Transcriptional targets of Pax3 during development

By

Benjamin Fenby

Thesis submitted for the degree of doctor of philosophy at the University of Edinburgh

May 2005
“...it seems we are converging on mutual indifference.”

R. Connolly, 2003
Disclaimer

I (Benjamin Fenby) performed all of the experiments presented in this thesis unless otherwise clearly stated in the text. No part of this work has been, or is being, submitted for any other degree or qualification.
Acknowledgements

Numerous people have been indispensable to the completion this work.

John Mason has tirelessly supervised me over the past 4 years. This is in spite of me foisting myself into his lab uninvited and proceeding to throw regular tantrums whilst spending quite a frightening quantity of cash. He receives top thanks and probably some kind of serious karmic upgrade for the next life I should imagine.

Karen Chapman, Lesley Forrester and Mike Clinton all kindly agreed to sit on my Thesis Committee; their advice and guidance was invaluable.

Technical support (a term which never properly reflects how impossible work like this would be without it) from Katy Gillies and Duncan McNeil must also be warmly applauded.

My contemporaries in the Price / Mason / Kind / Spears / West labs for kindness, help and above all tolerance over the years.

These data would not have been obtained without the breeding and sacrificing of mice. Whilst this was always kept to the barest minimum (partly due to my being rubbish at dissection), I acknowledge and state this was performed as humanely as possible.


T. Ellender gets a much bigger bit.
Table of Contents

DISCLAIMER 3
ACKNOWLEDGEMENTS 4
TABLE OF CONTENTS 5
ABBREVIATIONS 9
ABSTRACT 10

CHAPTER ONE: INTRODUCTION 12
Pax3 Biochemistry 13
The Pax Gene Family 14
The Pax3 gene and its isoforms 16
The Pax3 Protein 21
Pax3 Target Sequences 24
Pax3 In Neural Tube and Neural Crest Development 26
Wnt1 and Wnt1 Regulation 31
Pax3 and Cardiac Development 33
Hypothesis One 37
Hypothesis Two 38
Pax3 and Somite and Limb Muscle Development 39
Factors regulating Pax3 in the mesoderm 40
Pax3 function in mesoderm 45
Pax3 and Pax7 49
Hypothesis Three 53
Conclusion 55

CHAPTER 2: MATERIALS AND METHODS 56
Introduction 56
Organisms 56
Bacteria 56
Antibiotics 57
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammalian Cell Culture</td>
<td>58</td>
</tr>
<tr>
<td>Cell lines</td>
<td>58</td>
</tr>
<tr>
<td>Maintenance</td>
<td>58</td>
</tr>
<tr>
<td>Mouse Colony</td>
<td>58</td>
</tr>
<tr>
<td>Mouse Embryos</td>
<td>59</td>
</tr>
<tr>
<td>Genotyping of Mice</td>
<td>59</td>
</tr>
<tr>
<td>Nucleic Acids</td>
<td>60</td>
</tr>
<tr>
<td>DNA Extraction</td>
<td>60</td>
</tr>
<tr>
<td>Plasmid</td>
<td>60</td>
</tr>
<tr>
<td>PAC (P1 Artificial Chromosome)</td>
<td>60</td>
</tr>
<tr>
<td>Genomic DNA</td>
<td>61</td>
</tr>
<tr>
<td>RNA Extraction</td>
<td>62</td>
</tr>
<tr>
<td>Trizol</td>
<td>62</td>
</tr>
<tr>
<td>RNeasy Mini</td>
<td>62</td>
</tr>
<tr>
<td>RNeasy Midi</td>
<td>62</td>
</tr>
<tr>
<td>mRNA</td>
<td>63</td>
</tr>
<tr>
<td>NP-40 isolation of cytoplasmic RNA</td>
<td>63</td>
</tr>
<tr>
<td>DNA / RNA precipitation</td>
<td>64</td>
</tr>
<tr>
<td>DNA Quantification</td>
<td>64</td>
</tr>
<tr>
<td>RNA Quantification</td>
<td>65</td>
</tr>
<tr>
<td>DNA Sequencing</td>
<td>65</td>
</tr>
<tr>
<td>Agarose Electrophoresis</td>
<td>65</td>
</tr>
<tr>
<td>Polyacrylamide Electrophoresis (Denaturing)</td>
<td>65</td>
</tr>
<tr>
<td>Cloning</td>
<td>66</td>
</tr>
<tr>
<td>Restriction Digests</td>
<td>67</td>
</tr>
<tr>
<td>Phosphatasing</td>
<td>67</td>
</tr>
<tr>
<td>Kinasing</td>
<td>68</td>
</tr>
<tr>
<td>Ligations</td>
<td>68</td>
</tr>
<tr>
<td>TOPO-cloning</td>
<td>69</td>
</tr>
<tr>
<td>Heat-Shock Transformation</td>
<td>69</td>
</tr>
<tr>
<td>PCR (Polymerase Chain Reaction)</td>
<td>70</td>
</tr>
<tr>
<td>PCR Reaction Conditions</td>
<td>70</td>
</tr>
<tr>
<td>PCR Optimisation</td>
<td>71</td>
</tr>
<tr>
<td>TEPCR</td>
<td>71</td>
</tr>
<tr>
<td>RTPCR</td>
<td>72</td>
</tr>
<tr>
<td>QRTPCR</td>
<td>72</td>
</tr>
<tr>
<td>Site Directed Mutagenesis</td>
<td>77</td>
</tr>
<tr>
<td>Primer extension</td>
<td>78</td>
</tr>
<tr>
<td>5'RACE</td>
<td>79</td>
</tr>
<tr>
<td>Ribonuclease Protection Assay (RPA)</td>
<td>81</td>
</tr>
<tr>
<td>PAC Library Screen</td>
<td>82</td>
</tr>
<tr>
<td>Southern Blot</td>
<td>84</td>
</tr>
<tr>
<td>Transfections</td>
<td>85</td>
</tr>
<tr>
<td>Luciferase Assay Transfections (protocol for adherent NIH-3T3 and C2C12 cells)</td>
<td>85</td>
</tr>
<tr>
<td>Transfections for Western Blotting</td>
<td>88</td>
</tr>
<tr>
<td>Proteins</td>
<td>89</td>
</tr>
<tr>
<td>Choice of peptides used to raise α-Pax3 antibody</td>
<td>89</td>
</tr>
<tr>
<td>Protein Quantification</td>
<td>89</td>
</tr>
<tr>
<td>Western Blot</td>
<td>90</td>
</tr>
</tbody>
</table>
CHAPTER THREE: WNT1 REGULATION

