can be bound by several dimeric or monomeric orphan receptors. Finally, we have attempted to supershift footprints 2 and 3 with an antibody against the human AP-2, without success. However, at least three proteins are known which bind AP-2 sites: AP-2, AP-2β, and AP-2.2 [Chazaud *et al.*, 1996; Moser *et al.*, 1995; Williams *et al.*, 1988]. Thus the identification of all the binding activities on the nestin enhancer could be a slow process.

On the other hand, it has been possible to make a few successful guesses of candidate transcription factors. Class III POU genes were proposed to bind the nestin enhancer, based on recognition site homology, mRNA localization patterns, and co-expression in the P19 cell line. Supershift assays then confirmed that Brn-1 and Brn-2 bind to two footprinted sites of the second intron, and also found a minor interaction with Tst-1 in postnatal tissue. Interestingly, Brn-2 is observed in two forms: a transient binding activity which appears to be specific to the early forebrain, and a persistent activity found in all tested regions of the CNS during and after neurogenesis. In addition, a few unidentified binding activities

are detected on the same sites and are also present only during the embryonic period. Further work is necessary to determine if these are also POU proteins.

References

Alvarez-Bolado, G., M.G. Rosenfeld and L.W. Swanson (1995) Model of forebrain regionalization based on spatiotemporal patterns of POU-III homeobox gene expression, birthdates, and morphological features. J. Comp. Neurol. **355**, 237-295.

Andersen, B. and M.G. Rosenfeld (1995) Pit-1 determines cell types during development of the anterior pituitary gland. J. Biol. Chem. **269**, 29335-29338.

Andersen, B., M.D. Schonemann, R.d. Pearse, K. Jenne, J. Sugarman and M.G. Rosenfeld (1993) Brn-5 is a divergent POU domain factor highly expressed in layer IV of the neocortex. J. Biol. Chem. **268**, 23390-8.

Berry, M., A.W. Rogers and J.T. Eayrs (1964) Pattern of cell migration during cortical histogenesis. Nature **203**, 591-593.

Bodner, M., J.L. Castrillo, L.E. Theill, T. Deerinck, M. Ellisman and M. Karin (1988) The pituitary specific transcription factor GHF-1 is a homeobox containing protein. Cell **50**, 267-275.

Chazaud, C., M. Oulad-Abdelghani, P. Bouillet, D. Decimo, P. Chambon and P. Dolle (1996) AP-2.2, a novel gene related to AP-2, is expressed in the forebrain, limbs and face during mouse embryogenesis. Mech. Dev. **54**, 83-94.

Clerc, R.G., L.M. Corcoran, J.H. Lebowitz, D. Baltimore and P.A. Sharp (1988) The B-cell specific Oct-2 protein contains POU box and homeo-box type domains. Genes Dev. **2**, 1570-1581.

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

Dahlstrand, J., M. Lardelli and U. Lendahl (1995) Nestin mRNA expression correlates with the central nervous system progenitor cell state in many, but not all, regions of developing central nervous system. Dev. Brain Res. **84**, 109-129.

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

Finney, M., G. Ruvkun and H.R. Horvitz (1988) The *C. elegans* cell lineage and differentiation gene unc-86 encodes a protein

containing a homeodomain and extended sequence similarity to mammalian transcription factors. Cell **55**, 757-769.

Fujii, H. and H. Hamada (1993) A CNS-specific POU transcription factor, brn-2, is required for establishing mammalian neural cell lineages. Neuron **11**, 1197-1206.

Gerrero, M.R., R.J. McEvilly, E. Turner, C.R. Lin, S. O'Connell, K.J. Jenne, M.V. Hobbs and M.G. Rosenfeld (1993) Brn-3.0: a POU-domain protein expressed in the sensory, immune, and endocrine systems that functions on elements distinct from known octamer motifs. Proc. Natl. Acad. Sci. U S A **90**, 10841-5.

Ginty, D.D., J.M. Kornhauser, M.A. Thompson, H. Bading, K.E. Mayo, J.A. Takahashi and M.E. Greenberg (1993) Regulation of CREB phosphorylation in the suprachiasmatic nucleus by light and a circadian clock. Science **260**, 238-241.

Hara, Y., A.C. Rovescalli, Y. Kim and M. Nirenberg (1992) Structure and evolution of four POU domain genes expressed in mouse brain. Proc. Natl. Acad. Sci. USA **89**, 3280-3284.

He, X., M.N. Treacy, D.M. Simmons, H.A. Ingraham, L.W. Swanson and M.G. Rosenfeld (1989) Expression of a large family of POU-domain regulatory genes in mammalian brain development. Nature **340**, 35-42.

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

Jones-Villeneuve, E.M.V., M.A. Rudnicki, J.K. Harris and M.W. McBurney (1983) Retinoic acid-induced neural differentiation ofembryonal carcinoma cells. Mol. Cell. Biol. **3**, 2271-2279.

