The Regulation of the embryonic Transcription Factor Pax-3

Robert Harris

Ph.D. Thesis

September 2000

Department of Biochemistry and Molecular Biology

University of Southampton

ABSTRACT

UNIVERSITY OF SOUTHAMPTON
FACULTY OF SCIENCE
BIOCHEMISTRY AND MOLECULAR BIOLOGY
DOCTOR OF PHILOSOPHY

THE REGULATION OF THE EMBRYONIC TRANSCRIPTION FACTOR PAX-3

BY ROBERT GEORGE HARRIS

Pax-3 belongs to a family of nine murine transcription factors, the Pax genes. These all contain a highly conserved DNA binding domain, the paired domain and play key roles in embryogenesis in the mouse. Pax-3 is expressed in neural crest cells, developing spinal cord and brain from embryonic day 8.5 post coitus. Expression is temporally and spatially regulated and is restricted to mitotic progenitor cells before the onset of differentiation. Pax-3 has previously been shown to be necessary for continued proliferation of the sensory neuron derived ND7 cell line.

In this study the cell cycle transcription regulators E2F and the Myc/Max/Mad network have been shown to control the expression of Pax-3 in ND7 cells. In addition their binding sites in the Pax-3 promoter have been identified and it has been shown that binding to these sites is cell-cycle dependent.

This work has also looked at the subcellular localisation of Pax-3. Using fusions of Pax-3 and Enhanced Green Fluorescent Protein (EGFP) a Nuclear Localisation Sequence (NLS) has been identified in the Pax-3 protein. I have also shown that Pax-3 nuclear import in primary neural crest cells is serum dependent.

To date few target proteins of Pax-3 have been identified. Studies on an antisense Pax-3 cell line showed that suppression of Pax-3 expression led to the upregulation Snap-25 a protein involved in axonal outgrowth. I have shown that Pax-3 directly inhibits Snap-25 expression and located the Pax-3 binding site in the Snap-25 promoter.

Table of Contents

Table of Figures

Acknowledgments

There are many people without whom I never would have completed this work.

First and foremost I would like to thank Dr..Karen Lillycrop, my erstwhile supervisor without whose support and faith this work would never have been completed.

I would also like to thank my friends and colleagues at Boldrewood, Bob, Michele, Niamh, Pete, Phil, Richard, Graham, Neil, Paul and especially those in my Lab Emma, Fiona, Helen Jameson, Fiona, Helen Rodway, Holly, Thelma, Jo, Howard, Andy and Pete who have survived my repeated attempts to carve a career in demolitions and have remained friends in the face of danger.

Finally I would like to thank my family who have put upwith my bizarre passion for science and knowledge for many years.

ABBREVIATIONS

Amp	Ampicillin
ATP	Adenosine triphosphate
APS	Ammonium Persulphate
BSA	Bovine Serum Albumin
dATP	Deoxy Adenosine Triphosphate
dH_20	distilled water
DTT	Dithiothreotol
DMSO	Dimethyl Sulphoxide
DNA	Deoxyribonucleic acid
HBSS	Hanks Buffered Saline Solution
EDTA	Ethylenediaminotetraacetic acid
EtOH	Ethanol
IPTG	Isopropyl-β-D-thiogalactoside
Kbp	Kilobase pair
LB	Luria Broth
MeOH	Methanol
OD	Optical Density
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PEG	Polyethylene Glycol
PMSF	Phenylmethylsulphonyl Fluoride
PulydIdC	Polydeoxyinosinic-deoxycytidilic acid
RPM	Revolutions per minute
SDS	Sodium Dodecyl Sulphate
TBE	Tris Borate EDTA buffer
TEMED	N, N, N', N',-tetraethylmethylenediamine
TLC	Thin Layer Chromatography
Tris	Tris(hydroxymethyl)diamine
X-Gal	5-Bromo, 4-Chloro-3-Indolyl-β-galactopyranoside

Introduction

1 Introduction

1.1 Development

One of the key differences between multi-celled organisms and single cell eukaryotes such as yeasts is the evolution of different cell types to fulfill specific functions. Specialisation enables the higher degree of functional complexity that is the hallmark or the true multi-celled organism. This cellular diversification can range from the single digit number of cell types in the marine hydra to the thousands of different cell types present in mammals. Development is the sequence of events that lead from a single fertilised egg to the mass of different cells and cell types that makes up a mature adult.

In mammalian development fertilisation of the egg is followed by a series of quick cleavage divisions in which a single cell divides into 32 cells without a significant increase in mass. At the eight cell stage, each cell is equipotent, that is all cells have equal potential to form any cell lineage. At this point however the developing embryo starts developing axial organisation, that is anterio-posterior and dorso-ventral polarisation. At the 16 cell stage the outer cells become differentiated from the inner cells and the fates of the cells begin to diverge.

Cell division carries on apace as does progressive differentiation of cells. Morphogenesis is the start of major changes in the 3-D structure of the embryo.

1.1.1 Gastrulation

Gastrulation is the first part in this stage of embryonic development. The embryonic disc becomes elongated and is wider at the cranial end. The primitive streak is a dorsal invagination that runs down the anterio-posterior axis. Cells migrate through the primitive streak and form cells that become the endoderm and mesoderm germ layers. The remaining cells in the top layer form the ectoderm

layer. These germ layers each contribute to a specific group of differentiated cell types. These include

- Ectoderm – Central nervous system and epidermis.
- Mesoderm – muscles and skeleton in the trunk.
- Epithelium of the gut and epithelial lining of other organs.

At the end of gastrulation the future head and tail are well defined and the dorso-ventral axis is defined by the primitive streak, dorsally and the lateral edges of the embryonic disc ventrally.

1.1.2 Neurulation

Neurulation begins as gastrulation ends. Part of the mesoderm forms the notachord, in response to which, the dorsal region of the ectoderm thickens into the neural plate which then sinks down in the dorsal midline while two neural folds rise up on either side of the midline. These folds eventually meet at the future neck region and continue to merge cranially and caudally, to form the neural tube. The neural tube separates from the overlying ectoderm destined to form the epidermis. This will continue to develop into the spinal cord and brain. Cells which lie at the juncture between the closing neural tube and the surface ectoderm, the neural crest cells migrate to form the peripheral nervous system, sympathetic and parasympathetic systems and pigment cells.

1.1.3 The Neural Crest

Sometimes thought of as a fourth germ layer, the neural crest is a transient population of mulitpotent cells sandwiched between the neural tube and ectoderm that migrate extensively to give rise to a wide range of cell lineages including

those of the sympathetic, parasympathetic and enteric nervous systems, sensory neurons, melanocytes and melanoblasts (Le Douarin 1982).

1.1.3.1 Migration

Neural crest cells migration is unusual in both its extent and degree of patterning. While most developmental movement occurs as sheets of adhering cells, for example in gastrulation and neurulation, the cells of the neural crest migrate as individuals in an extensive and invasive manner within the early embryonic environment (Thiery *et al.*, 1982). They migrate away from the neural tube in well defined pathways at precisely timed periods of development until they settle in elected sites to produce such ordered structures as the complex networks of the autonomic and sensory nervous systems.

Migration starts concurrently with the closure of the neural tube, occurring in a wave, starting at the midbrain, slightly later in the forebrain and extending progressively down the rostro-caudal axis, following neural tube closure. From the initial single layer of cells in a narrow band down the neural tube they migrate laterally to cover the entire dorsal half of the neural tube. At this point some deviation occurs, with the trunk and brain sections following different patterns of migration. Migration in the trunk is characterised by a divergence of neural crest cell expansion into two spatio-temporally different streams (Erickson and Weston, 1983, Serbedzija, G. N., *et al.* 1990; 1992; 1994).

The Ventral Pathway (E10)

The first neural crest cells to reach the somites migrate ventrally between the neural tube and the somite mesenchyme. The cells following this pathway give rise to the autonomic nervous system and the cells of the adrenal medulla

The Dorso-lateral pathway (E12)

The later neural crest cells migrate into the extracellular space lateral to the neural tube between the ectoderm and dermomyotome. These cells migrate into the most ventral portions of the ectoderm and differentiate into melanocytes and melanoblasts, the pigment cells of the skin.

Neural crest emigration from the neural tube lasts approximately 24 hours and the cells halt in such a fashion that the more ventral regions are populated first (Fig. 1.01). (Serbedzija *et al.* 1989; 1990; 1994)

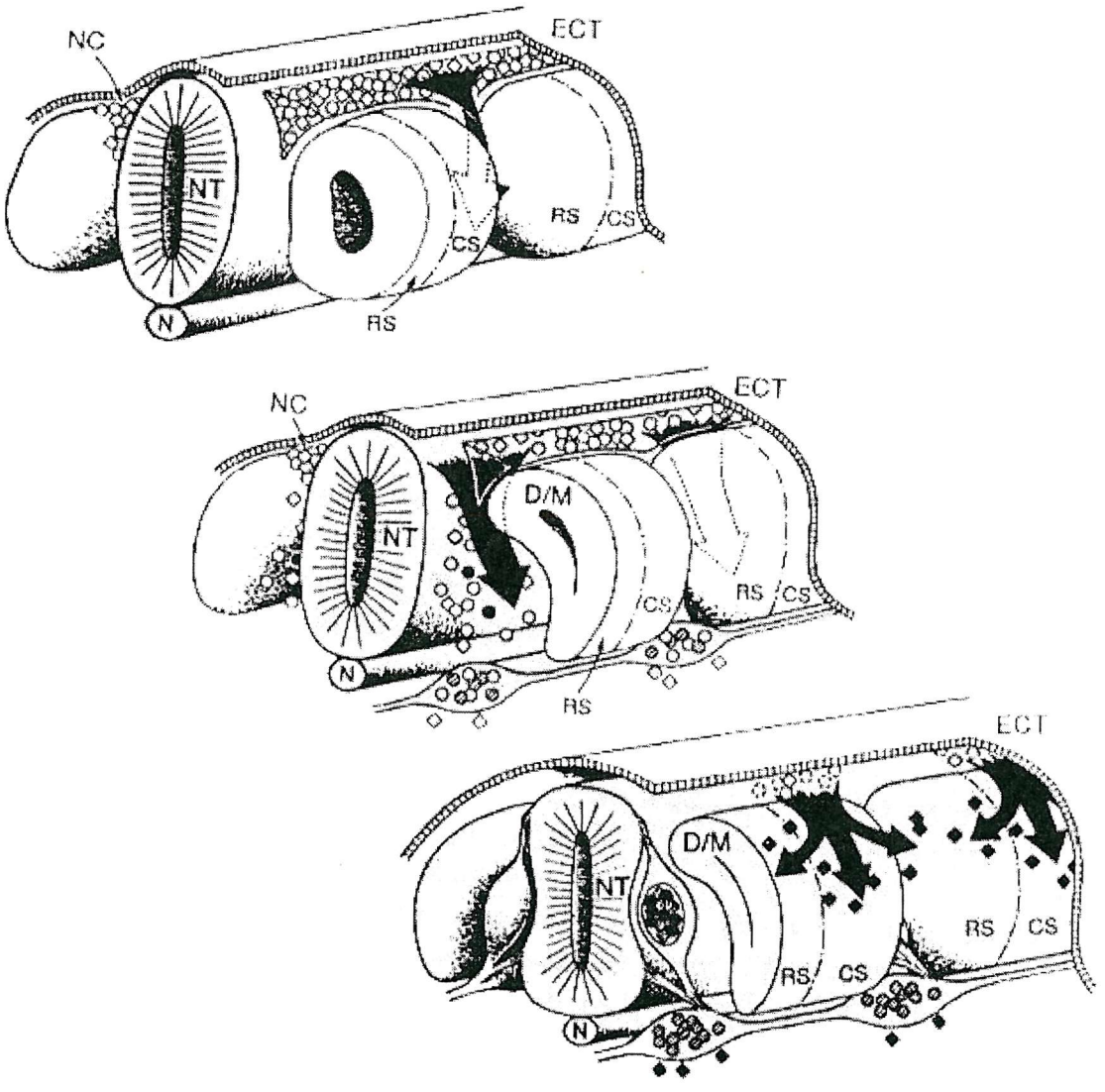


Fig. 1.01 Diagram depicting the migration of neural crest cells first through the ventral, between the neural tube and dermomyotome, and then through the dorso-lateral, between the dermomyotome and ectoderm, pathways. From Weston 1991

1.1.3.2 Control of migratory pathways

The degree of spatio-temporal regulation in the migration of the neural crest cells indicates a tight control of the migratory pathways.

Initiation of migration appears to be prompted several changes in the neural crest and surrounding cells. Firstly there is a decrease in the number of cell adhesion molecules such as N-CAM, N-Cadherin and E-Cadherin within the neural crest itself (Thiery *et al.* 1982, Newgreen and Gibbons 1982). This leads to decreased cell adhesion and increased motility. Secondly there is an increase of the glycosaminoglycan, hyaluronic acid, within the surrounding tissue; this increases the amount of extracellular matrix, which enables the neural crest cells to easily migrate. Thirdly there is an increase in fibronectin concentrations around the neural tube and somites around the time that neural crest cells begin migration.

The migratory pathway itself appears to be determined by four parameters

1. Extracellular space

   The extracellular spaces provide a pathway of least resistance through which the neural crest cells can easily migrate. This is increased by the presence of hyaluronic acid in the extra-cellular matrix, this is found in high concentrations along the migratory routes. (Derby 1978, Pintar 1978)

2. An adhesive substratum

   Neural crest cells appear to prefer to migrate in the presence of adhesive molecules such as fibronectin and the laminins. In vivo experiments show that they migrate away from the neural tube more avidly in the presence of fibronectin. In vivo this is found in much higher concentrations on the neural crest pathways than off (Rovasion *et al.* 1983, Newgreen and Thiery 1980).

3. Basal lamina boundaries

   The basal lamina associated with the neural tube, ectoderm and dermomyotome are impermeable to neural crest cells and as such act as barriers to migration.

4. Boundaries caused by lack of cell adhesion.

   Molecules that prevent cell adhesion, such as chondroitin sulphate, are found at borders of the migration pathways. Ultimately they counter the effect of fibronectin.

Other, perhaps, than the presence of increasing concentrations of fibronectin there appear to be no environmental factors that affect direction of movement of neural crest cells, rather contact inhibition appears to be the mechanism providing impetus to neural crest migration

1.1.3.3 Neural crest cell fate

The neural crest is made up of cells that exhibit an extraordinary range of potentialities, after migrating through the developing embryo they differentiate into a wide range of different cell types. It has been shown that individual neural crest cells are multipotent progenitor cells, that is each cell can terminally differentiate into a wide range of cell types depending on the environmental cues it receives during migration. (Le Douarin 1982, Bronner-Fraser and Fraser 1988; 1989, Frank and Sanes 1991, Fraser and Bronner-Fraser 1991). The developmental potential of a migrating neural crest cells becomes progressively restricted until it terminally differentiates into an appropriate cell type for its location, as defined by local environment.

However it has been found that there is some heterogeneity in potentiality within the neural crest cell population (For review see Selleck *et al.* 1993). This is most significantly noticed along the rostro-caudal axis where there is a distinct difference in derivatives formed by neural crest cells. The sections along the

neural axis have been designated, from rostral to caudal, as, cranial, vagal, trunk and lumbosacral.

- Cranial neural crest cells contribute to connective tissue and face skeletal cells in addition to facial sensory neurones and Schwann cells.
- Vagal neural crest cells populate the enteric nervous system.
- The trunk neural crest cells form melanocytes, sensory and sympathetic ganglia, Schwann cells and adrenomedullary cells.
- The lumbosacral neural crest populates the posterior enteric nervous system

Transplantation experiments in which neural tube and neural crest segments from one region were grafted onto another shows some restriction in potentiality from different regions. When trunk neural tube was grafted onto vagal regions of the host, graft derived neural crest cells contributed to the enteric nervous system and therefore changed their fate to that appropriate for their new location (Le Douarin *et al.* 1975). However when cranial neural tube and neural fold were grafted onto the trunk some cells continued to differentiate into cartilage although others contributed derivatives characteristic of their new position. Similarly when vagal neural tube was grafted onto the trunk some neural crest cells changed their fate and contributed to dorsal root ganglia, adrenal medulla and sympathetic neurones. These experiments have been performed on non-clonal populations of cells, and therefore it is not clear whether individual cells have true multipotency or whether it is merely the population of cells as a whole which exhibits the wide range of potencies. In vitro experiments, however, have shown that some individual neural crest cells do show true pluripotency (For review see Stemple and Anderson 1993).

Thus while there appears to be some heterogeneity in the subpopulation of neural crest cells for the most part their cell fates appear to be determined by environmental factors. While factors such as cell-cell contact are implicated in cell

fate, terminal differentiation itself appears for the most part to be controlled by a variety of extracellular factors. For example the Fibroblast growth factors (FGF) 1 and 2 have been implicated in stimulating proliferation of neural crest, inhibiting their differentiation. The polypeptide Endothelin 3, also promotes neural cell proliferation, but also mediates a vast increase in melanocyte number in culture (Lahav *et al.* 1996). This suggests that either it stimulates proliferation of neural crest cells already predetermined to be melanocytes or somehow mediates a limitation of cell fate to melanocyte lineages. Likewise Leukaemia inhibitory factor (LIF) has been show to stimulate the generation of sensory neurones *in vitro* (Murphy *et al.* 1991). Such extracellular factors are secreted by surrounding tissues, and act upon cell surface receptors in the neural crest cells. These receptors activate cell signalling cascades whose final targets are often transcription factors that are activated or de-activated and thus allow the changes in cellular transcription that accompany changes in cell function.

1.2 The control of Development

Development is a tightly regulated series of events that leads from a single undifferentiated cell to a complex organism made up of a mass of cells with a multitude of different types and specificities all acting synergistically to mutual benefit. Just as a cell's function is determined entirely by its internal biochemistry, due solely to the specific regulation of genes, so in turn development is fundamentally controlled by gene expression. Gene expression itself is controlled by regulatory transcription factors and it is at these that we must look in order to understand the molecular basis of development.

1.2.1 Transcription Factors

It is perhaps axiomatic to say that it would be impractical for a cell to contain all the types of protein it might ever need. Thus evolution has solved the problem by

storing the information required to manufacture any protein into the DNA carried by every cell. Thus any cell can theoretically synthesise every protein required by that organism. The conversion of this abstract coded information stored in DNA into concrete physiologically active proteins, called gene expression, is tightly controlled by transcription factors. As well as the information required to synthesise a protein, a gene also contains instructions on its spatial and temporal expression, this code is deciphered by the action of transcription factors.

1.2.1.1 A typical eukaryotic gene can be separated into three broad components

The transcribed region – From the transcriptional start site, this is the portion that is transcribed into DNA and contains the coding sequence that is eventually translated into protein

The promoter – This sequence of DNA contains the recognition elements required to assemble the basal transcriptional apparatus, necessary to start transcription, as well as regulatory elements that control initiation and rate of transcription.

The enhancer – This region of DNA, which can extend over a large region of DNA, and generally starts about 1kbp upstream of the start site contains binding elements for a large number of regulatory elements and is responsible for the spatio-temporal control of transcription.

1.2.2 Transcriptional Control

Eukaryote gene expression can be summarized as the transcription of DNA into RNA followed by the translation of RNA into a protein sequence. Thus the fundamental level of control is at the level of gene transcription from DNA into RNA. Transcription of mRNA from DNA is by the RNA polymerase II complex. There are three classes of transcription factors that can interact with RNA polymerase II and therefore affect gene transcription. These are

Basal Factors – These are required for initiating RNA synthesis at all promoters. They join with RNA polymerase to form the basal transcription complex which forms around the start point

Upstream Factors – These are DNA binding proteins that recognise specific sequences of DNA. The activity of these factors is not regulated, they are ubiquitous and act upon any promoter that contains the appropriate binding sites. They increase the efficiency of initiation and are required for a promoter to function at an adequate level.

Regulatory Factors – These function in a similar manner to upstream factors but their activity is controlled, such that they are only present in an active form in specific tissues or at specific times. They are therefore responsible for the control of spatio-temporal patterns of gene expression and are thus the most important class of transcription factors when studying the control of development.

1.2.3 Regulatory transcription factor structure and DNA binding

Transcription factors have a fairly discrete modular structure in which particular domains can be found that mediate different functions such as DNA binding, activation or inhibition of transcription. Analysis of domain swap experiments, in which the DNA binding region of one transcription factor is fused to the activation domain of a second has lead to the identification of function of several specific classes of domains.

1.2.3.1 DNA binding domains

There are several common types of DNA binding domain motif, many of which have been highly conserved throughout evolution. Examples include -

The Zinc Finger motif – This motif takes its name from its structure, in which a loop of 12 amino acids, a 'finger', containing conserved leucine and phenylalanine residues, projects from the surface of the protein. At its base

it is anchored by four amino acids that tetrahedrally co-ordinate with a zinc atom. There are two types of zinc finger. The 'classic' zinc finger and the steroid receptors. In the classic zinc finger the co-ordinating residues are two cysteine and two histidine residues. The zinc fingers form alpha helices that bind into the major groove and direct specificity, while the zinc atom forms stabilizing interactions with the phosphate backbone. The DNA binding site of a transcription factor is usually made up of a series of these motifs. There are three repeats for example in the ubiquitous basal transcription factor Sp1. In the steroid zinc finger the co-ordinating residues are all cysteine residues, mutation of histidine residues of classic zinc fingers into cysteines has abolished activation, showing the two types of finger are not interchangeable. The glucocorticoid receptor and oestrogen receptor have two fingers of this type in their DNA binding domain, which form alpha helices. These fold together to form a large globular domain that gives both major groove binding and recognition and a hydrophobic dimerisation portion that is required for steroid receptor binding..

The leucine zipper and basic region DNA binding domain - This motif has been detected in several major transcription factor families including Myc, Fos and Jun. This consists of a leucine rich region in which successive leucine residues occur every seventh amino acid. This means that when it forms an α-helix, leucine residues occur every two turns on the same side of the helix. The long side chains of the leucines would interdigitate with those of an analogous helix forming the leucine zipper, which would result in dimerisation of the factor. While it does not interact with DNA itself, in those factors in which it is present mutations rendering it inactive prevent DNA binding. Instead it appears to facilitate binding by a basic rich region which can directly interact with DNA Facilitation of dimerisation is thought to result in the correct positioning of the two basic direct DNA-binding domains in the dimeric molecule, so DNA binding can occur.

The Homeodomain motif – The homeodomain motif was first identified in the *Drosophila* homeotic developmental control genes. It consists of 3 α-helices and binds DNA via a helix-turn-helix (HtH) type motif. Structural studies of the homeodomain when in contact with DNA show that helix 3 lies in the major groove, making specific base contacts. Helices 1 and 2 are antiparallel to each other, lying over the DNA-helix 3 complex and perpendicular to helix 3. Helix 2 and 3 form the HtH motif against which helix 1 packs to provide stability. Helix 2 makes non-specific contact with the sugar phosphate backbone. Only a short central region of helix 3 contacts specific bases. It appears that a number of arginines in the N-terminal region of the homeodomain, upstream of helix 1 contact bases in the minor groove, so gripping the consensus sequence from two sides. It is often found in association with other DNA binding domains in a single protein. For example the Octamer binding proteins contain a domain similar to a homeodomain in addition to a POU DNA binding domain.

1.2.3.2 Other Transcription factor domains

Activation domains – Transcriptional activators work through activation domains that are hypothesised to work in a different number of ways. The common feature is probably an interaction with a basal factor of the initiation complex, TFIID or TFIIB being the most likely candidates. One theory is that the activation domain stabilises the pre-initiation complex and thus RNA transcription. The presence of a large number of negative charges in many activation domains has led to the theory that 'acid blobs' interact in a non specific manner with some protein involved in transcription. These acid blob activation domains enhance the ability of TFIIB to join the basal transcription apparatus. Other types of activating motifs are less well characterised but include a glutamine rich region in Sp1 and a proline rich region in CTF.

Inhibitory domains – These are thought to work in one of two ways, either by destabilising the pre-initiation complex or through recruitment of histone

deacetylases. Histone acetylation is a major regulatory mechanism for transcription. In tightly packed chromatin DNA is wound around histones to form a nucleosome. In this state the basal transcriptional apparatus is unable to bind and thus transcription is prevented. Many transcription factors however, including the TAF130/250 subunit of the TFIID complex, contain histone acetylase activity. Histones can be acetylated lysine residues on their N-terminal tails. This neutralizes the positive charge on the lysine side chain, abolishing ionic interaction with the phosphate backbone of the DNA. This has the effect of loosening the conformation of the DNA wrapped around the nucleosome and thus enabling transcription initiation. Inhibitory domains that recruit histone deacetylases therefore reverse this effect, deacetylation of histones cause tightening of the nucleosome structure and thus repressing transciption. (Struhl, K.)

1.2.4 Post transcriptional control of transcription factors

In development regulatory transcription factors will often act upon each other in a hierarchical manner, the expression of one leading to a cascade of expression of other transcription factors. In addition of course they ultimately act upon the promoter of other genes that are actually responsible for the formation of pattern and cellular biochemistry. However transcriptional control alone is undesirable – it is a relatively slow process, it takes time for active proteins to be synthesised once transcription has been initiated, and it takes time for protein synthesis to be stopped and activity to cease. Therefore some means of control after transcription is required. There are a number of ways in which a transcription factor can be regulated post transcriptionally.

1.2.5 RNA splicing

Numerous examples have been described in eukaryotes where a single RNA transcript from a gene can be spliced in alternative ways to form several different mRNA's that form proteins with different functions. This has been shown to be

the case in regulating some transcription factors. For example in the case of *era-1*, which encodes a transcription factor that mediates the induction of gene expression in early embryonic cells in response to retinoic acid. Two different mRNA's are produced, one encoding an active form of the molecule, while the other lacks the homeobox region that encodes for the DNA-binding homeodomain. In this case RNA splicing is used like transcriptional regulation to control the amount of protein expressed. In the case of CREM which binds to cAMP response elements, mRNA's are produced that do not contain the regions encoding the CREM activation domains. In this case the protein can still bind to DNA but does not activate transcription, it competes with the form that activates transcription. Alternative splicing also gives proteins with one or other of the two distinct DNA binding domains that the CREM gene encodes. (Calkoven and Ab 1996)

1.2.5.1 Degradation of mRNA

An important aspect in the control of gene expression is mRNA turnover; this can range from days to minutes depending on the gene and is often influenced by environmental signals. Rapid mRNA degradation provides an efficient mechanism for transient protein expression because it links protein expression directly with the gene transcription rate. Of transcription factors the most studied in turns of mRNA stability is *c-fos*. This has a UUAUUUA(U/A)(U/A) consensus sequence in its 3' untranslated region that triggers mRNA degradation. When transplanted into the 3' regions of usually stable mRNA's it destabilises the RNA, causing an increased rate of degradation.

1.2.6 Post-Translational control of activity

There are a number of mechanisms that are utilised to regulate post-translational activity of transcription factors (For Review see Calkoven and Ab 1996; Latchman 1995). These are often the final targets of signal transduction pathways in which

transient signals generated by stimulation of surface cells receptors, or changes in internal biochemistry are transmitted via phosphorylation cascades (Karin 1994). Phosphorylation can occur on all three alcoholic amino acid side chains. Its presence can be felt both through steric hindrance, which can lead to conformational change and through its negative charge. Despite being a covalent modification it is easily reversible. Thus single molecules can change their activation states many times within their lifetime – according to the need of the cell.

1.2.6.1 Direct Phosphorylation

Although most short term control mechanisms work via some sort of phosphorylation signal, direct phosphorylation on the transcription factor can affect its activation state. For example the AP1 binding factor c-jun has five phosphorylation sites that affect its activity. Three are just N-terminal to its DNA binding domain, phosphorylation of these sites prevent DNA binding and thus its interaction with the basal transcription apparatus (Boyle *et al.* 1991, Lin *et al.* 1992). Additionally it has two phosphorylation site in the C-terminal activation domain, phosphorylation of these sites result in elevated activation potential (Binetruy *et al.* 1991).

1.2.6.2 Protein-Ligand binding

One of the simplest methods for a stimulus to affect transcription is for an inducing ligand to bind to a transcription factor and alter it so that it becomes activated. Perhaps the classic example of this is the thyroid hormone receptor; this binds to DNA in the absence of thyroid hormone and inhibits gene expression via a specific inhibitory domain. Upon binding of the thyroid hormone the receptor undergoes a conformational change which exposes its activation domain and allows it to activate transcription. One of the classic examples of ligand-transcription factor interactions is that of the glucocorticoid receptor family. In this

case the glucocorticoid receptor is prevented from localising to the nucleus, and thus binding to DNA, by its association with the 90kDa heat shock protein (hsp90). When the steroid binds to the receptor a conformational change is induced resulting in its dissociation from hsp90.

1.2.6.3 Sequestration

The mechanism in which the glucocorticoid receptor is prevented from localising from the nucleus is a common one in transcription factor regulation. A transcription factor prevented from reaching the nucleus cannot of course activate or inhibit expression. The classic example of this is the B-cell specific transcription factor NF-κB a member of the Rel-related family. Although it is ubiquitously expressed in most cells it is sequestered in the cytoplasm as an inactive complex with the inhibitory protein IκB, probably masking the nuclear localisation sequence (NLS), a signal in the primary sequence that targets the protein to the nucleus (Bauerle and Baltimore 1988). In response to various signals such as mitogens, cytokines and oxidative stress, IκB is inactivated by phosphorylation, triggering its degradation by the ubiquitin-proteinase pathway. This liberates NF-κB which is then translocated to the nucleus, where it can activate gene expression.

1.3 The Pax Genes

An important class of developmental transcription factors and the focus of this research are the *Pax* genes. These make up a family of murine transcription factors, which are named for their evolutionarily conserved DNA binding motif, the Paired Box (Treisman *et al.* 1991; For reviews see Stuart *et al.* 1993, Chalepakis *et al.* 1993). They were originally identified by their sequence homology to the *Drosophila* segmentation genes *paired (prd)*, *gooseberry (gsb-p)*, *gooseberry-distal (gsb-d)*, and the tissue specific genes, *pox-meso* and *pox-neuro*

(Frigerio *et al.*, 1986, Baumgartner *et al.* 1987, Bopp *et al.*, 1986, Cai *et al.* 1994). *Pax* genes are also found in other organisms from zebrafish to quail and nematodes to man. They have been shown to be important in the development of the central and peripheral nervous systems (Daston *et al.*, Noll 1993). To date nine different *Pax/PAX* genes have been identified, designated *Pax-1* to *Pax-9* in both mouse and man (Deutsch and Gruss 1991, Walther *et al.* 1991). With the exceptions of *Pax-1* and *Pax-9*, which are expressed in the developing skeleton, they all show spatially and temporally restricted expression patterns during the course of neurogenesis and neural crest migration, suggesting that they play an important role in the tissue specification and pattern formation of the central and peripheral nervous systems. Mutations in the genes can lead to abnormalities in neuroectoderm, scleretome and myoderm derived tissue, all sites of *Pax* gene expression during development.

1.3.1 *Structural motifs in the* Pax *genes*

All of the Pax proteins share the common DNA binding motif, the paired domain. This binding domain of 128 amino acids consists, in effect, of a pair of homeodomain-like subdomains. Each triplet of α-helices can bind to DNA in the same manner as a homeodomain. This DNA binding domain has been shown to bind to a core motif of GTYMC or GTYMT (Chalepakis *et al.* 1991, For a more detailed study see later). Two other motifs are present in some but not all of the Pax proteins.

The octapeptide region is a conserved region of 8 amino acids, of unknown function, found downstream of the paired domain in seven of the Pax proteins. The final conserved motif is the paired-type homeodomain of 61 amino acids, located 3' to the paired domain and recognises an ATTA motif. In some of the Pax genes it is found in a truncated version containing only the first helix (For Reviews see Mansouri *et al.* 1994, Stuart *et al.* 1993).The nine *Pax* genes can be divided into four subfamilies based on their structure and expression patterns.

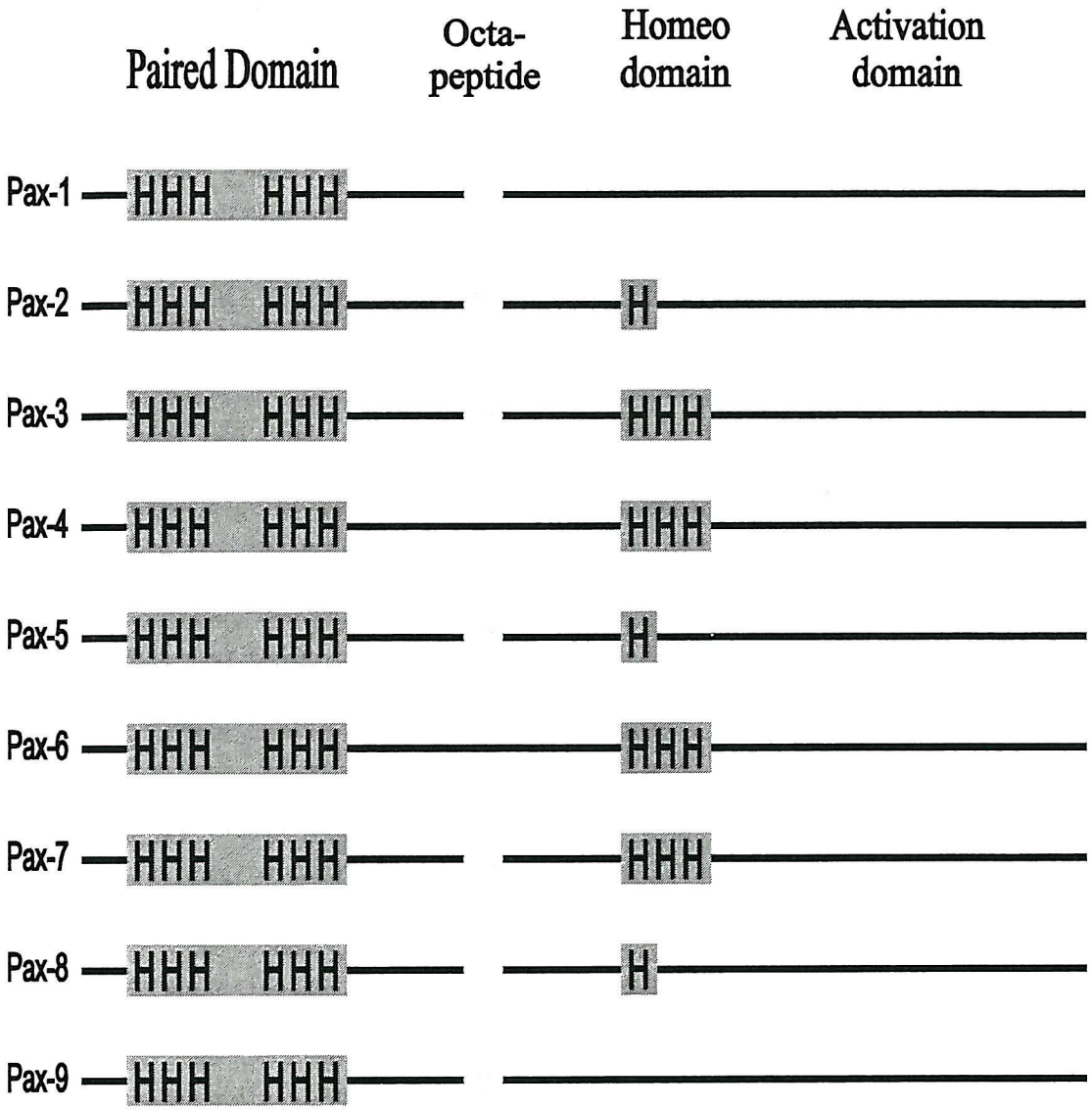


Fig. 1.02 Diagrammatic representation of the protein structure of the Pax genes. The Paired and homeoboxes are represented as striped boxes, α-helices represented by H and the octapeptide region by circles.

1.3.2 Expression patterns and classification of the Pax genes

The Pax genes can be divided into four subfamilies based on their structural homology and to a certain extent, expression patterns (Fig. 1.02, For Review see Stuart and Gruss 1995).

Pax-1 & 9 – Contain only the paired domain and octapeptide motif, no homeodomain at all. The only members of the *Pax* family not expressed in the developing nervous system, they are involved in skeletal development.

Pax-3 & 7 – These contain all the structural motifs, a paired domain, an octapeptide motif and a full homeodomain. The earliest of the *Pax* genes to be expressed they are switched on at around embryonic day 8.5 post-coitus (E 8.5 p.c.), before the onset of cellular differentiation. Initially expressed in the mesencephalon, the entire ventricular developing CNS, expression patterns are gradually limited as cells differentiate.

Pax-2, 5 & 8 – Contain the paired domain, the octapeptide and a partial homeodomain. Expressed later than Pax-3 and 7 they are first detected around E 9.5-10.5 p.c. Expression is limited to post-mitotic cells in the epichordal part of the neural tube, with a rostral limit at the midbrain-hindbrain boundary.

Pax-4 & 6 – Structurally in the same family, with a paired domain and full homeodomain only, their expression patterns are dissimilar.

Genes in the same family show evolutionary conservation (Balczarek *et al.* 1997), and with the exception of Pax-4 and 6 show very similar patterns of expression. Indeed it seems likely that with these paralogous genes there is some evidence for redundancy. In *Splotch* (*Pax*-3 compromised) mice axial muscle develops normally, which could be a compensation from *Pax-7*, the paralogous gene to *Pax-3*.

1.3.3 Expression patterns of the Pax *genes*

With the exception of *Pax-1* and *Pax-9* all the *Pax* genes are expressed in various restricted territories in the neural tube. Unlike the *Hox* genes, *Pax* genes are found in the more rostral domains of the brain (For Reviews see Mansouri *et al.* 1996; Chalepakis *et al.* 1993).

Pax-1 – Expression starts at E9 p.c. in the mesoderm of the developing vertebral column (scleretome), in the sternum and thymus (Deutsch *et al.* 1988). It carries on being expressed in the thymus of adults. It is mutated in *undulated* mice causing skeletal abnormalities, reduced thymus size, and affecting maturation of thymocytes. (Balling *et al.* 1988)

Pax-2 – Expression begins at E 9-10 p.c. and is detected in specific cells of the neural tube and hindbrain. Expression also occurs in the developing excretory system – specific kidney cells, Wolffian duct and ureter, developing eye and ear. (Nornes *et al.*1990, Dressler *et al. 1990*). In the eye it is expressed exclusively in the developing optic stalk and is inversely regulated compared to Pax-6. Overexpression has been described in Wilms tumour (nephroblastoma), a paediatric renal cancer.

Pax-3 – Expression begins at E 8.5 pc, in the dorsal ventricular region of the neural tube and neural crest. Pax-3 is expressed in the early brain and neural tube prior to cellular differentiation (Goulding *et al.* 1991)

Pax-4 – Expression is found in insulin producing B-cells in the pancreas. (Sosa-Pineda *et al.* 1997, Tao *et al.*1998)

Pax-5 – Expression is first detected at E 9.5 p.c. at the onset of differentiation and is present in the latero-medial portion of the developing brain and neural tube and in foetal liver cells, giving rise to B-lymphocytes (Adams *et al.* 1992). It is an important regulator of B-cell development and differentiation.

Pax-6 – Expression occurs in the developing central nervous system E 8.5 p.c., in the brain and neural tube. Expression is also found in the nose and eyecup, pituitary, and pancreas (Walther and Gruss, 1991). At the cellular level *Pax-6* mutations cause a delay in neuronal migrations. The *smalleye* (*sey*) mutation in mice, the murine equivalent of Aniridia, is due to mutations in the *Pax-6* gene. These give rise to cranio-facial abnormalities such as complete absence of eyes and olfactory organs (in homozygotes, which are embryonically lethal), and impaired development of cornea

Pax-7 – Expression is detected in the brain and neural tube just after Pax-3 in mitotically active cells. It may also be found in neural crest cells and dermomyotome (Jostes *et al.* 1991). Mutations in *Pax-7* are recessive, heterozygote mice appear normal, while homozygotes exhibit growth retardation, defects in nasal tubes and nasal capsule due to neural crest defects and frequent lethality (Mansouri *et al.* 1996). It is also involved in translocations that give rise to a Pax-7-FKHR fusion implicated in alveolar rhabdomyosarcoma (Davis *et al.* 1994)

Pax-8 – Expression of Pax-8 is found in specific cells of the neural tube and hindbrain, in the developing excretory systems, kidney and thyroid. Like Pax-2 it is also found overexpressed in Wilms tumour, it has also been implicated in some thyroid cancers. Unlike Pax-2, expression is undetectable after E 13.5 p.c. (Plachov *et al.* 1990, Zanini *et al.* 1992).

Pax-9 – Pax-9 expression is found in the developing vertebral column, but is absent from the central nervous system. (Neubuser *et al.* 1995). Pax-9 deficient mice show developmental abnormalities in the thymus, parathyroid glands, ultimobranchial bodies and in all limbs a supernumerary preaxial digit is present. In addition craniofacial and visceral skeletenogenesis is disturbed (Peters *et al.* 1998)

1.3.4 Pax *genes and oncogenesis*

Overexpression of *Pax* genes can transform fibroblasts into tumours in nude mice. (For review see Maulbecker and Gruss 1993). Similarly the expression of human *PAX2* and the paralogous gene *PAX8*, are abnormally upregulated in Wilm's tumour, a paediatric renal carcinoma of mesenchymal origin (Poleev *et al.* 1992, Dressler and Douglass 1992). Translocation of human *PAX-3* and *PAX-7* result in expression of a PAX-forkhead fusion protein carrying the intact binding domains of the PAX-3 or PAX-7 molecules, that are probably responsible for rhabdomyosarcoma (Davis *et al.* 1994, Shapiro *et al.* 1993). Deregulation of *PAX-5* has also been reported in malignant astrocytomas and medullablastomas.