Introduction

Hypothesis One

Wnt1 expression is decreased in Sp^{2h} homozygotes

Pax3 does not interact with the Wnt1 3' enhancer

The Wnt1 5' promoter is Pax3 responsive

Discussion (Hypothesis One)

Regulation of Wnt1 by Msx2

Hypothesis Two

Pax3 and Msx2: a regulatory relationship at the Wnt1 locus?

Discussion (Hypothesis Two)

CHAPTER FOUR: PAX7 REGULATION

Introduction

Pax7 regulatory regions: Bioinformatics

Pax7 regulatory regions: 5'UTR mapping

Pax7 5' Promoter: Cloning

Regulation of Pax7 by Pax3

Hypothesis Three

Discussion (Hypothesis Three)

CHAPTER FIVE: GENERATION OF A NOVEL α-PAX3 ANTIBODY

Introduction

Peptide Design and Antibody Production

Verification of α-Pax3 immunogenicity

Antibody Purification

α-Pax3 Immunohistochemistry
Discussion

CHAPTER SIX: IN VIVO CHROMATIN IMMUNOPRECIPITATION

Introduction

ChIP: an overview

Thesis findings: an overview

ChIP: technical considerations

Pax3 interacts with the Wnt1 proximal promoter in vivo

Msx2 may associate with the Wnt1 locus in vivo

Pax3 binds to the Pax7 I1R element in vivo

Conclusions

CHAPTER SEVEN: DISCUSSION

Pax3 and Wnt1
  Hypothesis One

Msx2 and Wnt1
  Hypothesis Two

Pax3 and Pax7
  Hypothesis Three

Pax3 ChIP: future directions

APPENDIX 1: CONSTRUCTS AND CLONING STRATEGIES

APPENDIX 2: PRIMER LIST

APPENDIX 3: PIPMAKER OUTPUT

APPENDIX 4: LUCIFERASE ASSAY DATA TREATMENT

BIBLIOGRAPHY
## Abbreviations

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'UTR</td>
<td>5' Untranslated Region</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>ddH2O</td>
<td>Double distilled, column purified water</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotides</td>
</tr>
<tr>
<td>LacZ</td>
<td>B-Galactosidase</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>NCAM</td>
<td>Neural Cell Adhesion Molecule</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>PAC</td>
<td>P1 Artificial Chromosome</td>
</tr>
<tr>
<td>PCIA</td>
<td>Phenol Chloroform Iso-amly Alcohol</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PEXT</td>
<td>Primer extension</td>
</tr>
<tr>
<td>QRTPCR</td>
<td>Quantified RTPCR</td>
</tr>
<tr>
<td>RA</td>
<td>Retinoic Acid</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid Amplification of cDNA Ends</td>
</tr>
<tr>
<td>RLU</td>
<td>Relative Light Units</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>RPA</td>
<td>Ribonuclease Protection Assay</td>
</tr>
<tr>
<td>rpm</td>
<td>Rotations per minute</td>
</tr>
<tr>
<td>RT</td>
<td>Room Temperature (18-20 °C)</td>
</tr>
<tr>
<td>RTPCR</td>
<td>Reverse Transcriptase PCR</td>
</tr>
<tr>
<td>SAP</td>
<td>Shrimp Alkaline Phosphatase</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>SDS Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>Sp²H</td>
<td>Splotch-null mutant allele of Pax3</td>
</tr>
<tr>
<td>TEPCR</td>
<td>Template Enriched PCR</td>
</tr>
<tr>
<td>TTLB</td>
<td>Tail Tip Lysis Buffer</td>
</tr>
</tbody>
</table>
Abstract

Transcription factors play multiple and important roles during embryonic development. The *Pax* gene family have been shown to be essential in these processes, and mutations which abrogate their function disrupt the development of a range of embryonic tissues. One member of this family, *Pax3*, is characterised by a semidominant mutant phenotype and is involved in the proper formation of the brain, central and peripheral nervous system, hypaxial musculature, and cardiac outflow vessels. *Pax3* is a transcription factor and its biological role involves the binding of specific sequences in genomic DNA to regulate the expression of a set of target genes. Defects observed in animals harbouring a mutation in *Pax3* are assumed to arise from the misregulation of the transcriptional targets of this gene.