Ko, H.S., P. Fast, W. McBride and L.M. Staudt (1988) A human protein specific for the immunoglobulin octamer DNA motif contains a functional homeodomain. Cell **55**, 135-144.

Le Moine, C. and W.S.I. Young (1992) RHS2, a POU domain-containing gene, and its expression in developing and adult rat. Proc. Natl. Acad. Sci. USA **89**, 3285-3289.

Lillycrop, K.A., V.S. Budharan, N.D. Lakin, G. Terrenghi, J.N. Wood, J.M. Polak and D.S. Latchman (1992) A novel POU family transcription factor is closely related to Brn-3 but has a distinct expression pattern in neuronal cells. Nucl. Acids Res. **20**, 5093-5096.

Mangelsdorf, D.J. and R.M. Evans (1995) The RXR heterodimers and orphan receptors. Cell **83**, 841-850.

Mathis, J.M., D.M. Simmons, X. He, L.W. Swanson and M.G. Rosenfeld (1992) Brain 4: A novel mammalian POU domain transcription

factor exhibiting restricted brain-specific expression. EMBO J. **11**, 2551-2561.
Meyer, T.E. and J.F. Habener (1993) Cyclic adenosine 3',5'-monophosphate response element binding protein (CREB) and related transcription-activating deoxyribonucleic acid-binding proteins. Endocr. Revs. **14**, 269-290.
Monuki, E.S., R. Kuhn, G. Weinmaster, B. Trapp and G. Lemke (1990) Expression and activity of the POU transcription factor SCIP. Science **249**, 1300-1303.
Moser, M., A. Imhof, A. Pscherer, R. Bauer, W. Amselgruber, F. Sinowatz, F. Hofstader, R. Schule and R. Buettner (1995) Cloning and characterization of a second AP-2 transcription factor: AP-2β. Development **121**, 2779-2788.
Muller, M., S. Ruppert, W. Schaggner and P. Matthias (1988) A cloned octamer transcription factor stimulates transcription from lymphoid specific promoters in non-B cells. Nature **336**, 544-551.
Nakai, S., H. Kawano, T. Yudate, M. Nishi, J. Kuno, A. Nagata, K. Jishage, H. Hamada, H. Fujii, K. Kawamura, K. Shiba and T. Noda (1995) The POU domain transcription factor Brn-2 is required for the determination of specific neuronal lineages in the hypothalamus of the mouse. Genes Dev. **9**, 3109-3121.
Okamoto, K., H. Okazawa, A. Okuda, M. Sakai, M. Muramatsu and H. Hamada (1990) A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell **60**, 461-472.
Okamoto, K., M. Wakamiya, S. Noji, E. Koyama, S. Taniguchi, R. Takemura, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, M. Muramatsu and H. Hamada (1993) A novel class of murine POU gene predominantly expressed in central nervous system. J. Biol. Chem. **268**, 7449-7457.
Rosner, M., M.A. Vigano, K. Ozato, P.M. Timmons, F. Poirier, P.W.J. Rigby and L.M. Staudt (1990) A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature **345**, 686-692.
Scheidereit, C., J.A. Cromlish, T. Gerster, K. Kawakami, C.G. Balmaceda, R.A. Currie and R.G. Roeder (1988) A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein. Nature **336**, 552-557.
Scholer, H.R., S. Ruppert, R. Balling, N. Suzuki, K. Chowdhury and P. Gruss (1990) New type of POU domain in germ line-specific protein Oct-4. Nature **344**, 435-439.

Schonemann, M.D., A.K. Ryan, R.J. McEvilly, S.M. O'Connell, C.A. Arias, K.A. Kalla, P. Li, P.E. Sawchenko and M.G. Rosenfeld (1995) Development and survival of the endocrine hypothalamus and posterior pituitary gland requires the neuronal POU domain factor Brn-2. Genes Dev. **9**, 3122-3135.

Shimazaki, T., H. Okazawa, H. Fujii, M. Ikeda, K. Tamai, R.D.G. McKay, M. Muramatsu and H. Hamada (1993) Hybrid cell extinction and re-expression of Oct-3 function correlates with differentiation potential. EMBO J. **12**, 4489-4498.

Sturm, R.A., G. Das and W. Herr (1988) The ubiquitous octamer binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes Dev. **2**, 1582-1599.

Suzuki, N., H. Rohdewohld, T. Neuman, P. Gruss and H.R. Scholer (1990) Oct-6: A POU transcription factor expressed in embryonal stem cells and in the developing brain. EMBO J. **9**, 3723-3732.

Treacy, M.N. and M.G. Rosenfeld (1992) Expression of a family of POU-domain protein regulatory genes during development of the central nervous system. Annu. Rev. Neurosci. **15**, 139-165.