Thus overexpression of *Pax* genes in tissues in which they are normally expressed may lead to tumorigenesis.

1.4 Pax-3

The *Pax-3* gene encodes a paired domain protein 479 amino acids in length with a molecular mass of 56kDa. It was first identified in a mouse E8.5 p.c. embryonic cDNA library using a 313bp fragment from the mouse *Pax-1* paired box (Goulding *et al.* 1991). The longest cDNA found was 2347 bp in length and contained an open reading frame of 1437bp with a 297nt 5' untranslated region and a 697nt 3'untranslated region containing a polyadenylation signal. Northern blots however showed two transcripts of 3.3kbp and 3.6kbp indicating that the cDNA clones may only contain part of the 5' UTR. The differences between, and the reasons for, the two transcripts are as yet unclear. Only one expressed protein has been found however.

1.4.1 Pax-3 *expression patterns during embryogenesis*

Pax-3 transcripts are first detected at E 8.5 p.c., rise to a high between days 9-12 and start to decline on day 13. By E 17 p.c. no *Pax-3* expression can be detected.

At E 8.5 p.c. Pax-3 transcripts are distributed throughout the dorsal half of the neural tube, in both the alar and roof plates along the entire rostro-caudal axis of the embryo. A clear ventral border of Pax-3 expression coincides with the sulcus limitans.

Pax-3 has also been found to be a key factor in the development of limb muscle. Specification of myogenic lineages begins prior to gastrulation and culminates in the emergence of determined myogenic precursor cells from the somites, which migrate to form limb muscle cells. The later steps of myogenic specification are controlled by the MyoD family of transcriptional regulators, which are in turn regulated by Pax-3.

Pax-3 is expressed in all the cells of the caudal segmental plate, the early mesoderm compartment that contains the precursors of skeletal muscle. Beginning at the time of segmentation Pax-3 becomes repressed in the ventral half of the somite, leaving expression only in the dermomyotome. Pax-3 expression then becomes repressed as the dermomyotome cells activate the MyoD family. Pax-3 continues to be expressed in migrating limb precursor cells, which do not activate MyoD until 2 days after they leave the somite. It has also been shown that artificial Pax-3 expression inhibits myogenic differentiation induced by MyoD. Pax-3 may suppress terminal differentiation ahead of migrating limb myoblasts.(See Ludolph and Konieczny 1995, Lobe 1992)

Neural crest cells that migrate along axons of the peripheral nervous system become Schwann cells. An inverse correlation between *Pax-3* expression and myelin basic protein has been observed in embryonic Schwann cells. Dysfuntional *Pax-3* is associated with a reduction of SC lineage and non-myelinated nerves in the adult.

In later development much of the evidence for Pax-3 expression comes from developmental abnormalities associated with Pax-3 mutations. Pax-3 mutations mainly affect tissues that receive contributions by neural crest cells. This leads to

pigmentary abnormalities of the iris, hair and skin, mild facial malformations and hearing problems.

In summary *Pax-3* transcripts are detected in the dorsal neuroepithelium just prior to closure of the neural tube and remain restricted to dorsally located mitotic neuroepithelial cells throughout early neurogenesis. This expression pattern may be a primary response to inductive signals from the notachord and floor plate that regionalise the neural tube.

1.4.2 The Splotch *mouse and developmental abnormalities*

In the house mouse (mus musculus) mutations in *Pax-3* are associated with the *splotch* phenotype (Russell and Greene 1947; Epstein *et al.* 1991; Walther *et al.* 1991). There are six semidominant *splotch* alleles in the mouse; *splotch (Sp)*(Russell and Greene 1947), *splotch-delayed* (Sp^{d})(Dickie 1964) arose spontaneously while *splotch-retarded* (sp^{r}), Sp^{1H}, Sp^{2H}, Sp^{4H} arose after X-irradiation (Beechey and Searle 1986; Franz 1993). Heterozygous *splotch* mutants display pigmentation defects that are visible as spotting upon the belly, back, tail and limbs and are due to defects in the neural crest derived melanocytes. The heterozygous Sp^{r} mutant embryos also show retarded growth. Homozygous *splotch* mutants are embryonically lethal and all mutants except Sp^{d} and Sp^{r} die at E 14 pc. Sp^{d} die at parturition and Sp^{r} mutants die before implantation due to a large chromosomal deletion encompassing *Pax-3* and other adjacent genes that are important for development. The developmental defects seen in homozygous embryos include spina bifida, exencephaly, meningocele, the development of certain skeletal muscles and neural crest associated deficiencies (Franz 1989; 1990; 1993; Franz *et al.* 1993; Steele and Smith 1992; Moase and Trasler 1992).

1.4.3 Waardenburgs Syndrome

Mutations in the Human PAX-3 gene give rise to Waardenburgs syndrome. This is an autosomal dominant disorder, symptomised by cleft palate, premature greying, pigment deficiency, deafness and lateral displacement of the corner of the eye. Based on the presence or absence of this lateral displacement of the eye it is classified as either WS I or WS II. These abnormalities are all due to defects in neural crest cell derivatives such as sympathetic neurons and melanocytes.

1.4.4 Pax-3 structure

The Pax-3 protein is a member of the class II subfamily of Pax proteins it contains five active elements within its structure: The paired domain, The homeodomain, The inhibitory domain and the activation domain.

1.4.4.1 The Paired Domain

The paired domain is a 128 amino acid conserved DNA-binding domain (Treisman *et al.* 1991, Chalepakis *et al.* 1991) first identified in the *Drosophila paired (prd)* and *goosberry (gsb)* genes (Bopp *et al.* 1986). The first 74 amino acids of the paired domain are basic rich and highly conserved throughout the Pax family, while the last 54 amino acids are more divergent. A high-resolution (2.5Å) x-ray crystal structure of the paired domain has been solved for the *Drosophila paired* protein (Xu *et al.* 1995). This structure (Fig. 1.03) indicates that the paired domain actually includes two structurally independent globular sub-domains. Each of these subdomains (Czerny *et al.* 1993) is a potential DNA binding domain containing a Helix-turn-Helix (HtH) motif similar to that of the homeodomain or the Hin recombinase. The co-crystal shows that the N-terminal or PAI subdomain is the only one that binds to the DNA site chosen. A β-sheet grips the sugar-phosphate backbone of the DNA. This is immediately followed by a β-turn that makes critical base contacts in the minor groove. The first helical region (residues

20-60) contains a HtH motif. Helix 3 sits in the major groove making extensive base contacts, and presumably determining site specificity, while helices 1 and 2 sit over it in an anti-parallel fashion, helix 2 making extensive sugar-phosphate contacts, non-specific anchoring interactions.

The c-terminal RED domain does not make any DNA contacts in the co-crystal shown although its structure can be superimposed reasonably well on the *engrailed* homeodomain and on Hin recombinase. So like the PAI domain it contains a potential DNA binding HtH motif. However there is support for non-binding from footprinting studies, which show only a 15bp site protection as in the co-crystal. However there is evidence that in some proteins such as Pax-5 and Pax-6 there is a larger binding site, which may include the RED domain. Additionally the footprinting studies were done with the paired domain alone and it may be that binding of the entire protein

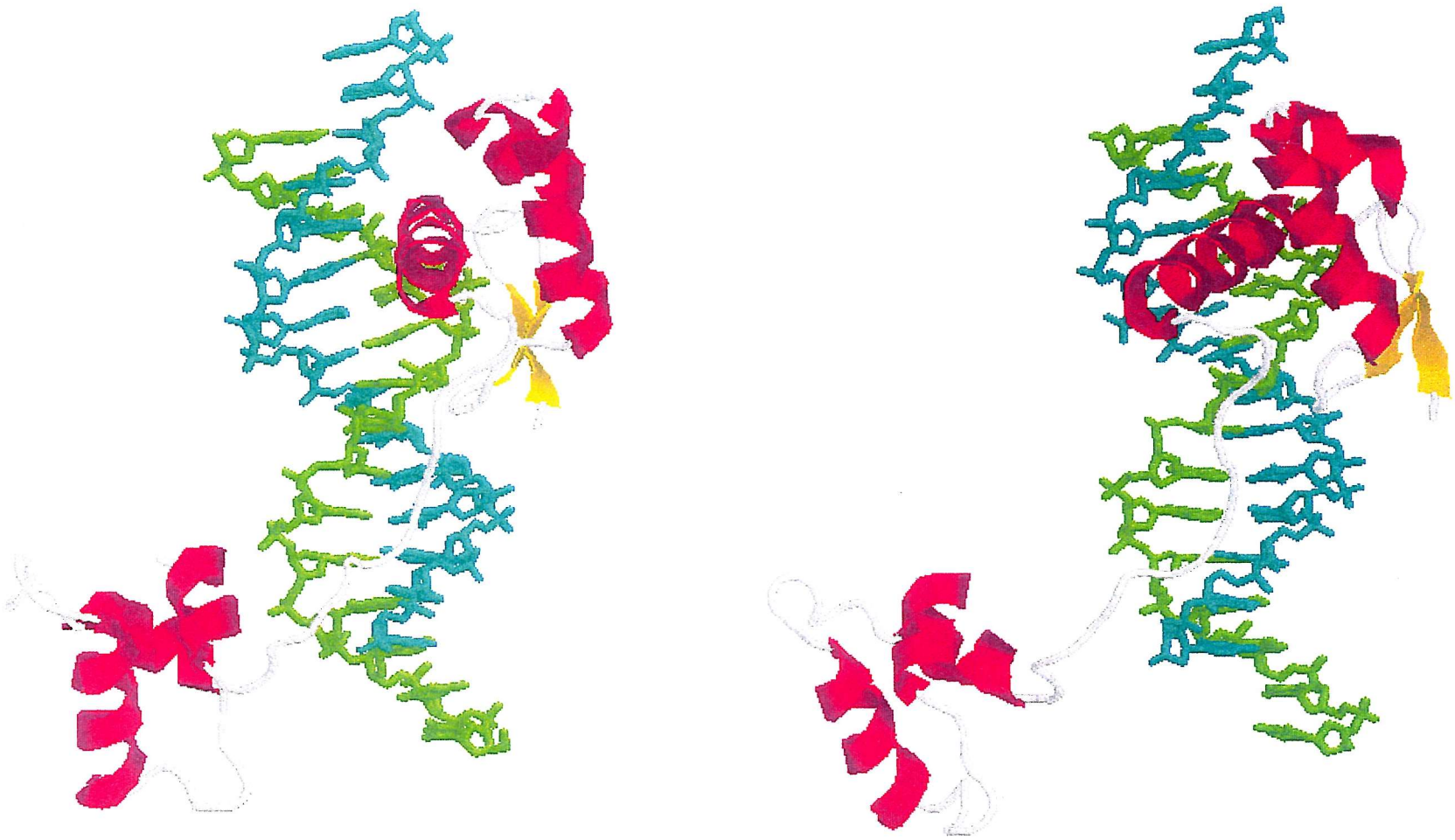


Fig. 1.03 Cartoon picture of the co-crystal structure of the *paired*, paired domain with its DNA binding site. Binding of the PAI subdomain, but not the RED subdomain can quite clearly be seen (Xu *et al.* 1995). Created using the Brookhaven protein databank file and the Chime molecular modeling program.

induces conformations that allow RED binding. The RED domain has a far less conserved sequence than the PAI domain, perhaps due to its lack of importance in site determination.

There are several paired domain mutations that are associated with developmental abnormalities in both mice and humans (Tassabehji *et al.* 1992, 1993; Baldwin *et al.* 1992; Balling 1988, Hoth *et al.* 1993). An alternative splicing event in Pax-3 has been found to generate isoforms of Pax-3 differing only by the presence of a single glutamine residue in the linker region of the two subdomains. Levels of these isoforms have been found to be similar in the developing mouse embryo. The isoform deficient in the glutamine residue had the ability to bind to the paired domain recognition site (GTCAC) with a two to five-fold higher affinity than the other isoform. The two isoforms may affect target site selection and mediate distinct functions within the embryo (Vogan *et al.* 1996).

1.4.4.2 The Homeodomain

The homeodomain is an evolutionarily conserved 60 amino acid DNA binding domain first found in genes that control *Drosophila* morphogenesis. Pax-3 contains a paired-type homeodomain. Homeodomains contain a classic Helix-turnHelix motif, interaction with DNA occurs principally through the third (recognition) helix which makes base specific contacts in the major groove (Wilson *et al* 1995). The Pax-3 homeodomain has a core recognition site of ATTA although it can only bind this sequence when the paired domain is bound to its recognition site. However it has been observed that almost all major homeodomain classes can dimerise upon DNA binding (Wilson and Desplan 1995), the *Pax* genes are no exception (Wilson *et al.* 1995). Using a palindromic double homeodomain motif (TAAT(N)ATTA) binding site induces homodimerisation and enables high affinity binding of Pax-3. Each homeodomain binds to one of the recognition sequences.

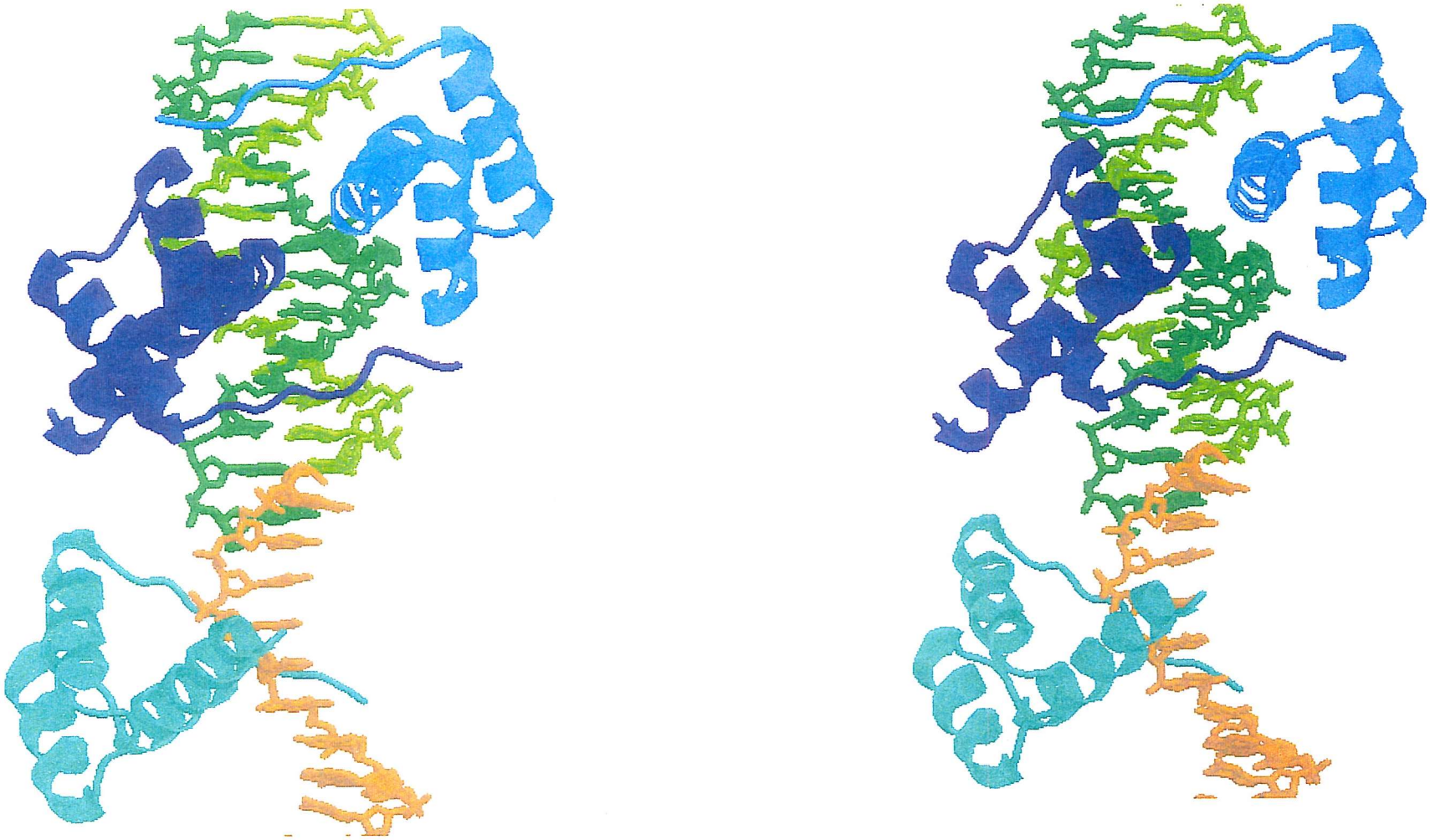


Fig. 1.04 Cartoon picture of the co-crystal structure of a paired-type homeodomain with its oligonucleotide binding site. Created using the Brookhaven protein databank file and the Chime molecular modeling program.

1.4.4.3 The activation domain

A transcription activating domain has been defined in Pax-3 that includes the 78 C-terminal amino acids. It is Serine, Threonine and Proline rich, similar to the domains in Oct 2 and CTF-1 that are obligatory for transcriptional activation. This transactivation domain can be transferred to a heterologous GAL4 DNA binding domain and still function as an activator (Chalepakis *et al.* 1994a).

1.4.4.4 The inhibitory domain

A domain responsible for strong transcriptional inhibition was found to be located within the first 90 amino acids of Pax-3 (Chalepakis *et al,* 1994a). This includes 57 amino acids of the paired domain, which constitutes the entire HtH motif of the DNA-binding PAI domain. This domain can function as an inhibitor independent of the rest of Pax-3, experiments transferring it to a heterologous GAL4 DNA binding domain shows that it continues to act as a transcriptional inhibitor. Pax-3 is the only member of the Pax family shown to have a transcription inhibiting activity and is the only one to have a significant number of amino acids in front of the paired domain. These 33 N-terminal amino acids have been shown to be essential for transcription inhibition (Chalepakis *et al.* 1994a).

1.4.5 Pax-3 recognition sequences and DNA binding

The DNA binding site for Pax protein is unusually large, up to 24 nucleotides in length. Initial electromobility shift assay (EMSA) experiments showed that Pax-3 could bind to the *Drosophila paired* binding site, *e5*, first identified by in vitro footprinting upstream of the *even-skipped* gene (Goulding *et al.* 1991). The core *e5* sequence, G**ATTA**GCACC**GTTCC**GCT contains recognition sites for the homeodomain (ATTA) and the paired domain (GTTCC), separated by a five amino acid spacer. Initially it was thought that both recognition sequences were necessary for efficient binding of Pax-3, however both the paired domain

recognition sequence or a double homeodomain motif alone allow effective Pax-3 binding (Epstein *et al.* 1994, Chalepakis and Gruss 1995, Wilson *et al.* 1995).

Later EMSA/PCR selection enrichment experiments showed that the ideal core Pax-3 paired domain DNA binding site (GTCAC(G/A)C) cannot be particularly critical considering actual target gene binding sites and other oligonucleotides that allow binding (Chalepakis *et al.* 1995). However these experiments used the paired domain alone and it seems likely that the entire protein alters DNA binding. Recent experiments (Phelan and Loecken 1998) showed that the paired box consensus sequence, when the EMSA/PCR enrichment techniques are performed using both the paired domain and homeodomain, is GTTAT. In order for optimal binding the paired domain binding site was 5 or 8 bases downstream of the homeodomain ATTA sequence. However it was also noted that most target genes found had low affinity binding sites indicating that they would only be activated when cellular levels were high (Phelan and Loecken 1998).

1.4.6 Transcriptional Regulation by Pax-3

Pax-3 is a bifunctional modulator of transcription; it contains domains for both transcription inhibition and activation (Chalepakis *et al.* 1994a). Thus upon binding of Pax-3 to DNA it may either activate or inhibit transcription of downstream genes. It has been suggested that this modulation may be concentration dependent. Using a $(PRS\text{-}1)_6$ artificial binding site in front of a reporter gene causes transcriptional activation with low concentrations of Pax-3 while high concentrations of Pax-3 lead to transcriptional inhibition (Chalepakis *et al.* 1994a). An alternative suggestion is that the regulatory effect is dependent on which domains bind (Jun and Desplan 1996; Sheng *et al.* 1997). There are three alternative modes by which Pax-3 can bind to the promoter or a target gene. Firstly the paired domain alone may bind to a paired recognition site. Secondly both the paired and homeodomain may bind simultaneously to their respective recognition sites and thirdly a Pax-3 can dimerise through its homeodomain to a

double homeodomain recognition sequence. Depending therefore on the binding mode used either activation or inhibition might take place.

1.4.7 Pax-3 Target genes

To date few target genes have been identified that are known to be directly transcriptionally regulated by Pax-3. The first gene that was found to be regulated by *Pax-3* was the *c-met* (hepatocyte growth/scatter factor tyrosine kinase). Like *Pax-3*, *c-met* is expressed in limb muscle progenitors. The absence of *c-met* expression in *splotch* (Pax-3 deficient) mice implies a role in activation of *c-met* by Pax-3. In addition the *c-met* promoter contains a Pax-3 paired domain binding site (GTTCC) to which Pax-3 can bind with high affinity.(Yang *et al.* 1996)

However *Pax-3* is expressed at the same time as myogenic determination factor genes and is essential for the transcription of *MyoD* during early muscle development (Epstein *et al.* 1995; Ludolph and Konieczy 1995). While the *MyoD* promoter is transcriptionally activated by Pax-3, a target recognition site for Pax-3 has not been found.

MITF (microphthalmia-associated transcription factor) is a tissue specific transcription factor, which transactivates the gene for tyrosinase, an important enzyme in melanogenesis. *MITF* is therefore important in melanocyte differentiation and loss of its function is known to lead to Waardenburg's Syndrome Type 2 (WS2). This is characterised by auditory-pigmentation symptoms as a consequence of an absence of melanocytes in the cochlea resulting from changes in *MITF* gene transactivation. *PAX-3* expression in neural crest cells, from which melanocytes derive, is earlier than *MITF* and has been shown to transactivate the *MITF* promoter. Failure of *MITF* regulation by *PAX-3* leads to auditory-pigmentation problems in Waardenburg's Syndrome Type 1 (WS1) (Watanabe *et al.* 1998).

Neural Cell Adhesion Molecules (N-CAM's) are also subject to some regulation by *Pax-3* since Pax-3 upregulates their promoters when attached to a reporter gene. However no binding to the promoter by Pax3 can be detected by EMSA (Chalepakis *et al.* 1994a). Neuronal migration and differentiation processes altered by *Pax* genes are also affected by cell adhesion molecules (Doherty *et al.* 1991). Pax-3 has been found to bind to the NgCAM (neuroglia – cell adhesion molecule) promoter, however another CAM molecule with two Pax binding motifs is not bound by Pax-3 (Kallunki *et al.* 1995).

Neural tube defects are among the most common abnormalities to be associated with diabetic embryopathy. Pax-3 has been found to be significantly reduced in the embryos of diabetic mice, and in areas of diabetic embryos where Pax-3 was under expressed high concentrations of cells undergoing apoptosis have been found (Phelan *et al.* 1997). Recently *Dep-1* (from *Diabetic Embryopathy*) has been identified as being regulated by *Pax-3* and may be used as a marker for altered gene expression during diabetic pregnancy (Cai *et al.* 1998). Although *Dep-1* is first expressed at E 6.5, before the onset of *Pax-3* expression, its expression has been found to be reduced following reduction of *Pax3* expression during formation of the neural tube (Cai *et al.* 1998). Reduction of *Dep-1* expression has also been found in homozygous *Splotch* embryos showing that *Dep-1* is regulated either directly or indirectly by Pax-3. Likewise induction of Pax-3 by retinoic acid in P19 embryonal carcinoma cells gives a corresponding induction of *Dep-1*.

1.5 Aims

The function of Pax-3 in the developing embryo is not completely clear. The aim of this work was to further elucidate the function of Pax-3 by :

- Investigating the regulation of Pax-3 transcription
- Investigation of the subcellular localisation of Pax-3
- Identification of potential target genes of Pax-3

Materials and Methods

2 Materials and Methods

2.1 General Laboratory Chemicals and buffers

The majority of general chemicals were purchased from Sigma UK Ltd. or Fisher Scientific Ltd., enzymes from Promega UK Ltd. and tissue culture products from Gibco BRL. Sequagel (concentrate and diluent) and Protogel were purchased from National Diagnostics and radiochemicals and membranes (PVDF and Hybond C, P or N) from Amersham International. Where necessary solutions were sterilised by autoclaving at 115lbs/sq. in. for 20 minutes. Antibiotics and other heat sensitive or low volume solutions were filter sterilised using .22 μm syringe end filters.

2.2 Equipment

Beckman -21 RPM x 1000 centrifuge for 50-400ml volumes and 2-30ml high speed volumes (>5000rpm). Sigma 3K10 Howe centrifuge for <50ml volumes (<5000 rpm). MSE Microcentaur microfuge (<1ml volumes). Uniscience univap – evacuated spinning evaporator. Techne progene polymerase chain reaction machine. UVP transilluminator and UV camera. R100 Lucham Rotatest orbital shaker. Sartorius balance. Hook and Tucker Rotamixer, Biorad mini-gel apparatus and semi-dry western transfer cell. Biorad Powerpac 200 and 300. Grant 1000C waterbath. Zeiss microscope 459315. Molecular Dynamics STORM phosphorimager, Biohit Biological safety class II cabinet; Biohit CO_2 incubator.

2.3 Buffers

Phosphate Buffered Saline (PBS) was either purchased as tablets from Sigma, or made up as follows; 140mm NaCl, 2.7mM KCl, 9.2 mM Na_2HPO_4, 1.84mM KH_2PO_4;

Tris-Borate EDTA (TBE); 450mM Tris-HCl, 400mM Boric acid, 10mM EDTA (pH 8.0)

Tris-acetate-EDTA 50X conc. (TAE); 2M Tris-HCl, 57.1% (w/v) glacial acetic acid, 50mM EDTA (pH 8.0)

Tris-glycine running buffer; 25mM Tris-HCl (pH 8.0), 250mM glycine, 0.1% (w/v) SDS

Laemli buffer (for protein gel loading); 50mM Tris-HCl (pH 6.8), 100mM DTT, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue 10% (v/v) glycerol

Agarose gel Loading dye (6X), 0.25% Cyanol yellow, 15% ficoll

Western Transfer Buffer; 39mM Glycine, 48mM Tris HCl (pH 8.0), 0.037% SDS, 20% (w/v) MeOH

Coomassie blue stain; 50% (w/v) Tri-chloroacetic acid, 0.25% (w/v) Coomassie blue.

Destain; 5% (v/v) acetic acid, 45% (v/v) MeOH

2.4 Bacterial Growth Media and Solutions

2.4.1 Luria-Bertani (LB) Leonard Broth Medium

This was prepared as instructed by BDH, 20g of their powdered LB mix being dissolved in 1L of d H_2O. This was immediately autoclaved.

2.4.2 LB agar

This was prepared as above with the addition of 1.5% Agar (BDH) before autoclaving. To make up LB agar plates on which E. coli were grown the agar was first melted, then cooled to approximately 50°C. This was then poured into sterile plastic petri dishes (20ml/90cm plate). These plates were allowed to cool to room temperature, then air-dried at 37°C. The dried plates were then stored at 4°C.

2.4.3 LB broth and LB agar supplemented with antibiotics.

Ampicillin was added to the growth media to select for the growth of bacteria containing plasmids encoding the Amp^r (ampicillin resistance) gene. Ampicillin (Sodium salt, Sigma) was dissolved in filter sterilised, AnalR H_2O to a final concentration of 100mg/ml. This solution was sterilised by filtration through a 0.2μm milli-pore filter, aliquoted (1ml) into eppendorf tubes, and stored at -20°C. This stock was then added to either the LB broth or LB agar (at 50°C) to give a final concentration of 100μg/ml. Kanamycin was added to the growth media to select for the growth of bacteria containing plasmids encoding the Kan^r (Kanamycin resistance gene.) A 1000X stock solution of 30mg/ml was made up in the same way as for ampicillin.

2.4.4 LB agar with X-Gal.

When using the blue/white selection plasmids pBluescript or pGEM-T-Easy ampicillin containing agar plates were coated with a solution of X-Gal (80μl, 100mg/ml) in DMSO 1 hour before addition of the bacterial culture.

2.5 Bacteria and Plasmids

2.5.1 Bacterial storage

Bacterial strain DH5α was used in all cases except where noted; its phenotype is ;

sup E44 D lac end U169 (φ80 lacZ ΔM15) hsd R17 rec A1 end A1 gyr A96 thi-1-rel A

Long term storage of bacteria was as a 50% bacterial culture in LB with 50% glycerol and stored at -80°C

2.5.2 Bacterial Cell growth

Routinely bacteria, with or without plasmids, were grown from a single colony on an agar plate. Using a flame sterilised platinum wire a single colony of bacteria was picked from the plate and transferred to a 10ml culture of LB media. Bacteria were grown at 37°C overnight on a rotary shaker.

2.5.3 Preparation of Competent E. coli.

A static culture of E. coli (DH5α) was grown up overnight at 37°C. An aliquot (100μl) of this starter culture was inoculated into LB broth (10ml). The bacteria were then grown at 37°C up to a cell density of $A_{600} \approx 0.5$ OD. The bacteria were then pelleted by centrifugation (2500 rpm, 4°C, 10mins). The pellet was resuspended in ice-cold $CaCl_2$ (100mM, 5ml). This was then re-centrifuged (2500rpm, 10mins) and the resulting pellet resuspended in 500μl ice-cold 100mM $CaCl_2$.

2.5.3.1 Transformation of competent cells

Competent E. coli were dispensed into 100 μl aliquots in ice cold eppendorf tubes for each transformation. Ice cold DNA/ligation mixtures were added to each tube

of competent cells and left on ice for 20 minutes. The uptake of plasmid into the cells was promoted by heat shocking at 42°C, in a water bath, for 90 seconds. Cells were allowed to recover on ice for a further 10 minutes before LB (500µl) was added and the transformation incubated at 37°C for 45 mins. This prior to spreading on LB agar (+ amp) plates.

2.6 Plasmids

Plasmid	**Relevant features**	References
PBluescript	Cloning vector ampicillin resistance	Stratagene
PGEM - T- Easy	TA cloning vector Ampicillin resistance	Promega
PCDNA3.1+	Mammalian expression vector CMV promoter	Invitrogen
pJ7Ω	Mammalian expression vector CMV promoter	Morganstein and Land 1990
pBL3CAT	Reporter vector Cloning site upstream of CAT gene	Luckow and Schutz 1987
pCAT-Basic	Reporter vector Cloning sites upstream of CAT gene	Promega
Snap-25 promoter –CAT deletions	CAT reporter, Snap-25 promoter	Ryabinin *et al.* 1995

pEGFP.	Kanamycin Resistance, CMV promoter Cloning site upstream of Enhanced Green Fluorescent Protein.	Clontech
E2F-1 expression	Ampicillin resistance, CMV promoter	Helin *et al.*1993
N-Myc expression, Max, and Mad expression vectors	CMV promoter, Ampicillin resistance	Tsuneoka *et* al 1990
C-Myc expression	CMV promoter, Ampicillin resistance	Shang *et al.* 1990
pRB, pΔRB expression vectors	CMV promoter, Ampicillin resitance	Furukawa *et al.* 1990

2.7 Preparation of DNA

2.7.1 Phenol/chloroform extraction.

An equal volume of phenol/chloroform (1 : 1 v/v) was added to the DNA. This is mixed well by vigorous shaking, which action denatures any remaining protein and causes its aggregation and precipitation. The whole is then centrifuged (13K rpm, 3mins or 5K rpm, 7mins). This causes separation of the phenol/chloroform and aqueous layers. The aggregated protein collects at the interface. The top aqueous layer is removed, carefully leaving behind the protein, this is then washed

using an equal volume of chloroform followed by centrifugation (13K rpm, 2mins or 5K rpm, 4mins). The aqueous layer is then removed and subsequent operations (usually ethanol precipitation) performed.

2.7.2 Ethanol precipitation

Sodium acetate was added to aqueous DNA to a final concentration of 0.3M by the addition of 0.1Volumes of a 3M stock. Two volumes of ethanol (-20)C) were added and the DNA precipitated at –20Oc for 10 mins. The DNA was then pelleted by centrifugation for 10mins and washed with 70% ethanol to remove excess salts. The DNA pellet was then air dried for 2-3 mins and resuspendend in the required volume of dH20.

2.7.3 Small scale extraction of DNA (Miniprep)

The following protocol is a modification of that of Sambrook *et al.* LB culture supplemented with ampicillin (10 ml) was inoculated from a single colony. This culture was then grown up overnight. The following day the bacteria were pelleted by centrifugation in a centrifuge (3000 rpm, 10 mins). The resulting pellets were resuspended in Solution I* (100µl) transferred to a clean, sterile eppendorff tube and left for 5 minutes at room temperature. To this was then added Solution II (200µl), mixing was by careful inversion, in order that the genomic DNA was not sheared. The resulting suspension was then cooled on ice for 5 minutes. Solution III (150µl) was added to precipitate protein and cell fragments. This precipitate was centrifuged (13K rpm, 10 mins) in a microcentrifuge. The supernatant was removed into a clean eppendorff tube. This should contain the required plasmid DNA. The pellet contains genomic DNA, protein and cell fragments.

Any remaining protein was removed from the supernatant by a phenol/chloroform extraction followed by chloroform wash. The resulting DNA solution was then treated with RNAse (5µl stock), to degrade the mRNA present. The DNA was

then precipitated using ethanol (100%, 3 volumes, -20°C). The precipitated DNA was pelleted by micro-centrifugation (13K rpm, 10mins). The pellet was then washed with ethanol (70%, 300µl) to remove excess salt. The pellet was air dried then dissolved in filter-sterilised AnalR H_2O (20µl).

2.7.4 Large scale preparation (Maxiprep)

LB Broth (400ml) was innoculated with either a static culture of bacteria (5ml), by preference, or with colonies from a plate. This was grown up overnight on an orbital shaker (180 rpm, 37°C). The bacteria were harvested by centrifugation (5000rpm, 15 mins) in the Sigma centrifuge. The bacteria were then suspended in 4ml Solution A and left on ice for 15 minutes. EDTA was then added to a final concentration of 10 mM. The suspension was held on ice for a further 15 minutes and Triton B then added (½ vol). The suspension was mixed gently to prevent shearing of genomic DNA and then left on ice for 30 minutes or until the solution thickened. This solution was centrifuged in a Beckman centrifuge (18K rpm, 4°C, 40 mins). The supernatent was transferred to falcon tubes and subjected to several phenol/chloroform extractions to remove any remaining protein, followed by a chloroform wash. The aqueous DNA solution was then precipitated using PEG 6000 (10% w/v). This was dissolved into the solution at 37°C followed by precipitation of the whole at 4°C for 1 hour. The precipitated DNA was then spun down in a Beckman centrifuge (12000, 20mins). The supernatant was discarded and the pellet redissolved in Tris buffer (500µl, 1M), with RNAse solution (10µl). This was incubated at 37°C to degrade all RNA. To the solution was then added Solution C (1:1 v/v) and the resulting solution left on ice for one hour. The precipitated DNA was then spun down (12K rpm, 4°C, 15 mins). The supernatent was poured off and the pellet resuspended in Tris (10mM) and NaCl (0.5M),(400µl). This solution then underwent a phenol/chloroform extraction and chloroform wash. To the now clean DNA solution was added NaCl (250mM, 0.1

vol) and ethanol (3vols, -20°C) to precipitate the DNA. This was left at -20°C for 10 minutes to ensure complete precipitation of the DNA subsequent to which it was microfuged (12K rpm, 15mins). The pellet was washed with 70% ethanol, all ethanol removed, followed by air-drying. The pellet was then resuspended in filter sterilised water (~200µl).

2.7.5 Preparation of Genomic DNA from mouse liver

Mouse liver (1g) was frozen in liquid nitrogen and fragmented in a pestle and mortar. The fragmented liver was then resuspended in 2ml extraction buffer (10mM Tris-HCl (pH 8.0), 0.1mM EDTA (pH 8.0), 0.5% SDS (w/v), 20mg/ml RNAse A) at 37°C for 20 minutes with constant agitation. Proteinase K was added to a final concentration of 200µg/ml and gently mixed before incubation at 50°C for 5 hours. An equal volume of phenol was added and very gently mixed to avoid shearing the DNA. After centrifugation (3000rpm, 20mins) the upper layer was removed and the process repeated with chloroform 3 times. The aqueous, DNA containing layer was then mixed with 0.2 volumes 10mM NH_4Ac and 2 volumes EtOH to precipitate the DNA. The DNA was recovered by gently pulling out of the solution using a hooked pipette. It was then washed in 70% ethanol, allowed to air dry for 30 minutes and then resuspended in sterile water.

2.8 Analysis and Manipulation of Nucleic acids

2.8.1 Restriction endonuclease digests

Enzymes and buffers (10X concentration) were stored at -20°C. Enzymes were added at an activity of ~5-10 units/µg DNA. Care was taken not to exceed 10% glycerol in the final reaction volume in order to prevent loss of enzyme specificity.

2.8.2 Ligation of DNA

The Covalent joining of molecuales of DNA was perfomred using T4 DNA ligase.A ligation reaction contained 10ng vector and a three to five molar excess of insert DNA.The other components of the reaction wre 1ul 10xligase buffer (700mMTris-HCLpH7.5, 70mMMgCl2,10mMDTT,1mMATP and 1ulT4 DNA ligase. The rection was incuated at 4OC O/N.

2.8.3 Synthesis and deprotection of oligonucleotides

All oligonucleotides were synthesised and subsequently deprotected by Oswel Ltd or Helena Bioscience Ltd.

2.8.4 Polyacrylamide gel purification of oligonucleotides.

Using manual sequencing plates and solutions from National Diagnostics, a 20% polyacrylamide gel buffered with TBE was poured. The oligonucleotides were loaded onto this gel and separated out by electrophoresis. (30W ~1000V, 2.5 Hrs). The oligonucleotides which were standardly ~20-25 bases in length, ran near the bromophenol blue marker.

The gel was removed from the plates then wrapped in clingfilm. The oligonucleotides were visualised by placing a UV activated TLC underneath the gel. The oligonucleotides cast a shadow on the TLC plate and could thus be excised from the gel. Only the slowest running i.e longest, portion of the shadow was excised then placed in an eppendorff with an equivalent amount of filter-sterilised dH_2O. This was placed at 4°C overnight allowing the oligonucleotide to elute from the gel into the water. The oligonucleotides were then ethanol precipitated and subsequently resuspended in an appropriate amount of water.

2.8.5 Agarose Gel Electrophoresis of DNA.

Agarose gel electrophoresis was used for the separation, quantitation and characterisation of DNA fragments and plasmids. Low melting point agarose gels were also used for the purification of specific DNA fragments. The concentration of agarose used to prepare gels was determined for the size of the DNA to be resolved and whether it was circular or linear. High strength ultrapure analytical grade agarose (Biorad)

Agarose (w/v)	Size of DNA fragments (kb)
0.6%	20 - 1
0.7%	10 - 0.8
0.9%	7 - 0.5
1.5%	6 - 0.4
2.0%	4 - 0.2

Agarose gels were cast in perspex trays of various sizes. The ends were sealed with tape and the tray placed on a level surface. A comb was used to form the sample wells. These were in several sizes, 2.5mm + 24µl, 5mm and 7.5mm. Agarose was melted in 1X TAE by heating in a microwave oven.

The melted agarose solution was cooled to ~ 50°C, ethidium bromide was added to a final concentration of 10µg/ml, and the solution poured into the gel mould. Once set the sealing tape and comb were removed and the gel submerged in 1X TAE buffer in a gel tank with platinum wire electrodes running the full width of the gel tank, at either end. The samples of DNA were loaded in the wells (~24µl/well) in addition an aliquot (2-4µl) of a 1Kb DNA ladder was added to an

adjoining well as a size marker. A current of between 75V and 120V was passed between the electrodes, across the gel. The DNA migrated down the gel away from the anode, toward the cathode, and separated according to size, the smallest fragments migrating the most quickly.The gel was then placed over a UV transilluminator. The DNA fluoresces brightly with its intercalated ethidium bromide.

2.8.6 Quantification of DNA

The concentration of a DNA sample was measured using quartz cuvettes at 260nm using the Gene Quant (Pharmacia) spectrophotometer. At 260nm an OD = 1 was taken to have a concentration of 50μg/ml.