In dissecting the precise biological function of *Pax3*, it is important to differentiate between primary and secondary phenotypes in the mutant animal; specifically, between defects caused by the direct misregulation of *Pax3* targets and those generated by the compounded knock-on effects of a dysfunctional master-regulator such as *Pax3*. A simple comparison of expression profiles between wild type and mutant animals, using microarray or differential display based methodologies, would not achieve this aim. To date, no attempt to perform a comprehensive screen for direct *Pax3* targets in development has been performed.

This thesis presents a development of the tools necessary to perform the task of identifying direct targets of *Pax3 in vivo*. Firstly, two candidate genes for direct *Pax3* regulation are considered, *Wntl* and *Pax7*. These candidates are implicated from previous work on *Pax3* although no direct transcriptional link has ever been
established. Firstly, differences in the expression of these genes between the Pax3 mutant and wild type embryos is quantitatively analysed. In the case of Pax7, little work had been performed to identify the regulatory elements controlling its expression in the mouse. Described here is a series of experiments mapping the 5' end of the transcript, the delineation of a putative promoter region and the identification and cloning of a highly conserved enhancer element within the first intron. The Wnt1 regulatory elements have been well described elsewhere. A series of experiments are then performed to confirm the interaction between these regulatory regions and Pax3 and the identification and testing of specific Pax3 binding sites is reported. To confirm these interactions in vivo, chromatin immunoprecipitation, using a novel mono-specific anti-Pax3 antibody designed for the purposes of this thesis, is used. This technique, having been verified on these two direct Pax3 targets in this way, could then be used to screen for novel direct targets of Pax3 regulation in vivo and in the wild type; expanding our understanding of the pathways and developmental function of this gene in future.
Chapter One: Introduction

The transcription factor *Pax3* is involved in a range of developmental and pathological processes, ranging from the patterning and specification of the central nervous system to the progression of cancer (Chi and Epstein, 2002). In animal models where the function of the *Pax3* gene has been ablated or modified, a number of interesting phenotypes have been observed. These include neural tube closure defects, which are similar to the condition *spina bifida* in humans, severe brain defects, such as anencephaly and exencephaly, problems in a range of tissues which are normally populated by a migratory population of stem cells originating from the embryonic spinal cord known as the neural crest (in humans these defects are categorised under the umbrella term neurocristopathies) and defects in the proper septation and morphology of the heart. In humans the condition Waardenburg’s Syndrome has been attributed to mutations in *Pax3* (Epstein, 1996). Patients suffering from this congenital disorder have a heterozygous mutation in the *Pax3* gene (only one *Pax3* homozygous mutant human has been reported at term (Zlogotara, 1995)) and exhibit characteristic pigmentation defects (including a white forelock reminiscent of the white belly spot found on *Pax3* heterozygous mice), deafness, craniofacial abnormalities, and even defects in limb musculature.