Wegner, M., D.W. Drolet and M.G. Rosenfeld (1993) POU-domain proteins: Structure and function of developmental regulators. Curr. Opin. Cell Biol. **5**, 488-498.

Williams, T., A. Admon, B. Luscher and R. Tjian (1988) Cloning and expression of AP-2, a cell-type-specific transcription factor that activates inducible enhancer elements. Genes Dev. **2**, 1557-1569.

Witta, S.E., V.R. Agarwal and S.M. Sato (1995) XlPOU2, a noggin-inducible gene, has direct neuralizing activity. Development **121**, 721-730.

Conclusions: A Model of the Nestin CNS-Specific Enhancer Complex

The experiments presented here lead to a model of the nestin neural enhancer and its associated transcription factors (see figure 23). The 257 base pairs of nestin's second intron which are sufficient to drive CNS-specific expression *in vivo* include five nuclear protein binding sites detected *in vitro*. Candidate *trans*-acting factors for each of these sites have been proposed on the basis of homology to known binding motifs.

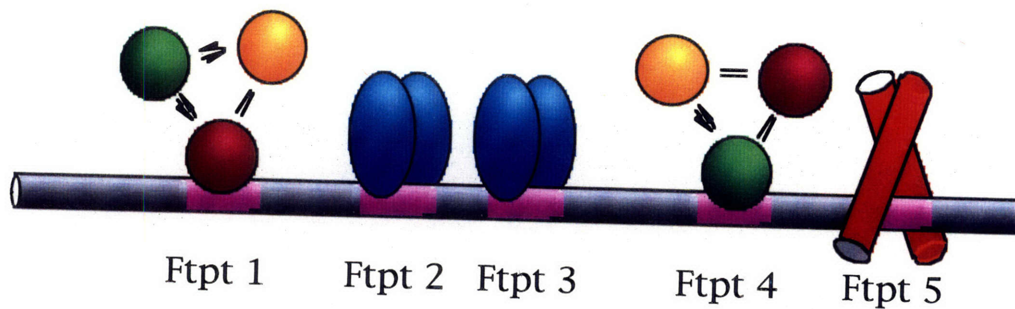
Three factors which bind the nestin enhancer in vitro have been positively identified by the use of antibodies. The POU family proteins Brn-1 and Brn-2 interact with a pair of sequences similar to the canonical octamer site. In postnatal brain, Tst-1 also binds these sites. Each of these proteins forms a separate complex with DNA, and probably competes with the others and at least one yet-unidentified POU protein for binding to DNA. The two octamer-like sites flank a pair of sites for an unknown factor which may be a member of the AP-2 family. Both AP-2-like sites are fully occupied only in posterior portions of the CNS such as spinal cord and cerebellum, which strengthens the argument that the binding protein is one of the caudally restricted AP-2α or β proteins. A much weaker binding activity for these two sites is detected in the E16 cortex; this could be the related AP-2.2 protein which is highly expressed in forebrain [Chazaud *et al.*, 1996].

The fifth binding site in the nestin enhancer has homology to two kinds of transcription factor motifs, the cAMP response element and the half-site for non-steroid nuclear hormone receptors. Heterodimers of thyroid hormone receptor (T3R) and 9-*cis*-retinoic acid receptors (RXRs) sometimes bind a single half-site [Marks *et al.*, 1992], as do the monomeric orphan receptors NGFI-B and Nurr1 [Mangelsdorf and Evans, 1995]. However, the inverted repeat structure of the binding site in the second intron augments its resemblance to the cAMP response element. Such an element can be bound by numerous proteins of the CRE-binding family or fos/jun families in a variety of homo- or heterodimeric complexes.

Figure 23. Model of the nestin CNS-specific enhancer

In vitro assays show that the nestin CNS-specific element is composed of five footprinted sites within a region of 257 base pairs.
Ftpt 1 and Ftpt 4 are octamer-containing sites. Brn-1 and Brn-2 compete with each other for binding to these sites, and also with Tst-1 and unidentified POU proteins. These sites are occupied in all tested regions of the CNS, but some binding activities, including a form of Brn-2, are found only in forebrain.
Ftpt 2 and Ftpt 3 are related to the consensus binding motif for dimeric AP-2 proteins. These two sites are fully occupied only in posterior CNS regions such as spinal cord and cerebellum.
Ftpt 5 is related to the cyclic AMP response element, and also to the half-site for non-steroid nuclear hormone receptors. Therefore it may be bound by dimeric CRE-binding proteins from the CREB or AP-1 families, or by dimeric hormone receptors or monomeric orphan receptors (nuclear hormone receptors, NHR in figure). One binding activity for this site is present throughout the CNS, another is found only in forebrain.

Model of the CNS-Specific Enhancer of the Nestin Gene



= Brn-1

= Brn-2

= Other POUs

= AP-2?