2.9 Polymerase chain reaction

2.9.1 Primers

Bluescript - GAAATTAACCCTCACTAAAGGG

Prd-1-Fwd - CCCGCTGGAAGTGTCCACCCC

Prd-1-Rev - CCACGTCAGCCGCGGTCACCTGC

Prd-2-Fwd - CAGGCTCCATGCGACCTGGTG

Prd-2-Rev - GCCTCCTCCTCGGATCCTTTTCC

Homeodomain Rev - TGACGGAATTCATCAGTTGATTGGC

Pax-Promoter Fwd - CTGAGCACCTTTGCCAGTAG

Pax-Promoter Rev - GGTGACGAGGCAGGAAC

β -Gal Fwd - GAGCGAAAAGTACATCGTCACC

β-Gal Fwd - GCAGACATGGCCAGCCCGG

2.9.2 Binding sites

e5 - CACCGCACG**ATTA**GCACC**GTTCC**GCTCAGGCTGTCCCGT

P2xHomeodomain - GATCCTGAGTC**TAATT**G**ATTA**CTGTACAGG

E2F - GCTTCTTTTATTTTTCCAGCAGTTTA

Myc - GCCCAATCAGCGCGTGTCTTTGCCAC

2.9.3 Primer design

Oligonucleotides used for PCR were bought from Gibco BRL. The sequence selection was determined using the following criteria:

- Primers should not contain palindromic sequences
- The primers should be between 17 and 35 bases in length
- Pairs of primers should not contain any more than 4 sequentially overlapping bases

2.9.4 PCR reaction

Amplification of template DNA was achievded in cycles comprising of three steps. 1)template denaturation 2) primer annealing 3) primer extension. Each step within the cycle was carried out at a defined temperature for a specific period of time. Each cycle was repeated 25-30 times.Denaturation of the DNA template was genrally perfomed at 94)C for 40 sec, the annealing temperature was selected based on the Tm of the two primers,, usually TM-2OC for 40 sec was used. The extension temperature was at 72OC for reaction carried out with taq ploymerase but 68OC for reactions carried out with accutaq,which contains a proof reading ability.

PCR reactions were done in a 50µl reaction volume, using either Taq polymerase or AccuTaq (Sigma). A typical reaction contained the following. 5ul of Reaction buffer, Template (0.5-1µg), Primers (300ng each), 12.5mM dNTP, 2 units Taq or AccuTaq.

2.10 Analysis of proteins

2.10.1 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE separates linear protein-SDS complexes by electrophoretic sieving through the polyacrylamide matrix. The proteins are denatured by SDS, which destroys secondary and tertiary structure and β-mercaptoethanol, which reduces disulphide bridges. The treated proteins have an approximately equal charge to mass ratios and thus are separated solely on the basis of size.

All glass plates and spacers used to form the gel mold were cleaned using a detergent, and ethanol. SDS-PAGE was performed with a 10% resolving and 4% stacking gel, unless otherwise stated.

The 10% resolving gel was cast first (3.3ml) allowing room for a strip (5-10mm) of stacking gel to be poured. The resolving gel was overlaid with water equilibrated butanol to ensure the gel surface was flat and the gel percentage remained consistent.

The required volume of Tris-Glycine running buffer (~350ml) was freshly made up from a 10X stock[i]. The well comb was removed and the equipment was assembled and lowered into the electrophoresis tank. The wells were washed thoroughly with tris-glycine buffer, using a syringe and needle. Protein samples were denatured using a 6X SDS loading buffer then boiling for 5 minutes. The samples were then loaded into the wells using a 20µl Gilson pipette or a Hamilton glass syringe. Electrophoresis was carried out at 35mA per gel. On completion the

gels were stained with Coomassie blue stain or transferred onto hi-bond C transfer filter for Western blotting.

2.10.2 SDS-PAGE Protein markers

Molecular weights and concentrations of electrophoresed proteins were estimated using proteins of known molecular weight. Rainbow markers (Biorad) were most extensively used.

2.10.3 Protein staining

The gel was removed from the glass plates and immersed in Coomassie blue staining solution for 30 minutes on a rotary shaker. It was then removed, rinsed in dH_2O and placed in destaining solution, which was regularly changed, until all the excess stain had been removed.

2.10.4 Electrophoretic Transfer of Proteins (Western blot)

The proteins on the polyacrylamide gel were transferred onto Hybond-C nitrocellulose membrane (Amersham International) in the presence of transfer buffer, using a Trans-Blot semi-dry blotting apparatus (Biorad) as described by the manufacturer. A sheet of nitrocellulose membrane and six pieces of 3MM chromatography paper (Whatman International) were cut to the size of the gel. These were soaked in the transfer buffer. The gel and membrane were then sandwiched between the six sheets of wet 3MM paper and any air bubbles removed by rolling the sandwich with a glass pipette. The sandwich was assembled on the horizontal cathode plate with the membrane on the cathode side of the gels. The proteins, being negatively charge will migrate away from the anode onto the membrane. The anode was placed on top of the sandwich and electrophoresis performed for 60 minutes.

2.10.5 Immunological Detection of Proteins. (Western blot)

The Hybond-C membrane was incubated overnight at 4°C, or for 1 hour at room temperature with a blocking solution (5% w/v Marvel/PBS-Tween) on a rotary shaker. The milk protein in the Marvel bound to any free protein binding sites on the membrane, thus preventing any subsequent non-specific binding of antibody. Excess marvel was then removed by 2 rinses of PBS-Tween-20 (20ml) followed by 2 x 5 minute washes (20ml) with the same. The protein of interest was then detected by incubating the membrane with the primary antibody (diluted 1:500 - 1:2000) in 5% (w/v) Marvel/PBS-Tween-20. Residual unbound antibody was removed by 2 rinses with PBS-Tween-20 (20ml) followed by 2 x 5 minute washes with the same. The primary antibody was detected by incubating the membrane in a secondary antibody/Marvel/PBS/Tween-20 solution made up as above with the appropriate dilution as recommended by the manufacturer. The secondary antibody is conjugated to the horseradish peroxidase enzyme which catalyses a reaction releasing light, which can be detected using X-ray, film. Again the membrane is rinsed twice, then washed twice using PBS/Tween-20. The membrane is finally rinsed in ddH_2O. Visualisation was then performed using the Enhanced Chemiluminescence (ECL) kit from Amersham.

2.10.6 Electrophoretic mobilility Shift Assay (EMSA)

The most common method of detecting the presence of transcription factors is by EMSA. This takes advantage of the fact that the transcription factor will bind very tightly and specifically to a short oligonucleotide sequence containing its recognition element. When run on a native gel a bound transcription factor will impede the progress of the DNA causing a so-called band shift. This can be detected by radio-labelling the oligonucleotide probe.

All samples were equalised for protein content. In addition each assay contained 3ng polydIdC, 10µl 2x Parker buffer and the whole made up to 20µl with distilled

water. In competition assays 100X concentration of specific or non-specific oligonucleotides were also added. The samples were then incubated for 10 minutes at room temperature. Radiolabelled oligonucleotide probe was added at 200 count per second and incubated for a further 10 minutes. The samples were then run on a 4% polyacrylamide gel containing 0.25X TBE for 2.5 hours at 150V, pre-run for half an hour. The results were then visualised by autoradiography.

2.10.7 Probe labelling

Complementary oligonucleotides containing the required recognition element were annealed by heating to 80°C and cooled slowly. 25ng double-stranded oligo was radio-labeled using 5 units polynucleotide kinase with 20μCi [γ-^{32}P] ATP, incubated for 30 minutes at 37°C in a final reaction volume of 50μl. This was then made up to 100μl and the probe separated from unincorporated ^{32}P using a G-25 Sephadex column.

2.10.8 Preparation of whole cell Extracts.

1x10^6 Cells were harvested then washed in PBS and subsequently lysed in 50mM Tris-HCl (pH 7.9), 50mM KCl, 2mM DTT, 2μg/ml PMSF, 25% (v/v) glycerol. Samples were freeze thawed in ethanol-dry ice slush and 37°C. They were then centrifuged for 20mins at 13000 rpm. The resulting supernatant was whole cell extract.

2.10.9 Preparation of Nuclear and cytoplasmic extracts

Nuclear and cytoplasmic extracts were made by the method of Dignam *et al.* 1983.

2.11 Microscopy and immunostaining

All cells for microscopy were grown on 22x22mm cover slips at a density of $1x10^4$ cells/coverslip.

Cells were prepared for immunostaining by washing twice with PBS followed by fixation using ice cooled MeOH for ½ an hour. The fixed cells were then incubated in Hanks Buffered Saline Solution (HBBS) and 1% BSA for 30 minutes. The cells on their coverslips were then incubated with an anti-Pax-3 antibody (1:200 in HBBS/1% BSA) for 1 hour. The cells were then washed with HBSS/1% BSA, three times. A secondary anti-rabbit IgG fluoroisothiocyanate (FITC) was added (1:50 HBSS/1%BSA) overnight at 4°C. The cells were again washed three timed with HBSS/1% BSA and then incubated with 25μg/ml propidium iodide for 20 minutes. After a final wash with HBSS/1% BSA the coverslips were mounted onto slides with DAKO fluorescent mounting medium, confocal microscopy was then carried out as standard.

For EGFP images the cells were not fixed, coverslips were washed in PBS followed by incubation in 20μg/ml Diamidinopropylindole (DAPI) in PBS. After rinsing in PBS the coverslips were mounted on slides and pictures taken using a digital fluorescence microscope. Images were taken using a Biorad MRC 600 confocal Microscope with a Krypton-Argon Laser or an Axan Imaging Lightning system with inverted fluorescence microscope attached to a cooled photonic CCD digital sensicam still video camera.

2.12 Cell Line Culture

2.12.1 ND7 cell lines

The ND7 neuronal cell line (Wood *et al.* 1990) was maintained in DMEM containing 10% (v/v) Fetal calf serum, 10 units/ml Streptomycin/penicillin and

1mM glutamine. Where serum starvation is indicated this refers to the same media without serum. All cells were maintained using a Class II biological safety cabinet and a 5% CO_2 incubator at 37°C.

2.12.2 Neural Crest cell culture

Female adult mice were mated with adult males and checked for copulation plugs at 12 hourly intervals. If a copulation plug was present, females were culled 8.5 days later (E 8.5 p.c) and embryos removed from the uterine sacks, using low power magnification (x2-20) microscopy. The embryos were treated with dispase (80μg/ml) in PBS for 15-20 minutes. The proteolysis was subsequently stopped by addition of 1:1 DMEM/F-12 Hams nutrient medium with 10% Fetal Calf Serum. The neural tube was subsequently excised from the embryo and cut into two sections (after the reomoval of the most anterior and posterior portions). These two sections of neural crest were then allowed to attach to culture plates coated with fibronectin. Plates or slide were coated with fibronectin by adding fibronectin in distilled water 20μg/ml for 2 hours, the excess removes and the plates or slides allowed to air dry. The neural tube and subsequently migrating neural crest cells were maintained in the above media.

2.12.3 DNA transfection

The DNA used was made by large scale DNA preparation and resuspended in H_2O under sterile conditions.

Transient transfections were carried out using the calcium phosphate method (Gorman 1985). ND7 cells were plated onto 90mm diameter tissue grade dishes (5 x 10^5 cells/dish) or onto 22mm x 22mm sterile glass coverslips coated in fibronectin (1.5 x 10^4 cell/plate). The next day they were transfected using the calcium phosphate method of Gorman 1985.

2.12.4 Chloramphenicol Acetyl Transferase (CAT) assay

After transient transfection by the calcium phosphate method of Gorman the cells were harvested 55 hours post transfection into ice cold PBS. The cells were pelleted by microcentrifugation (3000rpm, 3 mins) and washed twice with ice cold PBS. They were then resuspended into 100µl 0.25M Tris and lysed by freeze/thawing –50°C/37°C five times. The cell fragments were removed by centrifugation and the supernatent analysed for protein content (Pierce BCA assay kit see manufacturer for instructions). An equal quantity of protein from each sample in an experiment was added to 35µl Tris (0.5M, pH 7.8), 20µl Acetyl CoA (4mM) and [^{14}C]-Chloramphenicol (1µCi). This was made up to 180µl. It was incubated for 6 hours, the Chloramphenicol and derivatives extracted using ethyl acetate (1ml). This was then dried down using an evacuated rotary evaporator, resuspended in 15µl ethyl acetate and spotted onto a TLC plate. The chloramphenicol and derivatives were then separated by chromatography using chloroform/methanol (95:5). The chloramphenicol and derivatives were visualised using autoradiography and a STORM phosphorimager. Quantitation was by ImageQuant. The plasmid pCMV-βGal was used as an internal control in all transfections, to correct for differences in transfection efficiency.

2.12.5 [^{3}H]-Thymidine incorporation assay.

Cells cycle analysis using thymidine incorporation was as follows. Cells were grown as for normal experimental protocol. One hour before harvesting [^{3}H]-Thymisine (1µl/ml) was added to the media. After harvesting the cells were kysed using Tris (100mM) / 1% SDS. The DNA was precipitated from the solution using 10% TCA and filtered out using a GFC filter with a Buchner funnel and washed with TCA. The GFC filter was then added to scintillation fluid (2ml) and thus the radiation counted.

2.12.6 β-Galactosidase assay

The cell lysates from the above CAT assay were added (20µl) to 0.2 ml (2mg/ml ONPG, 60mM Na_2HPO_4, 40mM NaH_2PO_4) and 1µl (60mM Na_2HPO_4, 40mM NaH_2PO_4, 10mM KCl, 1mM $MgCl_2$, 50mM β-mercaptoethanol). The assays were incubated at 37°C for 2 hours, then 1M $NaHCO_3$ was added to stop the reaction. The colourimetric change was measured at 420nm and related to the results of the CAT assay.

Regulation of Pax-3

3 Regulation of Pax-3

3.1 Introduction

As with any major developmental control gene, expression of Pax-3 must be tightly governed. The spatio-temporal regulation of transcription of any gene is under the control of a large number of transcriptional regulators that act upon the promoter and enhancer, the fragment of DNA often present, 5′ to the start of transcription that carries the recognition elements for transcriptional activators and inhibitors.

Previous work on the *Pax-3* gene (Natoli *et al.* 1997) has shown that the elements required for cell specific expression and regulation during retinoic acid (RA) and DMSO induced differentiation of P19 embryonal carcinoma cells are located within a fragment 1.6kbp 5′ to the start site of transcription. When introduced into transgenic mice this is sufficient to mediate faithful expression of a reporter gene in the dorsal neural tube and somites. However while this fragment mimics ventral *in vivo Pax-3* expression up until E9.5, additional elements found in a 14kbp region 5 to the promoter, are required for expression in migrating myoblasts and neural derivatives anterior to the hindbrain. To gain further insights into the factors that regulate Pax-3 transcription, I have analysed the elements that are important in the regulation of Pax-3 expression within the region 1.6kbp upstream of the transcription start site of Pax-3. The published sequence (Fig. 3.01) was initially cloned (Natoli *et al.* 1997) by primer extension after using a 5' UTR probe on a 129/sv mouse genomic library.

GAGCTCTAATGCTCCTCCCCAAATGTGGGTGGGATCGGGGGTGGGGTGTCCAGACATTTT -1520

CTGGAGTCCCCAGAGGCCCGTCTAGCCTGCTCCCAGGGTCGATAAGCCCTTTTGACTTCA -1460

GGGGGAGCAAAGAGCTACCCAGCTTTCTCCACTGGGAGTGAACTCTTAAAGGGCCAGCTC -1400

AGGGACTTTGAACAACTACACGCACGCCCCCCCCCCTTGAGTTTGCCTTAATTAGGGGGT -1340

GGGCAAGTAGGATTGGGGGCCTGAGTTCACATCAAAACTGTCAAAGGAGATGGGGGGGGA -1280

GGACATGGAGAGAGGGGTGGAGGATTCAGGGGCAGGAAAAATTCAAATATCATAAAAAGA -1220

AAAGGGAACTTGAAGCCCAGAAGTCATTGTTTCCTTTATTTCAAAGGGGAAAAAACCTGC -1160

CTGATTCTCATCCTTTTGATTGATTAAGGCCCTGAATACCTCTCGGCCCCAGAGTGACAG -1100

TCCCTGAATAGACTTGAGGGAATAAAGAGCAGCGCAGAACGTGGGGCTGGCACGGGAGGG -1040

GGGTGTAAGGGAGGAGTGTTCGCTGGCACTTACCAAGTTATAAATAAAAGGCTAGGCACA -980

ATGGTACCTTCTCTAAGGACAGACAGTCTTTACAACACTCCTGGCGTCATATCCTGCTGG -920

GGACACTTCAGCTCCTAGCCAAGACGTTGCTTCTTTTATTTTTCCAGCAGTTTAGTCTGA -860

ATGCCATAATAAATTCCTGAGAACAAACGCTGCACCCGGGCAAAACCTCAACATATAGAT -800

GCAAGTGCATCGGGGATGAATGTGTACGTGGAGATTTAAAGTCCCCGCTTCTCAGAAGGG -740

TTTTTAAACCAAGAAAAGGAGCTGGCTTCCCCATTTCAGCTTGCAACTTGGAGCCCAGGG -680

GAGGCCTAACCTCTTCAGTCTCTGTCATGGCTATGCCACCTGTCTTTTCATTCCCAGCCT -620

AAGAAAGACAAAAGGACCTTCTGGGCACATCTTGGACTCTTCCTGAGAGATGCCTTTCTG -560

GGAGGCTCACATAGGGAGCTCAGGCCCAGAGGCAGACCAGAGAAAGGAAAAGGAAACACA -500

TCCGGAATTCCTCTGTGCCCCTTTCCAAATTTGCTCCACCGCCATCTCCGTATTAGTAAT -440

CCCGAAGCGAAATCCACAGGTGAAAAGCAGGGGCAGACAGAGGAAATAAGGGGCGGATAT -380

AGCAAGGTTTCCAGAGAAAGGCGGAAAGAGATTAGGACAGAGTAATAGAAAACCAAGATG -320

ATCAAGGCTGAATCTCATGTAGAGGGACAGGGACACCGTGAGCCTTTTTCCATCCTCACT -260

GGCTCTTGACACCCAAGCTGCCTCTCCTAGTTTCAACCTGTCCACCCTTCTCTTGAGAAA -200

GGGACACAGAGCTAGTCCCCCTCTGGGTCGGCGGAGGCCTGGGGACCGTCAGGGATGGGA -140

AGAGAAGTAGCTTCTGTCGCCCAATCAGCGCGTGTCTTTGCCACCCGGGACGGTCTCCTC -80

CTCGGCCAATCGCAGCTCAGGGCTCCTGATCAAGCTTTGGGTGAAAGAACTAATAAATGC -20

TCCCTAGTCCGGATCCCTGCACTCGGTGTCACGACGGGAGGAGACTTGGGACGTGTTCCT +42

GCCTCGTCACCACCCTCCCTCACCCGAT +70

TAATAAATG – TATA Box

C – Possible start of transcription.

TATTTTTC – Potential E2F binding site

CAGCGCGTGTCT – Potential Myc/Mad/Max binding site.

GGTACC – Kpn I cleavage site

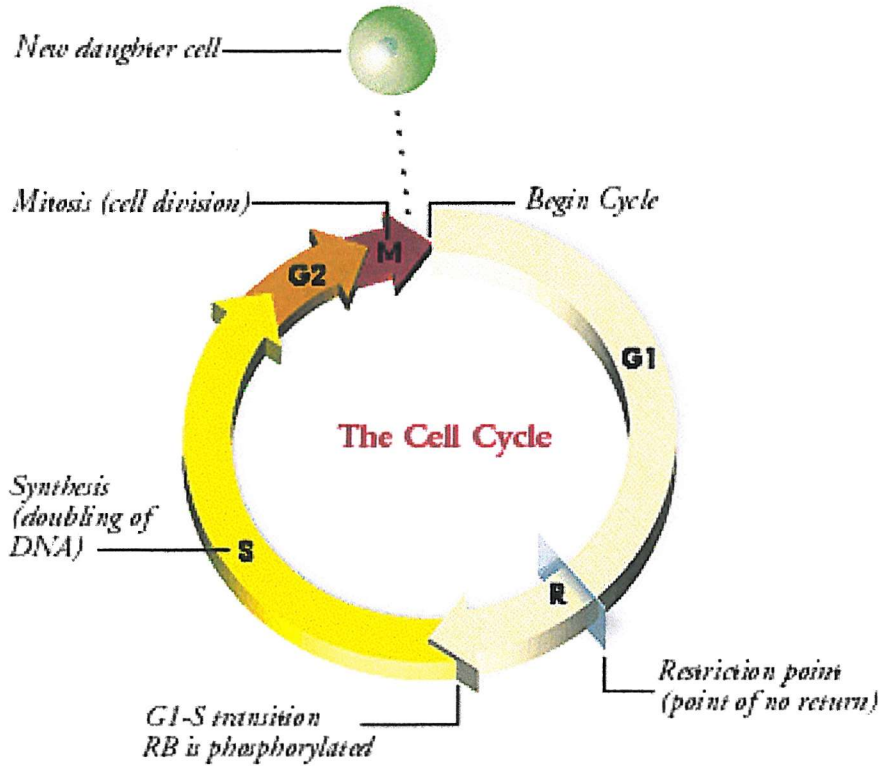
GAGCTC – Sac I cleavage site

GGACCG – Ava II cleavage site

Fig. 3.01 Nucleotide sequence of the 5' flanking region of *Pax-3*. Only the coding strand is shown. The TATA box is underlined and the start site of transcription is shown in red. A potential Myc/Max E box element is shown in yellow, and E2F elements in blue.

3.2 Pax-3 and the Cell Cycle

In proliferating cells the act of division is the culmination or a series of events that have occurred either continuously or discretely since the last time a cell divided. The events are traditionally broken into the four phases of the cell cycle.



Diagrammatic representation of the phases of the cell cycle – illustration by George Eades http://www.geocities.com/CollegePark/Lab/1580/cycle.html

Gap 1 (G1) – This is a resting phase after mitosis characterised by gene expression and protein synthesis. It is the only phase of the cell cycle affected by external stimuli such as mitogens, growth factors and adhesion molecules. It is during this phase that the cell commits to further division or withdrawal from the cell cycle. Once the commitment to enter into replication occurs, at the restriction point division is inevitable.

Synthesis (S-phase) – Shortly after the restriction point is passed the cell starts replicating its DNA in a diploid cell 2n DNA is replicated to 4n. The start of S-

phase is marked by initiation of DNA synthesis. Gene expression and protein synthesis continues.

Gap 2 (G2) – During this phase the cell has two full diploid sets of chromosomes. The cell is readying itself for cell division, cell mass and cellular proteins must be sufficient to support two daughter cells. Gene expression and protein synthesis continues.

Mitosis (M-phase) – Cell division occurs. Gene expression stops.

In addition there is a fifth phase of the cell cycle, withdrawal or G0. This is similar to G1 but cells cannot enter into S-phase from it. Cells in G0 can however be reactivated and stimulated into G1 where they can progress into S-phase. Cells that start to differentiate do so in G0. As a transcription factor involved in the continued proliferation of neural crest cells Pax-3 may be controlled by, and/or have a part in controlling the cell cycle.

3.3 Results

3.3.1 Examination and dissection of the Pax-3 promoter

In order to study the effect of various transcription factors on *Pax-3* expression a region of DNA encompassing 1.6kbp upstream and 50bp downstream of the transcription start site was cloned by PCR (using AccuTaq (Sigma) with an error rate of < 1 in 3000) from mouse genomic DNA using the primers Pax-prom-Fwd and Pax-prom-Rev, derived from the published sequence. The resulting PCR fragment of 1.6kbp was TA cloned onto pGem-T-Easy. The sequence of the Pax-3 promoter sequence was then verified by sequencing. The 1.6kbp Pax-3 promoter fragment was then excised from pGem-T-Easy using the restriction enzymes *SphI* and *SalI* into the reporter vector pCATBasic which was linearised using *SphI* and *SalI*.

The mouse strain used for generating the promoter was mf1. Three clones from two separate PCR products were sequenced. All of these three clones had identical sequences, however they showed some polymorphisms when compared to the published sequence. A sequence alignment of the mf1 sequence and the 129/sv Pax-3 promoter sequences is shown in Fig. 3.03. None of the five polymorphisms occur within the binding sites of E2F or c-myc, which are the transcription factors examined in this chapter.

Fig. 3.02 Sequence alignment from the cloned mf 1 mouse strain (top line) and the published sequence (Natoli *et al.* 1997) from the 129/sv mouse genomic library (bottom line).

C – Polymorphism

TATTTTTC – Potential E2F binding site

CAGCGCGTGTCT – Potential Mvc/Mad/Max binding site.

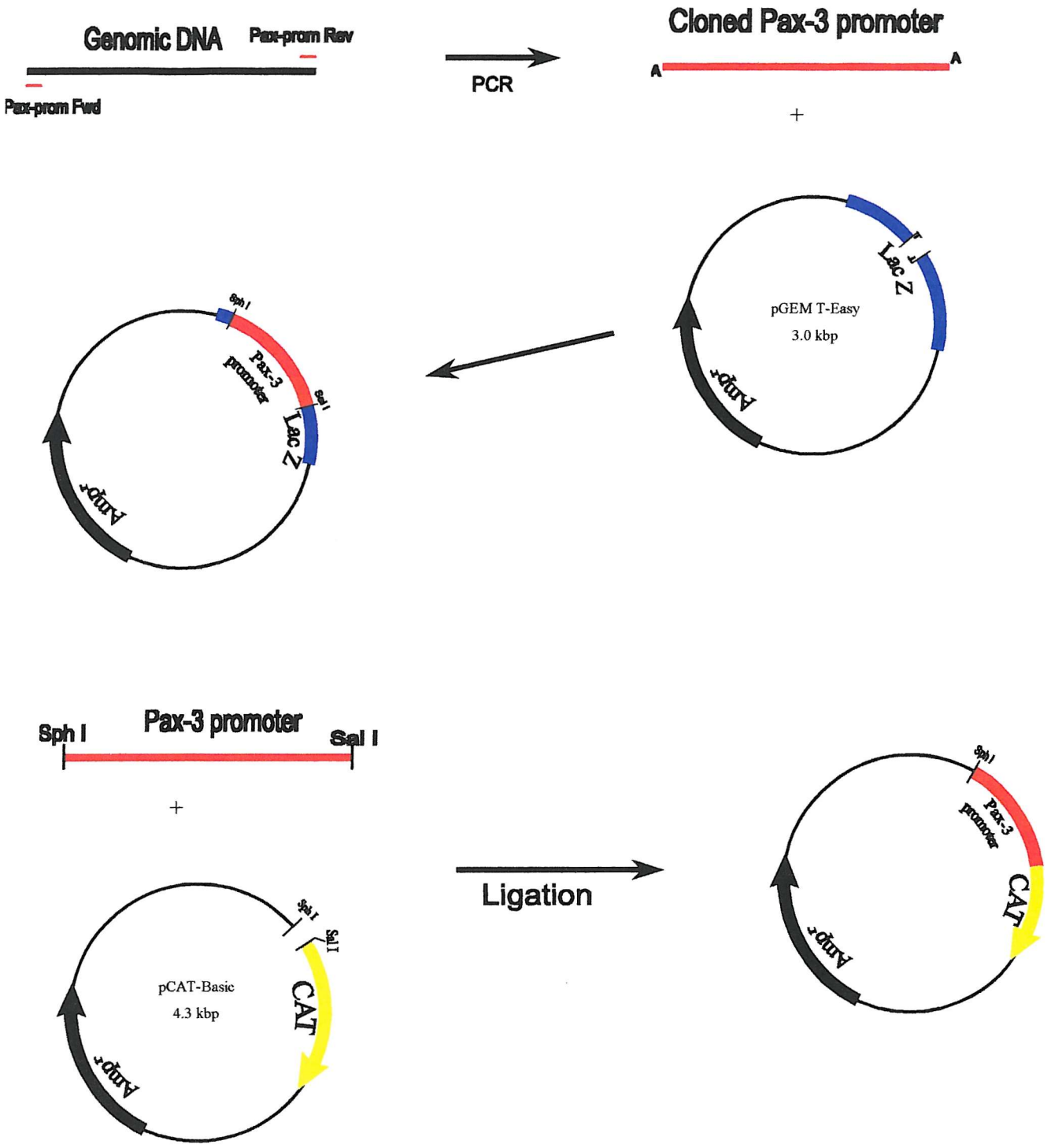


Fig. 3.03 Schematic representation of the steps used in cloning the Pax-3 promoter from mouse (mf1) genomic DNA.

3.4 The *Myc* family of transcription factors.

There are a number of proteins required for cell cycle progression, such as cyclins, cyclin dependent kinases (cdk's), replication factors and some enzymes of nucleotide metabolism that are not present in quiescent cells, suggesting that cellular proliferation is clearly controlled at the level of gene expression. Some of the key proteins in this control are the *Myc* family of oncogenic transcription factors. Consisting of four closely related proteins, c-myc, N-myc, L-myc and s-myc and a fifth, B-myc, which has some homology but lacks several crucial functional domains, this family is responsible for the expression of many key cell-cycle proteins. The first member discovered, and the most ubiquitous, *c-myc*, has emerged as a central oncogenic switch in many human cancers.

There is a wealth of biological data pointing to the central role of *c-myc* in control of proliferation. Its expression is solely proliferation dependent, its constitutive expression in cells induces progression through the G1/S-phase boundary and into the cell cycle even in the complete absence of growth factors (Eilers *et al.* 1989). While conversely antisense *c-myc* severely impaired the growth factor stimulated proliferation cells in culture. Together with the observation that homozygous myc deficient mutant fibroblasts showed a marked delay in the G1/S transition, this led to the speculation that *c-myc* was centrally involved in transcriptional expression of genes expressed in G1 and required for the G1/S transition. Subsequent identification of target genes such as cyclins, CDK's, E2F, p53 proved this to be substantially correct.

c-myc expression and thus control of cell proliferation appears to be tightly controlled at several levels by signal-transduction cascades, especially the Ras-Raf cascade. It is down regulated in differentiating cells and up-regulated in dividing cells, the importance of its tight control means that its *in vivo* half-life is extremely short. (Chin *et al.* 1995)

The *myc* genes are members of the basic Helix-loop-Helix/Leucine Zipper (bHLH/LZ) superfamily of transcription factors that bind to DNA as obligate dimers with a variety of possible partners. In the case of the *myc* genes they are the transcriptional activating members of the Myc/Max/Mad network of

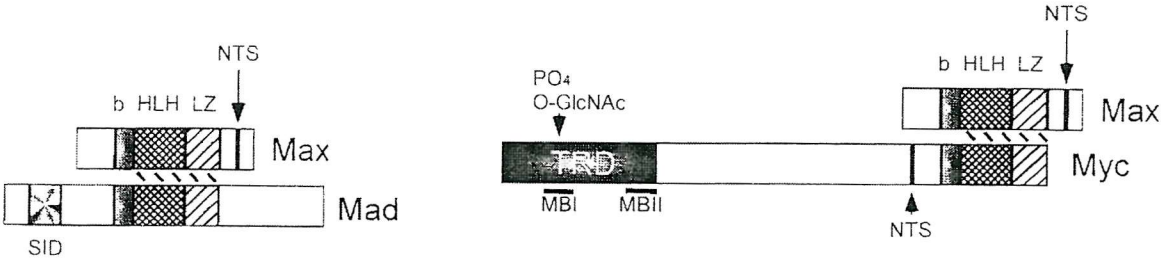


Fig. 3.04 Graphical representation of the interaction of Max with either Mad or Myc. The basic (b), Helix-Loop-Helix (HLH) and Leucine Zipper (LZ) regions are all shown as are the myc transregulatory domain (TRD), its phosphorylation and glycosylation sites, the Nuclear Targetting signal (NTS) of Myc and Max, and the Sin-3 interacting domain (SID) of MAD. MB I and MB II are conserved sequences of the Myc family, Myc Box I and II.

transcriptional regulators. The common binding partner in this network is Max. In order for either Myc to bind to DNA they must heterodimerise to Max, Max however can homodimerise and in doing so competitively reduces binding of the heterodimer (Amin *et al.* 1993). This is shown in the diagram Fig. 3.04

These heterodimers bind to myc type E-box recognition elements. The Myc/Max heterodimer activates transcription through interaction of the myc transregulatory domain (TRD) with the transcriptional machinery, and possibly through histone acetylation. The central component of the network, Max is very stable and is almost constitutively present in all cells, very little regulation of its expression has been reported. The Max-Max homodimer represses Myc transcription by competition with the Myc/Max heterodimer for E-box binding. This is compounded by the fact that although Max-Max and Myc-Max have similar DNA binding specificities and constants, the interchange rate of Myc/Max is significantly faster than that of Max-Max. However the dimerisation of Max to myc is preferred to the homodimerisation of Max, thus Max need to in significant excess in order to repress Myc function. (Sommer *et al.* 1998).

The third partner in the Myc/Max/Mad network, is a family of four proteins Mad1, Mxi1, Mad3 and Mad4 collectively referred to as the Mad family. These can heterodimerise with Max and bind to the same consensus sequence as the other dimers in the network. They act to antagonise c-myc and are present in very low amounts in dividing cells, but are upregulated in differentiating, quiescent and terminally differentiated cells. The Max/Mad heterodimer binds to the Myc E-box recognition element, recruits Sin3, co-repressors and histone deacetylases to the Sin-3 interacting domain (SID) to actively repress transcription and stops cell cycle progression (Cultraro *et al.* 1997).

3.4.1 The N-myc gene and neuronal development

While *c-myc* is undoubtedly the most ubiquitous member of its family, in the context of neuronal proliferation, development and differentiation it appears that the *N-myc* gene is far more important. First discovered as a gene that was often amplified in neuroblastomas, it shows sequence similarities to c-myc and is essential to neuronal proliferation. N-myc down-regulation has also been shown to be essential to neuronal differentiation in the spinal cord (Shimono *et al.*1996). While the targets of N-myc are not as well defined, it certainly binds to Max in much the same way and appears to fulfil at least some of the same functions of *c-myc* and is found to be constitutively expressed in most neuroblastoma cell lines. During neural crest development, N-myc is initially expressed in the entire population, it is then differentially expressed in migrating, and subsequently differentiating neurons (Wakamatsu *et al.*1997). Overexpression of N-myc leads to massive ventral migration. While it was still present in maturing neurons it became localised to the cytoplasm and therefore inactive as far as transcriptional activation goes (Wakamtsu *et al.* 1993).

3.4.2 The regulation of Pax-3 *expression by N-myc and c-myc*

Examination of the 5' promoter sequence of Pax-3 revealed a Myc/Max binding site. To determine if the Myc/Max/Mad network of proteins could regulate Pax-3 transcription the Pax-3 promoter (-1.6kbp – +120bp) was inserted upstream of the Chloramphenicol Acetyl Transferase gene in the reporter vector pCAT-Basic.

To determine if c-myc or N-myc can regulate Pax-3 promoter activity, the Pax-3 promoter construct (Pax-3CAT) was co-transfected into ND7 cells with either C-myc or N-myc expression vectors. After 60 hours, the cells were harvested, and the CAT activity assayed. When Pax-3CAT is transfected into ND7 cells, chloramphenicol acetyl transferase expression will be under the control of the Pax-3 promoter. The CAT enzyme is responsible for the conversion of chloramphenical into the mono- and then di-acetylated forms in the presence of acetyl Co-A. By calculating the percentage conversion, using densitometry the amount of CAT activity and thus promoter activity, may be measured. Using this technique the effect of N- and c-myc on the basal level of *Pax-3* transcription in ND7 cells was measured.

The results were then correlated to the basal expression level of the Pax-promoter co-transfected with a null expression vector. These are displayed graphically in Fig. 3.05.

N-myc and C-myc produced very high levels of activation of the 1.6kbp Pax-3 promoter, in the case of N-myc an 18-fold increase in Pax-3 promoter activity was observed and with C-myc a 10-fold increase was seen. In contrast the addition of expression vectors containing Max led to the inhibition of Pax-3 promoter activity. No effect on the expression of pCATBasic or pSVβGal was observed in the presence of N-myc, C-myc or Max expression vectors.

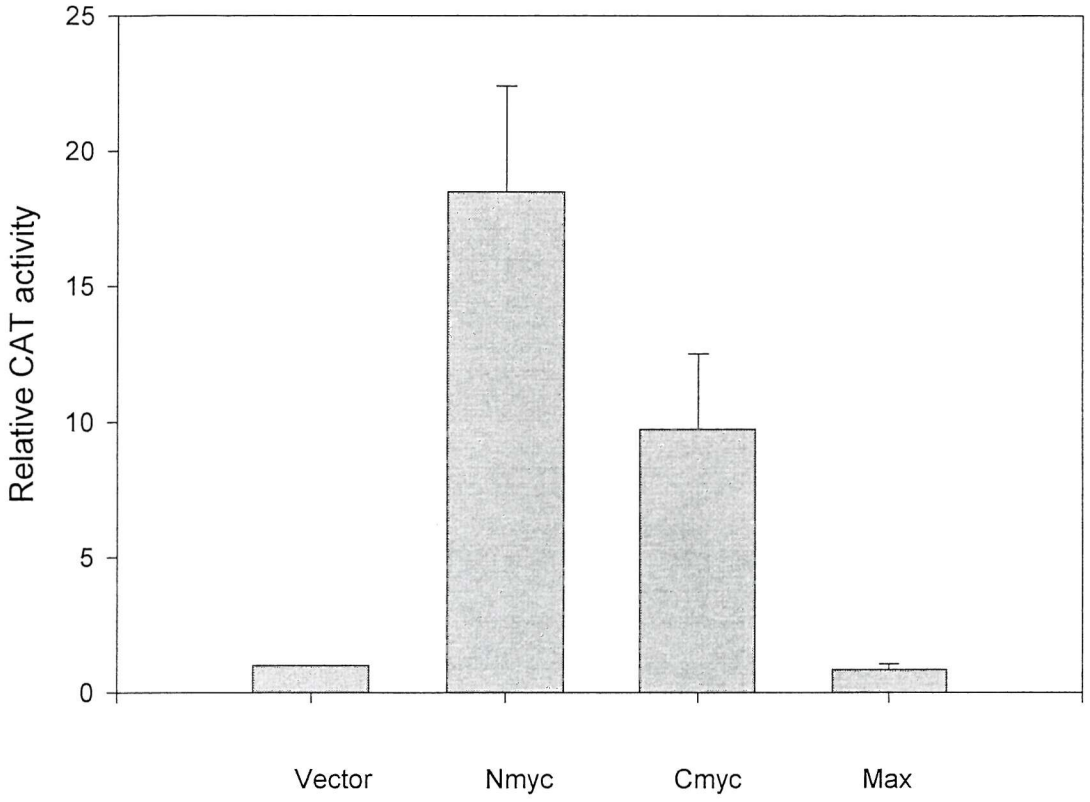


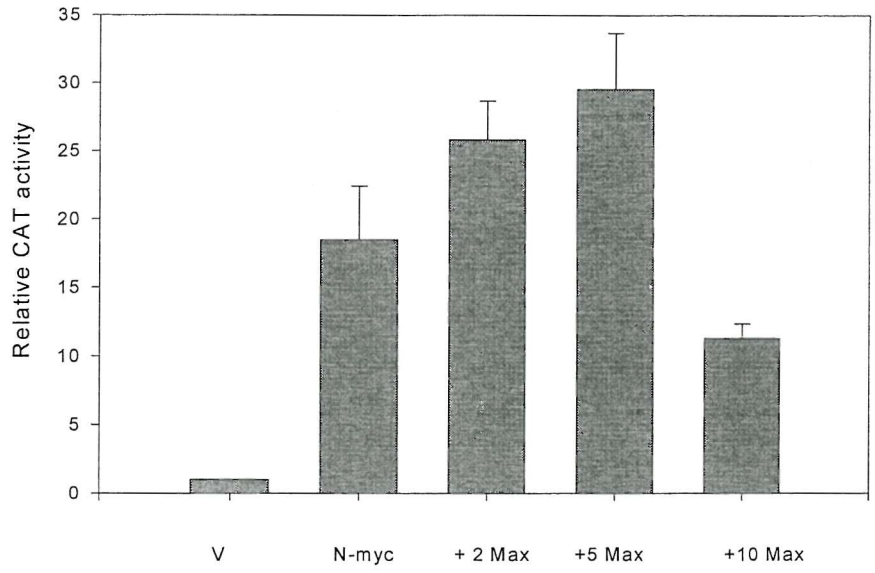
Fig: 3.05 CAT assay showing ND7 cells co-transfected with PaxCAT and either 5μg empty expression vector (vector), 5μg N-Myc expression vector, 5μg C-myc expression vector or with 5μg Max expression vector. Each value is plotted relative to the CAT activity in ND7 cells transfected with PaxCAT and the empty expression vector. Each value is the mean of three independent experiments ±SEM

3.4.3 The effect of the Myc/Max/Mad network on the Pax-3 promoter.

As C-myc and N-myc are known to heterodimerise with Max, ND7 cells were co-transfected with the Pax-3 promoter construct together with N-*myc* expression vectors (5μg) and increasing amounts of the *Max* expression vector. (0, 2, 5 and 10μg; Fig. 3.06). The effect of adding increasing amounts of Max expression vector led to a further increase in the Pax-3 promoter activity. Similarly co-transfection of C-myc with Max also led to a greater activation of Pax-3 promoter activity than with C-myc alone. However at high concentrations of Max (10μg), with either C-myc or N-myc a decrease in Pax-3 promoter activity was seen.