Understanding such developmental processes, and using the *Pax3* gene as a tool to probe these further, should ultimately enable the better understanding and more appropriate treatment of a range of diseases in future.
The *Pax3* protein acts as a transcription factor, binding to DNA and regulating the expression of other genes (Epstein et al., 1991). Whilst much work has been done to describe the role *Pax3* plays in different developmental processes, relatively few direct targets of *Pax3* have been established and no comprehensive screen for *Pax3* target genes has been performed. This is important since, without an understanding of the direct targets of *Pax3*, any appreciation of the specific pathways this gene modulates is difficult. Furthermore, the interpretation of data obtained from manipulations of the *Pax3* gene, whether knockouts, overexpressions, or even the use of *Pax3* as a developmental marker, will be enhanced by a comprehensive understanding of the gene’s transcriptional function.

This thesis therefore aims to identify novel targets of direct regulation by *Pax3* and defines the specific elements through which *Pax3* is acting. These targets are then confirmed *in vivo* through the optimisation of a technique which can then be used as a tool for performing a more comprehensive screen for *Pax3* target genes in future.

**Pax3 Biochemistry**

The paired DNA binding domain is found in several *Drosophila* genes, such as *Paired, gooseberry-proximal/distal* (Mansouri, 1996). These are involved in the segmentation patterning and development of the embryo. The paired domain is found throughout evolution, and represents a highly conserved 128 amino acid motif that forms a helix-turn-helix structure to contact DNA. In vertebrates, the first paired domain containing proteins identified were the *Pax* family.
The Pax Gene Family

Pax genes encode a family of DNA binding proteins. These proteins are expressed in a temporally and spatially dynamic manner in the mammalian embryo. The disruption or misregulation of these factors has been shown to have serious consequences for the process of development and ultimate function of many tissues. Pax genes are 9 in number in the mouse and are all thought to bind DNA and act as transcription factors, regulating the expression of downstream target genes. These proteins are thought to bind to specific sequences in the promoters and enhancers (cis-regulatory sequences) of their targets and regulate their expression via the recruitment of either inhibitory or activatory transcriptional complexes (for general reviews, see (Chi and Epstein, 2002; Epstein, 1996; Mansouri, 1996)). The known roles of the Pax genes in mouse development are summarised in Figure 1.1.

<table>
<thead>
<tr>
<th>Pax Gene</th>
<th>Paired, homeodomain, or octapeptide</th>
<th>Name and Phenotype of mutant</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pax1 (I)</td>
<td>Paired + octapeptide</td>
<td>Undulated (Un). Axial skeletal defects.</td>
<td>(Wallin, 1994)</td>
</tr>
<tr>
<td>Pax2 (III)</td>
<td>Paired + octapeptide + truncated homeodomain</td>
<td>k/o defects in urogenital tract, optic nerve, cochlea defects</td>
<td>(Torres, 1995)</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Phenotype/Isoform</td>
<td>Reference</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------------------------</td>
<td>-------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Pax3</td>
<td>Paired + Octapeptide + Homeodomain</td>
<td>Splotch (Sp) Open neural tube, lethal by E15 due to cardiac outflow tract phenotype</td>
<td>(Epstein et al., 1991)</td>
</tr>
<tr>
<td>Pax4</td>
<td>Paired + Homeodomain</td>
<td>k/o defects in pancreatic α-cells</td>
<td>(Habener and Stoffers, 1998)</td>
</tr>
<tr>
<td>Pax5</td>
<td>Paired + Octapeptide + Truncated Homeodomain</td>
<td>k/o defects in midbrain and ablation of B cell development</td>
<td>(Urbanek, 1994)</td>
</tr>
<tr>
<td>Pax6</td>
<td>Paired + Homeodomain</td>
<td>Small eye (Sey) Forebrain defects, microphthalmia lethal at birth</td>
<td>(Hill, 1991)</td>
</tr>
<tr>
<td>Pax7</td>
<td>Paired + Homeodomain</td>
<td>k/o defects in cephalic neural crest derivatives</td>
<td>(Mansouri et al., 1996)</td>
</tr>
<tr>
<td>Pax8</td>
<td>Paired + Octapeptide + Truncated Homeodomain</td>
<td>k/o absent follicular cells of Thyroid gland</td>
<td>(Mansouri, 1998)</td>
</tr>
<tr>
<td>Pax9</td>
<td>Paired + Homeodomain</td>
<td>k/o absent thymus,</td>
<td>(Peters,</td>
</tr>
</tbody>
</table>
Figure 1.1: The roles of Pax genes in development, reviewed in (Chi and Epstein, 2002; Eccles, 2002; Mansouri, 1996). Numbers in roman numerals denote the subfamily to which the factor belongs; i.e. Pax1 and Pax9 are within sub-family I (classification from (JacksonLabs))

The Pax genes are divided into subfamilies on the basis of their domain composition, as shown in Figure 1.1. Members of the same subfamily have highly conserved DNA binding domains, with sequence divergence occurring towards the 3' end of the gene (C-terminus of the protein). This implies similar DNA sequence recognition by members of the same sub-family and may account for functional redundancy between family members in vivo. It is important to note that Pax proteins may also have other functions, aside from their roles as transcription factors, which are not highlighted in these studies.