= CREB or NHR or AP-1?

Table I: Summary of binding activities on the nestin CNS-specific enhancer

This table summarizes information on the *in vitro* binding activities for each of the five footprinted sites of the 257 bp CNS enhancer, and the effects of deletion of sites *in vivo*. The five footprinted regions are represented by five columns in the table.
Row 1 represents transgenic mice made with the construct 257gIITKlacZ (see figure 7). All five footprinted sites are contained in the nestin genomic fragment of this construct. The pattern of reporter expression in these mice is summarized at the right; lacZ staining is seen throughout the CNS in all embryos which express the enzyme.
Row 2 represents transgenic mice produced with a 70 bp deletion from the enhancer (ΔRIRI NsiSalgIITKlacZ, see figure 4). The deletion removes footprints 1 and 2, so the mice carry only footprints 3-5 of the enhancer. Expression in these embryos is CNS-specific, but fills the entire CNS in only one mouse. All others express in portions of the CNS, mainly in midbrain or anterior CNS.
Row 3 represents transgenic mice made with an upstream portion of the region most highly conserved between rat and human second intron. This construct (gIICons5'TKlacZ) shares 136 bp with the 257 bp enhancer, including footprints 1-3 but not 4 or 5. Transgenic embryos (see figure 6) show various ectopic patterns of expression; half show staining of the spinal cord or spinal cord and midbrain in addition to ectopic patches.

Row 4 summarizes results of EMSAs. Each of the five footprints is bound by activities which are present in all tested regions of the CNS as well as by activities with more limited expression domains. Binding activities for footprints 2 and 3 are most prevalent in cerebellum and spinal cord, but not limited to these areas. Footprints 1, 4, and 5 are recognized by binding activities which are limited to the cortex.
Row 5 lists the best sequence homologies found for each footprint based on methylation interference at the core binding sites. The octamer is the consensus recognition motif for POU proteins; AP-2 is the consensus site for a family of at least three AP-2 genes; CRE is the cyclic AMP response element, and NHRE refers to a half-site which mediates binding by nuclear hormone receptors.
Row 6 summarizes the results of antibody supershift assays. By this method, the POU proteins Brn-1, Brn-2, and Tst-1 have been shown to recognize the nestin CNS-specific enhancer *in vitro*.

Table I: Summary of Binding Activities on the Nestin CNS-Specific Enhancer

		Footprint1	Footprint2	Footprint3	Footprint4	Footprint5	LacZ Expression:
1	257gII-TKlacZ	Yes	Yes	Yes	Yes	Yes	Strong throughout CNS
2	ΔRIRINS-gIITKlacZ	No	No	Yes	Yes	Yes	CNS-spec. Brain > SpC
3	gIICons5'-TKlacZ	Yes	Yes	Yes	No	No	Var. Ectopic some SpC or Mid
4	Tissue-spec. Binding	•All CNS •Cortex	•All CNS •Cb/SpC	•All CNS •Cb/SpC	•All CNS •Cortex	•All CNS •Cortex	
5	Sequence Homology	Octamer	AP-2	AP-2	Octamer	CRE NHRE	
6	Binding Proteins	**Brn-1 Brn-2**			**Brn-1 Brn-2 Tst-1**		

An interesting feature of the binding sites in the nestin CNS enhancer is the occurrence of octamers and AP-2-like sites in pairs. This suggests there may be cooperativity between the similar factors. Some POU proteins such as Oct-2 and Pit-1 display cooperative binding among multiple sites. There may also be interactions between different transcription factor families on the enhancer, just as Pit-1 functions synergistically with estrogen receptors or thyroid hormone receptors on the enhancers of prolactin or growth hormone genes [reviewed by Wegner *et al.*, 1993]. Therefore, the correct regulation of the nestin gene may depend on interactions between similar and dissimilar proteins bound to the enhancer.

The 257 base pair enhancer which contains these five binding sites is sufficient for CNS-specific expression at all embryonic ages tested. Does the identification of potential trans-acting factors indicate how this tissue-specific pattern is generated? Several of the transcription factors being considered here can be induced by intercellular signaling molecules, which suggests that nestin expression may be regulated by such signals.

Retinoic acid is a good example of a signaling molecule which is involved in CNS development and also induces some of our candidate transcription factors. In P19 embryonal carcinoma cells, retinoic acid induces expression of AP-2 and Brn-2, as well as nestin. The expression of these *trans*-acting factors is required for differentiation of P19 to a neuronal fate [Fujii and Hamada, 1993; Philipp *et al.*, 1994]. Retinoic acid can also alter CNS fates *in vivo*, leading to anterior-posterior transformation of cell types in the Xenopus neural tube [Durston *et al.*, 1989; Sive *et al.*, 1990] (see introduction).