The third member of the network, Mad, is known to have an inhibitory effect on promoters. It has also been shown to be upregulated during differentiation of a number of types of cells. Having shown that *Pax-3* expression is affected by Myc/Max, we also investigated whether Pax-3 transcription would be repressed by Mad, in light of the fact that Pax-3 downregulation has been shown to cause differentiation. The *Pax-3* promoter construct was co-transfected with expression vector containing *Max* and *Mad.* (Fig. 3.07). As shown in Fig. 3.07 co-transfection of Mad with Max significantly inhibits the expression of the Pax-3 promoter. In contrast transfection of Mad alone had no effect on Pax-3 promoter activity. The ability of Mad to inhibit Pax-3 promoter activity in conjunction with Max appears to be dependent upon the leucine zipper motif of the Mad protein. Since deletion of the Leucine zipper region of Mad abolished the ability of Max/Mad to repress Pax-3 expression.

A)



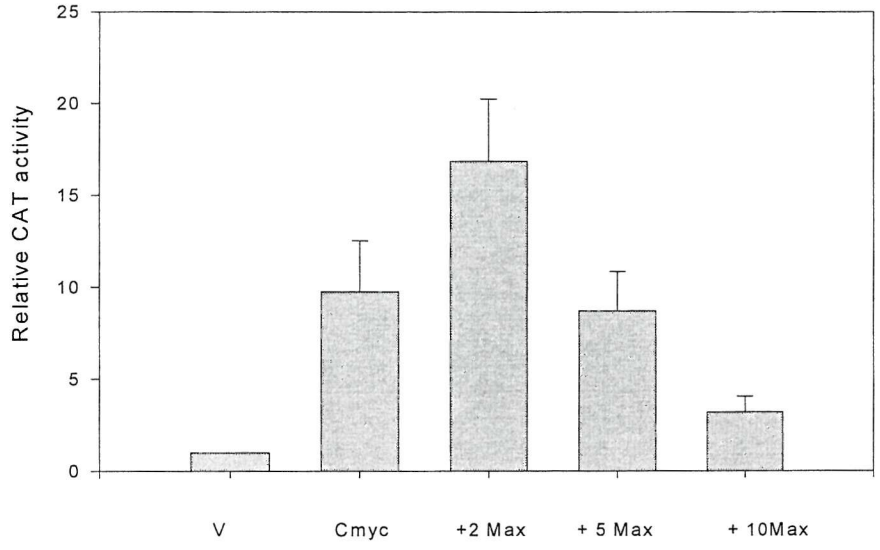


Fig 3.06 CAT assay showing the effect on Pax-3 promoter activity of adding increasing amounts of Max together with either N-myc (A) or C-myc (B). ND7 cells were co-transfected with PaxCAT (5μg) with either 5μg empty expression vector (V), or with 5μg myc expression vector , 5μg myc with 2μg max expression vector (+2 max), 5μg myc with 5 μg max (+5 max) or with 5μg myc with 10μg Max. Each value is plotted relative to the CAT activity in ND7 cells transfected with PaxCAT and the empty expression vector. Each value is the mean of three independent experiments ±SEM.

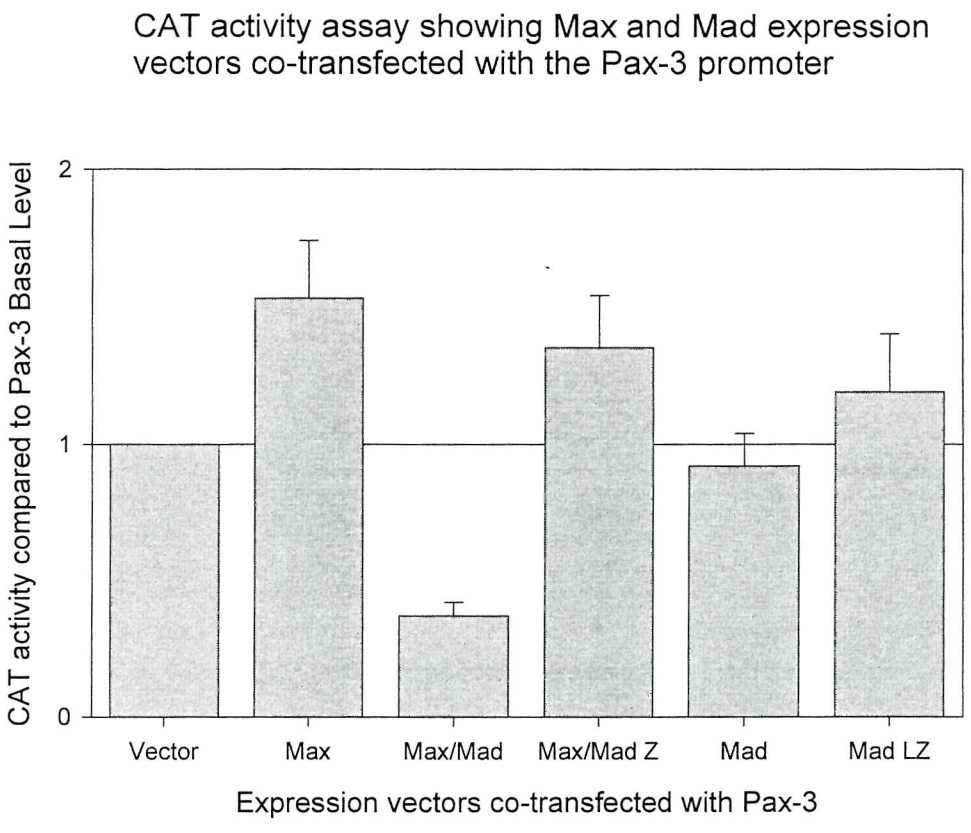


Fig. 3.07 CAT activity of ND7 cell extracts co-transfected with the Pax-3 promoter, a *Max* expression vector and either the *Mad* expression vector or a vector expressing a Mad deletion missing the Leucine Zipper domain (Mad LZ) plotted with respect to the basal level of expression of the Pax-3 promoter CAT co-transfected with a null expression vector.

3.4.4 The Myc binding site in Pax-3

To localise the site of Myc binding within the Pax-3 promoter, several deletions of the *Pax-3* promoter were made. These are shown schematically in Fig.3.08. The truncated promoter constructs were made by digesting PaxCAT with *SphI* and either *KpnI*, *SacI*, *AvaII* or *XmaI*. The digested plasmids were then purified, the ends filled in with either Klenow or T4 DNA polymerase and religated and sequenced.

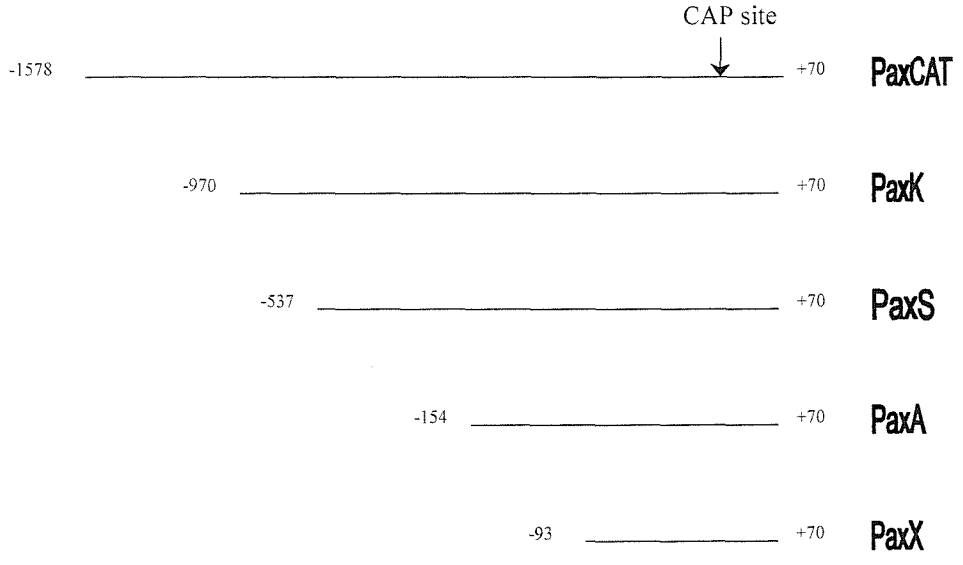
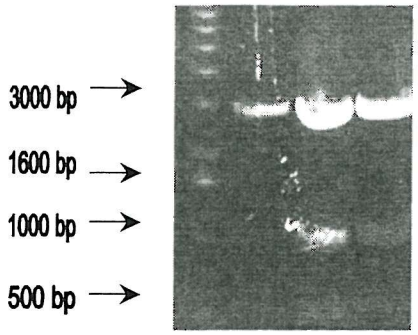
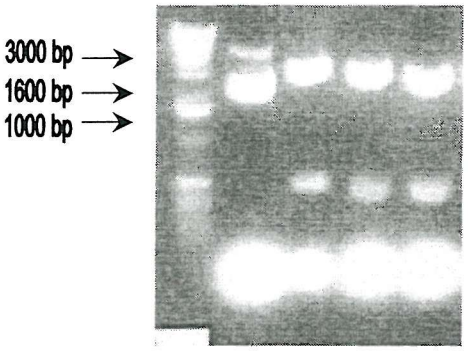


Fig 3.08 Schematic diagram showing the truncated Pax-3 promoter constructs used to identify the sequence through which Myc/Max activates Pax-3 transcription

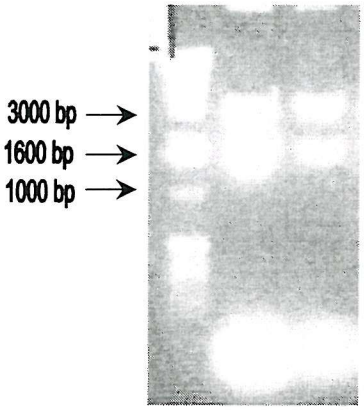
Agarose gels showing the analysis of these constructs are shown in Fig. 3.09, while schematics showing their construction are shown in Fig. 3.10.



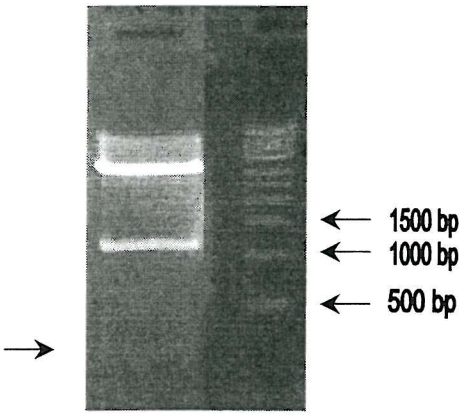
As expected the full length Pax-3 promoter in pCAT-Basic produces two fragments of 1079bp and 550bp when cut with EcoRI



As expected the Kpn I deletion produced fragments of 540bp and 3110 when cut with EcoR I



As expected the Sac II deletion clone produced fragments of 1660bp and 3240bp when cut with EcoR I



As expected the Ava II deletion gave a 106 bp fragment when cut with Hind III

Fig. 3.09 Agarose gel pictures of the Pax-3-promoter and its deletion mutants in pCAT-Basic.

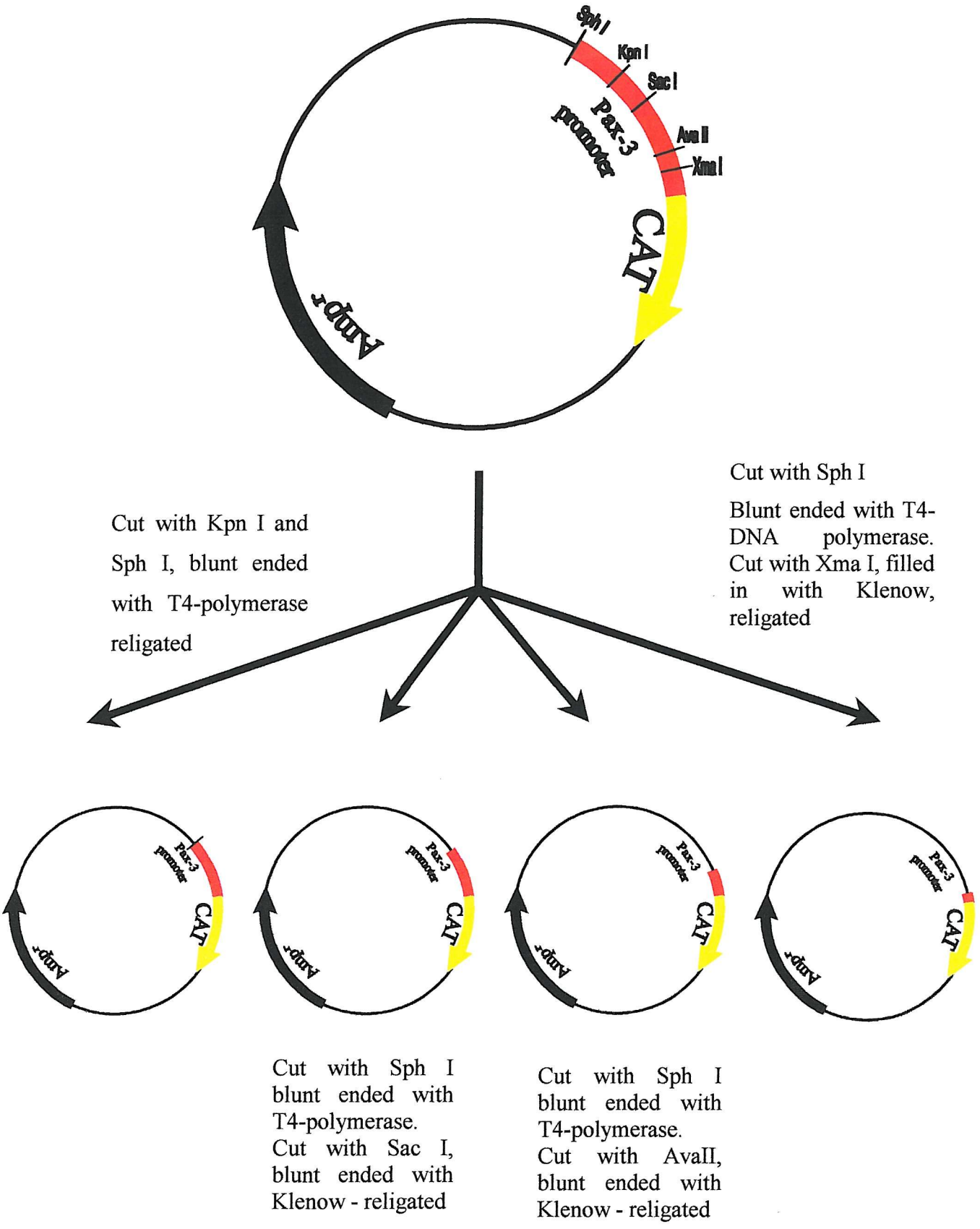


Fig. 3.10A Schematic diagram showing the procedures used for the generation of Pax-3 promoter deletion constructs.

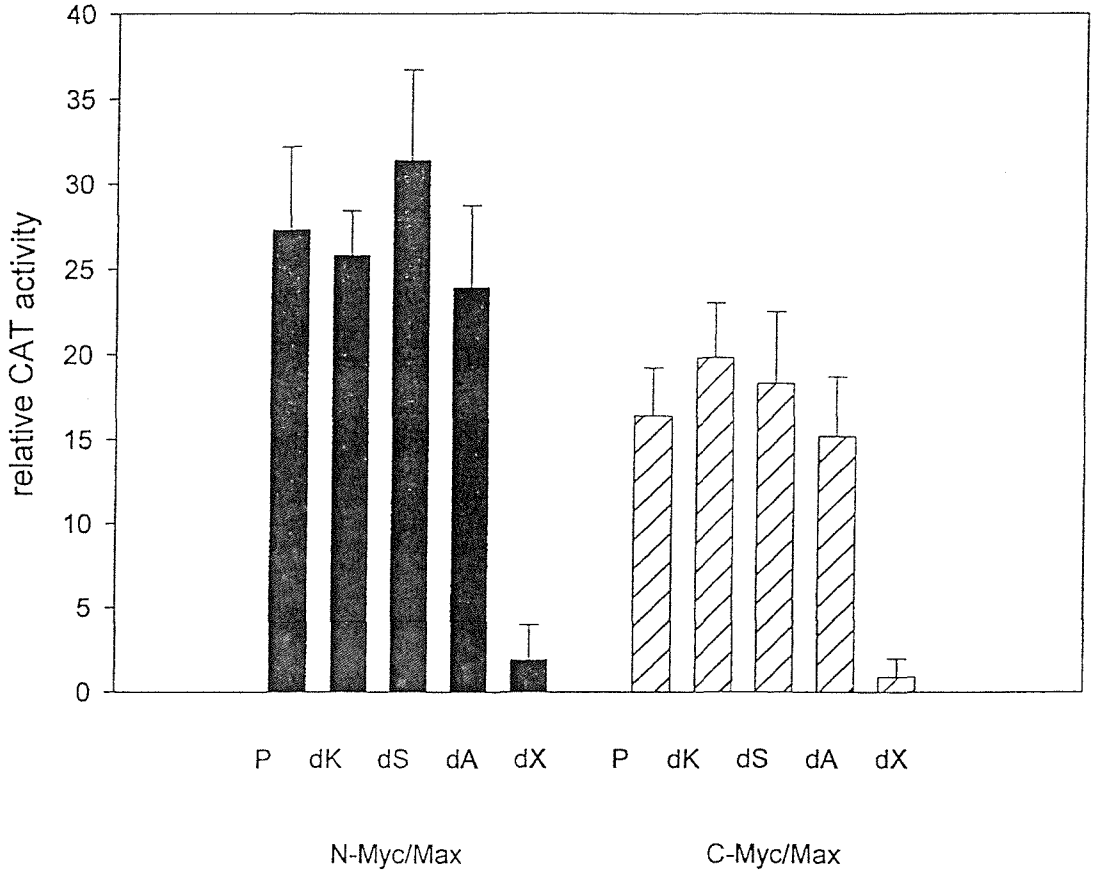


Fig. 3.10B CAT activity of ND7 cell extracts co-transfected with the Pax-3 promoter deletions (5μg) and either N-myc and Max or c-Myc and Max, with respect to the basal level of expression when cotransfected with a null expression

Co-transfection of C-myc or N-myc with the truncated Pax-3 promoter constructs revealed that all the truncated promoter constructs were activated by both C-myc and N-myc apart from the construct PaxX. This suggests that the myc E box site is located between -157 and –93 with respect to the transcription start site of the Pax-3 promoter. This is the region of the Pax-3 promoter that contains a putative E box sequence CACGTC which differs from the consensus E box site CACGTG by 1 base pair. In order to ascertain whether myc could bind to this sequence, the Pax-3 E box site was synthesised as an artificial oligonucleotide and used as a probe in electrophoretic mobility assays. ND7 nuclear extracts were incubated with the ^{32}P labelled oligonucleotide and run down a 4% non-denaturing polyacrylamide gel. As shown in Fig 3.11 one high molecular weight complex was seen binding to the E box motif from the Pax-3 promoter in dividing ND7 cells. To determine the identity of this complex, the ND7 nuclear extracts were incubated with antibodies raised against C-myc, N-myc, Max and MyoD. (Fig. 3.11 A & B). These autoradiographs show that binding to the oligonucleotide is abolished by the N-myc, c-myc and Max antibodies, but not by MyoD. These results suggest that the protein complex binding to the Pax-3 E box sequence in ND7 cells comprises a combination of N-myc/Max and C-myc/Max.

3.4.5 The myc-binding site and the cell cycle.

To investigate whether in vitro binding to the E box sequence changes during the cell cycle, ND7 cell extracts were serum starved for 24 hours which induces them to exit the cell cycle into G0. The cells were then restimulated to re-enter the cell cycle by the addition of serum back to the cells. ^{3}H thymidine incorporation and E box binding was assessed after 0, 1, 3, 6, 10, 12 18 and 24 hours after serum addition/. The incorporation of ^{3}H thymidine dramatically increased between 12 and 16 hrs after the addition of serum to the cells, suggesting that the cells are entering S phase. However binding to the E box motif in the Pax-3 promoter increased 3-6 hours after serum addition prior to the onset of S phase (Fig. 3.12).

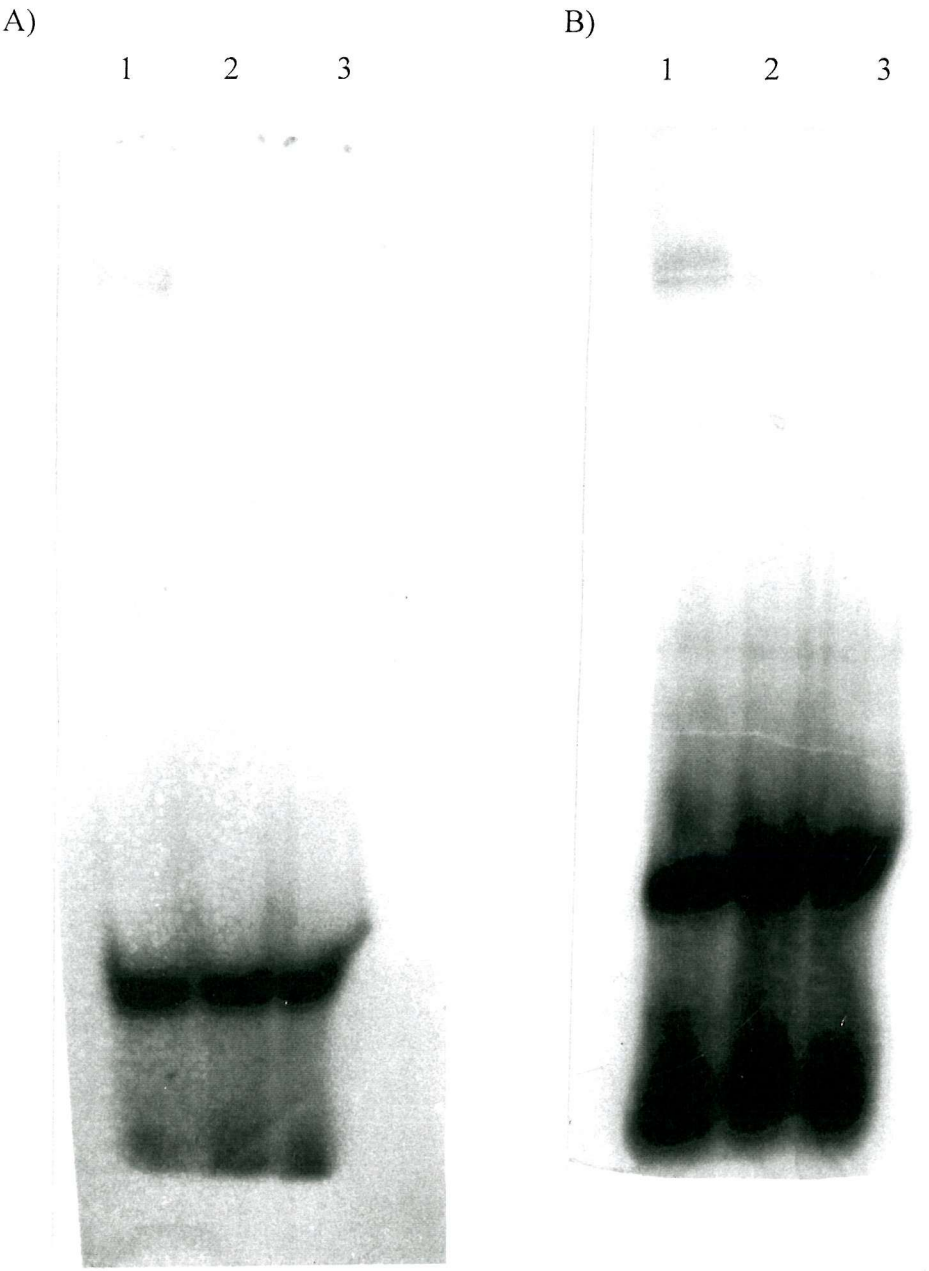
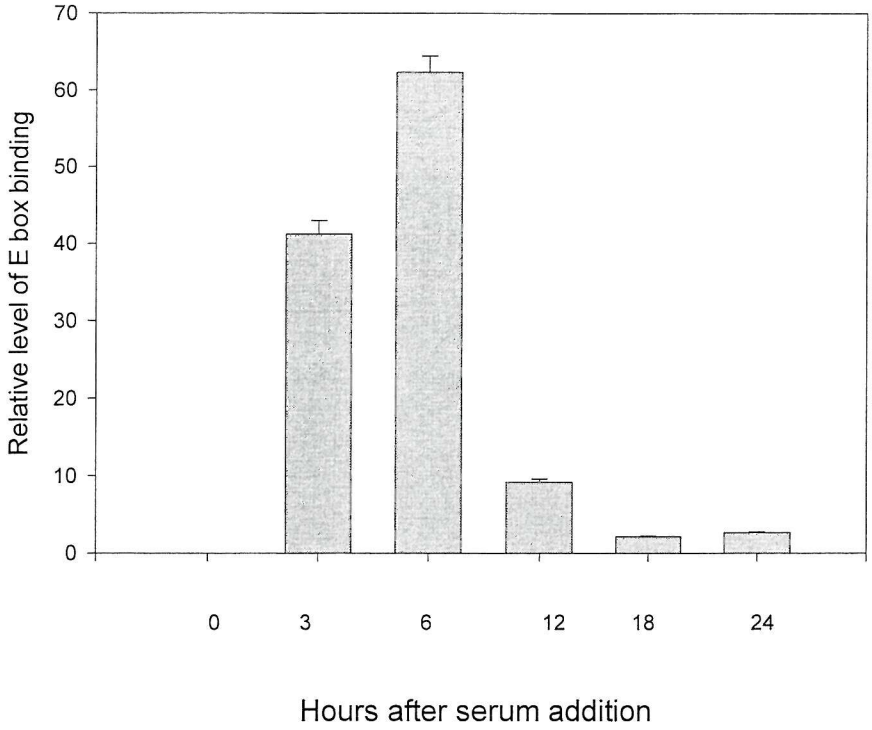


Fig. 3.11 EMSA using the E-box/Myc site in the *Pax-3* promoter. Nuclear extracts from ND7 cells were incubated with a radio-labelled oligonucleotide containing the E-box/myc site either alone (A: Lane 1, B:Lane 1) or with 1μg of N-myc antibody (A:Lane 2), or with 1μg MyoD antibody (A:Lane 3), or with 1μg C-myc antibody (B:Lane 2) or with 1μg Max antibody (B:Lane 3).

A)



B)

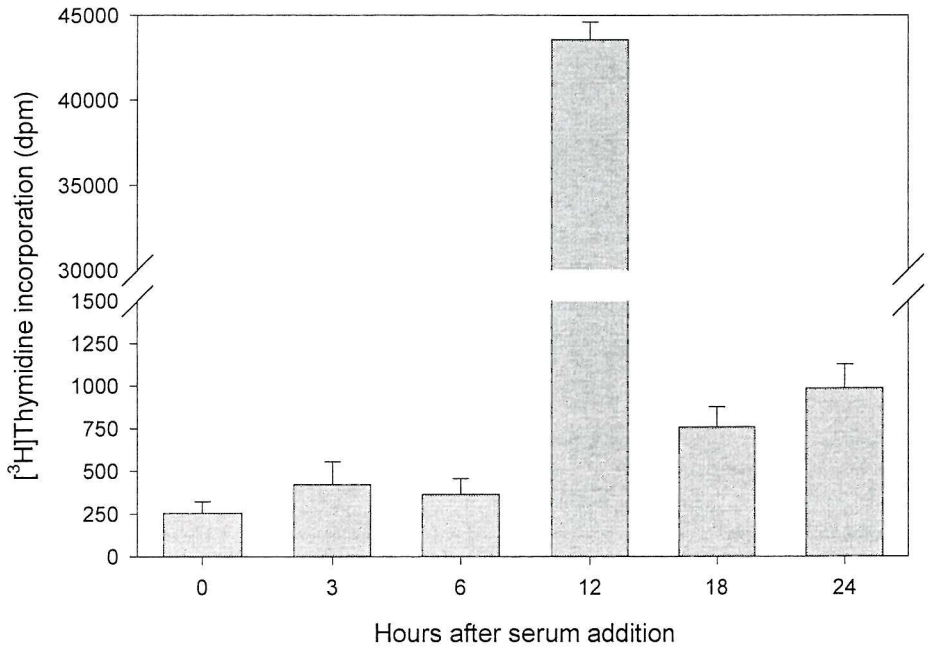


Fig.3.12 A) Graph showing the relative level of E box binding in ND7 cells, 0.3.6.12.18. and 24 hrs after serum stimulation. Each value represents the average of three independent experiments ±SEM **B**) Graph showing ^{3}H thymidine incorporation in ND7 cells 0, 36, 12, 18 and 24 hrs after serum addition.

3.5 The E2F's and Rb.

The E2F family, (E2F-1-6), is a class of cellular transcription factors that play a pivotal role in the regulation of cell division. Its responsive genes are either strongly implicated in or directly linked to the induction of cellular proliferation. It appears to limit transcription of its target genes to the phases of the cell cycle during which their products act.

The E2F's can act as both transcriptional activators and, in association with the Retinoblastoma protein (pRB), transcriptional repressors. In both these roles E2F must, obviously, be localised to the nucleus. In addition high affinity binding of E2F to its response elements is dependent on heterodimerisation of E2F with their binding partners, DP (DP1, 2 or DPδ3), although this does not appear to be a major point of control.

The regulation of E2F activity during the cell cycle is almost entirely due to its interactions with pRB and its close homologues p107 and p130. During G1 phase pRB is hypophosphorylated and as such can bind to the E2F/DP/DNA complex. pRB in turn recruits the Histone Deacetylase Complex (HDAC), which as its name implies, deacetylates nearby histones, tightening the chromatin structure and therefore preventing the binding of RNA polymerase and thus repressing transcription.

Cyclin dependent kinases regulate the phosphorylation of pRB. At the G1/S transition pRB becomes hyperphosphorylated, at which point it can no longer bind to the E2F DP complex and transcriptional repression is abolished. This is illustrated in Fig. 3.09

Although E2F-4 is the most abundant cellular E2F it appears to act only in a repressive role. Unlike E2F 1-3 it lacks a Nuclear Localisation Sequence (NLS) and therefore must rely on binding partners p107, p130 (homologues of pRB) or

occasionally DP3δ (an NLS containing form of DP3) to provide an NLS for it to enter the nucleus (Puri, et al 1998). It is only found in the nucleus during G1 when it is bound to one of the pRB family that repress transcription. In contrast E2F-1 almost exclusively nuclear during all phases of the cell cycle, and is therefore the most abundant nuclear E2F in the S- and G2-phases of the cell cycle.

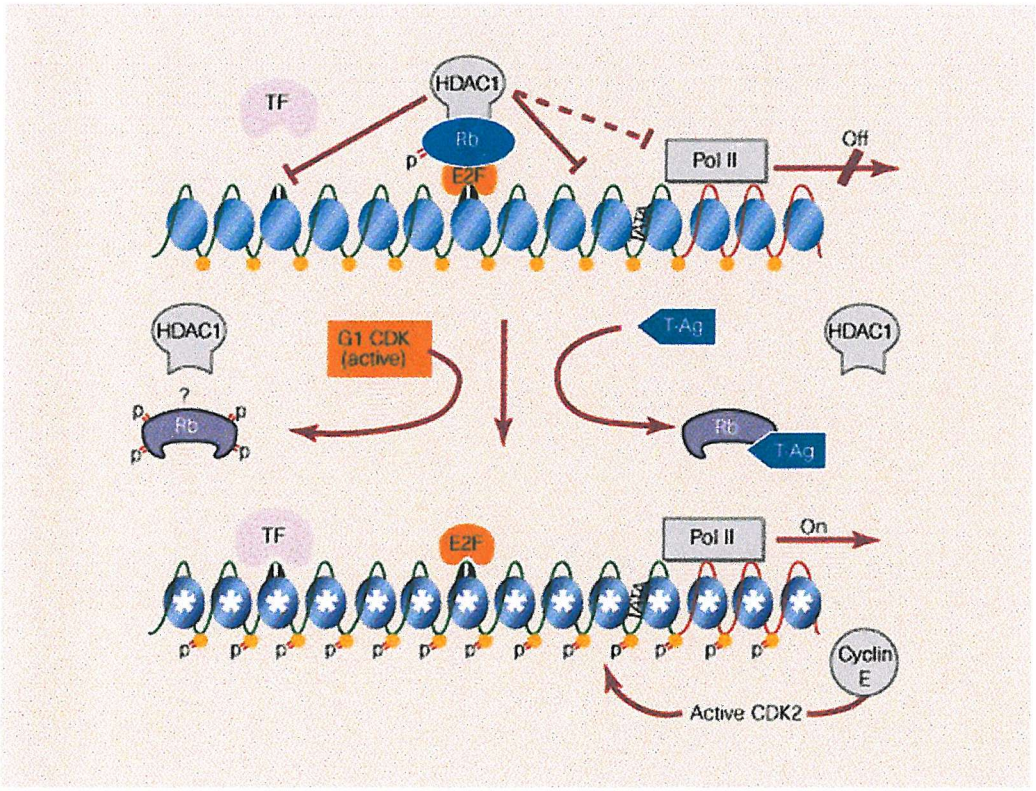


Fig. 3.13 In G1 Rb is hypophosphorylated and can bind to E2F, recruiting HDAC which deacetylates histones leading to a closed chromatin structure and thus preventing transcription. During the G1/S transition Rb becomes hyperphosphorylated and is released from E2F which then activates transcription of cell cycle proteins.

http://www.geocities.com/CollegePark/Lab/1580/pRB.html

3.5.1 The effect of E2F on Pax-3 *transcription*

Analysis of the Pax-3 promoter revealed two E2F consensus sequences, therefore to investigate whether the E2F-Rb pathway can regulate Pax-3 transcription, the Pax-3 promoter CAT construct was co-transfected with combinations of E2F-1 and pRb expression vectors. Preliminary experiments showed that ND7 cells co-transfected with the Pax-3 promoter CAT construct and E2F-1 had significantly higher levels of CAT activity than those transfected with a null expression vector in place of E2F. E2F had no effect however on the expression of pSVβgal or pCATBasic.

In order to determine the potential site of action of E2F-1on the Pax-3 promoter, the E2F expression vector was co-transfected with the PaxΔK, PaxΔS and PaxΔA deletion constructs. The ND7 cells were harvested after 60 hours and the CAT activity measured.

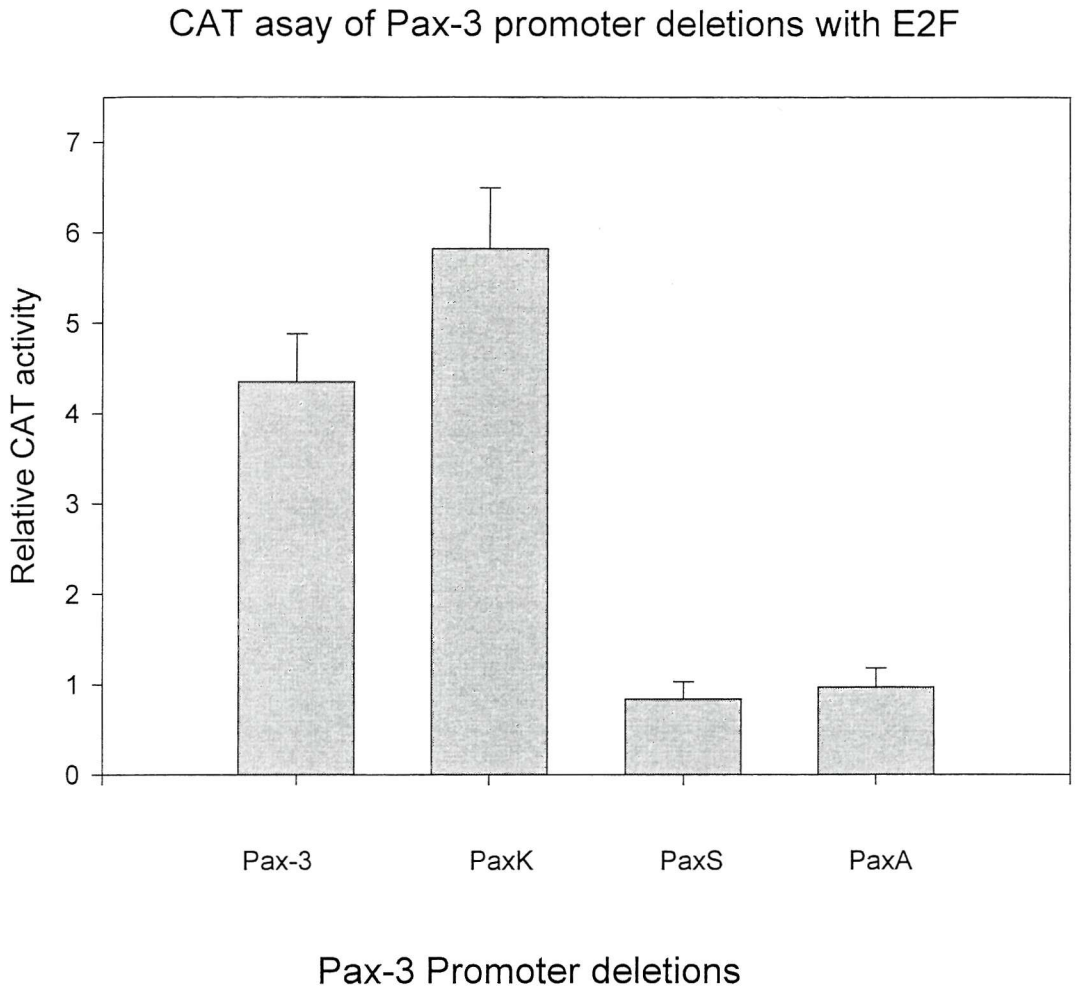


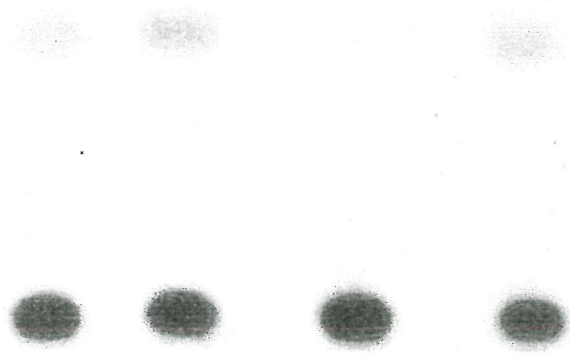
Fig. 3.14. There is a significant activation of the PaxCAT and PaxΔK deletions while there is no observed activation in either the PaxΔS or PaxΔA deletions. Examination of the Pax-3 promoter sequence shows two potential E2F sites (Fig, 3.01) the second of which is just upstream of the PaxΔK deletion. Thus it seems that this second site is the more important of the two.

3.5.2 Effect of Rb upon the E2F activation of the Pax-3 promoter

In order to further determine the effects of E2F and its co-factor RB upon the Pax-3 promoter, an expression clone of RB and an expression plasmid of RB without its 'pocket' domain (ΔRb), were co-transfected with the Pax-3 promoter CAT. CAT assays upon ND7 cells transfected with E2F, E2F and Rb, E2F and ΔRb, Rb and ΔRb with the Pax-3 promoter CAT. As can be seen from Fig. 5.11, the co-transfection of E2F with Rb leads to repression of Pax-3 promoter activity and abolishes the ability of E2F to activate Pax-3 transcription. In contrast the co-transfection of the ΔRb expression plasmid, missing the 'pocket domain' does not repress transcription compared to the basal level, however it does abolish the activation of expression by E2F.

The transfection of ND7 cells with Rb alone appears to inhibit slightly the basal level of expression of the Pax-3 promoter. The ΔRb vector, which lacks the pocket domain, does not affect the basal transcriptional level at all.

A)



Null Vector | E2F | E2F +Rb | E2F +ΔRb

B)

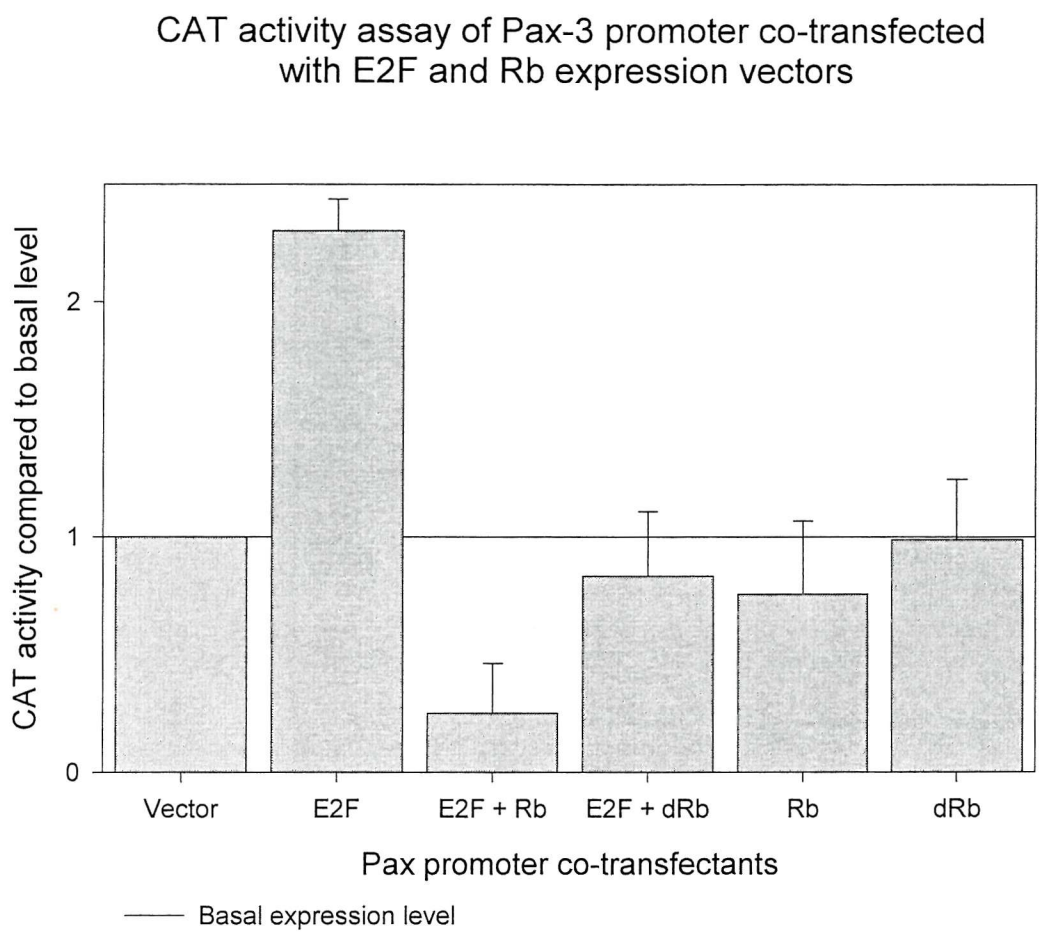


Fig. 3.14 A) A representative CAT assay showing ND7 cells co-transfected with Pax-3 promoter CAT construct (Pax-CAT) together with the null expression vector or with E2F, E2F + Rb or with E2F + ΔRb.