The Pax3 gene and its isoforms

Pax3 was first identified in 1991 in the mouse and was shown to have a dynamic expression pattern during embryogenesis. Clones were isolated from a cDNA library using the Pax1 paired box as a probe. The Pax3 gene was shown to be expressed from the onset of neurulation (E8.5) in the developing CNS (Goulding et al., 1991). This
has since been confirmed in several other studies (Bang et al., 1999; Natoli et al., 1997) and Pax3 is further expressed in the adult hair follicle (Lang, 2005). The Pax3 gene maps to mouse chromosome 1 (Epstein et al., 1991) and encodes a protein product of 479 amino acids (53 KDa). The Pax3 locus spans 95kb (ensembl) in the mouse and homologous loci have been identified in humans, Xenopus, Fugu and rat. A schematic of the Pax3 protein, highlighting the major domains, is illustrated in Figure 1.2.

The Pax3 locus encodes up to 10 exons (Barber et al., 1999) with transcripts varying in length from 900 bp (Tsukamoto et al., 1994) to 3.6 kb (Goulding et al., 1991). Six exon splice variants (named PAX3a to Pax3f, respectively) were originally reported between human and mice (Barber et al., 1999). A recent study has identified a new isoform generated through the use of a novel splice site formed at the junction of exons 7 and 8 (Pax3(Δ8)) in the mouse. This isoform removes part of the transactivation site of Pax3 (encoded by exon 8) and has been shown to result in a transcriptionally dominant negative protein in vitro (Pritchard, 2003).

The biological importance of many of these isoforms has yet to be determined experimentally. Indeed PAX3a and PAX3b, whilst thought to potentially also encode dominant negative forms of Pax3 (Tsukamoto et al., 1994), have yet to be confirmed as biologically functional at the protein level or even existant in the mouse. Similarly, Pax3e and Pax3f have not been characterised in any
Figure 1.2: Scale drawing of mouse Pax3 protein illustrating major domains and features. NT = Amino-terminus, CT = Carboxy-terminus
more detail than the sequencing of RTPCR clones (Barber et al., 1999). Pax3c was the original full length transcript reported for the mouse by (Goulding et al., 1991). Pax3d is identical to Pax3c, but has an extra C-terminal coding exon (exon 9). Because no functional difference between Pax3c and Pax3d could be detected, (Barber et al., 1999), the major Pax3c isoform has been used throughout this thesis.

Another major source of transcript variation for this gene is the inclusion or exclusion of a glutamine (Q) codon generated by an alternative splicing event at the exon 2 - exon 3 junction, see Figure 1.3E (Vogan et al., 1996). All 8 of the Pax3 transcripts denoted above also therefore have a Q± status, bringing the total number of splice variants to 16 recorded so far.

The inclusion of a glutamine residue at position 76 in the Pax3 protein was first identified by (Vogan et al., 1996). This glutamine residue is positioned between the two DNA binding sub-domains of the paired domain, see (Czerny, 1993). This places the splice variation in the C-terminal sub-domain of the Pax3 paired domain in a similar position to exon 5a in the Pax6-5a isoform (shown to be important in changing the sequence specificity of Pax6 see (Kozmik, 1997) and below). The inclusion of the Pax3 Q residue was shown in vitro to mediate changes in binding affinity of the paired domain to artificial consensus sites (the Q- isoform binding with around 5 times greater affinity to specific sequences as the Q+). In vivo the Q+ isoform is by far the most abundant of the isoforms (Vogan et al., 1996) and no dynamic change in the relative abundance of the isoforms was observed during development. Later
Figure 1.3: Genomic Pax3 locus, from www.ensembl.org.
A) Pax3 in context of mouse chromosome 1
B) expanded contig view
C) exon / intron schematic
D) protein domain structure of transcript
E) illustration of Q± alternative splicing between exon 2 and exon 3 in Pax3. Blue = exon sequence, green = splice donor / acceptor sites, red = glutamate codon, arrows = alternative splice acceptor sites
studies have revealed that, when selected for using *in vitro* Selex\(^*\) based methodologies, *Pax3Q+* and *Pax3Q-* may also have affinities for slightly different sequences (Vogan and Gros, 1997).