In addition to being induced by retinoic acid, AP-2 is post-translationally activated by phorbol ester stimulation of protein kinase C or cAMP activation of protein kinase A [Imagawa *et al.*, 1987]. AP-2 is thus a part of two signal transduction pathways. These two pathways also activate factors which may bind another site in the nestin enhancer: CREB family members are activated by cAMP-dependent phosphorylation, while AP-1 activity is increased by phorbol ester [Kruijer *et al.*, 1985]. In the human proenkephalin

gene, an AP-2 site synergistically interacts with two CREs to increase the responses to cAMP and phorbol esters [Hyman *et al.*, 1989].

The protein kinase C pathway which activates AP-1 responds to a number of extracellular signals, including basic fibroblast growth factor (bFGF). Basic FGF induces expression of another CNS intermediate filament, vimentin, via AP-1 sites [Carey and Zehner, 1995] and also stimulates proliferation of nestin- and vimentin-expressing neural precursor cells [Vicario-Abejon *et al.*, 1995]. AP-1 may therefore be a transcriptional activity which regulates both the nestin gene and the proliferation state of neuroepithelial stem cells in response to bFGF.

On the other hand, the expression patterns of Brn-1 and Brn-2 nearly parallel the pattern of nestin expression in the early neural tube. The exception is the absence of these POU genes from the dorsal hindbrain, where nestin is transcriptionally active. One might conclude that the CNS-specific activity of the nestin second intron is largely established by Brn-1 and Brn-2, with help from other factors in the hindbrain. However, nestin expression is shut off in newborn neurons while Brn-1 and Brn-2 persist in many neurons [Alvarez-Bolado *et al.*, 1995]. The down-regulation of nestin may be due to changes in other transcription factors or members of the enhancer-binding complex; the shutoff is not post-transcriptional, because the heterologous reporter lacZ is similarly affected.

Another possibility is that the apparent competition between Brn-1 and Brn-2 for binding to the two octamer sites is involved in shutting off transcription. In P19 cells, Brn-2 and nestin are co-expressed upon differentiation with retinoic acid [Fujii and Hamada, 1993], suggesting a positive effect of Brn-2 on nestin transcription. Brn-1 is not expressed in P19, but *in vivo* has virtually the same expression pattern as Brn-2 from the neural tube stage to adulthood [Alvarez-Bolado *et al.*, 1995]. Perhaps Brn-1 opposes the action of Brn-2, with the balance of activity between the co-expressed proteins determined by association with additional factors or changes in phosphorylation. On the other hand, the two POU proteins may be merely redundant in most cells, as deletion of the Brn-2 gene affects

only a few nuclei of the hypothalamus [Nakai *et al.*, 1995; Schonemann *et al.*, 1995].

The balance of Brn-1, Brn-2, and other POU-III genes present in a neuroepithelial cell may also have important implications for that cell's developmental potential. [Alvarez-Bolado *et al.*, 1995] have painstakingly mapped expression of the four class III POU genes in the developing forebrain, and conclude that these genes are involved in dividing the neural tube into numerous subregions. For example, Brn-4 expression distinguishes the diencephalon from the telencephalon before an anatomical boundary is apparent. Later, both of these regions are fragmented into small areas expressing distinct combinations of Brn-1, Brn-2, Tst-1 and Brn-4 which often foreshadow anatomical divisions. However, these regions defined by combinations of POU-III genes do not correspond to the neuromeric divisions of the forebrain described elsewhere [Figdor and Stern, 1993; Puelles and Rubinstein, 1993].

If combinatorial patterns of Brn-1 and Brn-2 establish the neural expression domain of the nestin gene, deletion of their octamer recognition sites ought to block CNS-specific expression. Generation of transgenic mice with such a mutated enhancer is underway, but previous *in vivo* results with partial nestin enhancers could be informative. In chapter 2, transgenic mice were made with the upstream portion of a sequence highly conserved between rat and human nestin second introns (gIICons5'TKlacZ, figure 6). This construct contains the first three footprinted sites of the nestin enhancer and lacks footprints 4 and 5 (see Table I). The effect on tissue-specific expression of removing one octamer site and one CRE or nuclear receptor half-site is severe. Four embryos stained for lacZ, and all four expressed it in different regions outside the CNS. However, two embryos also had expression in the CNS. One shows strong staining of the spinal cord and hindbrain only, and the other has relatively weak staining in the midbrain and portions of the spinal cord (figure 6). Therefore, footprints 4 and 5 are necessary for CNS-specific expression, although the other three *cis*-elements can sometimes produce expression in the posterior CNS.