B) Graph showing CAT assay of ND7 cells co-transfected with Pax-3 promoter CAT construct (Pax-CAT) together with the null expression vector or with E2F, E2F + Rb or with E2F + ΔRb, or Rb or ΔRb.

3.6 E2F binding during the cell cycle

The E2F binding site identified as active in the *Pax-3* promoter was synthesised as an artificial oligonucleotide and used in a cell cycle electromobility shift assay (EMSA). ND7 cells were serum starved for 24 hours to induce cell cycle arrest, then serum added back to stimulate cell-cycle re-entry. Extracts were then made 0, 3, 6 12 and 24 hours after serum was added back. These were then incubated with ^{32}P labelled E2F binding site oligonucleotide and run down a non-denaturing polyacrylamide gel. Protein retardation of the oligonucleotide was then visualised using autoradiography.

The presence of a large number of possible E2F/Rb/p107 complexes means that there are several retarded bands. However the lower band is non-specific since it was competed out by an excess of cold oligonucleotide containing either the E2F site or a scrambled E2F site (not shown). Binding however of the three other complexes was specific since binding was competed out by an excess of oligonucleotide containing the E2F site but not by an excess of oligonucleotide containing a scrambled E2F site (personal communication K.A.Lillycrop). Furthermore the addition of an E2F antibody which recognises all members of this family abolished binding of complexes 1, 2 and 3. Binding of complexes 1,2 and 3 also varied however upon serum stimulation of serum starved cells. Little binding was observed after 24 hours of serum starvation, but 3-6 hours after serum stimulation an increase in binding of all three complexes is seen. After 12 hours when cells have entered S-phase there is a reduction in binding. (Fig. 3.16).

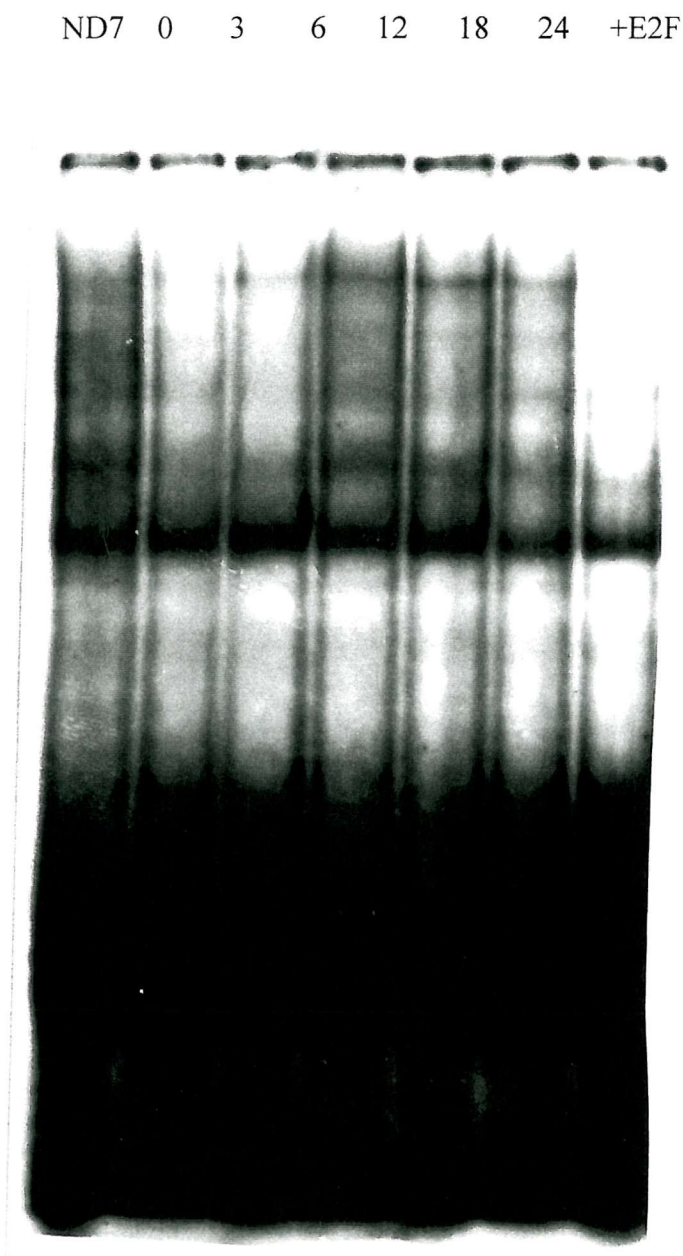


Fig. 3.16 EMSA showing binding to the putative E2F site in the *Pax-3* promoter. The radio-labelled oligonucleotide containing the E2F site was incubated with nuclear extracts from ND7 cells and 0, 3, 6, 12, and 24 hrs after serum stimulation. The last lane shows ND7 nuclear extracts incubated with a pan E2F antibody.

3.7 Discussion

Using co-transfection assays both N-Myc and C-Myc have been shown to affect the activity of the Pax-3 promoter in cells. The activation of Pax-3 promoter activity in the presence of N-Myc or C-Myc is further enhanced by the addition of their dimerisation partner Max. Although at high concentrations Max inhibits Pax-3 expression. This is probably due to competition from Max-Max homodimers that bind to the same E box site and thus prevent activation by the Myc/Max heterodimers. It is interesting to note that N-Myc activates Pax-3 promoter activity to a greater extent than C-Myc. This may reflect differences in DNA binding specificities or differences in transactivation ability of the two proteins. Although there is considerable conservation of amino acid sequence motifs among the Myc proteins, there are N-Myc specific domains in the N-myc gene suggesting differences in activities between the two proteins.

In contrast Mad/Max heterodimers strongly repressed Pax-3 promoter activity, suggesting that Pax-3 transcription could be regulated by the ratio of Myc/Mad present in the cell. In support of this hypothesis *in situ* hybridisation studies have shown that transcripts of Pax-3 are restricted to mitotically active progenitor cells within the developing nervous system. Furthermore high levels of Pax-3 expression are also associated with mitotically active immature neuronal cell lines. Pax-3 expression however in these cell lines is rapidly downregulated upon differentiation. Interestingly the fall in Pax-3 expression occurs prior to the onset of cell cycle arrest and prior to any overt changes in cell morphology, suggesting that the downregulation in Pax-3 expression upon neuronal differentiation is not merely a consequence of differentiation but may be a necessary requirement. This association of Pax-3 with proliferating immature progenitor cells is consistent with a model whereby Pax-3 expression is regulated by the Myc/Max/Mad network of transcriptional regulators.

The ability of Myc/Max/Mad to regulate Pax-3 promoter activity is probably mediated through an E box motif in the Pax-3 promoter. Myc E box elements have the potential to be recognised by at least two other members of the BR-HLH-ZIP family of transcription factors - USF and TFE-3. Specificity of Myc versus USE or TFE-3 binding to these elements may involve several mechanisms depending on promoter distance. For the α prothymosin gene, the E box element proximal to the start site of transcription is recognised by USE and TFE-3 as well as Myc, whereas distal elements are recognised by Myc/Max heterodimers. In ND7 cells, the E box motif was not bound by USF or TFE-3 but antibodies directed against N-Myc, C-Myc and Max dramatically reduced binding, suggesting that in ND7 cells the E box motif is bound by a combination of N-Myc/Max and C-Myc/Max. Both N-Myc and C-Myc are expressed in ND7 cells (personal communication. E White).

Pax-3 expression was also affected by the E2F family of transcription factors. E2F factors also play an important role in cell cycle control. Overexpression of E2F in ND7 cells led to a large increase in Pax-3 promoter activity. This increase in Pax-3 promoter activity was abolished however by the addition of an expression vector containing pRB but not by an expression vector containing pRB lacking the pocket domain. These results suggest that Rb is capable of inhibiting the transcription of E2F responsive genes. However the ability of Rb to inhibit E2F responsive genes appears to be dependent upon the pocket domain of the protein which mediates the interaction between E2F and Rb.

While there are two potential E2F binding sites in the Pax-3 promoter deletion studies show that the site at −845bp upstream of the start of transcription is responsible for the E2F activation of Pax-3 expression. In electromobility shift assays when nuclear extracts from ND7 cells were incubated with an oligonucleotide comprising the E2F binding site, four complexes were seen. The lower complex is non-specific as it was competed out by an excess of cold oligonucleotide containing a scrambled E2F site (personal communication K.Lillycrop). However the complexes 1,2,3 were specific as these complexes were

not competed out by an excess of cold oligonucleotide containing a scrambled E2F site but were by an oligonucleotide containing the E2F site. Addition of an antibody raised against all members of the E2F family abolished binding of complexes 1, 2 and 3. As E2F can interact with members of the pRB family, it is likely that these complexes represent either free E2F or E2F/pRb, E2F/p107 or E2F/p130. The addition of specific antibodies raised against the different members of the pRB family will determine if this is the case. Interestingly however binding of all three complexes varied through the cell cycle. Binding was very low in serum starved cells but increased 6 hours after the addition of serum back to the cells. This variation in binding to the E2F element suggests that binding to the E2F site increases just prior to the onset of S phase. Interestingly binding to the E box element also increased just prior to S phase. It will therefore be interesting to determine if the expression of Pax-3 does vary during the cell cycle. This could be tested by fusing the promoter of Pax-3 to the reporter gene luciferase. The Pax-3 luciferase reporter construct could then be transfected into ND7 cells. The cells would be serum starved for 24 hrs to synchronise the cells. Serum could then be added back to the cells and cells harvested and assessed for luciferase activity after 0,3,6,12,18 and 24 hours. This would establish whether Pax-3 promoter activity was cell cycle dependent.

The finding that both Myc and E2F transcription factors can regulate Pax-3 expression strongly suggests that Pax-3 does play a role in cell proliferation within the embryo.

Gene Targets of Pax-3

4 Gene Targets of Pax-3

4.1 Introduction

The tightly controlled spatio-temporal regulation of cellular events that are crucial to the proper development of an embryo is, as has been discussed before, ultimately due to the strict control of gene expression by transcription factors. However the study of the target genes of these transcription factors will give us clues as to how they effect the cellular changes that eventually lead to normal development.

Pax-3, like all Pax genes, contains a paired domain of 128 amino acids. In addition, Pax-3 also contains a second DNA binding domain, a paired type homeodomain. As, to date, very few target genes of Pax-3 have been identified, the binding properties of Pax-3 and the contribution each DNA binding domain makes to DNA site specificity has been most extensively studied *in vitro* using the e5 binding sequence which is derived from the promoter of the *Drosophila* even skipped gene (Treisman et al., 1989). Nuclease protection experiments have shown that the homeodomain of Pax-3 binds to the ATTA sequence within the e5 sequence and the paired domain of Pax-3 binds to the sequence GTTCC, which is located 5 bp downstream of the ATTA sequence in e5. The length of the spacer sequence between the elements ATTA and GTTCC has been shown to be important for high affinity binding, spacing the two elements 1 or 13bp apart conferred low affinity binding, whereas spacing the elements 5 or 8 bp apart conferred high affinity binding (Phelan and Loecken 1998). Recent studies have shown that Pax-3 regulates the expression of the hepatocyte growth/scatter factor receptor tyrosine kinase (c-met). C-met, like Pax-3, is expressed in limb muscle progenitors and is also required for mouse muscle development. The lack of c-met expression in *splotch* mice (Yang *et al.*) led to the suggestion that Pax-3 regulates the c-met receptor by transcriptional activation (Epstein *et al.* 1996). The promoter of the c-

met gene contains only a consensus Pax-3 paired domain binding site - GTTCC and no appropriately spaced homeodomain element, however it has been shown that Pax-3 can bind to this single paired domain element with high affinity.

The Myogenic transcription factors MyoD and Myf5 have also been shown to be upregulated upon ectopic expression of Pax-3 (Maroto *et al.* 1997). Infection of tissue with a virus expressing Pax-3 leads to the activation of their expression. While a GTTCC and ATTA consensus sequence was found in the chicken MyoD promoter, to date no experiments have been done to show that these sites are responsible for activation by Pax-3. Another potential target is the Melanocyte-specific Tyrosinase related protein-1 (TRP-1), which is also up-regulated by Pax-3 (Galibert *et al.* 1999).

There is increasing evidence that Pax-3 can also inhibit the expression of its target genes. Kioussi *et al.* (1995) have shown that Pax-3 was able to repress a 1.3kbp myelin basic protein promoter fragment in co-transfection studies. Furthermore Pax-3 has been shown to repress N-CAM promoter basal level activity in NIH 3T3 cells, however no direct target sequence can be identified (Chalepakis *et al.* 1994a)

Recently however another possible Pax-3 target gene has been brought to light. The elimination of constitutively expressed Pax-3 in the neuronal ND7 cell line, by induction of antisense Pax-3 RNA, leads to the rapid morphological differentiation of the cells, the extension of neurite processes and the expression of genes such as Snap-25 which are characteristic of a mature neuronal phenotype (Reeves *et al.* 1999). With the evidence that Pax-3 contains domains for both activation and repression, it is possible that Pax-3 directly inhibits the expression of Snap-25 in immature neuronal cells.

4.2 Expression and function of the synaptosomal protein of 25-kDA (Snap-25)

Transmission of information across synapses is the culmination of the electrochemical gradient down a neuron. The key event in this process is the Ca^{2+} triggered exocytosis of the synaptic vesicle by membrane fusion. Snap-25 is a presynaptic protein that has been identified as a key component in assembling a protein complex responsible for the docking of vesicles at nerve terminals.

4.2.1 The synaptic vesicle cycle.

The synaptic vesicle cycle can be divided into 9 steps.

1) **Endosome fusion**. Recycling SVs fuse with early endosomes.

2) **Budding**. SVs are primarily regenerated by budding from endosomes.

3) **Neurotransmitter uptake**. SVs accumulate neurotransmitters by active transport driven by an electrochemical gradient created by a proton pump.

4) **Translocation**. SVs filled with neurotransmitters translocate to the active zone either by diffusion or by a cytoskeletal based transport process.

5) **Docking**. SVs filled with neurotransmitters dock at the active zone.

6) **Priming**. After docking SVs must be primed before they can become competent for the fast Ca^{2+}-triggered fusion step. This priming step is rate limiting.

7) **Fusion/exocytosis**. Primed SVs are stimulated for rapid exocytosis by a Ca^{2+} spike during an action potential.

8) **Endocytosis**. Empty SVs are internalised probably by way of clathrin coated pits.

9) **Translocation**. Coated vesicles shed their coats, acidify and translocated into the interior, becoming recycled SVs

SNAP-25 is one of the key components of the priming step. During priming a core complex is formed between two plasma membrane proteins, SNAP-25 and syntaxin, and one SV membrane protein, synaptobrevin/VAMP. First the two plasma membrane proteins bind tightly to each other, forming a high affinity binding site for synaptobrevin/VAMP. This core complex forms the anchor for a cascade of protein-protein interactions required for exocytosis to occur. (Sollner et al. 1993; For reviews see Bark and Wilson, 1994; Sudhof 1995).

However prior to synapse formation, Snap-25 is localised in the cell bodies and fibres of neurons within the neonatal brain (Oyler et al., 1991). This has led to the suggestion that Snap-25 may participate in axonal growth by directing the fusion of construction vesicles that supply general membrane components needed for elongation throughout the neurite, as well as for the plasmalemma expansion of growth cones. The hypothesis that Snap-25 plays a key role in axonal growth is supported by the observation that the selective inhibition of Snap-25 expression in rat cortical neurons and in PC12 cells has been shown to prevent axonal elongation (Osen-Sand et al., 1993).

4.3 Transcriptional regulation of the Snap-25 gene

Much work has already been done on the transcriptional activation of the *Snap-25* gene, including its control by Oct-2, (Deans *et al.* 1997), the Brn family of POU-domain transcription factors (Lakin *et al.* 1995; Morris *et al.* 1997) and a dissection of the promoter regions required for its neurospecific expression (Hess *et al.* 1994; Ryabinin *et al.* 1995).

4.3.1 The effect of Pax-3 on Snap-25 expression

Previous experiments on the neuronal cell line ND7 have shown that the elimination of Pax-3, results in the substantial induction of *Snap-25* transcription (Reeves *et al.* 1999), raising the possibility that the Snap-25 gene may be a directly regulated by Pax-3 levels. To investigate whether Snap-25 is a direct target for Pax-3 regulation, a 2.2 kbp fragment of the 5' upstream region of the *Snap-25* gene (-2061 to +113bp with respect to the transcription start site of Snap-25 - a kind gift of M. C. Wilson; Ryabinin *et al.* 1995) was fused to the reporter gene CAT in the vector pCATBasic. This construct will subsequently be referred to as pPPCAT. The 2.2kbp fragment of the 5' upstream region of Snap-25 (Fig. 4.01) contains all the elements required for basal transcription, such as the TATA box and transcriptional start sites. In addition it also contains enough of the promoter and enhancer elements to promote basal neural- specific expression in tissue culture cells.

CTGCAGAATGAGGCTCCCTCAATGACGTACTTAAAAAAAAAAGGGGGGGGGGGGAAATCC -1984
CCTGCCCCCATTCTTTCACAAACTTTCTGCCTGAATGATTTCTAGTTGGAGGGCCCTATA -1924
CATATTTTCTTTCTCCTTCTTGTAGGGTAGTTTGTCTGTCTCTGCCAACTCTGCATGCTG -1864
CCTTTTTTTTCCCAAGTCATATCATAACTGAAAAATATTACATTCTTAAGCCTATTCCTG -1804
CGAGCTTTATTATAATGAAAATATGAGCTGGAATGGCAGCCTAGTGAGGGCCAAGCATAT -1744
TTTTTTCTGTCGTACAAAAATTTAAATGTTGAATAACTCAGTAAGCAAATTGCAGCAGCA -1684
ATATTTCAGGAATCCGGTTTTAATCAAAACCCAAACACCATCATCATCGCAAACATTAAG -1624
ATTTCAAAATATCCTTTGGATCTTGGTATCTGTTAGAAAAAGGAAATTCTGCTCCTCTCC -1564
TATGTACATCCACCCAACCCCCTCCTCCCTCAGCTCCCCGTCTGTCTTCGCCTACAATTA -1504
AAAGCTTTATTACCTAAGCTTGCCTTTAAAACCAGCAAGCGTGGTGTGTGTGTGTGTGTG -1444
TGTGTGTGTGTGTGTGTGTGTGTGTTATCGAGACCGGGGCGGGGGGCGTTTACGCGTT -1384
GCTTTTCAGATCCAGTTTTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG -1324
TGTGTGCGCGTGTGTGTCTCTAAGAGAAAAGGATTTTCCATCGCCCAAGAGGCCATCAAG -1264
TTTTAAGATTGCGAGGCACGACTTCTAATCGCATCCTCCGCGGAAAAGTACAAACAGTTC -1204
CTGAGACGAGGTCCCCCACCTCCCACGCGCTCCCCAGCTCTCCCTCCCTACCTAGAACCC -1144
GGACAACCTCCCCAACACCTGTGAATCATCTAGAGGCTGCACCACCGAACTATTTGCATA -1084
TCGCCGCCTGCCCTACCCAGCTCCTGACTGCTCCTGACGCACGACAAGCAAAGACAGGAT -1024
TATATTGTGTGGTTATTACTTCTAATCTTTCATCTGAACCAGCAGAAGTAGACATTCTTG -964
CTCATACCAGACCCAGAGGTTCAGAGGGTCGGCCACTGAAAGTGCTTGTGCAATGAGTGA -904
CCCTCCAGATTTCAGTACTGTCCTAGCCACCTTCCCCCCACCCCCACCCCCGGATGTCTG -844
GACACGTTCTAAGAAACCCGAATACCGCTCTAGGCCCAGACTGCAAATTCAACCCGCCAC -784
AGAGTCATCTATAAGCCCCGGCCAAACAACTAGCAACTCAAAGCCCTAGCCCTGCAAACC -724
CAAGCTAATAGGGCTAACTAGCTTCCCTCTCCCTTTCCCCAACTCAGGTGTTATCTCGGT -664
GATAATTGCTCTGACGTACTTTAAAAGAAAGTGTGAGCCTTTCCCAAACTGGCCATGATT -604
AAATGAGAACTCTAAACCTGCAATGCTGCTTTAATACAATAATTTTCATAAAAATAGCAA -544
TGTGAGTAACCTGGTCACGAAAGGAGCTGGTGAGAGCGTGTGCTGAGGGGAATGCCTCTT -484
CGGACCCTCTTCTCTGGAGAAAATAAATCACAACAGAAATCTATGCCAATCCCTCACATC -424
ATTACCCTTCTATTCCTTTACCTTCCTCTCGTGTCTGTGACTCAACCAAAGACAAATCAT -364
CAGGGGAACGTTTTACATATAAATATGTATATGTTTACATTATTATAATTGTATATACAT -304
TCCAAACACACGTATTTACATACATATATTCACATATGCATGCACACGTAAGCAAACGCA -244
TAGGCCATCAGGAAAACAGTCGCAAGAAATATGTATTTCCCAAGACGGTGCTAAAATGGA -184
GTCAAAAGAATCGCTCCATTTGGGTGTTTATAAAAGATTGAAATAAAAAGGGTTGCTGCA -124
AAGCGCTCAGGGAGCTGAGTCAAGCGCTCCGCACCTCCCGGCTCCTCCCCTACCGCCCCC -64
AGCCACCTTCCTCCCTCCGCCCTCCCCCATCACACTTCCTCCCCGCCCCTCGAGGAGCCG -34
GCTATTAAAAGCTGCTGACGTCAAAGGAACGGCCATCTTTGATGAGGGCAGAGCTCACGT +26
TGCATTGAAGAAGAAACCTCGGGGAGGTCTGGAGCTGTCTTTCCTTCCCTCCCTACCCGG +86
CGGCTCCTCCACTCTTGCTACCTGCAG +113

ATTAA – TATA Box
C – Start of transcription
GCTT – Hind III site - pHP
ATGC – Sph I site - pSHP
TCCG – Hae I site - pHAP
CACG – Sac I site - pSP

The whole fragment was inserted with Pst-Pst sites.

GTTCC – potential Pax-3 paired domain binding motif

TATTATAAT – potential Pax-3 homeodomain binding motif.

Fig. 4.01 Nucleotide sequence of the 5' flanking region of Snap-25. Only the coding strand is shown. The TATA box is underlined and the transcription start site is enclosed by a box. Potential Pax-3 binding sites are highlighted in yellow.

To investigate the effect of Pax-3 upon Snap-25 transcription, the reporter plasmid containing the 5' promoter region of Snap-25 (pPPCAT) was transfected into the ND7 neuronal cell line.

4.3.2 The effect of increasing Pax-3 concentration on the SNAP-25 promoter

As a preliminary experiment to determine whether Pax-3 regulates Snap-25 promoter activity, the pPPCAT construct was co-transfected into ND7 cells with differing amounts of the pJ7Ω expression vector containing a full length cDNA clone of mouse Pax-3. ND7 cells were transiently transfected with 5μg of the pPPCAT construct and varying amounts of the Pax-3 expression vector (0-5μg). The total amount of DNA in each experiment was kept constant by the addition of the empty expression vector pJ7Ω. A representative CAT assay and a bar chart showing the combined result of four experiments is shown in Fig. 4.02 and 4.03 respectively. These experiments show that Snap-25 promoter activity decreases with increasing amounts of Pax-3, in a non-linear fashion, suggesting that Pax-3 inhibits Snap-25 promoter activity. No effect on the expression of pCATbasic or pTKCAT were seen regardless of the input of Pax-3 expression vector (Fig. 4.04).

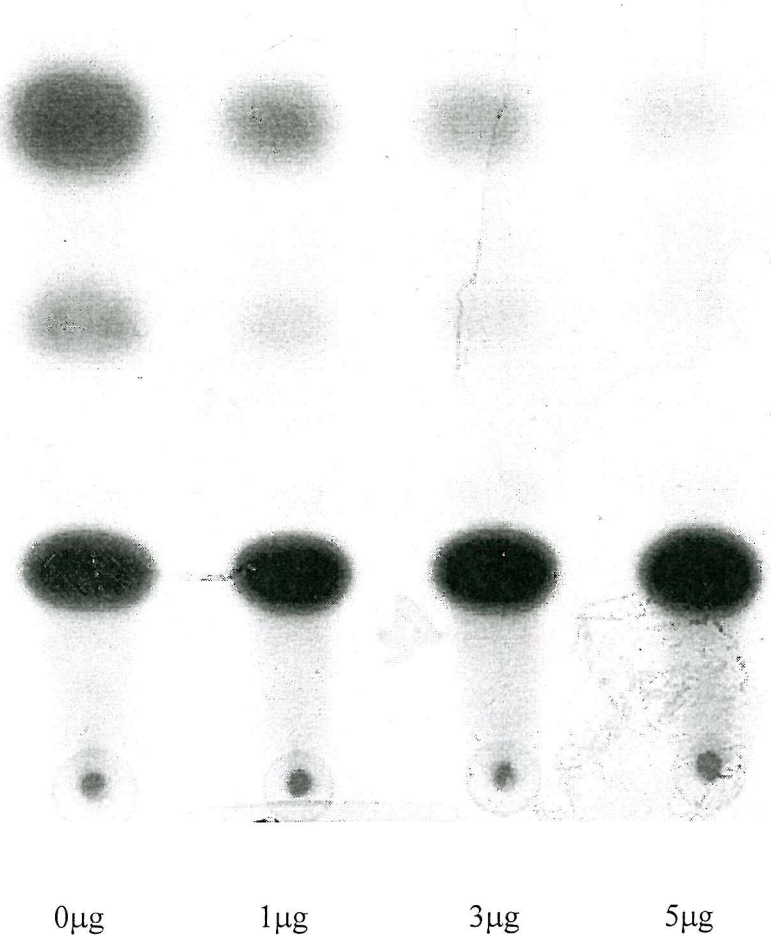
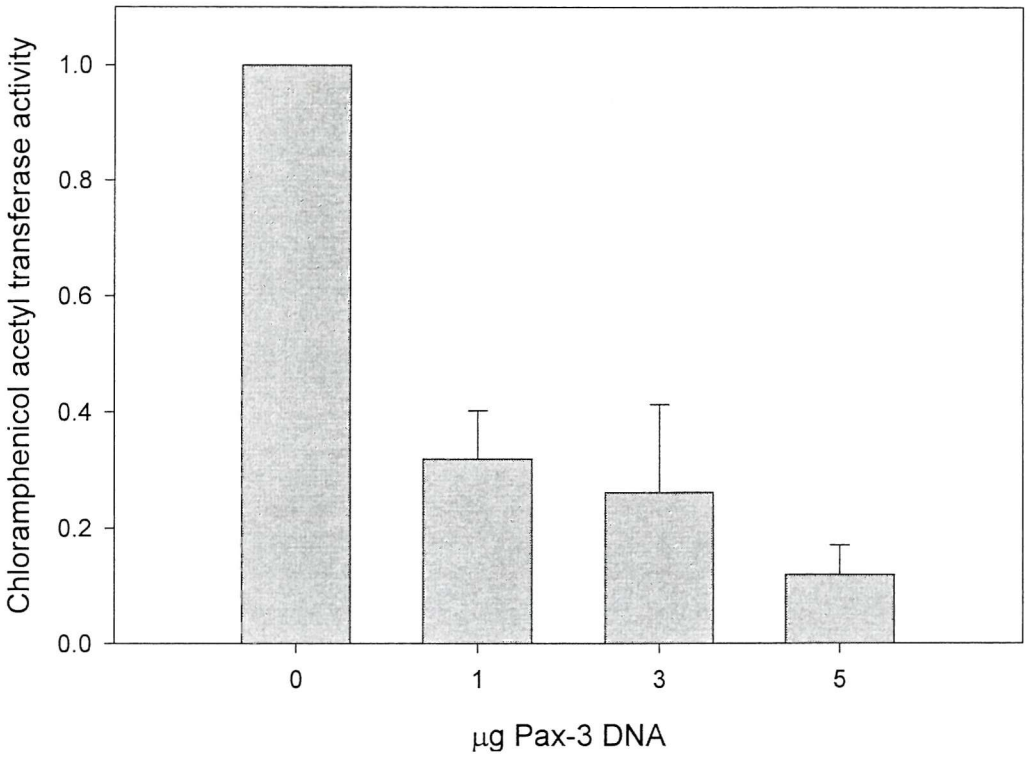


Fig 4.02: A representative CAT assay showing the effect of co-transfecting ND7 cells with pPPCAT and with increasing amounts of Pax-3 expression vector (0μg, 1μg, 2μg and 5μg).

CAT activity assay of the Snap-25 promoter with increasing concentrations of Pax-3 in ND7 cells.



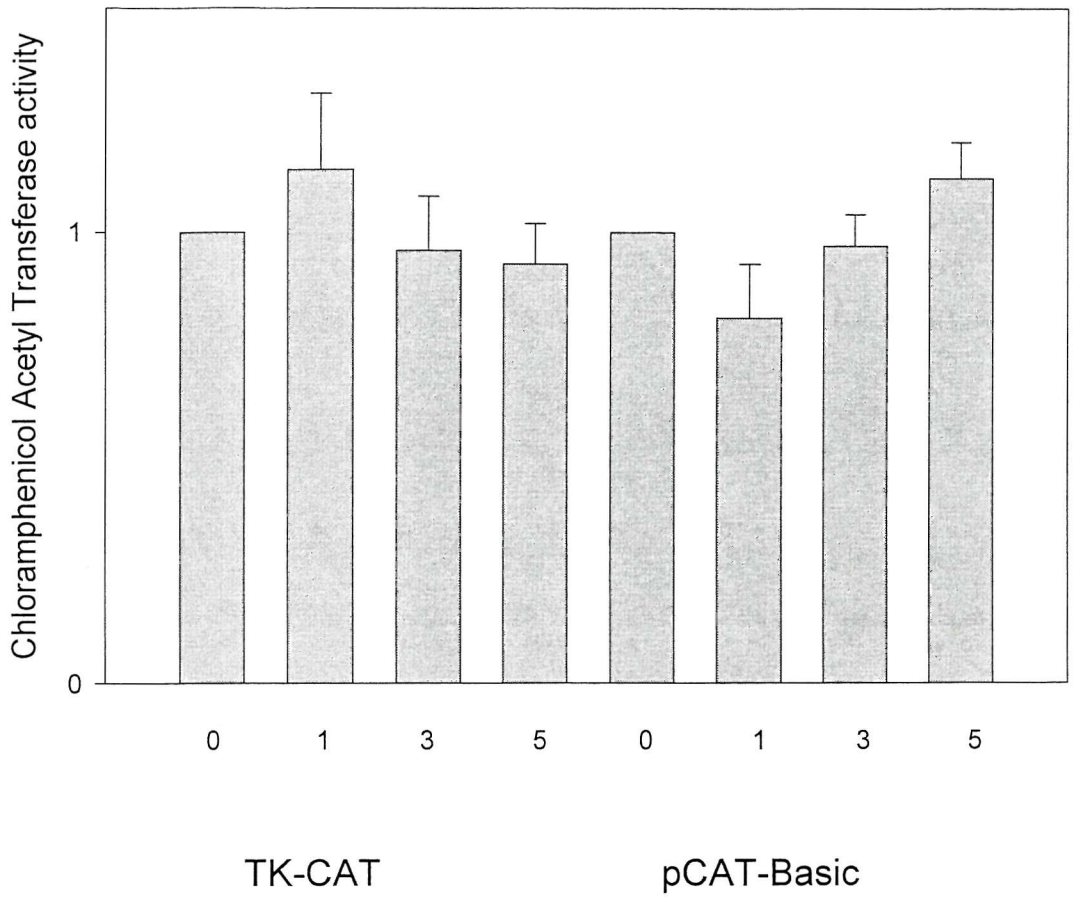
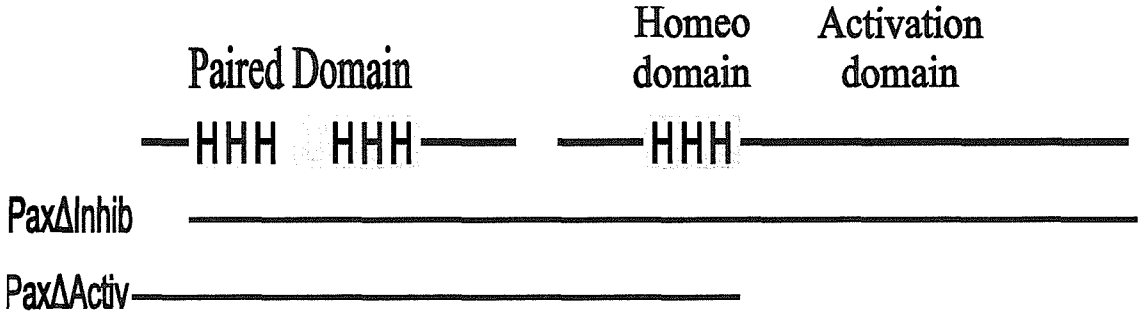


Fig. 4.04 Graph showing the effect of co-transfecting 0, 1, 3, and 5μg of a Pax-3 expression vector with 5μg of TK-CAT or with 5μg of pCAT-Basic into ND7 cells. Each value is expressed relative to the level of CAT activity in cells transfected with the reporter construct together with 5μg of pJ7 and represents the average value of 3 independent experiments +/- SEM.

4.4 The Pax-3 inhibitory and activation domains

In order to study further the effects of Pax-3 repression upon Snap-25 promoter activity, two Pax-3 expression plasmids were manufactured that were missing a) the first 33 amino acids of Pax-3 and b) the C terminal activation domain. The N terminal region of Pax-3 has been reported to be essential for Pax-3 inhibition of target genes. The PaxΔInhib deletion clone was prepared using PCR. A Prd-1-Fwd primer which encodes a methionine start codon just 5' to the start of the paired domain, and a Pax-3-Rev primer which encodes an EcoR1 site just after the stop codon of Pax-3 were used to amplify a fragment of Pax-3 DNA which lacks the sequences encoding for the first 33 amino acids. In addition, a clone lacking the C terminal activation domain of Pax-3 was also made using PCR primers that binds just in front of the BamHI cleavage site 5' to Pax-3 and a Homeodomain-Rev primer which binds at the end of the homeodomain. These mutant constructs are shown graphically below.



Initially the PCR reaction was ligated into pGEM-T-Easy using T/A cloning. They were subsequently cloned into the pJ7 expression vector under the control of the CMV promoter.

The clones were identified by cleavage with EcoR1, which cut on either side of the inserted fragment. The PaxΔInhib coding fragment was 1300 bp in length while the PaxΔActiv was 840 bp in length.

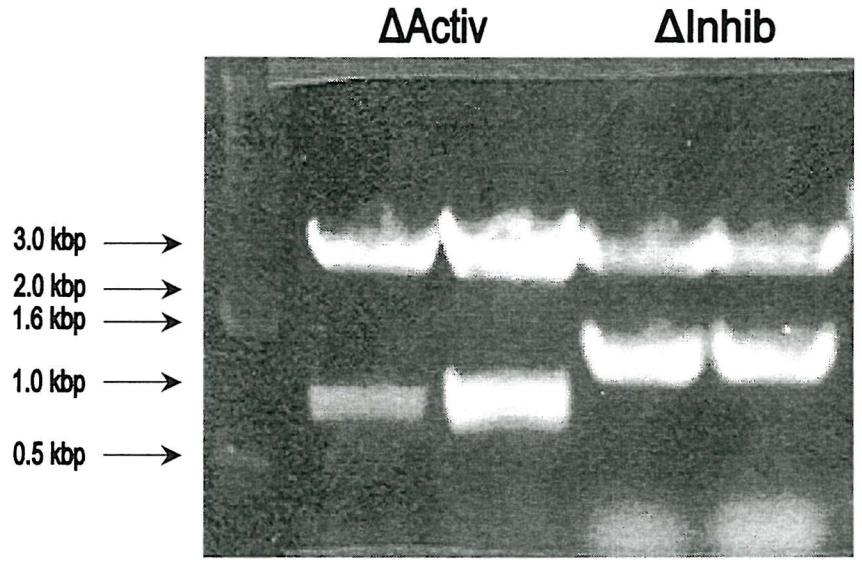


Fig. 4.05 Ethidium bromide stained agarose gel showing the DNA of Pax PaxΔInhib and PaxΔAct cleaved out of pJ7 using EcoRI.

Additional evidence came from restriction maps of the clones.

Co-transfection of pPPCAT with PaxΔInhib abolished the ability of Pax-3 to inhibit Snap-25 promoter activity. Co-transfection of PaxΔInhib resulted instead in the activation of Snap-25 promoter activity (Fig. 4.06 and 4.07). In contrast, the Pax-3 clone lacking the C terminal activation domain was able to inhibit Snap-25 promoter activity. Co-transfection of PaxΔInhib or PaxΔAct had no effect on the expression of pCATBasic or pSVBgal activity. These results suggest that the N terminal region of Pax-3 is essential for the ability of Pax-3 to inhibit Snap-25 promoter activity.

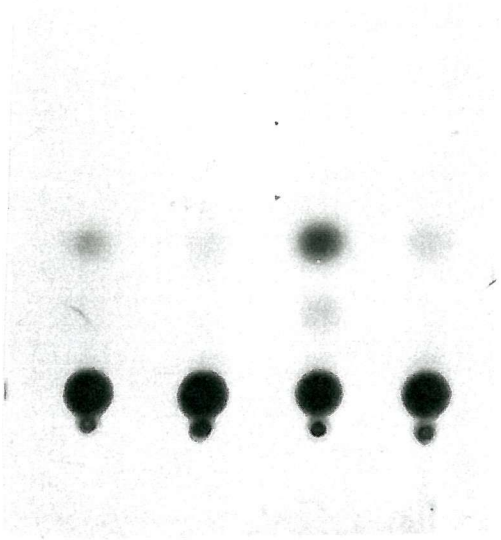


Fig 4.06: A representative CAT assay showing the effect of co-transfecting pPPCAT with pJ7 (vector alone), Pax-3 expression vector (Pax-3), PaxΔInhib, and Pax-3ΔAct.