It is plausible that the two isoforms are generated by stochastic splicing events in the cell due to the existence of similarly high affinity splice acceptor sites. Since both isoforms do recognise identical sequences, it is possible that the difference generates such a minor change in functionality (and target sequence specificity (Vogan and Gros, 1997; Vogan et al., 1996)) that evolution has not selected for one or the other isoform and hence both are observed. This view may be supported by the fact that, to date, no differences in transcriptional targets have been observed for these isoforms *in vivo*. In this thesis the most abundant, Q+, isoform has been used for all experiments.

**The Pax3 Protein**

The *Pax3* protein has been studied extensively and its functional domains mapped in detail. The N-terminus of the protein holds the paired and homeodomain, separated by an octapeptide motif, and the transactivating domain is found towards the C-terminal end of the protein (Figure 1.2). The functional significance of the conserved octapeptide motif is unclear, in spite of its presence in most *Pax* family proteins (with the exception of Pax4/6).

---

\(^*\) Selex is an *in vitro* method for selecting oligonucleotides which bind to a protein with high affinity using several rounds of PCR amplification
One group (Ziman and Kay, 1998) has proposed that this region is of regulatory importance at the DNA level, and that the octapeptide motif is in fact a site directing methyltransferase activity to regulate the transcription of Pax3 (i.e. DNA sequence is conserved over codon use between species). In Pax5, however, the octapeptide has been shown to interact with the Groucho family of co-repressors (Chi and Epstein, 2002) and Ets co-regulators (Wheat, 1999). Similarly the ability of Pax3 to interact with the hDaxx repressor protein has been shown to be partially compromised on deletion of an N-terminally extended octapeptide domain (Hollenbach et al., 1999). The functional significance of this in vivo was not investigated by this group, however.

The binding by the N terminal domains of the protein to specific sequences in genomic DNA during development is thought to represent the core biochemical function of the Pax3 protein*. Structural studies on the related human Pax6 and Drosophila paired proteins indicate a common mechanism for DNA interaction by this domain family (Xu, 1995; Xu, 1999). Much work has been done in vitro to dissect the precise components of these domains and how they interact with one another. It is generally accepted that, in Pax3, the paired domain imparts much of the sequence specificity of the protein and indeed plays an important role in modulating the DNA binding affinity of the homeodomain and its ability to dimerise on palindromic DNA sequences (Fortin et al., 1998; Underhill, 1997). This finding is reinforced by the Splotch delayed (Sp²) mutant, a missense Pax3 mutant harbouring a

* Interestingly, the regulation of p53 by Pax3 may be post-transcriptional since differences in p53 protein and not mRNA can be observed between a Sp²H⁻/⁻ and wild type background, see Pani, L., Horal, M. and Loeken, M. R. (2002). Rescue of neural tube defects in Pax-3-deficient embryos by p53 loss of function: implications for Pax-3- dependent development and tumorigenesis. Genes Dev 16, 676-80.
glycine to alanine substitution within position 9 of the paired domain. This mutation abrogates DNA binding by both domains and confers a null phenotype *in vivo*. Therefore, correct DNA – paired domain interactions are essential for the proper selection of Pax3 targets during development.

The role of the homeodomain of Pax3 in binding DNA is also complex. (Underhill, 1997) demonstrate that the paired domain of Pax3 affects the sequence specificity of the homeodomain and the spacing of the ATTA palindromic recognition sequence for this part of the protein. Spacing between the homeodomain and paired domain recognition sequences has been shown to affect the DNA binding affinity of the *Pax3* protein (Phelan and Loeken, 1998). The DNA binding of the paired and homeodomain are also reciprocally related, and abrogation of DNA binding in one can lead to loss of DNA binding in both (Apuzzo, 2004; Apuzzo and Gros, 2002). The paired and homeodomain of Pax3 are therefore considered functionally interdependent.