The effect of deleting footprinted sites 1 and 2 was also reported in chapter 2 (Δ RIRINsiSal gIITKlacZ, figure 4). Deletion of 70 base pairs from the enhancer region in this reporter construct removed one octamer site and one AP-2-like site (see Table I). The lacZ reporter in these mice is still expressed specifically in the CNS, perhaps because one octamer, one AP-2-like site, and footprint 5 remain. However, most of the mice express lacZ at much lower levels in the spinal cord than in the head. This suggests that expression in the spinal cord may depend on different *cis*-elements than does expression in the anterior CNS. However, both transgenic constructs reported here include second intron sequences from outside the 257 bp element. These may partially compensate for missing binding sites.

The five identified binding sites of the nestin enhancer seem to recognize different sets of *trans*-acting factors in different brain regions (see Table I). DNAse I footprinting and gel shift assays show variations in the enhancer binding activities detected in the cortex, midbrain, cerebellum and spinal cord. Some proteins, such as Brn-1 and Brn-2, are detected in similar quantities in all four tissues. In contrast, an AP-2-like activity is strongly present in cerebellum and spinal cord but weak or undetectable in midbrain and cortex. The AP-2 genes α and β have been reported to be expressed in the E13.5 mouse CNS at all levels except the forebrain [Moser *et al.*, 1995], while a third family member, AP-2.2, is expressed specifically in forebrain at the same age [Chazaud *et al.*, 1996]. Two strong binding activities are detected on the CRE-like footprint 5; one is present in all four brain regions tested and the other is detected only in extracts from cortex (see figure 13). Similarly, a group of three binding activites for the two octamer-containing footprints 1 and 4 are found only in cortex during neurogenesis (figure 12). One of these activities is a form of Brn-2 (figures 19-21). Thus the complex of proteins formed on the nestin enhancer *in vivo* may have different components in various brain regions.

This requirement for distinct sets of transcription factors in different regions of the brain to produce continuous nestin expression throughout the neural tube could reflect the evolutionary

history of the brain. The size and complexity of the mammalian forebrain has increased with time. However, many of the transcription factors expressed in the mammalian forebrain are homologs of the genes which pattern the fly head, such as *orthodenticle*, *empty spiracles*, and *distalless* [Finkelstein and Boncinelli, 1994]. Perhaps the transcriptional complex on the enhancer of an ancestral nestin-like gene evolved in parallel, adding new factors in the forebrain which increase the proliferation of forebrain precursors.

Finally, *in situ* hybridization studies have shown that nestin expression does occur in some areas of the adult brain. These areas include the subventricular zone, which is the proliferative region of the forebrain in the late embryonic period; the piriform cortex; the olfactory bulb and primary olfactory cortex; the arcuate and paraventricular nuclei of the hypothalamus; and subsets of basal forebrain nuclei [Dugich-Djordjevic *et al.*, in preparation]. In addition, nestin can be found in the adult brain in EGF-responsive precursors, reactive astrocytes, and some tumor types [Clarke *et al.*, 1994; Craig *et al.*, 1996; Dahlstrand *et al.*, 1992]. Nestin marks proliferating populations in the adult subventricular zone and hypothalamus, while cell division is not detected in the other nestin-expressing regions of the normal adult brain. Many of these same adult brain regions also express the POU-III genes Brn-1, Brn-2, or Brn-4 [Alvarez-Bolado *et al.*, 1995]. We are currently investigating whether these POU proteins co-localize with nestin expression in the adult, and whether this nestin expression is dependent on the second intron enhancer. Furthermore, we wish to know whether the POU proteins are part of the pathway which controls the proliferation of precursors in the adult brain in response to growth factors or brain injury.

The results presented here could be extended in a number of interesting directions by additional experiments. First, several nestin enhancer-binding proteins remain to be identified, and these are being pursued by supershift assays with antibodies to the three known AP-2 family members and other proteins. For footprint 5, which is similar to binding motifs of several transcription factor

families, the list of candidate proteins may be narrowed down by cross-linking protein to its bromodeoxyuridine-substituted binding site with ultraviolet light and determining the molecular weight of the bound protein on a denaturing protein gel [Lin and Riggs, 1974]. Some of the unidentified binding activities may be novel factors; it could be worthwhile to screen an embryonic brain expression library with binding site probes.

Another group of experiments which is underway follows up on the idea that individual binding sites of the nestin enhancer may be required for expression in only some portions of the central nervous system. Transgenic mice are being made from a 1.8 kb nestin second intron-TK promoter-lacZ reporter construct with single binding sites mutated. Analysis of these mice at various ages may also suggest whether the extinction of nestin expression is simply the reverse of transcriptional initiation or requires a separate silencing site. Silencing of expression could also require new negative transcriptional regulators which may compete with positive regulators for binding to the enhancer. It is possible to test the positive or negative effects of the identified enhancer-binding factors by co-transfecting cDNAs for these into cell lines along with a nestin enhancer-lacZ or -chloramphenicol acetyltransferase (CAT) reporter construct to quantitatively measure the effects of these factors on transcription. Of course, different factors may be rate-limiting at different times in the life of a precursor cell. It would therefore be best to perform the cotransfection assay not only in a nestin-expressing cell line such as ST15A, but also in committed cell lines and in embryonic stem cells, which do not express nestin but can be induced to do so [Shimazaki *et al.*, 1993].