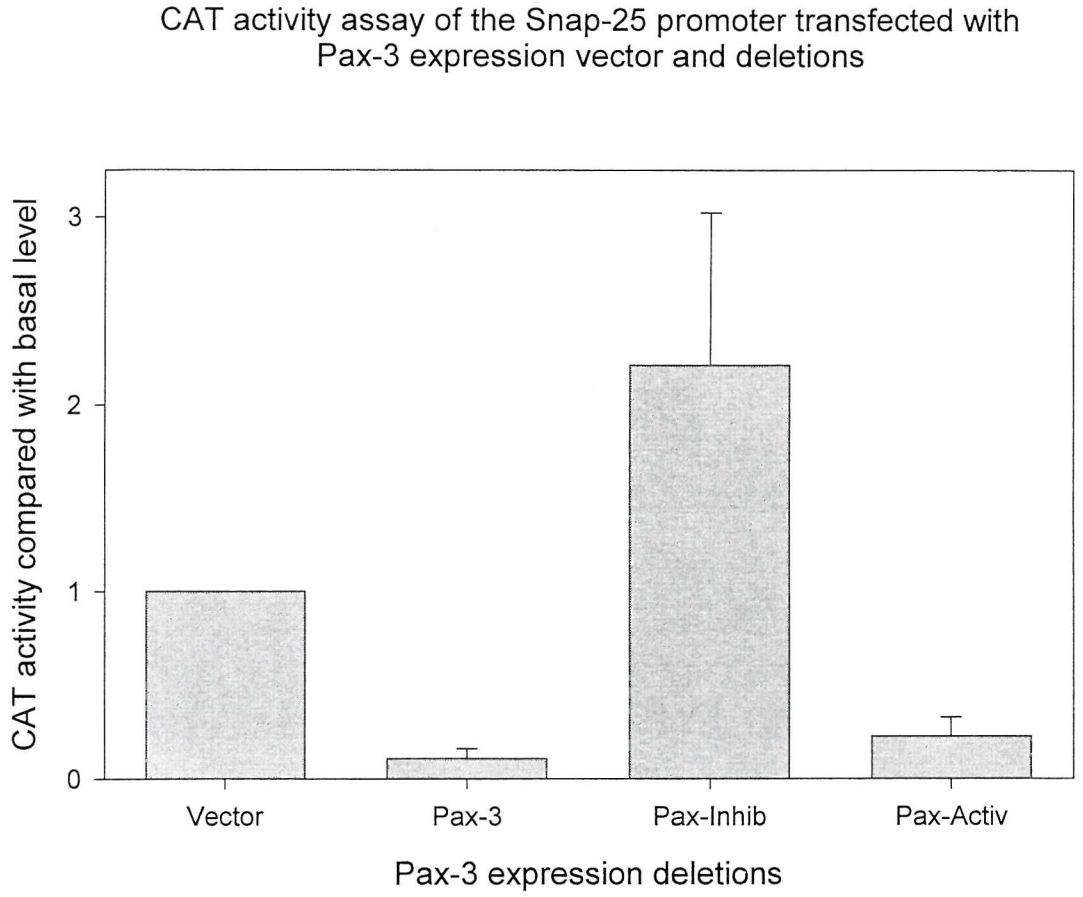


Fig 4.07 Graph showing the effect of co-transfecting pPPCAT with PJ7 (vector), Pax-3 expresion vector (Pax-3), Pax-3Δinhib (Pax-inhib) or PaxΔAct (Pax-Act). Each value is expressed relative to the level of CAT activity observed in ND7 cells transfected with pPPCAT together with the empty expression vector pJ7 and represents the average value of three independent experiments ± SEM.

4.5 The Pax-3 recognition site

Examination of the 2.2-kbp proximal promoter sequence of Snap-25 yielded two potential Pax-3 binding sites. The first is a consensus paired domain binding motif (GTTCC) at –1200bp and the second at –1811bp is a palindromic homeodomain motif (ATTATAAT). In order to determine whether these sequences are necessary for Pax-3 inhibition of Snap-25 promoter activity, a series of truncated Snap-25 promoter constructs were used. These clones and their position are shown in Fig. 4.08. Each deletion construct was co-transfected into ND7 cells with either the empty expression vector pJ7Ω or the pJ7Ω Pax-3 expression vector. As shown in Fig. 4.09, only the Snap-25 promoter construct containing sequences from -2061 to + 113 was strongly inhibited by Pax-3. A small degree of inhibition was seen with the truncated Snap-25 constructs in the presence of Pax-3, however Pax-3 also inhibited to a similar degree the expression of the thymidine kinase promoter (TK-CAT), which contains no Pax-3 binding site, and also the expression of a basal promoter construct (E1BCAT) which contains only a TATA box upstream of the CAT gene. This suggests that the strong inhibition of Snap-25 promoter activity by Pax-3 is dependent upon sequences between –2061 and –1518 of Snap-25. The palindromic homeodomain recognition site (TATTATAAT) in the promoter falls into this region, suggesting that Pax-3 may mediate the inhibition of Snap-25 activity by binding to sequences within the Snap-25 promoter.

In order to determine if Pax-3 binds to the palindromic homeodomain sequence at –1811bp in vitro, this putative binding site was synthesised as an oligonucleotide (Snap-25 B-S), and incubated with nuclear extracts from ND7 cells. Binding to this sequence was then assessed in an electromobility shift assay (EMSA). As shown in Fig. 4.09, one retarded complex with low mobility was observed. The binding was specific since it was competed out by the addition of 100X cold excess of the binding site, but not by a 100X fold excess of an oligonucleotide containing a scrambled TATTATAAT site. To determine the identity of this

protein complex, nuclear extracts from ND7 cells were incubated with antiserum raised against Pax-3 and then analysed by EMSA. We found that the addition of Pax-3 antibody abolished binding suggesting that Pax-3 is able to bind to the palindromic TATTATAAT site in the Snap-25 promoter.

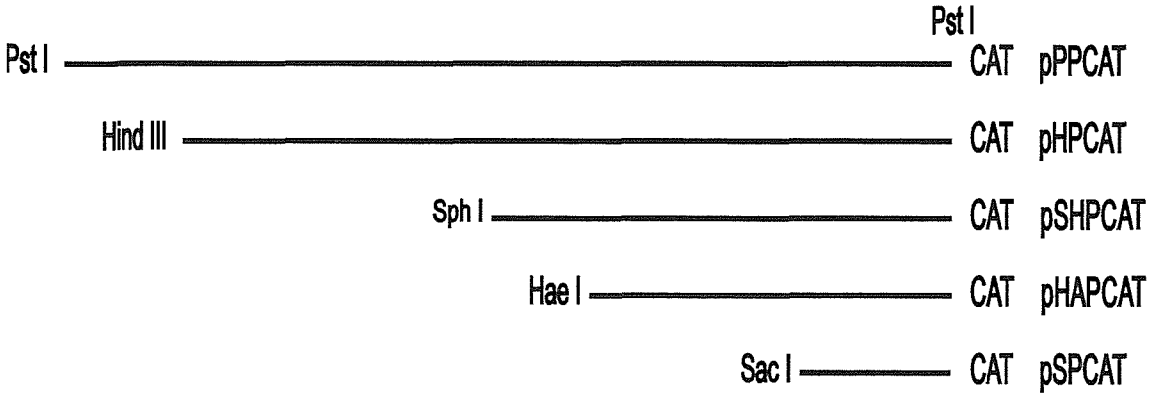


Fig. 4.08 Schematic diagram of the constructs used to locate the Pax-3 binding site within the Snap-25 promoter. The restriction enzymes used to create the truncated contructs are shown

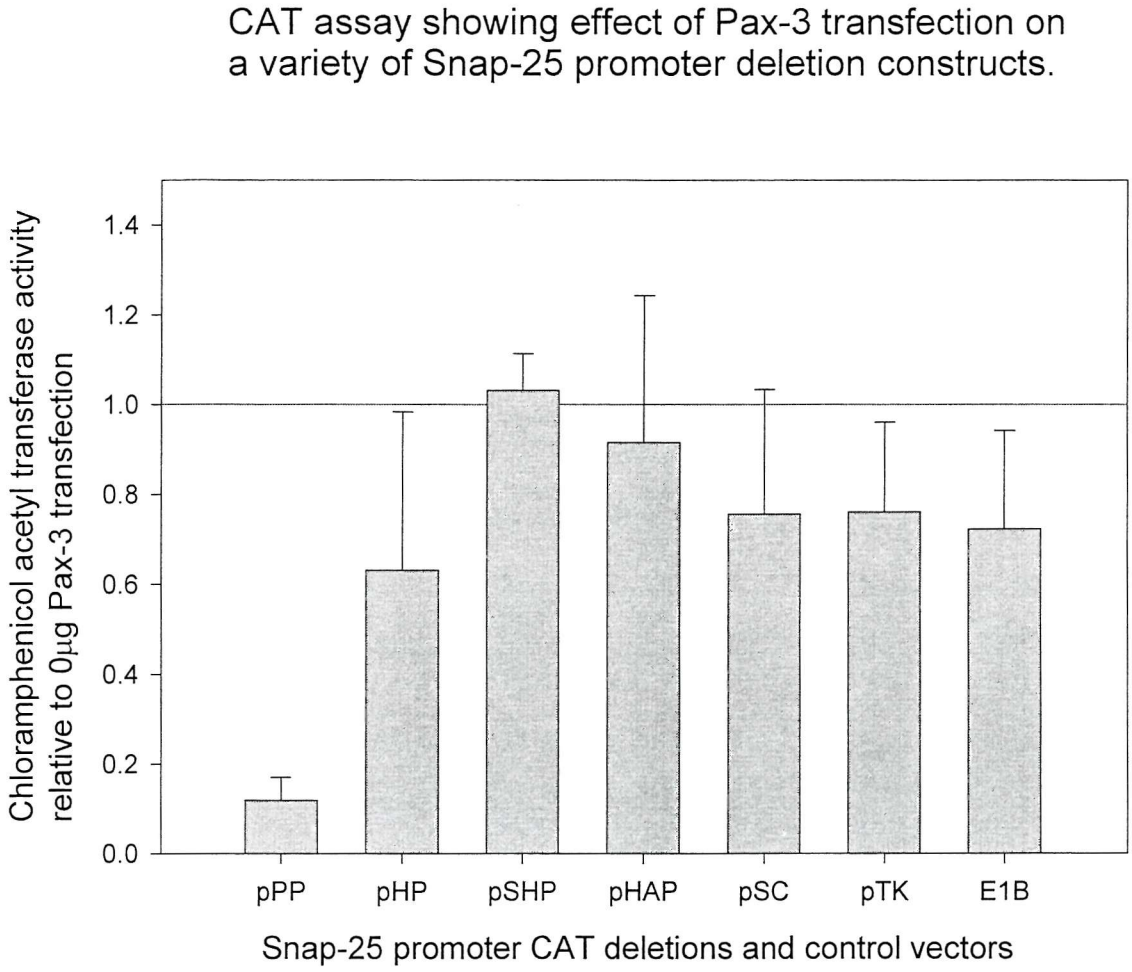


Fig. 4.09 CAT assay of ND7 cells co-transfected with Pax-3 expression vector (3ug) together with 5ug of either pPPCAT (-2061 to +113), pHP (-1516 to +113), pSHP (-292 to +113), pHAP(-127 to+113), pSC (-41 to+113), pTK-CAT and pEIBCAT. Each value represents the average of three independent experiments ±SEM.

Fig. 4.10 Electromobility shift assay using potential Pax-3 binding site in *Snap-25* promoter. Nuclear extracts from ND7 cells were incubated with a radio-labelled oligonucleotide, Snap-25 B-S. Either alone (Lane 1), in the presence of 100X excess cols Snap25 B-S (Lane 2) or with 100X excess cold non-specific oligonucleotide (Lane 3) or in the presence of 1μg Pax-3 specific antibody (Lane 4).

4.6 Discussion

The finding that *Snap-25* expression is upregulated in neuronal cells lacking Pax-3 has led to the suggestion that Pax-3 may negatively regulate the expression of Snap-25. To determine if Pax-3 directly inhibits Snap-25 promoter activity, a reporter construct containing the Snap-25 promoter was co-transfected with expression vectors containing Pax-3. This resulted in the strong inhibition of Snap-25 promoter activity. This repression was further enhanced by the addition of increasing amounts of Pax-3. The ability of Pax-3 to inhibit Snap-25 promoter activity appeared to be dependent upon a region of the Snap-25 promoter between –2061 to –1518 with respect to the transcription start site. A computer aided search of the promoter sequence of *Snap-25* for potential Pax-3 binding sites revealed that this region of the Snap-25 promoter contained one potential Pax-3 binding site, a palindromic homeodomain recognition site, positioned at –1811bp upstream of the start of transcription.

Pax-3 has been shown to bind with high affinity to the E5 sequence which contains both a paired domain and homeodomain recognition sequence. Pax-3 can also bind to a single GTTCC motif, which is found within the c-met promoter. In contrast Pax-3 binding to a single homeodomain recognition site ATTA is very weak. However Pax-3 has been shown to bind with high affinity to the P2 sequence TAATn2ATTA in vitro. This palindromic TAAT homeodomain motif enables dimerisation to occur. In view of the similarity between the double homeodomain site – P2 to which Pax-3 has been shown to bind to in vitro and the potential Pax-3 binding site in the Snap-25 promoter, we tested whether Pax-3 could bind to this sequence. We found that in ND7 cells the sequence TAATATTA from the Snap-25 promoter was bound by one low mobility complex. Binding was shown to be specific by competition with specific and non-specific probes. Furthermore binding was abolished by addition of a Pax-3 antibody, raising the possibility that Pax-3 binds to this site within the Snap-25

promoter to inhibit Snap-25 promoter activity. It should be noted here that the Pax-3/TAATATTA complex was retarded to a greater extent than Pax-3/e5. A likely explanation for this observation is that dimerisation of Pax-3 on the double binding site has occurred, despite the lack of the 2-3 amino acid spacer region found in published onligonucleotide enrichment experiments (Wilson *et al.* 1996).

The ability of Pax-3 to inhibit Snap-25 promoter activity was also dependent upon the N terminal region of the Pax-3 protein. Pax-3 has previously been shown to contain an inhibitory domain (Chalepakis *et al.* 1994a), which was located in the first 90 amino acids of Pax-3. When these 90 amino acids were linked to the DNA binding domain of the yeast transcription factor Gal4, they inhibited transcription from a reporter plasmid containing three Gal4 binding sites upstream of the thymidine kinase promoter., demonstrating that this domain can inhibit transcription, independent of the rest of Pax-3. To ascertain whether the inhibition of *Snap-25* expression by Pax-3 is due to this N terminal region, an expression construct of Pax-3 missing the first 33 amino acids was manufactured. This deletes the first 1/3 of the inhibitory domain, but leaves the paired domain intact and does not interfere with DNA binding. This PaxΔinhib plasmid was co-transfected with the Snap-25-CAT reporter vector. The results from this experiment show that deletion of these first 33 amino acids abolishes inhibition of *Snap-25* expression by Pax-3. Moreover deletion of the first 33 amino acids of Pax-3 converts Pax-3 from an inhibitor of Snap-25 promoter activity to a weak activator. The ability of Pax-3 to inhibit Snap-25 promoter activity appears therefore to be dependent only upon the inhibitory domain of Pax-3, for which the first 33 amino acids appear to be necessary. Only recently has there been significant progress towards identifying gene specific transcriptional repressors and their mode of action. The first transferable repression motif was identified in the N terminus of the product of the *Drosophila* gene Kruppel and was initially characterised as alanine rich region. However alanines in the N terminus are not essential for repression instead they probably serve a structural role in the formation of a alpha helix containing a

glutamine residue critical for repression activity (Margolin *et al.* 1994). Proline rich sequences are also characteristic of several inhibitory domains including the *even skipped* inhibitory domain (Biggin and Tjian 1989) and the Pax-3 N terminal inhibitory domain. It will be interesting to determine whether the first 33 amino acids of Pax-3 can also inhibit transcription when fused to a heterologous DNA binding domain and whether this region of Pax-3 inhibits transcription by recruiting a co-repressor complex. This could be tested by incubating the transiently transfected cells either in the absence or presence of Trichostatin A.

The ability of Pax-3 to directly inhibit the expression of the Snap-25 promoter is intriguing, as Snap-25 is a protein, which in the early embryo plays a role in axonal outgrowth. It has been suggested that Pax-3 plays a role in maintaining the undifferentiated phenotype of cells until appropriate environmental cues trigger differentiation and a downregulation in Pax-3 expression. Kioussi *et al.* 1995, have shown that in Schwann cells high levels of *Pax-3* expressed in Schwann cell precursors, but noted that upon their terminal differentiation into non proliferative myelinated Schwann cells, *Pax-3* expression was downregulated. Furthermore they also showed that Pax-3 was able to repress a 1.3 kb myelin basic protein promoter fragment in co-transfection assays, raising the possibility that Pax-3 may function to maintain the predifferentiated state of these cells by downregulating genes like the myelin basic protein which are necessary for the terminal differentiation of these cells.

Pax-3 may also play an analogous role in sensory neurons as high levels of Pax-*3* transcripts have been detected in the developing sensory ganglia at E10-12, however after E14, when the majority of the sensory neuron precursors have undergone cell cycle arrest and morphological differentiation into mature sensory neurons, *Pax-3* expression ceases (Goulding *et al.* 1991). Furthermore the induction of antisense Pax-3 expression in the sensory neuron cell line ND7 cells leads to the rapid morphological differentiation of the cells, the normally rounded cells exhibit increased cell to substratum adhesion, extend neurite processes and

upregulate genes such as Snap-25 (Reeves *et al.* 1998, 1999). The repression of Snap-25 expression by Pax3 might therefore by important for Pax-3 to maintain the undifferentiated state of the sensory neuron precursor cells

UNIVERSITY OF SOUTHAMPTON LIBRARY

Nuclear Localisation

of Pax-3

5 The Nuclear Localisation of Pax-3

5.1 Introduction

The defining feature of the eukaryotic cell is the nucleus. It contains the genomic DNA and transcriptional apparatus of the cell and is enclosed by a double membrane contiguous with the endoplasmic reticulum. This double membrane separates the nucleoplasm from the cytoplasm and thus offers an extra level of gene control unavailable to prokaryotes.

The only means of entering or exiting the nucleus is via the Nuclear Pore Complex (NPC), which contains about 100 polypeptides and is approximately 125 MDa in size. The NPC contains a passive diffusion channel of about 9nm in diameter. Small proteins less than 40kDa such as cytochrome C (13kDa) can freely pass through the NPC, whereas diffusion of larger protein between 40 and 60kDa, such as ovalbumin (43kDa) is delayed, while proteins larger than 60 kDa, such as BSA (66 kDa), are excluded from the nucleus. Virtually all proteins required in the nucleus however, irrespective of size, are actively transported into the nucleus once they are synthesized in the cytoplasm (Gorlich and Mattaj 1996, Luo *et al.* 1997, Chatterjee *et al.* 1997). The active transport of a protein into the nucleus is generally dependent on a Nuclear Localisation Signal (NLS) encoded in the primary sequence of a protein.

5.1.1 The Nuclear Localisation sequence

A number of different consensus sequences for the NLS have been reported (Chelsky *et al.* 1989; for Review see Dingwall and Laskey 1991), but examination of the full list of nuclear targeting sequences reveals only two common features. Nuclear localisation signals are short, generally less than twelve amino acids and contain a high proportion of positively charged residues – Argenine and Lysine.

However it has been shown that neither of these characteristics is fundamentally necessary for active translocation to occur. The NLS of the influenza glycoprotein, (AAFEDLRVLS) for example, only contains a single basic residue, and there is evidence that the targeting signal of the yeast GAL-4 protein is very large, possible up to 74 amino acids in length.

The most studied nuclear localisation sequence and the paradigm for the short, basic, NLS is that of the SV40 Large T antigen. This sequence, PKKKRKV fulfils both of the requirements of a true NLS. Firstly it is necessary for the nuclear localisation of the protein, and secondly it is sufficient to direct the nuclear localisation of a non-nuclear protein such as serum albumin (Goldfarb *et al.* 1986, Lanford *et al.* 1986). It also seems to be the minimal sequence requirement for nuclear localisation. Other nuclear proteins with this type of NLS have one or more non-basic amino acids inserted in their NLS but in general they fulfil this requirement. A second class of NLS is the bipartite targeting sequence in which two basic amino acids are followed by a non-basic region of 6-10 amino acids and then a basic cluster in which the next three out of five amino acids must be basic. An example of this class is the nucleoplasmin NLS **KRPAATKKAAGQAKKKK**. This bipartite sequence can be found in half of all nuclear proteins and is a considerably more reliable marker of nuclear localisation that the SV40 sequence of five consecutive basic amino acids.

5.1.2 The Mechanism of Active transport

Active transport of proteins across the NPC can accommodate rigid structures up to 25nm in diameter. Gold particles of this size covered with nuclear localisation signal polypeptides are imported. Since these cannot change their shape there must be conformational change of the NPC rather than of the translocated protein. The active transport of nuclear proteins has been shown to require energy, is signal dependent and can be saturated by the microinjection of large amounts of NLS peptide conjugates into the cytoplasm, thus indicating a carrier mediated process.

Subsequent experiments have show that the carrier in this case is a heterodimer, known as Importin-α (60 kDa) and Importin-β (90 kDa). Importin-α binds to the NLS while Importin-β later docks at the Nuclear Pore Complex. Experiments using chimeric fusions of proteins with Importin-β show that only Importin-β is required for nuclear translocation. It has been shown that Importin-β binds to Gly-Cys-Phe-Gly or Phe-x-Phe-Gly repeats within the nucleoporins of the NPC. The microinjection of wheat germ extract (WGA) to the cytoplasm binds to the N-acetyl glucosamine on the face of the nuclear pores which prevents binding of importin-β and thus translocation. Passive diffusion of small proteins, which do not require any docking, is unimpeded.

Once docking has been accomplished the small GTPase Ran then binds the NPC and the hydrolysis of GTP drives the active translocation of proteins across the NPC. The whole complex of Importin-α, Importin-β, and nuclear protein is translocated. Once in the nucleoplasm the components separate, probably by some sort of conformational (conversion to low affinity) change by Importin-α that releases the nuclear protein and prevents binding of the NLS of any of the high concentration of nuclear proteins, and thus their export with importin-α. The export of importin-α is so fast that higher concentrations are found on the cytoplasmic side of the nuclear envelope, however high sensitivity experiments (electron immunomicroscopy) have detected it on both sides. Examination of the amount of energy required for the ATP driven myosin powerstroke of 10nm leads to a figure of 10 GTPs to cross the 100nm NPC.

5.1.2.1 The nuclear localisation cycle can thus be separated into the following steps

1. Docking of Importin-α subunit to the NLS
2. Heterodimerisation of the Importin-α subunit with Importin-β
3. Docking of the Importin-β subunit to the Nuclear pore Complex

4. Hydrolysis of GTP by Ran giving energy driven translocation.
5. Dissociation of Importin-α from NLS
6. Export of Importin-α and importin-β

5.1.3 Regulation of Nuclear Translocation

The active translocation of regulatory transcription factors into the nucleus from their site of synthesis in the cytoplasm leads in turn to the potential regulation of this transport and thus giving rise to a further level of post-translational modulation of proteins necessary for transcriptional control. While some of the structural nuclear protein, such as histones are constitutively translocated to the nucleus, it seems that nuclear localisation of almost all transcription factors is regulated in some manner.

The first evidence for this regulation was found in the analysis of the nuclear translocation kinetics of the SV40 large T-antigen. While the classic NLS discussed above is sufficient to localise non-nuclear proteins, it was found that a flanking sequence containing a phosphorylation site for CKII could enhance the rate of translocation b 40-fold (Rihs *et al.* 1991). A further flanking sequence containing a phosphorylation site for the cell-cycle dependent kinase cdc2 also modulates SV40 T-ag nuclear localisation, phosphorylation at this site substantially reduces the extent of nuclear import. Since CKII is known to be integral in mediating cellular responses to mitogens such as insulin and EGF, while cdc2 is important in regulating cell-division, it seems likely that these exert antagonistic effects on nuclear translocation during the cell cycle (Jans 1995). Indeed such phosphorylation sequences have been shown to be active in such cell-cycle transcription factors as p53, c-myc, RB-1.

In addition to direct phosphorylation near the NLS there are two other important means by which nuclear entry can be controlled.

5.1.3.1 NLS masking

A number of proteins possessing an apparently functional NLS have been found to be predominantly cytoplasmic due to the masking of their NLS, generally by another protein or an external factor. This prevents the binding of importin-α and thus nuclear import. Upon dissociation of the masking polypeptide or domain the revealed NLS can bind to the importin and thus imported into the nucleus. This mechanism allows the presence of an inactive form of transcription factor, which can then be quickly activated, and imported into the nucleus without the need to wait for transcription.

The classic example of this form of regulation is that of NF-κB and its interacting partner I-κB. In this case hypophosphorylated I-κB binds to the NLS of NF-κB and thus preventing nuclear translocation. Subsequent phosphorylation of I-κB leads to dissociation of the two proteins and fast degradation of the phosphorylated I-κB.

5.1.3.2 Cytoplasmic retention factors.

A related method of regulation is that of cytoplasmic retention, whereby an anchoring protein binds to the nuclear protein, but not necessarily masking the NLS, and preventing nuclear translocation. This mechanism has been described for c-fos, GlucR and cyclin-B1. Perhaps the best studied case is that of the glucocorticoid receptor, GlucR, which is retained in the cell membrane by the heat-shock protein HSP-90. Upon ligand binding GlucR dissociates from HSP-90, dimerises and is translocated to the nucleus.

Thus it can be seen that nuclear localisation and its careful regulation can be an important method of gene control in the nucleus. The identification of the NLS and its regulatory co-factors or phosphorylation sites is therefore an important part of elucidating the control of Pax-3.

5.2 Nuclear Export

For molecules larger than 60kDa export from the nucleus must also be active, microinjection of large fluorescently tagged proteins into the nucleus has shown that they will be retained in the nucleus until degraded, unless they contain a Nuclear Export Signal (NES). One of the first examples found of the NES was that of the HIV-1 protein Rev, with the sequence LPPLERLTL. It appears that hydrophobic residues, especially leucines are required for nuclear export (Wen *et al.* 1995). Nuclear export is via the same pores as nuclear import. Proteins ranging from NF-κB to MAPKK have potential NES polypeptides. Unlike the importins there are a large number of proteins involved in nuclear export, which are presumably active under different circumstances (For review see Ullman *et al.* 1997).

As with the NLS the NES can be regulated by a variety of extra- and intracellular signals such as stress. However not all nuclear factors contain nuclear export signals, and there are some transcription factors such as c-Fos where regulation is by nuclear import. In cells grown in the presence of serum c-Fos appears to be exclusively nuclear, after serum starvation c-Fos is progressively lost from the nucleus over a period of 4-5 hours and there is a concurrent build up in the cytoplasm (Roux *et al.* 1990). It has been shown that this is not due to nuclear export but to a very short half-life, leading to quick degradation of c-Fos.

The very short half-lives of these proteins is due to the presence of PEST domains which expedite degradation of proteins. These were first found in sequence searches of very rapidly degraded proteins, such as Jun, Myc, Fos, p53 etc. These were found to contain regions highly enriched in Proline (P), Glutamate (E), Serine (S), Threonine (T), hence the sobriquet, PEST domains (Rogers *et al.* 1986, Rogers and Rechsteiner 1996). At the moment evidence points toward PEST domains promoting degradation via the ubiquitin-26S proteosome pathway, although there seem to be some contradictory reports. Sequence searches of Pax-3

have shown the presence of a strong PEST domain in the primary sequence of Pax-3, but no evidence of a Nuclear Export signal. The program PEST-FIND developed by Rogers and Rechsteiner gives the Pax-3 PEST domain a similar score to that of c-Fos.

5.3 Pax-3 localisation

Many ways have been reported to examine the nuclear localisation of transcription factors. These range from microinjection of fluorescently tagged proteins, to immunofluorescence. These methods however suffer from the disadvantage that the fixation or cell intrusion required might produce artifactual results. Recently however an *in vivo* method has been developed (Chatterjee *et al.* 1997). This takes advantage of the stimulated fluorescence of the Aequorae Victoria Green Fluorescent Protein (GFP). In natural circumstances this protein emits light at 509nm when energy is transferred to it from the Ca^{2+}-activated photoprotein aequorin. However it will also emit light after being stimulated at 400nm. A mutated GFP that has enhanced light emission from stimulation at 488nm has been manufactured (EGFP, Clontech) and its cDNA has been inserted into a plasmid under the control of a CMV promoter. This plasmid (see Fig 5.01) allows insertion of cDNA upstream of the EGFP moiety thus allowing production of a fusion chimera of the protein plus EGFP. The progress of this chimera can then be tracked using fluorescent microscopy or Laser Confocal microscopy. Assuming that the addition of the small (28kDa) EGFP does not affect the localisation of Pax-3 its localisation in the cell can be ascertained.

5.3.1 The localisation of Pax-3 in proliferating ND7 cells.

In order to manufacture a Pax-3-EGFP fusion protein in the vector pEGFP-N3, the Pax-3 cDNA must be inserted upstream of the coding region of EGFP. Therefore the 3' sequences within the Pax-3 full length cDNA that encode for a

stop codon must be removed. To do this, the full length mouse Pax-3 cDNA was excised from pBSPax-3 using a BamH1 site at the 5' end of the Pax-3 cDNA and the downstream EcoR1 site. The stop codon and the 3' untranslated were then excised by cutting the Pax-3 moiety with EcoN1 at 1685bp. The vector pEGFP-N3 was cut with Bgl II and EcoR1. The ends of the Pax-3 cDNA fragment and vector were blunt-ended and the vector and pax-3 fragment ligated. This produces a plasmid that produces a Pax-3-EGFP fusion protein under the control of a CMV promoter in mammalian cells (Fig. 5.01). In order to ascertain whether the EGFP fusion will affect the nuclear localisation of Pax-3, the construct was transfected into ND7 cells, incubated for 40 hours and the cells then examined by digital fluorescence microscopy. In order to visualise the nucleus, the cells were stained with 2,4 Diamidinopropylindole (DAPI) a live cell stain that intercalates with DNA, and when stimulated by UV light it emits in the blue spectrum and thus does not interfere with the signal from EGFP. As a control, the EGFP plasmid alone was transfected into the ND7 cells. The results from these experiments can be seen in Fig. 5.02.

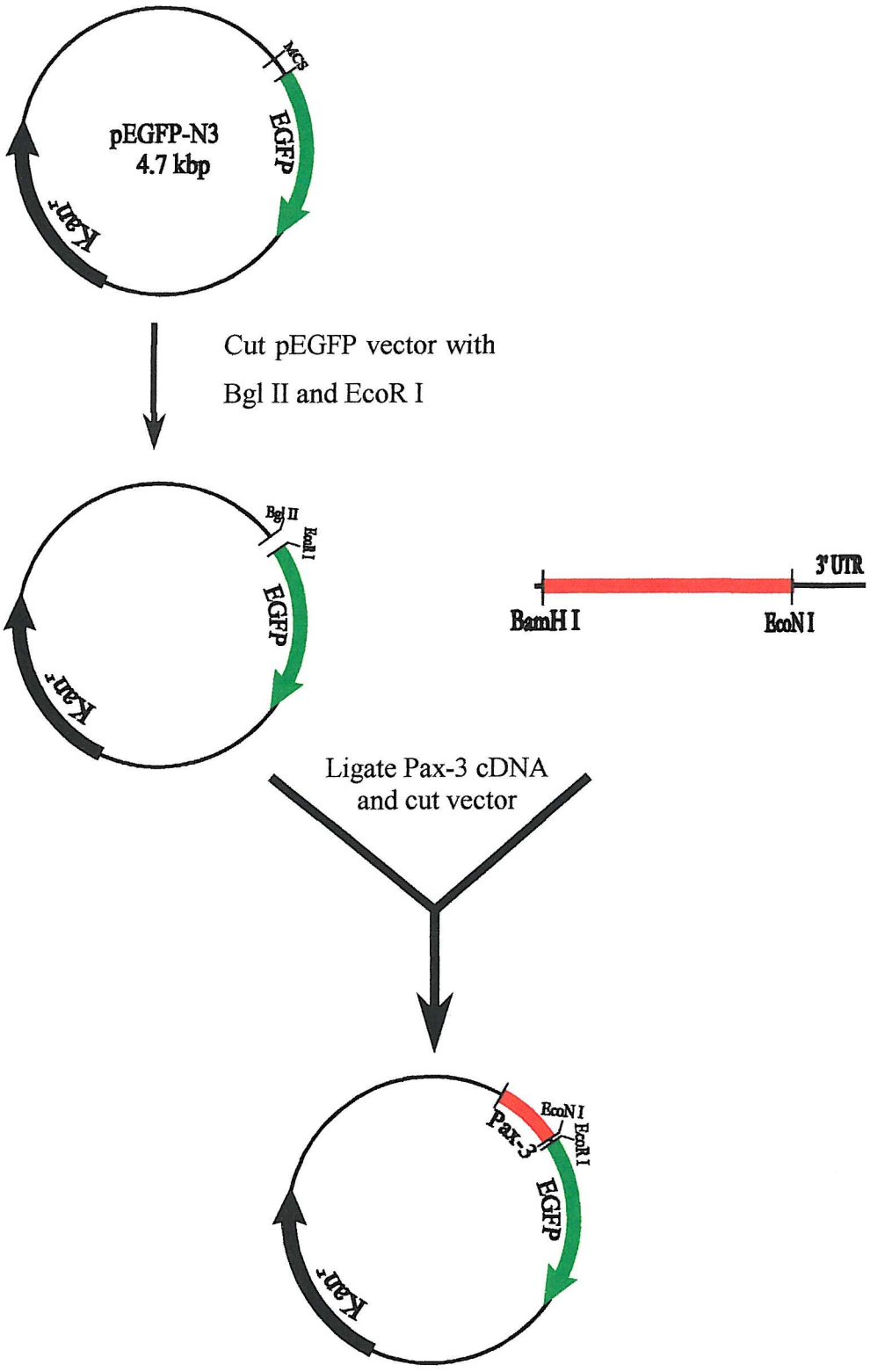


Fig. 5.01 Diagrammatic representation of the cloning of Pax-3 cDNA into the pEGFP vector, and the subsequent removal of the stop codon and 3' untranslated region.

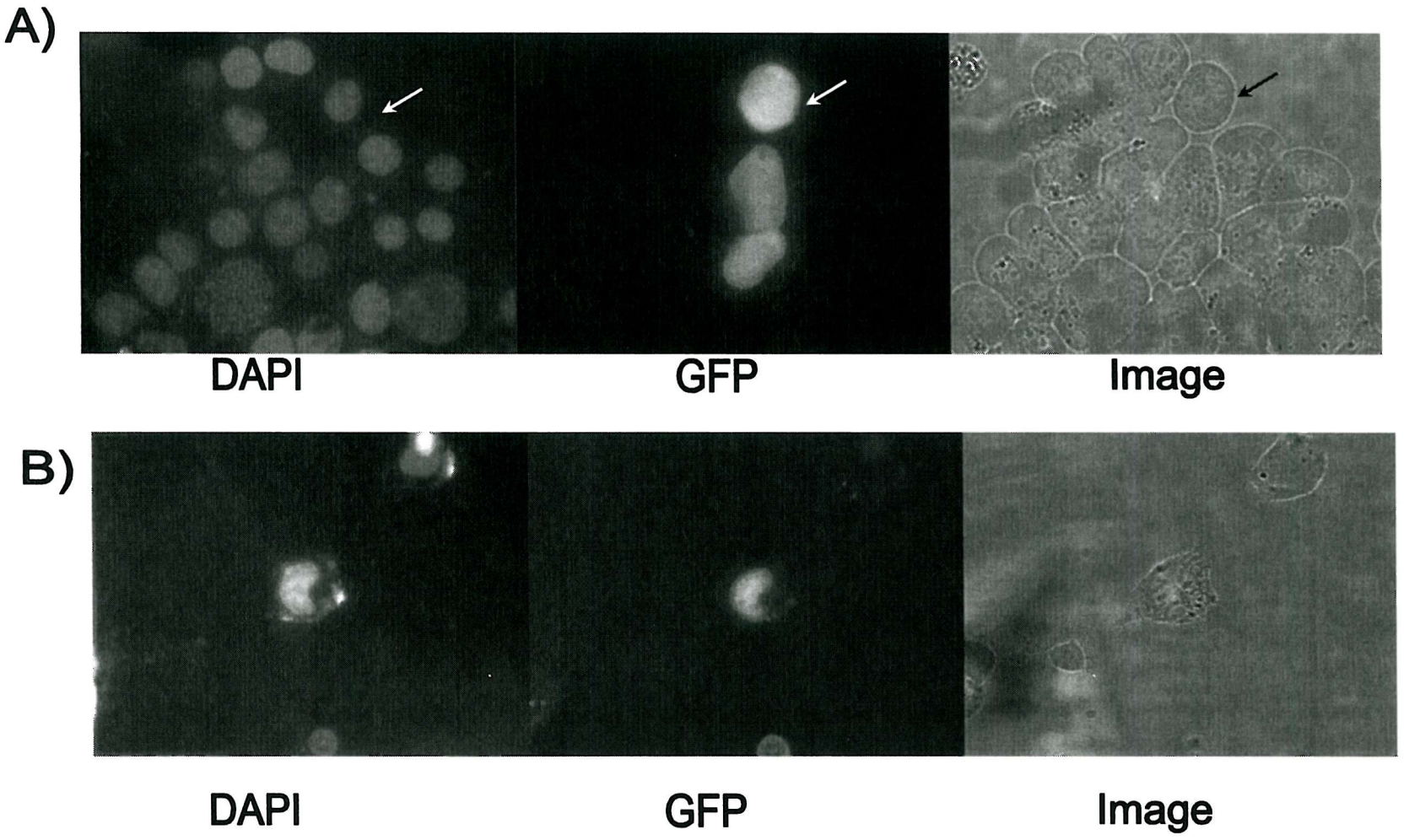


Fig. 5.02 Digital images showing the cellular localisation of A) EGFP and B) Pax-3-EGFP. ND7 cells were transfected with pEGFP and pPax-3-EGFP (5μg) for 40 hours after which cells were washed in PBS, stained with the nuclear stain DAPI and examined under a fluorescence microscope. Pictures show nuclear staining by DAPI, GFP localisation, and transmission images.

As shown in Fig 5.02 transfection of pEGFP alone results in fluorescence being localised throughout the cell, including the nucleus, as might be expected of a 28kDa protein, free to passively diffuse through the nucleus. The Pax-3 EGFP fusion protein however can be seen to be localised only in the nucleus, demonstrating that Pax-3 is indeed actively transported into the nucleus.

In order to show that the entry of EGFP into the nucleus is purely by passive diffusion, a further fusion of EGFP with β-Galactosidase was manufactured. This produces a fusion protein of approximately 130kDa, which because of its molecular weight is excluded from the nucleus. The β-Galactosidase cDNA was amplified from the plasmid pCH110 with the primers β-Gal-Fwd and β-Gal-Rev. These encode a Tat I site at the 5' end and a Not I site at the 5' end. This was then inserted 3' to the EGFP cDNA in the plasmid pEGFP-N3, by cutting with BsrG1 and Not I, which excises the stop codon and allows the in frame insertion of β-Galactosidase. This resulting plasmid retains the multiple cloning site and produces an EGFP-β-Galactosidase fusion protein. A diagrammatic representation of this cloning procedure is shown in Fig. 5.03.

This plasmid, GFP-β-Gal was then transfected into ND7 cells grown on cover slips, using the calcium phosphate method. After 40 hours, the cells were stained with DAPI and examined using fluorescence microscopy (Fig. 5.04). These pictures show that the GFP-β-Gal fusion protein is excluded from the nucleus, further confirming that the localisation of Pax3EGFP to the nucleus must be due to the presence of an active NLS within the Pax-3 protein.

Fig. 5.03 Diagramatic representation of the cloning of β-Galactosidase, downstream of the EGFP coding cDNA.

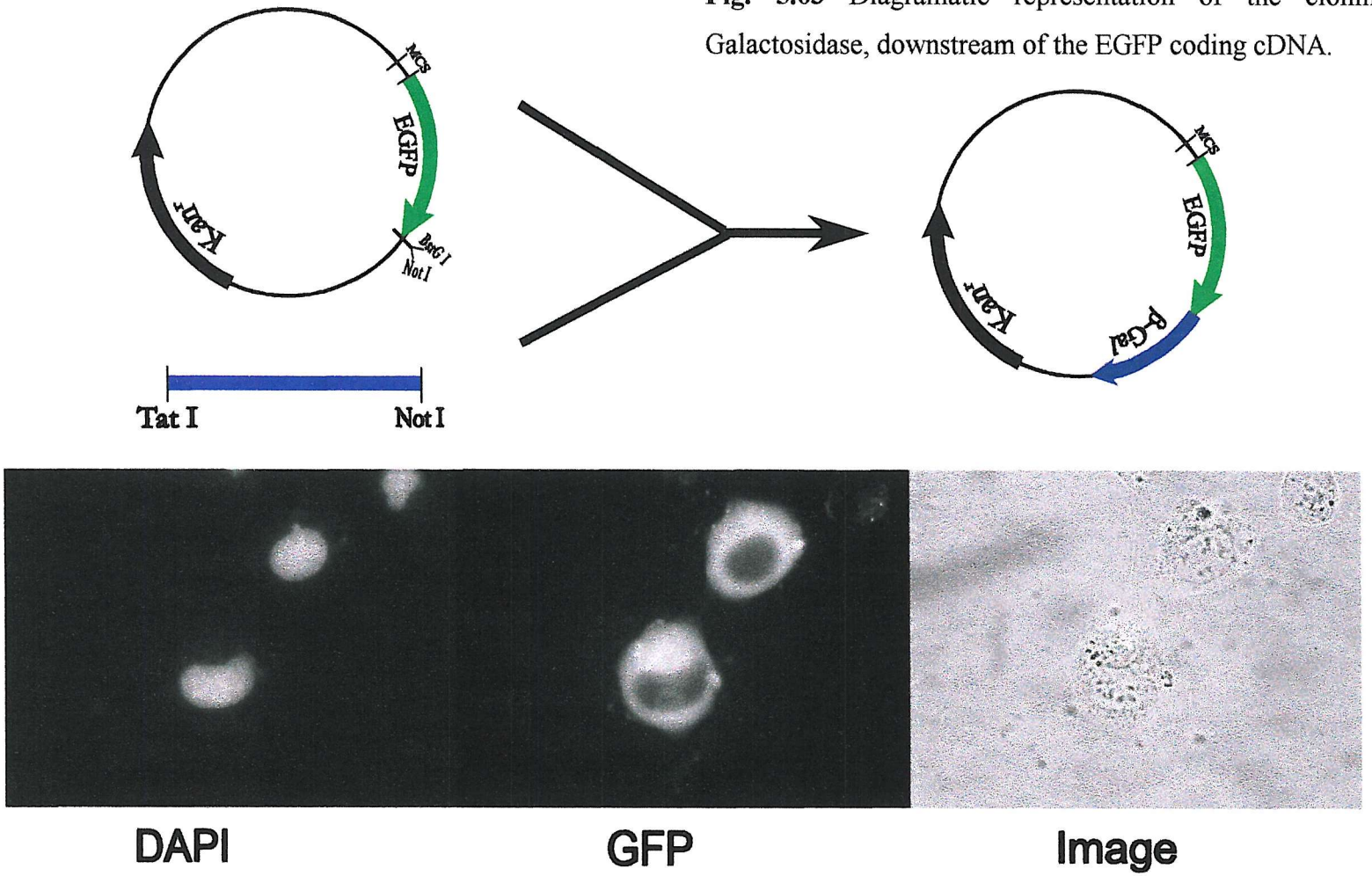


Fig. 5.04 Digital images showing the cellular localisation of EGFP-β-Galactosidase. ND7 cells were transfected with pEGFP-β-Gal (5μg) for 40 hours after which cells were washed in PBS, stained with the nuclear stain DAPI and examined under a fluorescence microscope. Pictures show nuclear staining by DAPI, GFP localisation and transmission images

5.4 Identification of the Nuclear Localisation Sequence of Pax-3

Positive identification of the NLS by sequence searching revealed a number of putative NLS sequences, therefore two deletion mutants of the Pax-3-EGFP fusion were made to narrow the potential region within which the NLS might be found. These were manufactured using the restriction endonuclease sites, Hind III and Cla I in the Pax-3 cDNA. To make the former, the Pax-3-EGFP plasmid was cut with Hind III, and then blunt ended. This was then cut with Nhe1 which gave a fragment of Pax-3 cDNA with a Nhe I overhang at the 5' end, and a blunt end at the 3' end. This fragment was then gel purified and ligated into a pEGFP-N3 plasmid that had been cut by Sma I and Nhe I (Fig. 5.05). When transfected into ND7 cells, this results in an EGFP fusion protein containing the first 257 amino acids of Pax-3. This Pax-3 fusion therefore contains the paired domain, the octapeptide region and the N terminal region of the homeodomain.