The C-terminal portion of the *Pax3* protein is responsible for the transactivatory functions (Chalepakis et al., 1994) and is not thought to be directly involved in DNA binding. The postulated role for this part of the protein in mediating transactivation of target genes has been borne out by functional dissections of the *Pax3* protein *in vitro* (Chalepakis et al., 1994). Pax3 has clearly been shown to be able to up-regulate as well as down-regulate target genes, presumably depending on the transcriptional context within a particular cell type and target locus. Examples of *Pax3* having both transactivatory and transrepressive functions *in vitro* (Chalepakis et al., 1994; Lang, 2005) and *in vivo* (Kwang et al., 2002; Lang and Epstein, 2003) can be found in the
literature. The C terminal domain of Pax3 has also been shown to modulate the binding activity of the N terminal domains of the protein (Cao and Wang, 2000). This is supported by the in vivo observation that the Pax3 protein, and the Pax3-FKHR fusion oncogene product (generated by a chromosomal translocation mutation common in alveolar rhabdomyosarcoma (Lam et al., 1999) replacing the Pax3 C-terminus with the transactivation domain of one of the forkhead family of transcription factors) have the potential to transactivate different target genes in spite of their identical DNA binding protein domains (Barber et al., 2002; Begum, 2005; Du, 2005). The phenomenon of long-range structural alterations affecting the functionality of transcription factors has also been observed in a completely heterogenous context; the glucocorticoid nuclear receptors undergo similar long-range conformational changes on DNA binding (Pearce, 1998; van Tilborg, 2000).

Pax3 Target Sequences

The root functionality of a transcription factor comes from its ability to bind to cis-regulatory sequences of target genes and hence mediate either up- or down-regulation of their expression. To understand the role Pax3 plays during vertebrate development, it is therefore necessary to identify the target sequences to which it binds. Several attempts have been made using in vitro techniques to identify the consensus sequences recognised by Pax3. This was first performed by Chalepakis et al (Chalepakis and Gruss, 1995) using a PCR selection based Selex analysis of the Pax3 paired domain to generate a 13bp consensus sequence. One of the most interesting results of this early work was the observation that, with the exception of the four 5' nucleotides, the Pax3 consensus sequence obtained was extremely similar to other
binding consensus sites derived for Pax2 and Pax6. Furthermore, the sequences of previously described high affinity paired domain binding oligo sequences CD19-2A and PRS9 (from the Drosophila e5 promoter) were conserved to the 3' end of the consensus sequence. These data were interpreted as meaning that the 3' end of this recognition sequence represented a high affinity Paired domain binding site (hence the homology to both other Pax sites and high affinity oligo sequences), and the 5' nucleotides may impart the actual sequence specificity to the Pax3 protein. Recent software designed to identify transcription factor binding sites (MatInspector) cites these first four base pairs as the core binding motif for Pax3. Some published Pax3 consensus sites are summarised in Figure 1.4.

This consensus sequence is thought to represent binding of the N-terminal sub-domain of the Pax3 protein only. In the Pax6 protein, an alternative splicing event in the DNA encoding the paired domain acts a molecular "toggle" to shift binding preference from the N to C-terminal sub-domain of the Pax6 paired domain, through the inclusion of the 5a exon (Kozmik, 1997). This dramatically alters the sequence specificity of the Pax6 protein. The presence of an alternatively spliced glutamine residue between the N and C terminal sub-domains of the paired domain in Pax3 obviously represents an attractive model for a similar post-transcriptional level of regulation. Several papers have addressed this in vitro (Vogan and Gros, 1997; Vogan et al., 1996) and have obtained target sequences with a higher affinity for one or other isoforms. No absolute changes in binding specificity or in vivo target selection has yet been identified for these isoforms, however.
Figure 1.4:
Comparison of Pax3 paired domain consensus sites previously reported in the literature. Predicted core consensus sites are highlighted in red.

<table>
<thead>
<tr>
<th>Experimental Source</th>
<th>Consensus Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>in vitro</em>, Selex methodology</td>
<td>TCGTCACTHYHA</td>
<td>Chalepakis and Gruss, 1995</td>
</tr>
<tr>
<td><em>in vitro</em>, Selex methodology</td>
<td>GTCACGCTT</td>
<td>Vogan et al, 1996</td>
</tr>
<tr>
<td><em>in vitro</em>, Selex methodology</td>
<td>TCGTCACTTT-A</td>
<td>Vogan and Gross, 1997</td>
</tr>
<tr>
<td><em>in vivo</em>, c-RET enhancer</td>
<td>TGTCACACTGCC</td>
<td>Lang and Epstein, 2003</td>
</tr>
</tbody>
</table>
The above studies enable the identification of putative Pax3 DNA binding sites \textit{in vivo} via sequence analysis. As pointed out by many authors, however (Wasserman and Sandelin, 2004), the presence of a Pax3 consensus does not automatically imply a Pax3 responsive site \textit{in vivo}. Functional studies are always essential to confirm the significance of any potential Pax3 binding site. Indeed, (Phelan and Loeken, 1998), observed that many of the Pax3 responsive genes confirmed so far actually contain low affinity Pax3 consensus sites, and not the high affinity sites identified by \textit{in vitro} methodologies. This suggests that a simple sequence based approach to identifying new Pax3 targets may not be appropriate and that a more direct method for identifying Pax3 target genes \textit{in vivo} may be required.