Furthermore, the positive or negative influence of any single factor on nestin transcription may be less important than the combinatorial interactions among multiple transcription factors. The data presented here suggests that some of the nestin enhancer-binding factors, such as the POU proteins Brn-1, Brn-2, and Tst-1 may compete with each other for binding to a single octamer site; also, cooperative interactions may occur between POU proteins bound to the two octamer sites, between proteins bound to the two AP-2-

like sites, or between members of different transcription factor families. Using purified protein, accurate equilibrium and kinetic binding constants of protein-DNA interactions can be measured by a quantitative gel shift assay [Chodosh *et al.*, 1986]. In addition, cooperative binding interactions among multiple DNA sites can be assessed using a DNAse I footprint titration method [Brenowitz *et al.*,].

One further level of complexity in nestin regulation suggested by this work is the post-translational regulation of transcription factor activity. A second form of Brn-2 was found in E13 and E17 telencephalon; this may represent a covalently modified form of the factor, or indicate an associated cofactor. The difference between the two forms of Brn-2 can be studied by immunoprecipitation with an antibody to Brn-2, followed by electrophoresis and western blotting to detect additional proteins or Brn-2 phosphorylation. Again, quantitative binding assays could determine the effect of such modification on Brn-2 activity.

In conclusion, the goal of this work has been to identify transcription factors which regulate the nestin gene in the central nervous system. Because nestin expression is very tightly linked to proliferation, it was anticipated that this study would detect transcription factors important in specifying multipotential precursor cells and driving their division or differentiation. The results indicate that expression of nestin in the developing CNS is directed by a 257 base pair enhancer containing recognition sites for factors of the POU domain, AP-2, and CRE-binding or nuclear hormone receptor families. *In vitro*, the enhancer is bound by the POU domain proteins Brn-1, Brn-2, and Tst-1, as well as unknown factors.

The identification of these POU proteins supports the hypothesis that nestin is transcriptionally linked to significant developmental events. A requirement for Brn-2 has been demonstrated in the terminal differentiation of specific neurons *in vitro* [Fujii and Hamada, 1993] and *in vivo* [Nakai *et al.*, 1995; Schonemann *et al.*, 1995], and all four class III POU proteins have also been proposed to function in forebrain patterning [Alvarez-Bolado *et al.*, 1995]. Thus the transcription factors identified in this

analysis illustrate that proteins which interact with the nestin enhancer can regulate not only processes common to all precursor cells, such as proliferation or terminal differentiation, but also the regionalizing processes which distinguish the progeny of one neural precursor from those of its neighbor. In addition, nestin-expressing precursor cells have now been identified in the adult brain. Further study of the transcriptional regulators of this intermediate filament could one day explain the processes which separate the quiescence of precursor cells in the normal adult brain from the rapid proliferation of injury-responsive cells or brain tumors.

References

Alvarez-Bolado, G., M.G. Rosenfeld and L.W. Swanson (1995) Model of forebrain regionalization based on spatiotemporal patterns of POU-III homeobox gene expression, birthdates, and morphological features. J. Comp. Neurol. **355**, 237-295.

Brenowitz, M., D.F. Senear and R.E. Kingston (1987) DNAse I footprint analysis of protein-DNA binding, in F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith and K. Struhl, Current Protocols in Molecular Biology New York: John Wiley & Sons

Carey, I. and Z.E. Zehner (1995) Regulation of chicken vimentin gene expression by serum, phorbol ester, and growth factors: identification of a novel fibroblast growth factor-inducible element. Cell Growth Diff. **6**, 899-908.

Chazaud, C., M. Oulad-Abdelghani, P. Bouillet, D. Decimo, P. Chambon and P. Dolle (1996) AP-2.2, a novel gene related to AP-2, is expressed in the forebrain, limbs and face during mouse embryogenesis. Mech. Dev. **54**, 83-94.

Chodosh, L.A., R.W. Carthew and P.A. Sharp (1986) A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. Mol. Cell. Biol. **6**, 4723-4733.

Clarke, S.R., A.K. Shetty, J.L. Bradley and D.A. Turner (1994) Reactive astrocytes express the embryonic intermediate neurofilament nestin. Neuroreport **5**, 1885-8.

Craig, C.G., V. Tropepe, C.M. Morshead, B.A. Reynolds, S. Weiss and D. van der Kooy (1996) In vivo growth factor expansion of endogenous subependymal neural precursor cell populations in the adult mouse brain. J. Neurosci. **16**, 2649-2658.