The Cla I deletion was manufactured in a similar manner, the Pax-3 EGFP construct was cut with Cla I instead of Hind III and the fragment inserted into the an EGFP plasmid that had been cut with Sal I, blunt ended and then cut with Nhe I. This construct contains a Pax-3 cDNA region encoding only the first 189 amino acids i.e. It contains the paired domain, a portion of the octapeptide region but none of the homeodomain. A graphical representation of these deletion mutants is shown in Fig. 5.05. These deletions were transfected into ND7 cells grown on cover slips. After 40 hours, they were imaged using fluorescence microscopy, producing the results shown in Fig. 5.07.

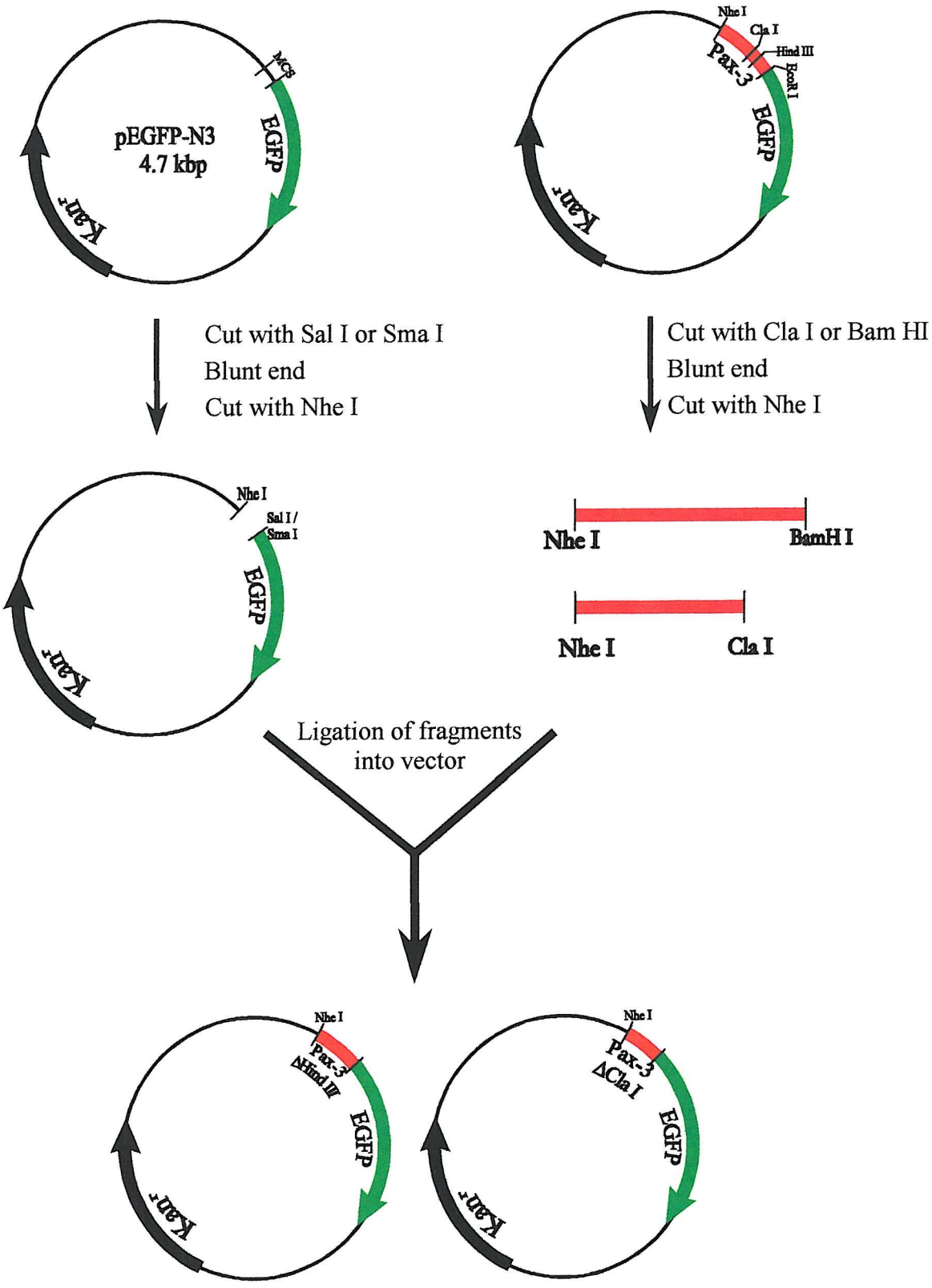
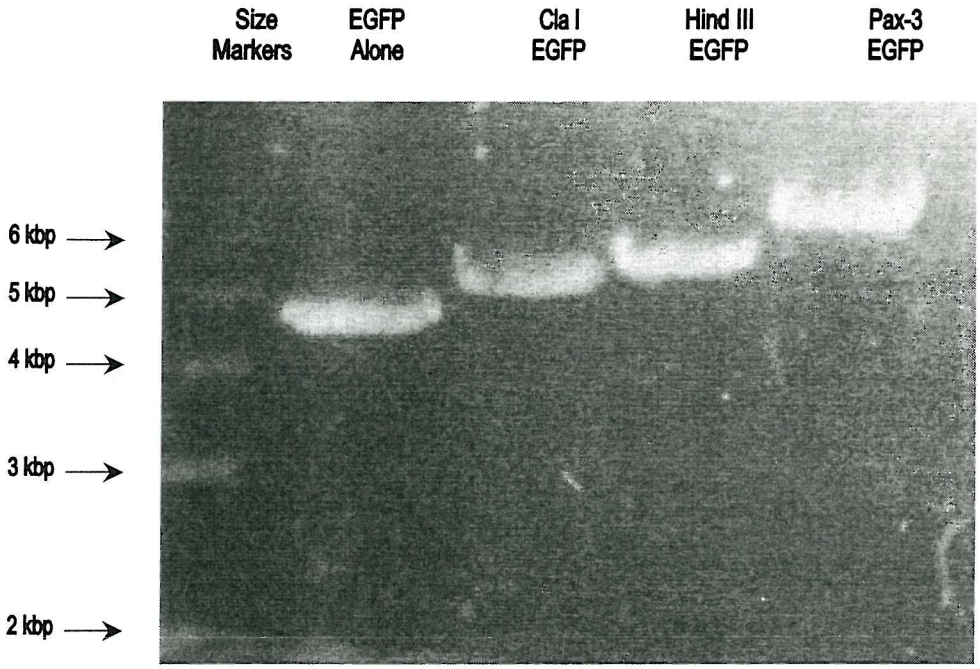


Fig. 5.05 Diagram showing the cloning of Pax-3 truncated at Hind III and Cla I restriction sites, into the pEGFP vector.

All of the above Pax-3 EGFP deletion mutants were confirmed by restriction mapping. One such restriction digest is shown in Fig. 5.06. In this example pEGFP, ClaI-EGFP, Hind III-EGFP and full-length Pax-3 were all cut using BamHI. This cuts at a single site immediately 5' to the start of the coding sequence of EGFP. The cut plasmids were run on a 0.7% agarose gel at 100V for 45 minutes. Taking into account the non-linear running of the gel both absolute and relative sizes of the clones are as expected.



The expected sizes of the plasmids are :-

PEGFP	4.7 kbp
Cla I –EGFP	5.2 kbp
HindIII-EGFP	5.5 kbp
Pax-3 EGFP	6.1 kbp

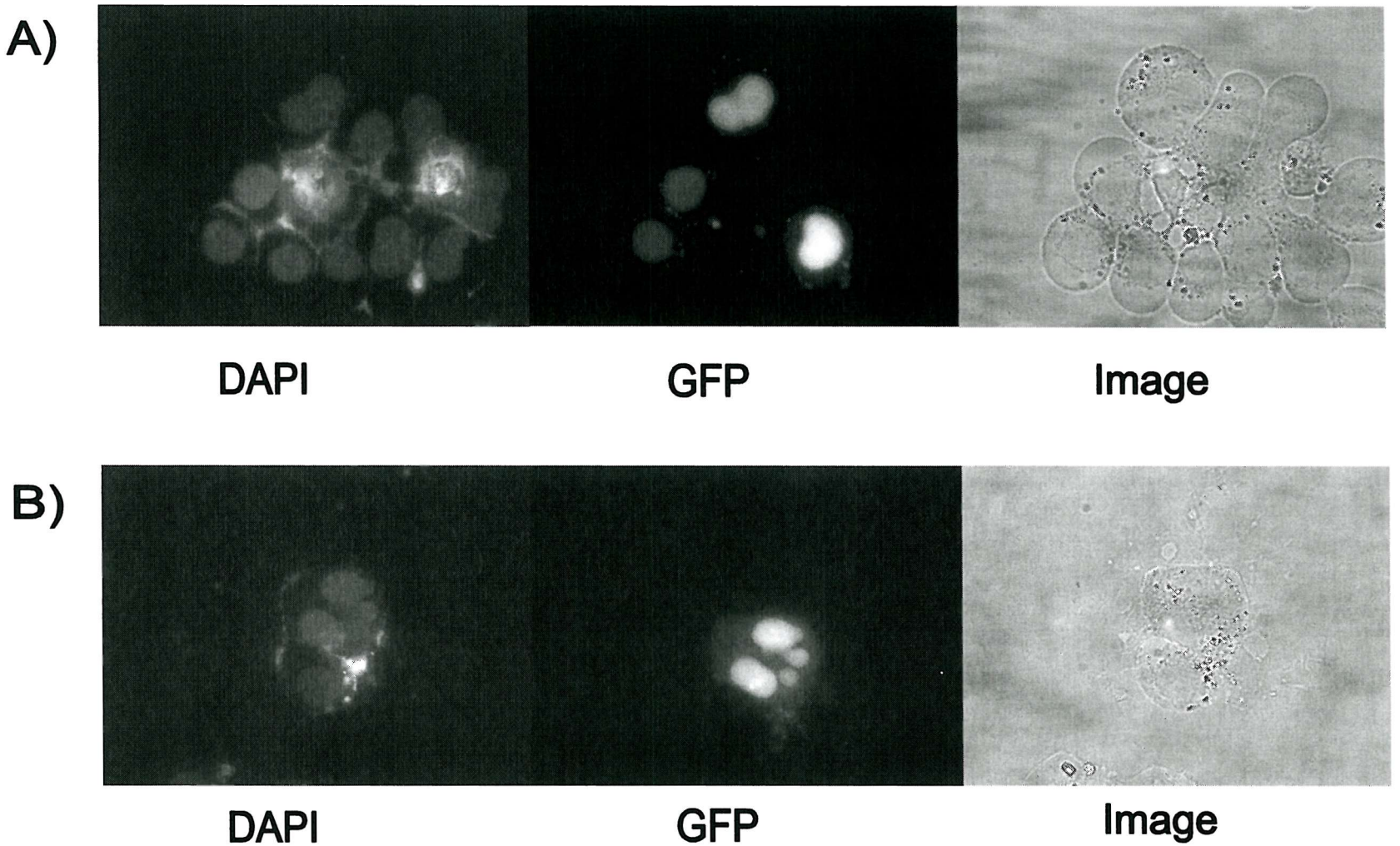


Fig. 5.07 Digital images showing the cellular localisation of A) PaxΔHind-EGFP and B) PaxΔCla-EGFP. ND7 cells were transfected with plasmids (5μg) for 40 hours after which cells were washed in PBS, stained with the nuclear stain DAPI and examined under a fluorescence microscope. Pictures show nuclear staining by DAPI, GFP localisation and transmission images.

Additionally, nuclear and cytoplasmic fractions were made from ND7 cells transfected by the various clones. These fractions were run on a 10% polyacrylamide gel, transferred to membrane and immuno-blotted using polyclonal anti-EGFP antibody (Clontech). Fig. 5.08.clearly shows the presence of all constructs in high abundance in the nucleus. There is no trace of the full length Pax-3-EGFP in the cytoplasm, while the Cla I and Hind III constructs show a reduced presence in the cytoplasm. It should be noted that with 5µg of protein in each lane no attempt has been made to correct for nuclear : cytoplasmic ratios. This may account for the presence of the Hind III construct in the western blot but

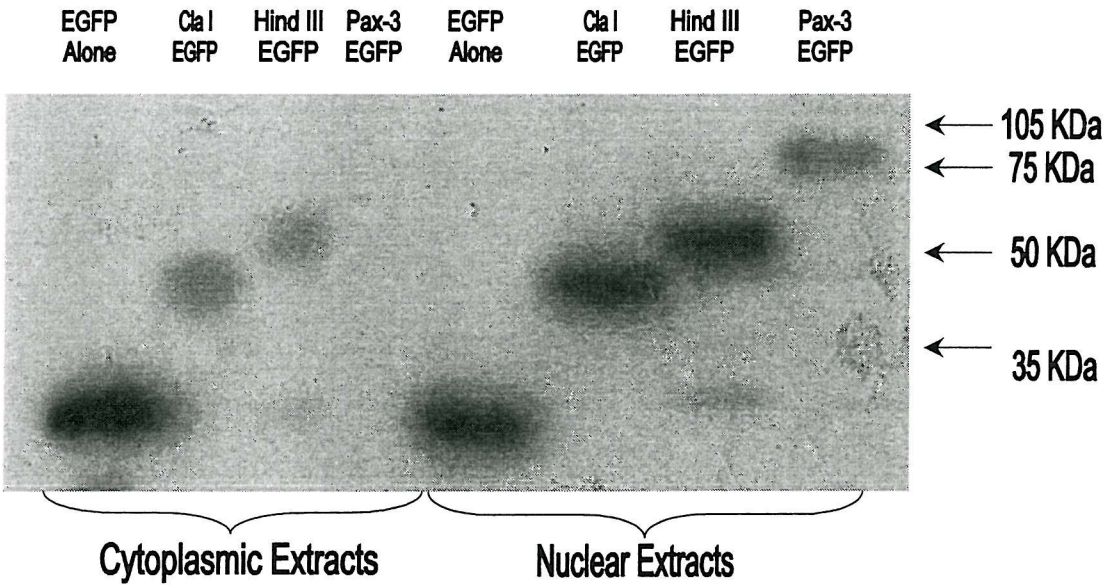


Fig. 5.08 Western blot of Nuclear and Cytoplasmic extracts (5µg) from ND7 cells transfected with EGFP, Cla I-EGFP, Hind III-EGFP and Pax-3 EGFP expression constructs and blotted with an anti-EGFP polyclonal antibody.

not in the fluorescence pictures.

The expected sizes of the fusion proteins are as follows

PEGFP	27 kDa
Cla I –EGFP	48 kDa
HindIII-EGFP	55 kDa
Pax-3 EGFP	80 kDa

The digital fluorescence images (Fig. 5.07) show that transfection of the PaxΔHindIII-EGFP mutant into ND7 cells resulted in the fusion protein being localised exclusively to the nucleus, suggesting that this deletion mutant still retains an active NLS sequence. However transfection of the ClaΔEGFP mutant resulted in fluorescence being found both in the nucleus and cytoplasm suggesting that part of the sequences important for nuclear localisation of Pax-3 have been deleted in this mutant.

There must therefore be some form of nuclear localisation signal between amino acid 189 and amino acid 257.

The approximate portion of Pax-3 that these deletion mutants cover and their subcellular localisation is shown in Fig. 5.09.

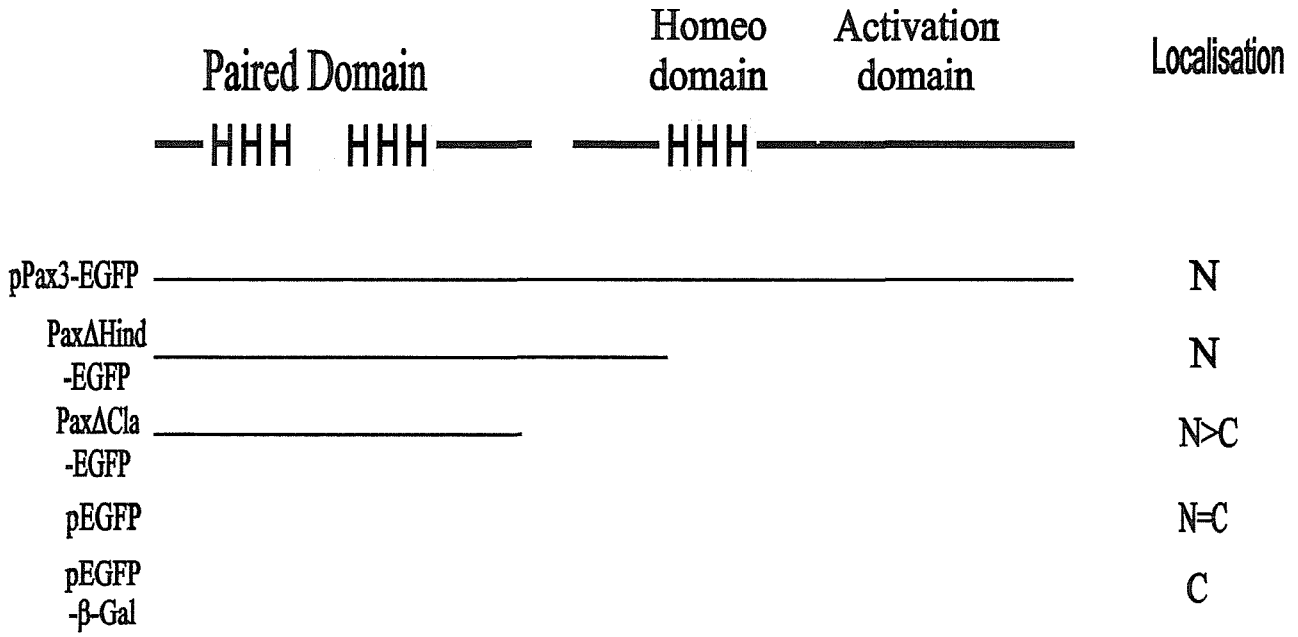


Fig. 5.09 Diagram showing the comparative localisations of the full length Pax-3 cDNA, the Hind III, Cla I truncations and GFP alone and GFP-β-Galactosidase fusion proteins respectively.

A manual search of the sequence between the Cla I site and the Hind III site, pinpoints one particular basic rich sequence.

Pro-Leu-Lys-Arg-Lys-Gln-Arg-Arg-Ser-Arg

A homology search of this region shows that it is unusually conserved in the other Pax proteins (see Fig. 5.10). The surrounding region is not particularly conserved and therefore it seems likely therefore that this sequence may be a functional NLS.

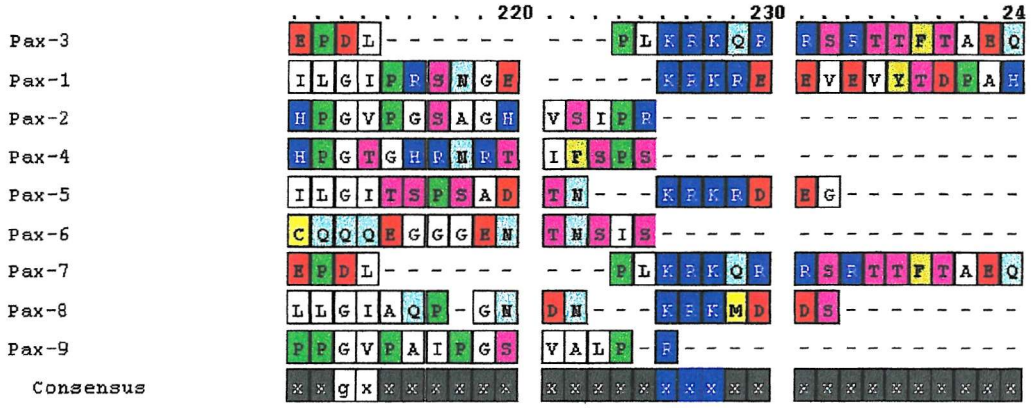


Fig. 5.10 Sequence alignment showing conserved regions of basic amino acids (blue) and putative NLS sequence, within the Pax family of proteins

In order to verify that this sequence can indeed function as an NLS, oligonucleotides encoding the sense and antisense of the above polypeptide sequence were manufactured, annealed and ligated into the pEGFP-N3 plasmid. When transfected into ND7 cells, this sequence alone was able to localise EGFP to the nucleus (Fig. 5.10), suggesting that this sequence may act as a NLS for Pax-3.

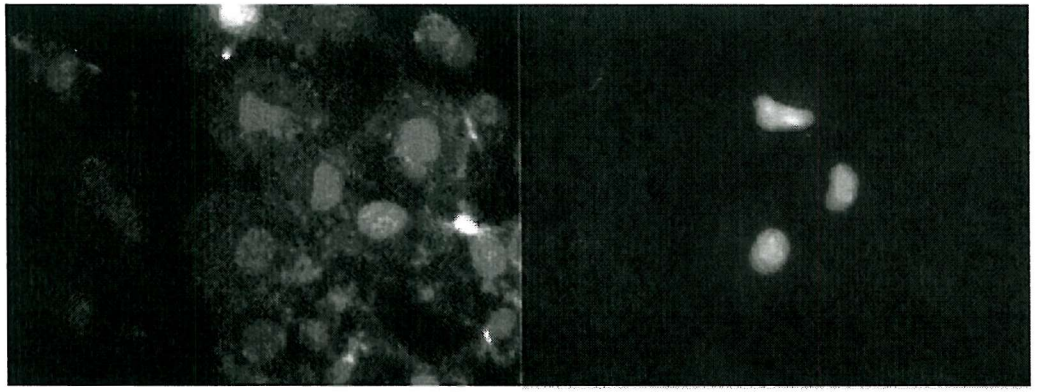


Fig. 5.11 Digital images showing the cellular localisation of NLS-EGFP. ND7 cells were transfected with pNLS-EGFP (5μg) for 40 hours after which cells were washed in PBS, stained with the nuclear stain DAPI and examined under a fluorescence microscope. Pictures show nuclear staining by DAPI, GFP localisation and transmission images

However as the Pax-3ΔClaI-EGFP fusion protein located both to the cytoplasm and nucleus, this suggests that the Cla I mutant although it lacks the NLS sequence identified above must contain other sequences that promote nuclear localisation. To determine whether Pax-3 contains a second NLS sequence upstream of amino acid 189, (this comprises mainly the paired domain) the two subdomains of the paired domain were separately amplified using PCR and then inserted into the EGFP-β-Galactosidase plasmid.

The first helix-turn-helix subdomain was amplified using the primers Prd1-Fwd and Prd1-Rev, this inserts a methionine start codon immediately upstream of Ser-30, and fuses to the EGFP moiety at Glu-115. This also encodes a Sac II site at the 3' end. After amplification the product was gel purified and inserted into the pGEM-T-Easy A/T cloning vector. The fragment was then excised using SacII and Sal I and inserted into pEGFP cut with Xho1 and SacII.

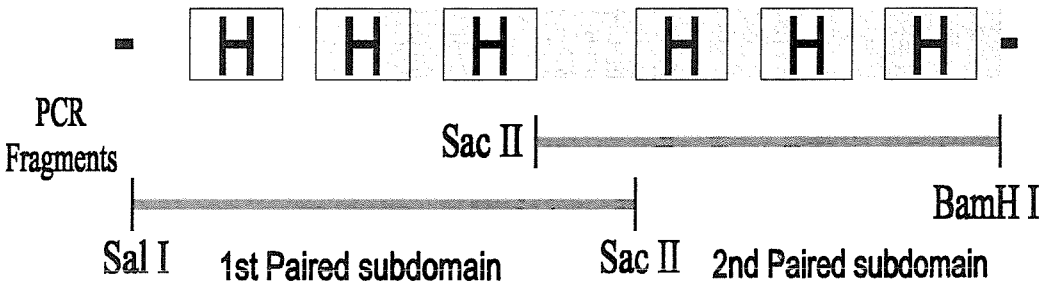


Fig. 5.12 Diagrammatic representation of the two paired subdomain expression clones manufactured.

The second helix helix-turn-helix subdomain was amplified using the primers Prd2-Fwd and Prd2-Rev. This inserts a methionine residue upstream of Arg-97 and fuses to the EGFP moiety at Glu 169. This fragment has a BamHI cleavage site at its 3' end. After amplification the product was gel purified, then inserted into the pGEM-T-Easy T/A cloning vector. It was then excised using BamHI and Sac II and inserted into the p-EGFP-β-Galactosidase vector.

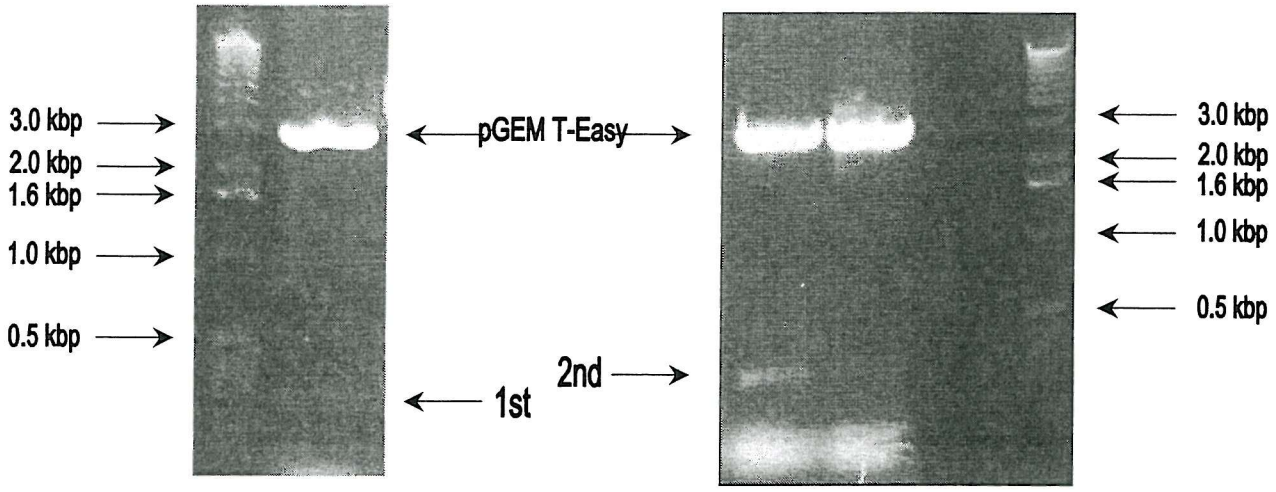


Fig. 5.13a Ethidium bromide stained agarose gels showing the 1st and 2nd paired subdomains cleaved out of pGEM-T-Easy with EcoRI

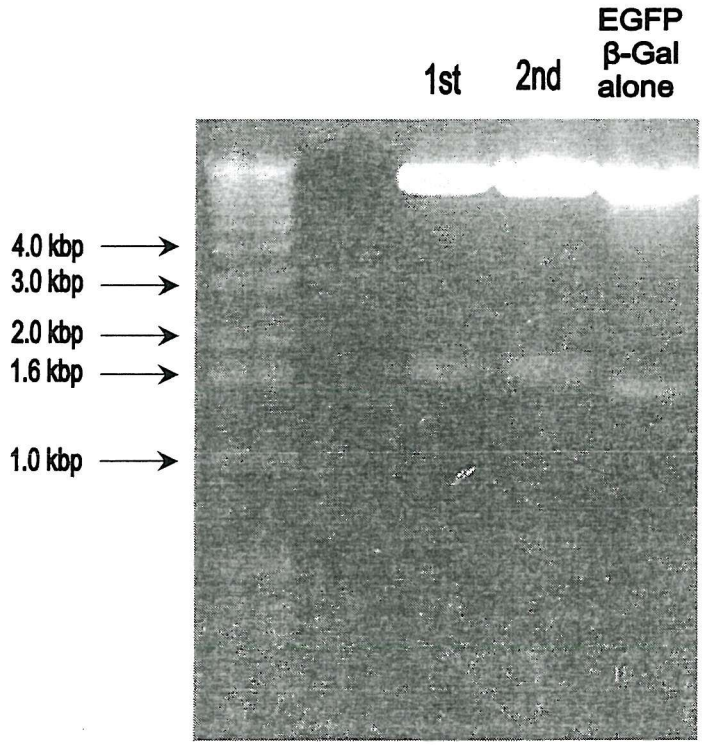


Fig. 5.13b Recombinant pEGFP-β-Gal paired domain constructs. Plasmid DNA was cut with EcoR1 which cleaved once in front of the inserted fragments from Pax-3, and once 600bp into β-Galactosidase.

These two expression plasmids were then transfected into ND7 cells for 40 hours stained with DAPI and then pictures taken with a digital camera attached to a fluorescence microscope. (Fig. 5.14)

As shown in Fig 5.14, no part of the paired domain is capable of localising a protein to the nucleus on its own. This suggests that the second NLS is not located within the paired domain but is located either in the first 30 amino acids of Pax-3 or between amino acids 167 and 188. Alternatively it is possible that the second NLS is a large bipartite NLS encoded within these regions.

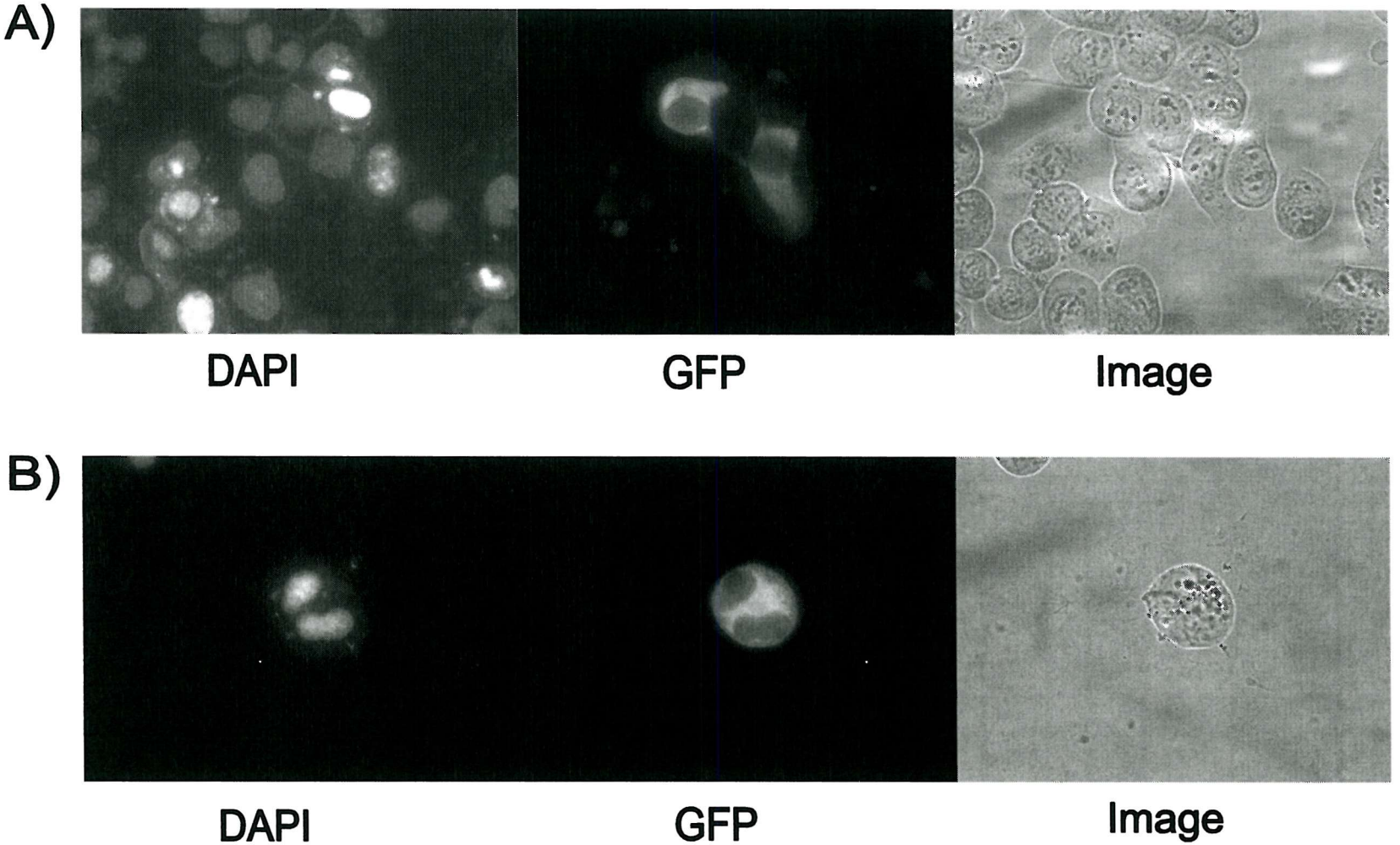


Fig. 5.14 Digital images showing the cellular localisation of A) Prd-1-EGFP and B) Prd-2-EGFP. ND7 cells were transfected with plasmids (5μg) for 40 hours after which cells were washed in PBS, stained with the nuclear stain DAPI and examined under a fluorescence microscope. Pictures show nuclear staining by DAPI, GFP localisation and transmission images.

5.5 Regulation of Pax-3 Nuclear Localisation

As has previously been mentioned the regulation of Nuclear Localisation is a common feature of control of transcription factor activation. Nuclear localisation has been found to be affected by many intracellular and extracellular factors from oxidative stress to serum and growth factors. To determine if the nuclear localisation of Pax-3 was regulated we examined the effect on serum on the localisation of Pax-3. Indirect immunofluorescence experiments were used using a Pax-3 specific antibody (Fredericks et al., 1993). In proliferating ND7 cells grown in the presence of 10% serum Pax-3 protein co-localised with propidium iodide to the nucleus of the cells (Fig. 5.15). However upon the removal of serum Pax-3 staining now localised exclusively to the cytoplasm of the cell. No staining was seen when the secondary antibody was used alone (Fig.5.15 C & E) or when the primary antibody was used To determine whether the nuclear localisation of Pax-3 in primary cells is also regulated by serum we examined Pax-3 localisation in primary cultures of mouse neural crest cells Neural crest cells are a transient population of neuronal precursor cells that develop into many peripheral and central nerve cell lineages. They are relatively easy to explant and culture *in vitro*. Immunofluorescence experiments using neural crest cells with the Pax-3 antibody showed that in neural crest cells grown in 10% serum Pax-3 protein was localised to the nucleus, however upon serum starvation pax-3 staining became predominantly cytoplasmic. These preliminary experiments suggest that the nuclear import of Pax-3 may be regulated and may be dependent upon serum. (Fig. 5.16).

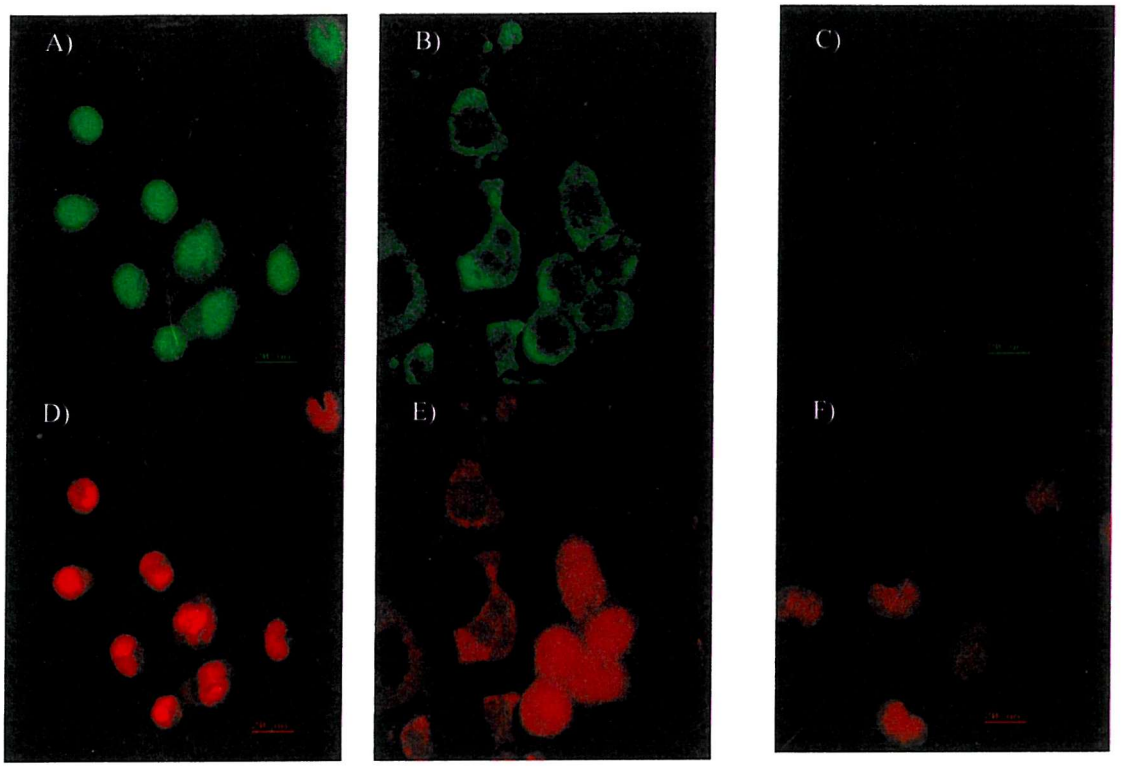


Fig. 5.15 Indirect immunofluorescence studies in ND7 cells using a Pax-3 specific antibody. ND7 cells were grown in the presence of 10% serum (A/D,C/E) or in 0% serum (B/D). Cells were stained using a Pax-3 antibody and a secondary anti rabbit FITC antibody (top panel) and with propidium iodide (bottom panel). Cells shown in C and E were grown in 10% serum and stained with the secondary antibody and propidium iodide only.

A)

B)

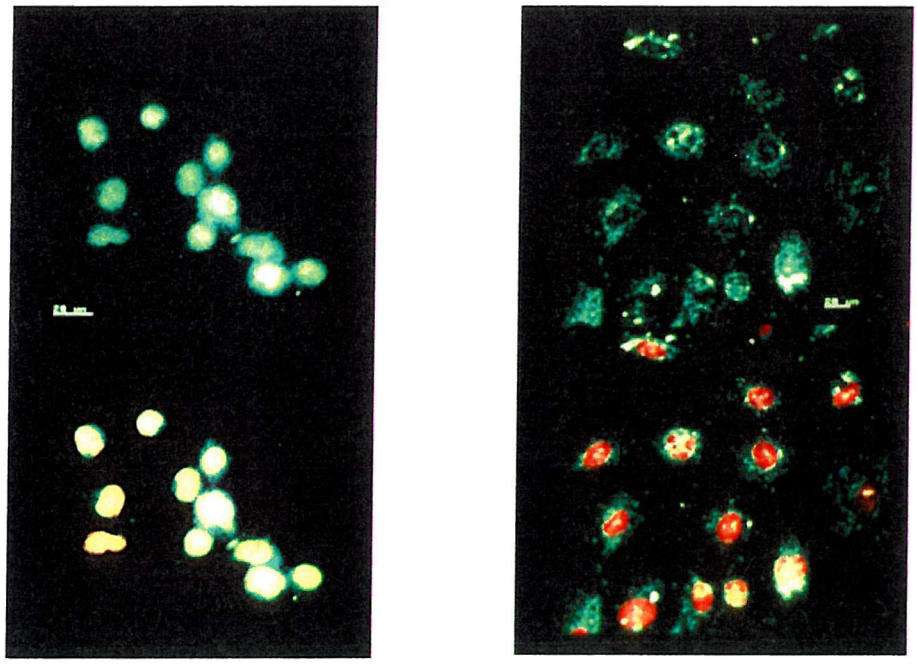


Fig. 5.16 Immunofluorescence studies of neural crest cells. Neural crest cells were grown either in 10% serum (A) or 0% serum (B). Top panel (green) shows Pax-3 immunofluorescence. Bottom panel shows Pax-3 Immunofluorescence (green) overlaid with nuclear staining (red). Co-localisation shows as yellow.

5.6 Discussion

To study Pax-3 nuclear localisation the Pax-3 cDNA was fused to the enhanced green fluorescent protein (EGFP) in the vector pEGFP-N3. Transfection of the empty vector pEGFP-N3 into ND7 cells resulted in the equal distribution of fluorescence throughout the cell. This is consistent with the small size of EGFP (27kDa) which can quite easily passively diffuse into the nucleus with ease. When Pax-3 is fused with EGFP the fusion protein is found exclusively in the nucleus of the cell. As the Pax-3 EGFP fusion protein has a molecular weight of greater than 80 kDa the nuclear localisation must be entirely due to active transport through the Nuclear Pore Complex (NPC). Despite the high expression levels expected due to transcriptional control through a CMV promoter the import apparatus was able to localise all the protein produced to the nucleus. The efficient translocation of Pax-3-EGFP to the nucleus indicates that binding of importins to Pax-3 is not affected in any way by the addition of EGFP to the C-terminus of Pax-3. Experiments also showed the efficient localisation of mutant Pax-3ΔHind III fusion protein to the nucleus, as this fusion protein is approximately 55kDa in size and too large for easy passive diffusion, this suggests that this deletion mutant contains the active NLS. The Pax-3 ΔCla I fusion protein is 47kDa in size and while passive diffusion can occur with molecules of this size it is generally agreed to be slow and inefficient. The localisation of this fusion protein is primarily nuclear, indicating some degree of active transport, concentration gradients once more ruling out passive diffusion. However some cytoplasmic localisation of the fusion protein is observed, suggesting that a portion of the nuclear localisation activity has been removed. The analysis of the sequence between the Cla I and Hind III sites revealed a potential NLS. A homology search of the primary sequence of other Pax proteins in this region revealed that this potential NLS was well conserved amongst other members of the Pax gene family.. When this 11 amino acid

sequence was fused to EGFP and transfected into cells complete nuclear localisation of the fusion protein was seen.

Pax-3 Ser-Asp-Ile-Asp-Ser -Glu-Pro-Leu-Lys-Arg-Lys-Gln-Arg-Arg-Ser-Arg-Thr-Thr
SV40 Ser-Ser-Asp-Asp-Glu-$(X)_8$-Thr-Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val

Lys – Basic residues

Ser – CKII phosphorylation site

Thr – cdc2 phosphorylation site

In the above comparison with the classic SV-40 large T-antigen NLS it can be seen that there is a fairly high degree of homology. The presence of a CKII site in the Pax-3 sequence near to the NLS, as well as in the SV40 Large T Ag (Rihs *et al.* 1991), raises the possibility that pax-3 nuclear import might also be regulated by CKII. Although there is no obvious cdc2 site this kinase has a less characterised consensus sequence and additionally many active cdc2 sites are nowhere near the NLS in the primary sequence.

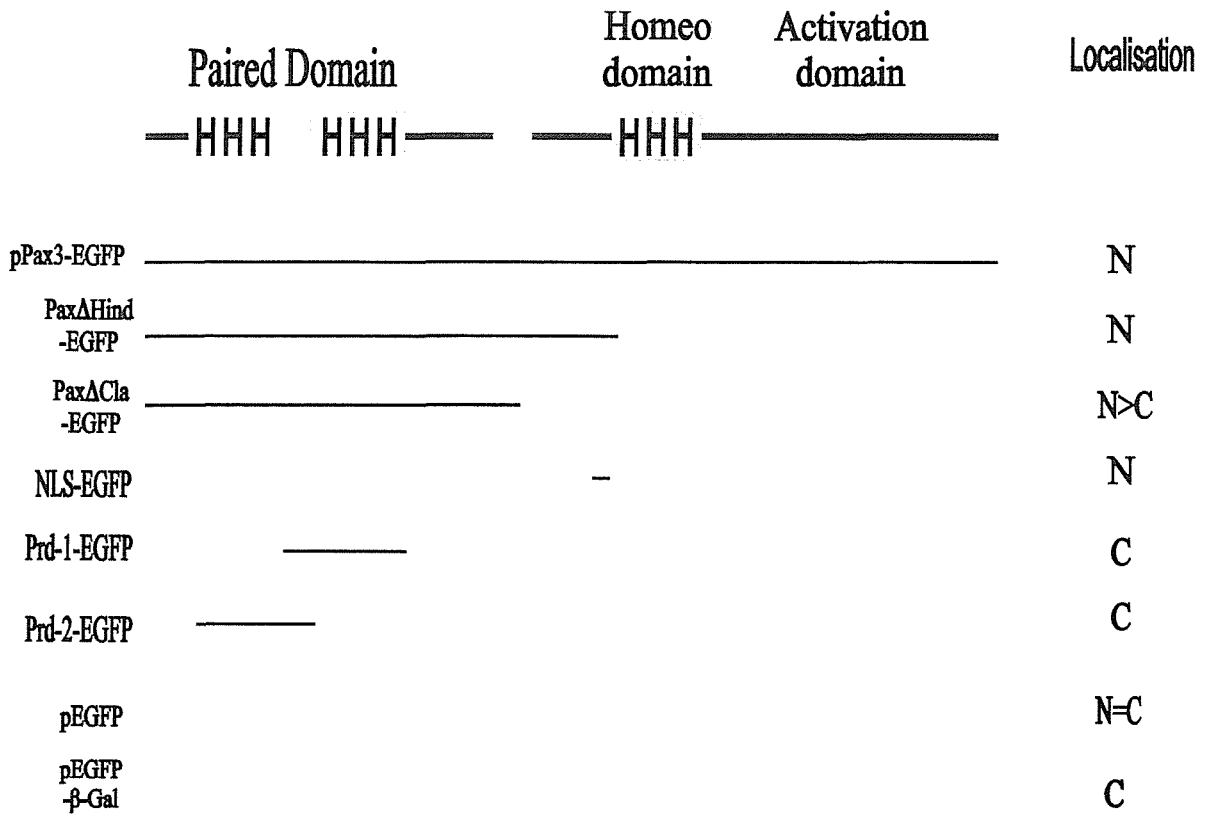


Fig. 5.17 Diagram showing the comparative localisation of the full length Pax-3 cDNA, the Hind II and Cla I truncations, the paired subdomains and EGFP alone and EGFP-β-Galactosidase fusion proteins respectively.

The presence of a nuclear localisation site or sites has been identified in Pax-8 (Poleev *et al.* 1997). However they found that the NLS remaining in the paired box region was capable of total nuclear localisation and that it could not be located to any specific sequence within the paired domain. This concurs with the results found that neither of the two subdomain of the Pax-3 paired domain were able to localise GFP to the nucleus in ND7 cells. It has been postulated that the nuclear localisation signal contained within this N-terminal region of Pax-8 is similar to that of the zinc-finger DNA binding domain of NGFI-A (Matheny *et al.* 1994). This has been found to contain a novel NLS that is dependent on the overall structure of the DNA binding domain and not on its highly basic nature. This would account for the inability to find a small polypeptide region capable of some nuclear localisation.

Serum starvation of both ND7 cell and neural crest cells show that Pax-3 localisation is quite clearly regulated by serum. This is not unusual, both c-fos and MAPK have been shown to be serum responsive in their nuclear localisation. The c-Fos nuclear localisation signal, KRRIRRIRNKMAAAKCRNRRRL is a classic bipartite signal quite unlike that of Pax-3 although its serum responsive phosphorylation site has not been found. It will be interesting to determine whether serum dependent nuclear import of Pax-3 is regulated by phosphorylation or by interacting factors that mask the Pax-3 NLS.

Discussion

6 Discussion

Pax-3 is a member of the paired box family of transcriptional regulators. These are highly conserved DNA binding proteins that share a DNA binding motif of 128 amino acids, the paired domain. They play an important role in the early development of the central and peripheral nervous systems. Pax-3 itself is expressed from E 8.5 in the neural crest, developing spinal cord and brain and in developing limb musculature. Expression is restricted to mitotically active cells in these lineages and is downregulated upon differentiation.

Loss of Pax-3 function leads to abnormalities such as limb deformities, changes in pigmentation, loss of hearing, exencephaly and spina bifida. These are due the loss of Pax-3 function in their ancestral cell lineages. Although the function of Pax-3 is as yet unclear, there have been suggestions that pax-3 may play a roles in proliferation, migration and differentiation. Analysis of Pax-3 expression in neural crest cells has shown that Pax-3 is required for maintenance of an undifferentiated phenotype. Terminal differentiation only occurs after a downregulation of Pax-3 expression. Likewise suppression of Pax-3 expression leads to terminal differentiation of the sensory neurone derived ND7 cell line.

To gain further insights into the role of Pax-3 in development we have studied the regulation of Pax-3 and the potential gene targets of this transcription factor. Examination of the Pax-3 promoter showed binding sites for two transcription factors that play an important role in the cell cycle and thus in proliferation. The Myc/Max/Mad network of transcriptional regulators act in concert to control the cell cycle and differentiation. Myc activates transcription when heterodimerised with Max and it has been shown to induce progression through the G1/S-phase boundary of the cell cycle even in the complete absence of growth factors, while Myc downregulation suppresses cell cycle progression. Experiments co-

transfecting c-Myc and N-myc with the Pax-3 promoter showed that they were both able to upregulate the Pax-3 promoter in cells. This upregulation is increased by ectopic expression of Max up to a limit after which Max was shown to have an inhibitory effect. The Myc E-box site is also recognised by Mad/Max dimers and it was shown that ectopic expression of Mad and Max significantly repressed activity of CAT linked to the Pax-3 promoter. These observations led to further experiments to identify the Myc/Max/Mad binding site in the Pax-3 promoter.

The E2F family of transcription factors also play a pivotal role in the cell cycle and cellular proliferation, it limits the transcription of genes to the phases of the cell cycle during which their products are required. It has been shown that overexpression of E2F in ND7 cells led to a large increase in Pax-3 promoter activity. However this was abolished by the co-transfection a pRB expression vector and it was shown that the Pax-3 promoter was effectively regulated by E2F and its associated cell-cycle dependent co-repressor. Once again the effective E2F binding site in Pax-3 was identified and it was shown that E2F could bind to its sequence *in vitro*.

We have shown that both of these important cell cycle transcription factors regulate the transcription of Pax-3. In addition we have located the exact sites of binding of E2F and Myc in the Pax-3 promoter and have shown that binding to these site is cell cycle dependent.

These experiments therefore strongly suggest that Pax-3 expression may be cell cycle controlled, implying a role for Pax-3 in cellular proliferation in the embryo. This is consistent with its expression pattern as Pax-3 is restricted to mitotically active progenitor cells. It also seems likely that Pax-3 is involved in regulating Myc since antisense Pax-3 can cause differentiation in ND7 cells, which have large amounts of endogenous myc – unusual since ectopic expression of myc has usually been shown to cause G1/S progression. This seems like a fruitful course on which to pursue further enquiry.

There is increasing evidence that transcription factor activity is not only controlled at the level of synthesis but also at the level of subcellular localisation. Studies using a Pax-3-EGFP (Enhanced green fluorescent protein) fusion have located a potential NLS within the protein. This is near to a Casein Kinase II (CKII) site. Phosphorylation of CKII sites near to the NLS have been shown in other proteins such as large T antigen to regulate subcellular localisation. Preliminary experiments have shown that Pax-3 import into the nucleus is serum dependent. However whether this is due to regulation by the CKII site has not been determined. This could be tested by mutating the potential CKII site and determining whether the nuclear localisation of Pax-3 is still serum dependent. The finding that Pax-3 nuclear localisation is serum dependent also correlates with the hypothesis that Pax-3 plays a role in cell proliferation. As these preliminary immunofluoresence experiments would suggest that in the absence of growth factors pax-3 is localised in the cytoplasm and is hence inactive, however upon the addition of serum/growth factors Pax-3 is imported into the nucleus. By regulating its subcellular localisation this may be a means of increasing the rate of response to growth factor stimulation.

To date very few target genes of Pax-3 have been identified. Studies on Pax-3 antisense cell lines have shown that Snap-25, a protein involved in axonal outgrowth is upregulated upon suppression of Pax-3. Transient transfection assays have shown that this inhibition of Snap-25 expression by Pax-3 is dependent on the N-terminal inhibitory domain of Pax-3 and this inhibition is lost if the 33 N-terminal amino acids are deleted. We have located the Pax-3 binding site in the Pax-3 promoter as a palindromic double homeodomain binding site. The ability of Pax-3 to inhibit Snap-25 expression raises the possibility that Pax-3 may in immature neuronal cells suppress the expression of genes required for the mature neuronal phenotype until appropriate environmental cues signal a downregulation

in Pax-3 and neuronal differentiation. Pax3 may therefore function to link the onset of cell cycle arrest with morphological differentiation of the cell.

References

7 References

Adams, B., Dorfler, P., Aguzzi, A., Kozmik, Z., Urbanek, P., Maurer-Fogy, I. And Busslinger, M. (1992), *Pax-5* encodes the transcription factor BSAP and is expressed in B lymphocytes in the developing CNS and adult testis, *Genes Dev.*, **6**, 1589-1607

Amin, C., Wagner, A. J., and Hay, N. (1993), Transcriptional activation by Myc and repression by Max, *Molecular and Cellular Biology*, **13**, 423-426

Balczarek, K. A., Lai, Z. C., and Kumar (1997), S., Evolution and functional diversification of the paired box (*Pax*) DNA binding domains, *Mol. Biol.,Evol.*, **14**, 829-842

Baldwin. C. T., Hoth, C. F., Amos, D. A., da-Silva, E. O., and Milunsky, A. (1992), An exogenic mutation in the HuP2 paired domain gene causes Waardenburg's Syndrome., *Nature*, **355**, 637-638

Balling, R., Deutsch, U., and Gruss, P. (1988), *Undulated*, a mutation affecting the development of the mouse skeleton has a point mutation in the paired box of *Pax-1*, *Cell*, **55**, 531-535

Bark I.C., and Wilson, M. C., (1994), Regulated vesicular fusion in neurons-snapping together the details., *Proc. Nat. Acad. Sci. USA*, **91**, 4621-4624

Bark, I. C., Hahn, K. M., Ryabinin, A. E., and Wilson, M. C., (1995), Differential expression of SNAP-25 protein isoforms during divergent vesicle fusion events of neural development., *Proc. Nat. Acad. Sci. USA*, **92**, 1510-1524

Baeuerle, P. A. and Baltimore, D. (1988), Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-κB transcription factor., *Cell*, **53**, 211-217

Baeurle, P. A. and Baltimore, D. (1989), A 65kD subunit of active NF-κB is required for inhibition of NF-κB by IκB., *Genes Dev.*, **3**, 1689-1698

Baumgartner, S., Bopp, D., Burri, M. and Noll, M. (1987), The structure of two genes at the *gooseberry* locus related to the *paired* gene and their spatial development during embrogenesis, *Genes Dev.*, **1**, 1247-1267

Binetruy, B., Smeal, T., and Karin, M., (1991), *Nature*, **351**, 122-127

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G., and Noll M. (1986), Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes of *Drosophila*, *Cell*, **47**, 1033-1040

Boyle, W. J., Smeal, T., Defize, L. H., Angel, P., Woodgett, J. R., Karin, M. and Hunter, T., (1991), *Cell*, **64**, 573-584

Bronner-Fraser, M., and Fraser S., (1988), Cell lineage analysis shows multipotentiality of some avian neural crest cells., *Nature*, **335**, 161-164

Bronner-Fraser M., and Fraser, S., (1989), Developmental potential of avian trunk neural crest cells in situ, *Neuron*, **3**, 755-766

Cai, J., Lan, Y., Appel, L. F. and Weir, M. (1994), Dissection of the *Drosophila* paired protein: Functional requirements for conserved motifs, *Mech. Dev.*, **47**, 139-150

Cai, J. Phelan, S. A., Hill, A. L. and Loeken, M. R. (1998), Identification of *Dep-1* a new gene regulated by the transcription factor Pax-3, as a marker for altered embryonic expression during diabetic pregnancy, *Diabetes*, **47**, 1803-1805

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

Chalepakis, G., Fritch, R., Fickenscher, H., Deutsch, U., Goulding, M. and Gruss, P. (1991), The molecular basis of *undulated/Pax-1* mutation, *Cell*, **66**, 873-884,

Chalepakis, G., Goulding, M., Read, A., Strachan, T., and Gruss, P. (1992), Molecular basis of splotch and Waardenburg's Syndrome., *Proc. Nat. Acad. Sci. USA*, **91**, 3685-3689

Chalepkakis, G., Tremblay, P. and Gruss, P. (1992), *Pax* genes, mutants and molecular function, *J. Cell Sci.*, **Suppl. 16**, 61-67

Chalepakis, G. Stoykova, A., Wijnholds, J., Tremblay, P. and Gruss, P. (1993), *Pax*: Gene regulators in the developing nervous system., *J. Neurobiology*, **24**, `1367-1384

Chalepakis, G., Jones, F. S., Edelman, G. M., and Gruss, P., (1994a), Pax-3 contains domains for transcriptional activation and transcriptional inhibition., *Proc. Nat Acad. Sci USA*, **91**, 12745-12749

Chalepakis, G., Wijnholds, J. and Gruss, P. (1994b), Pax-3 – DNA interaction: flexibility in the DNA binding and induction of DNA conformational changes by paired domains., *Nucleic Acids Research*, **22**, 3131-3137

Chalepakis, G. and Gruss, P. (1995), Identification of DNA recognition sequences for the Pax-3 paired domain, *Gene*, **162**, 267-270

Chatterjee, S., Javier, M. and Stochaj, U. (1997), *In vivo* analysis of nuclear protein traffic in mammalian cells, *Exp. Cell Res.*, **236**, 346-350

Chelsky, D., Ralph, R., Jonak, G. (1989), Sequence requirements for synthetic peptide mediated translocation to the nucleus, *Mol. Cell. Biol.*, 2487-2492

Cultraro, C. M., Bino, T., and Segal, S., (1997), Function of the c-myc antagonist Mad1 during a molecular switch from proliferation to differentiation, *Molecular and Cellular Biology*, **17**, 2353-2359

Czerny T., Schaffner, G., and Busslinger, M. (1993), DNA sequence recognition by Pax proteins – bipartite structure of the paired domain and its binding site., *Genes Dev.*, **7**, 2048-2061

Davis, R. J., DiCruz, C.M., Lovell, M. A., Biegel, J. A., and Barr, F. G. (1994), Fusion of *PAX7* to *FKHR* by the variant t(1:13)(p36:q14) translocation in alveolar rhabdomyosarcoma, *Cancer Res.*, **54**, 2869-2872

Deans, Z. C., Dawson, S. J., Kilimann, M. W., Wallace, D., Wilson, M. C., and Latchman, D. S., (1997), Differential regulation of genes encoding synaptic proteins by the Oct-2 transcription factor, *Mol. Brain Res.*, **51**, 1-7

Deutsch, U., Dressler, G. and Gruss, P. (1988), *Pax-1*, a member of a paired box homologous murine gene family, is expressed in segmented structures during development, *Cell*, **53**, 617-625

Deutsch, U. and Gruss, P. (1991), Murine paired domain proteins as regulatory factors of embryonic development., *Semin. Dev. Biol.*, **2**, 413-424

Doherty, P., Rowett, L., Moore, S., Mann, D. and Walsh, F. (1991), Neurite outgrowth in response to transfected N-CAM and N-Cadherin reveals fundamental differences in neuronal responsiveness to CAMs, *Neuron*, **6**, 247-258

Dressler, G. R., Deutsch, U., Chowdhury, K., Nornes, H. O., and Gruss, P. (1990), Pax-2, a new murine paired-box containing gene and its expression in the developing nervous system., *Development*, **109**, 787-795

Dressler, G. R., and Douglass, E. C. (1992), Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor., *Proc. Nat. Acad. Sci. USA.*, **89**, 1179-1183

Eilers, M., Picard, D., Yamamoto, K. R., and Bishop, J. M., (1989), *Nature*, **340**, 66-68

Hess, E. J., Collins, K. A., and Wilson, M. C., (1994), Deletion mutation of the *Snap* gene in mouse models attention deficit hyperactivity disorder, *Soc. Neurosci. Abst.*, **20**, 171

Epstein, D. J., Vekemans, M. and Gros, P. (1991), Splotch (Sp^{2H}), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3., *Cell*, **67**

Epstein, J., Cai, J., Glaser, T., Jepeal, L., and Maas, R. (1994), Identification of a *Pax* paired domain recognition sequence and evidence for DNA conformational change, *J. Biol. Chem.*, **269**, 8355-8361

Epstein, J. A., Lam, P., Jepeal, L., Maas, R. L., and Shapiro, D. N. (1995), Pax-3 inhibits myogenic differentiation of myoblast cells., *J. Biol. Chem.*, **270**, 11719-11722

Ludolph, D. C., and Konieczy, S. F., (1995), Transcription factor families: muscling in on the myogenic program., *FASEB J.*, **9**, 1595-1604

Epstein, J. A., Shapiro, D. N., Cheng, J., Lam, P. Y. P. and Maas, R. L., (1996) Pax-3 modulates expression of the c-met receptor during limb muscle development, *Proc. Nat Acad. Sci. USA*, **93**, 4213-4218

Erickson, C. A., and Weston, J. A. (1983), An SEM analysis of neural crest migration in the mouse., *Journal of Embryologu and Experimental Morphology*, **74**, 97-118

Frank, E., and Sanes, J. R. (1991), Lineage of neurons and glia in chick dorsal root ganglia: Analysis *in vivo* with recombinant retrovirus, *Development*, **111**, 895-908

Franz, T. (1989), Persistant truncas arteriosis in the splotch mutant mouse., *Anat. Embryol.*, **180**, 457-464

Franz, T. (1990), Defective ensheathment of motoric nerves in the *splotch* mutant mouse., *Acta. Anat.*, **138**, 246-253

Franz, T. (1993), The *Splotch* (Sp^{1H}) and *Splotch-delayed* (Sp^{d}) alleles; Differential phenotypic effects on the neural crest and limb musculature., *Anat. Embryol.*, **187**, 371-377

Franz T., Kothary, R., Surani, M. A., Halata, Z., and Grim, M., (1993), The *Splotch* mutation interferes with muscle development in the limbs., *Anat. Embryol.*, **187**, 153-160

Fraser, S. E. and Bronner-Fraser, M. E. (1991), Migrating neural crest cells in the trunk of the avian embryo are multipotent, *Development*, **112**, 913-920

Frigerio, G., Burri, M., Bopp, S., Baumgartner, S., and Noll, M. (1986), Structure of the segmentation gene *paired* and the *Drosphila* PRD gene sets a part of the gene network, *Cell*, **47**, 735-746

Galibert, M-D., Yavuzer, U., Dexter, T. J., and Goding, C. R., (1999), Pax-3 and the regulation of the Melanocyte-specific Tyrosinase related protein-1, *Journal of Biological Chemistry*, **374**, 26894-26900

Gorman, C., (1985), High efficiency gene transfer into mammalian cells, *DNA cloning : A practical approach. Cambridge Univ. Press,* **Chap. 6** , 143-180,

Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. R., and Gruss, P. (1991), Pax-3 a novel murine DNA binding protein expressed during early neurogenesis, *EMBO J.*, **10**, 1135-1147

Hess, E. J., Collins, K. A., and Wilson, M. C., (1994), Deletion mutation of the *Snap* gene in mouse models attention deficit hyperactivity disorder, *Soc. Neurosci. Abst.*, **20**, 171

Hoth, C. F., Milunsky, A., Lipsky, N., Sheffer, R., Clarre, S. K., and Baldwin, C. T. (1993), Mutations in the paired domain of the human Pax-3 gene causes Klein-Waardenburg's Syndrome (WS-111) as well as Waardenburg's Syndrome Type 1 (WS-1), *Am. J. Hum. Genet.*, **52**, 455-462

Jans, D. A. (1995), The regulation of protein transport to the nucleus by phosphorylation, *Biochem. J.*, **311**, 705-716

Jun, S. and Desplan, C. (1996), Cooperative interactions between the paired domain and homeodomain, *Development*, **122**, 2639-2650

Jostes, B., Walther, C. and Gruss, P. (1991), The murine paired-box gene, *Pax-7*, is expressed specifically during development of the nervous and muscular system., *Mech. Dev.*, **33**, 27-38

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

Karin, M. (1994), Signal transduction from the cell surface to the nucleus through the phosphorylation of transcription factors, *Current Opinion in Cell Biology*, **6**, 415-424

Lahav, R., Ziller, C., Dupin, E., and Le Douarin, N., Endothelin 3 promotes neural crest proliferation and mediates a vast increase in melanocyte number in culture, *Proc. Nat. Acad. USA,* **93**, 3892-3897

Lakin, N. D., Morris, P. J., Theil, T., Sato, T., Moroy, T., Wilson, M. C., and Latchman, D. S., (1995), Regulation of neurite outgrowth and SNAP-25 expression by Brn-3a, *J. Biol. Chem.*, **270**, 597-600

Latchman, D. S. (1995), Eukaryotic transcription factors 2nd Ed., *Academic Press Ltd.*

Le Douarin, (1973), A biological cell labeling technique and its use in experimental embryology., *Dev. Biol.*, **30**:217-222

Le Douarin, N. M., and Teillet M. A. (1974), Experimental analysis of the migration an differentiation of neuroblasts in the autonomic nervous system and neuroectodermal mesenchymal derivatives using a biological cell marking technique, *Developmental Biology*, **41**, 162-184

Le Douarin, N. M., Renaud, D., Teillet, M. A., and Le Douarin, G. (1975), Cholinergic differentiation of presumptive adrenergic neuroblasts in interspecific chimeras after heterotopic transplantations., *Proc. Nat. Acad. Sci. USA*, **72**, 728-732

Le Douarin, N. M. (1982), The Neural Crest, *Cambridge University Press*

Le Douarin, N., M., Ziller, C. and Couly, G. (1993), Patterns of neural crest derivatives in the avian embryos: *In vivo* and *in vitro* studies, *Developmental Biology*, **159**, 24-49

Lillycrop, K. A., Budhram, V. S., Lakin, N. D., Terrenghi, G., Woord, J., N., Polak, J. M., and Latchman, D. S., (1992), *Nucleic Acids Res.*, **20**, 5093-5096

Lin, A., Frost, J., Deng, T., Smeal, T., Al-Alawi, N., Kikkawa, U., Hunter, T., Brenner, D., and Karin, M. (1992), *Cell*, **70**, 777-789

Mansouri, A., Stoykova, A., and Gruss, P. (1994), *Pax* genes in development, *J. Cell Science*, **Suppl. 18**, 35-42

Mansouri, A., Hallonet, M., and Gruss, P. (1996), *Pax* genes and their roles in cell differentiation and development., *Curr. Opin. Biol.*, **8**:851-857

Maroto, M., Reslef, R., Munsterberg, A., Koester, S., Goulding, M. and Lassar, A. B. (1997), Ectopic Pax-3 activates MyoD and Myf5 expression in embryonic mesoderm and neural tissue., *Cell*, **89**, 139-148

Matheny, C., Day, M. L., and Millbrandt, J. (1994), The nuclear localisation signal of NGFI-A is located within the zinc-finger DNA binding domain, *J. Biol. Chem.*, **269**, 8176-8181

Maulbecker K., and Gruss, P. (1993), The oncogenic potential of Pax genes, *EMBO J.*, **12**, 2361-2367

Moase, C. E. and Trasler, D. G. (1992), *Splotch* locus mouse mutants: models for neural tube defects and Waardenburg Syndrome Type 1 in humans, *J. Med. Genet.*, **29**, 145-151

Morris, P. J., Lakin, N. D., Dawson, S. J., Ryabinin, A. E., Kilimann, M. W., Wilson, M. C., and Latchman, D. S., (1996), Differential regulation of genes encoding synaptic proteins by members of the Brn-3 subfamilay of POU domain transcription factors, *Mol. Brain Res.*, **43**, 279-285

McEvilly, R. J., Erkman, L., Luo, L., Sawchenko, P. E., Ryan, A. F., and Rosenfeld, M. G. (1996), Requirement for Brn 3.0 in differentiation and survival of sensory and motor neurons., *Nature*, **384**, 574-577

Murphy, M., Reid, K., Hilton, D. J., and Bartlett, P. F., (1991), Generation of sensory neurons is stimulated by Leukemia Inhibitory factor, *Proc. Nat. Acad. Sci. USA*, **88**, 3498-3501

Natoli, T. A., Ellsworth, M. K., Wu, C., Gross, K. W., and Pruitt, S. C., (1997), Positive and negative sequence elements are required to establish the pattern of Pax-3 expression., *Development*, **124**, 617-626

Neubuser, A., Koseki, H., and Balling, R., (1995), Characterisation and developmental exprssion of *Pax-9*, a paired-box containing gene related to *Pax-1*, *Dev. Biol.*, **170**, 701-716

Noll M. (1993), Evolution and role of Pax genes., *Curr. Opin. Genet. Dev.*, 595-605

Nornes, H. O., Dressler, G. R., Knapik, E. W., Deutsch, U., and Gruss, P. (1990), Spatially and temporally restricted expression of Pax-2 during murine neurogenesis, *Development,* **109**, 797-809

Nothias, J-Y., Majunder, S. Kaneko, K. J., and DePamphilis, M. L. (1995), Regulation of gene expression at the beginning of mammalian development, *J. Biol. Chem.*, **270**, 22077-22080

Peters, H., Neubuser, A., Kratochwil, K. and Balling, R. (1998), *Pax-9* deficient mice lack phrangeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities, *Genes Dev.*, **12**, 2735-2747

Phelan, S. A., Ito, M. and Loeken, M. R. (1997), Neural tube defects in embryos of diabetic mice. Role of *Pax-3* gene and apoptosis., *Diabetes*, **46**, 1189-1197

Phelan, S. A. and Loeken, M. R. (1998), Identification of a new binding motif for the paired domain of Pax-3 and unusual characteristics of spacing of bipartite

recognition elements on binding and transcriptional activation., *J. Biol. Chem.*, **273**, 19153-19159

Plachov, D., Chowdhury, K., Walther, C., Simon, D., Guenet, J. L., and Gruss, P. (1990), *Pax-8*, a murine paired box gene expressed in the developing excretory system and thyroid gland., *Development*, **110**, 643-650

Poleev, A., Fickensher, H., Mundlos, S., Winterpacht, A., Zabel, B., Fidler, A., Gruss, P. and Plachov, D. (1992), *PAX8*, a human paired box gene: isolation and expression in developing thyroid, kidney and Wilms tumor., *Development*, **116**, 611-623

Poleev, A., Okladnova, O., Musti, A-M., Schneider, S., Royer-Pokora, B. and Plachov, D. (1997), Determination of functional domains of the human transcription factor PAX-8 responsible for its nuclear localisation and transactivating potential, *Eur. J. Biochem.*, **247**, 860-869

Puri, P. L., Cimino, L., Fulco, M., Zimmerman, C., La Thangue, N. B., Giordano, A., Graessmann, A., and Levrero, M., (1998), Regulation of E2F mitogenic activity during terminal differentiation by its heterodimerisation partners for Nuclear Translocation, *Cancer Research*, **58**, 1325-1331

Rihs, H-P., Jans, D. A., Fan, H., and Peters, R. (1991), The rate of nuclear cytoplasmic transport is determined by the casein kinase II site flanking the nuclear localisation sequence of the SV40 T-antigen, *EMBO J.*, **10**, 633-639

Reeves, F. C., Fredericks, W., Rauscher, F., and Lillycrop, K. A. (1998), The DNA binding activity of the paired box transcription factor, Pax-3, is rapidly doqnregulated during neuronal differentiation., *FEBS Letts.*, **422**, 118-122

Reeves, F. C., Burdge, G. C., Fredericks, W. J., Rauscher., F. J., and Lillycrop, K . A., (1999), Induction of antisense Pax-3 expression leads to the rapid morphological differentiation of neuronal cells and an altered response to the mitogenic growth factor bFGF, *Journal of Cell Science*, **112**, 253-261

Rogers, S., Wells, R. and Rechsteiner, M. (1986), *Science*, **234**, 364-368

Rogers, S. W. and Rechsteiner, M. (1996), PEST sequences and regulation by proteolysis, *TIBS*, **21**, 267-271

Roux, P., Blanchard, J-M., Fernandez, A., Lamb, N., Jeanteur, P., and Piechaczyk, M. (1990), Nuclear Localisation of c-Fos but not v-Fos proteins, is controlled by extracellular signals, *Cell*, **63**, 341-351

Russell, W. C., Greene, L. (1947), Splotch, a new mutation in the house mouse – mus musculus., *Genetics*, **32**, 107

Ryabinin, A. E., Sato, T. N., Morris, P. J., Latchman, D. S., and Wilson, M. C., (1995), Immediate upstream promoter regions required for Neuro specific expression of SNAP-25, *J. Mol. Neurosci.*, **6,** 201-210

Sambrook, J., Fritsch, G., and Maniatis, T. (1989), Molecular cloning, A laboratory manual, *Cold Spring Harbor Press,* New York.

Selleck, M. A. J., Scherson, T. Y., and Bronner-Fraser, M. (1993), Origins of neural crest diversity, *Developmental Biology*, **159**, 1-11

Serbedzija G.N., Bronner-Fraser, M. and Fraser, S. E., (1989), A vital dye analysis the timing and pathways of avian neural crest cell migration, *Development*, **106**, 809-816

Serbedzija G.N., Fraser, S. E., and Bronner-Fraser, M. (1990), Pathways of trunk neural crest cell migration in the mouse embryo as revealed by vital dye labelling, *Development*, **108**, 605-612

Serbedzija G.N., Bronner-Fraser, M. and Fraser, S. E., (1992), Vital dye analysis of cranial neural crest cell migration in the mouse embryo, *Development*, **116**, 297-307

Serbedzija G.N., Bronner-Fraser, M. and Fraser, S. E., (1994), Developmental potential of trunk neural crest cells in the mouse, *Development*, **120**, 1709-1718

Shapiro, D. N., Sublett, J. E., Li, B., Downing, J. R., and Naeve, C. W. (1993), Fusion of *PAX3* to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma, *Cancer Res.*, **53**, 5108-5112

Sollner T., Bennett, M. K., Whiteheart S. W., Scheller R. H., and Rothman J. E., (1993) A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of the synaptic vesicle docking, activation and fusion., *Cell*, **75**, 409-418

Sommer, A., Bousset, K., Kremmer, E., Austen, M., and Luscher, B. (1998), Identification and characterization of specific DNA-binding complexes containing members of the Myc/Max/Mad network of transcriptional regulators, *Journal of Biological Chemistry*, **273**, 6632-6642

Sosa-Pineda, B., Chowdhury, K., Torres, M., Oliver, P. and Gruss, P. (1997), The Pax-4 gene is essential for differentiation of insulin producing β-cells in the mammalian pancreas., *Nature,* **386**, 399-402

Steele, K. P. and Smith, R. J. H. (1992), Normal hearing in splotch, the mouse homologue of Waardenburg's Syndrome Type 1., *Nat. Genet.*, **2**, 75-79

Stemple, D. L. and Anderson, D. J. (1993), Lineage diversification of the neural crest: *In vitro* investigations, *Developmental Biology*, **159**, 12-23

Stuart, E. T., Kioussi., C. and Gruss, P. (1993), Mammalian Pax genes, *Ann. Rev. Genet.*, **27**, 219-236

Stuart, E. T. and Gruss P., PAX genes: What's new in developmental biology and cancer?, *Hum. Mol. Gen.*, **4**, 1717-1720

Stuart, E. T. and Gruss, P. (1996), PAX : Developmental control genes in cell growth and differentiation, *Cell Growth and Diff.*, **7**, 405-412

Sudhof, T. C., (1995), The synaptic vesicle: a cascade of protein-protein interactions., *Nature* **375**, 645-653

Takashi, Y. (1996), Organisation of spina bifida neural tube in splotch (Pax-3 deficient) mouse embryos, *Dev. Growth Diff.*, **28**, 23-31

Tao, M., Wasson, J., Mizrachi, E. B., Benn, P. S., Chayen, S., Duprat, L., Meyer, J., Glaser, B., and Permutt, M. A. (1998), Pax-4 in pancreatic β-cells; isolation and characterisation of the human *PAX-4* gene., *Diabetes*, **47**, 1650-1653

Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P. and Strachan, T. (1992), Mutations in the *Pax-3* paired box gene causing Waardenburg syndrome type 1 and type 2., *Nature*, **355**, 635-636

Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P. and Strachan, T. (1992), Mutations in the *Pax-3* gene causing Waardenburg syndrome type 1 and type 2., *Nat. Genet.*, **3**, 26-30

Thiery, J. P., DuBand, J. L. and A Delouvee, (1982), Pathways and Mechansims of Avian Trunk Neural Crest Cell Migration and Localisation, *Developmental Biology*, **93**, 324-343 (1982)

Tosney (1978), The early migration of neural crest cells in the trunk region of the avian embryo : an electron microscope study, *Dev. Biol.*, **62**, 317-33

Triesman, J., Harris, E., and Desplan, C. (1991), The paired box encodes a second DNA binding domain in the paired homeodomain protein., *Genes Dev.*, **5**, 594-604

Ullman, K. S., Powers, M. A., and Forbes, D. J. (1997), Nuclear export receptors: From importin to exportin, *Cell*, **90**, 967-970

Underhill, D. A., and Gros, P. (1997), The paired domain regulates DNA binding by the homeodomain within the intact Pax-3 protein, *J. Biol. Chem.*, **272**, 14175-14182

Underhill, D. A., Vogan, K. J. and Gros, P. (1995), Analysis of the *splotch-delayed* mutant indicates that the Pax-3 Paired Domain can influence homeodomain-DNA binding activity, *Proc. Nat. Acad. Sci USA*, **92**, 3692-2696

Vandromme, M., Gauthier-Rouviere, C., Lamb, N., and Fernandez, A. (1996), Regulation of transcription factor localisation: fine tuning of gene expression, *Trends in Biol. Sci.*, **21**, 59-64

Vogan, K. J., and Gros, P. (1997), The C-terminal subdomain makes an important contribution to the DNA binding activity of the Pax-3 paired domain, *J. Biol. Chem.*, **272**, 28289-28295

Wakamatsu, Y., Watanabe, Y., Nakamura, H., and Kondoh, H.,(1997), Regulation of neural crest cell fate by N-myc: promotion of ventral migration and neuronal differentiation., *Development*, **124**, 1953-1962

Walther, C. and Gruss, P. (1991), *Pax-6*, a murine paired box gene, is expressed in the developing CNS, *Development*, **113**, 1435-1449

Walther, C., Guenet, J., Simon, D., Deutsch, U., Jostes, B., Goulding, M. D., Plachov, D., Balling, R., and Gruss, P. (1991), Pax : A murine multigene family of paired box containing genes., *Genomics*, **11**, 424-434

Watanabe A, Takeda, K., Ploplis, B. and Tachibana, M. (1998), Epistatic relationship between Waardenburg syndrome genes *MITF* and *PAX-3*, *Nat. Genet.*, **18**, 283-286

Wen, W., Meinkoth, J. L., Tsien, R. Y., and Taylor, S. S. (1995), Identification of a signal for rapid export of proteins from the nucleus, *Cell*, **82**, 463-473

Williams, B. A. and Ordahl, P. C. (1994), Pax-3 expression in segmental mesoderm marks the early stages in myogenic cell specification, *Development*, **120**, 785-796

Wilson, D. S., Guenther, B., Desplan, C. and Kuriyan, J. (1995), High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA., *Cell* **82**, 709-720

Wilson, D., Sheng, G., Lecuit, T., Dostani, N., and Desplan, C. (1996), Co-operative dimerisation of paired class homeodomains on DNA, *Genes Dev.*, **7**, 2120-2134

Wood, J. N., Bevan, S. J., Coote, P. R., Dunn, P. M., Harmar, A., Hogan, P., Latchman, D. S., Morrison, C., Ruogon, G., Theveniau, M., and Wheatly, S., (1990), Novel cell lines display properties of nocioceptive sensory neurons., *Proc. R. Soc. Lond. B*, **241**, 187-194

Xu, W., Rould, M. A., Jun, S., Desplan, C. and Pabo, C. O. (1995), Crystal structure of a paired domain-DNA complex at 2.5Å resolution reveals a structural basis for Pax- developmental mutations, *Cell*, **80**, 639-650

Yang, X-M., Vogan, K., Gros, P. and Park, M., (1996), Development, **122**, 2163-2171

Zanini, M., Francis-Lang, H., Plachov, D. and DiLauro, R. (1992), Pax-8, a paired box binding protein binds to a sequence overlappin the recognition site of a homeodomain and activates transcription from two thyroid specific promoters., *Mol. Cell Biol.*, **12**, 4230-4241.