Dahlstrand, J., V.P. Collins and U. Lendahl (1992) Expression of the class VI intermediate filament nestin in human central nervous system tumors. Cancer. Res. **52**, 5334-41.
Dugich-Djordjevic, M.M., R. Somogyi, X. Wen and R.D.G. McKay (in preparation) The distribution of the neurofilament protein nestin in the adult rat brain.
Durston, A.J., J.P.M. Timmermans, W.J. Hage, H.F.J. Hendriks, N.J. de Vries, M. Heideveld and P.D. Nieuwkoop (1989) Retinoic acid causes an anteroposterior transformation in the developing central nervous system. Nature **340**, 140-144.
Figdor, M.C. and C.D. Stern (1993) Segmental organization of embryonic diencephalon. Nature **363**, 630-634.
Finkelstein, R. and E. Boncinelli (1994) From fly head to mammalian forebrain: the story of otd and Otx. Trends. Genet. **10**, 310-315.
Fujii, H. and H. Hamada (1993) A CNS-specific POU transcription factor, brn-2, is required for establishing mammalian neural cell lineages. Neuron **11**, 1197-1206.
Hyman, S.E., M. Comb, J. Pearlberg and H.M. Goodman (1989) An AP-2 element acts synergistically with the cyclic AMP- and phorbol ester-inducible enhancer of the human proenkephalin gene. Mol. Cell. Biol. **9**, 321-324.
Imagawa, M., R. Chiu and M. Karin (1987) Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell **51**, 251-260.
Kruijer, W., D. Schubert and I.M. Verma (1985) Induction of the proto-oncogene fos by nerve growth factor. Proc. Natl. Acad. Sci. USA **82**, 7330-7334.
Lin, S.-Y. and A.D. Riggs (1974) Photochemical attachment of lac repressor to bromodeoxyuridine-substituted lac operator ultraviolet radiation. Proc. Natl. Acad. Sci. USA **71**, 947-951.
Mangelsdorf, D.J. and R.M. Evans (1995) The RXR heterodimers and orphan receptors. Cell **83**, 841-850.
Marks, M.S., P.L. Hallenbeck, T. Nagata, J.H. Segars, E. Appella, V.M. Nikodem and K. Ozato (1992) H-2RIIBP (RXRβ) heterodimerization provides a mechanism for combinatorial diversity in the regulation of retinoic acid and thyroid hormone responsive genes. EMBO J. **11**, 1419-1435.
Moser, M., A. Imhof, A. Pscherer, R. Bauer, W. Amselgruber, F. Sinowatz, F. Hofstader, R. Schule and R. Buettner (1995) Cloning and characterization of a second AP-2 transcription factor: AP-2β. Development **121**, 2779-2788.

Nakai, S., H. Kawano, T. Yudate, M. Nishi, J. Kuno, A. Nagata, K. Jishage, H. Hamada, H. Fujii, K. Kawamura, K. Shiba and T. Noda (1995) The POU domain transcription factor Brn-2 is required for the determination of specific neuronal lineages in the hypothalamus of the mouse. Genes Dev. **9**, 3109-3121.

Philipp, J., P.J. Mitchell, U. Malipiero and A. Fontana (1994) Cell type-specific regulation of expression of transcription factor AP-2 in neuroectodermal cells. Dev. Biol. **165**, 602-614.

Puelles, L. and J.L.R. Rubinstein (1993) Expression patterns of homeobox and other putative regulatory genes in the embryonic mouse forebrain suggest a neuromeric organization. TINS **16**, 472-479.

Schonemann, M.D., A.K. Ryan, R.J. McEvilly, S.M. O'Connell, C.A. Arias, K.A. Kalla, P. Li, P.E. Sawchenko and M.G. Rosenfeld (1995) Development and survival of the endocrine hypothalamus and posterior pituitary gland requires the neuronal POU domain factor Brn-2. Genes Dev. **9**, 3122-3135.

Shimazaki, T., H. Okazawa, H. Fujii, M. Ikeda, K. Tamai, R.D.G. McKay, M. Muramatsu and H. Hamada (1993) Hybrid cell extinction and re-expression of Oct-3 function correlates with differentiation potential. EMBO J. **12**, 4489-4498.

Sive, H.L., B.W. Draper, R.M. Harland and H. Weintraub (1990) Identification of a retinoic acid-sensitive period during primary axis formation in Xenopus laevis. Genes Dev. **4**, 932-942.

Vicario-Abejon, C., K.K. Johe, T.G. Hazel, D. Collazo and R.D.G. McKay (1995) Functions of basic fibroblast growth factor and neurotrophinsin the differentiation of hippocampal neurons. Neuron **15**, 105-114.

Wegner, M., D.W. Drolet and M.G. Rosenfeld (1993) POU-domain proteins: Structure and function of developmental regulators. Curr. Opin. Cell Biol. **5**, 488-498.