Pax3 also interacts with other proteins \textit{in vivo}. These interactions can occur via both the paired domain (Hollenbach et al., 2002; Lang and Epstein, 2003; Wheat, 1999) and homeodomain (Hollenbach et al., 1999; Stamatakis et al., 2001) and can elicit both negative and positive transcriptional responses from the target genes in question. Thus it is important to appreciate that, whichever sequences the Pax3 protein binds to during development, the binding of co-factors and ultimately the recruitment of the RNA polymerase machinery is core to the function of this transcription factor.

\textbf{Pax3 In Neural Tube and Neural Crest Development}

The induction of neural tube from ectoderm occurs early in the development of the mouse and involves a complex array of signals. The formation of the neural tube and neural crest cells probably represents one of the earliest events in the evolution of vertebrates and therefore a high degree of conservation in the mechanisms and signals
used between species is thought to exist (Meulemans and Bronner-Fraser, 2004). After gastrulation, the ectoderm is organised into neuroectodermal and ectodermal regions through the action of the Spemann's Organiser (in amphibians) or equivalent (the node in mice and chick). These structures secrete neural inducers, such as chordin and noggin, which enable the establishment of neural cell fates from non-neural ectoderm through the inhibition of the high levels of BMP signalling (LaBonnie and Bronner-Fraser, 1998; LaBonnie and Bronner-Fraser, 1999). Once the neuroectoderm has been specified in this manner, the tissue invaginates to form the neural fold and ultimately closes completely to generate the neural tube (Figure 1.5).

It is the interface between neural and non-neural ectoderm which stimulates the generation of the neural crest population (Selleck and Bronner-Fraser, 1995). In mice, noggin mutants are still able to express many of the dorsal neural tube markers by E10.5, and some migratory neural crest cells are still observed (McMahon, 1998), implying that other signals must be involved in this process, such as Wnt and FGF (Monsoro-Burq, 2003).

Intermediate levels of BMP signals from the ectoderm are however essential in the early stages of this process, and help to induce the expression of Msxl at the neural plate boundary (LaBonnie and Bronner-Fraser, 1999; Liem, 1997; Monsoro-Burq, 2004). Msxl and a Wnt8 signal from the underlying non-axial mesoderm (Bang, 1997; Bang et al., 1999) then induce Pax3 expression in the lateral edges of the neural plate. This is then restricted dorsally by Shh signals from the notochord and floor plate (Goulding, 1993) as neural tube closure proceeds, generating a Pax3 expression pattern restricted to the dorsal half of the developing neural tube. Another important gene, Zic1, is expressed in the dorsal neural tube after induction by BMP4 and BMP7.
Figure 1.5: Schematic of neural tube closure and neural crest induction highlighting the expression domains of the major genes outlined in the text. (Source image taken from Bronner-Fraser, 2002)

Dorsal

**NEURAL PLATE**

- BMP
- Msx1
- Pax3
- Shh
- Migrating NC (Pax3, FoxD3, etc)

**NEURAL FOLDS**

- BMP
- Wnt8

**NEURAL CREST**

- NT
- SOM
- No
from adjacent ectoderm and is restricted to dorsal neural tube by Shh in a manner analogous to Pax3 (Aruga, 2002). The activities of these genes converge to stimulate the expression of neural crest markers such as Slug, Snail and Foxd3 which are important in establishing the neural crest lineage (Dottori et al., 2001; Kos, 2001; LaBonne and Bronner-Fraser, 1999; Mennerich and Braun, 2001). One interesting study has proposed that a major function of Pax3 is the inhibition of p53 mediated apoptosis (Pani et al., 2002). This is based on the rescue of neural tube defects on a Sp2H -/- p53 -/- double mutant background due to reduced cell death. It is important to consider the role of Pax3 as a cell survival signal therefore, in addition to the regulation of developmental target genes.

The precise role of Pax3 in the process of switching on the genes specifying the neural crest lineage is still unclear. Whilst Pax3 is genetically upstream of Foxd3 (Dottori et al., 2001), direct interaction between Pax3 and the regulatory elements of Foxd3 has not been shown. Furthermore, the induction of Slug by the combined activities of Zic1 and Pax3 (Monsoro-Burq, 2004) have been shown to be mediated by Wnt signalling. These observations are in keeping with those made on Pax3 chimeras (Mansouri et al., 2001) that, whilst Pax3 null cells cannot be rescued by wild type counterparts in somites, a non-cell autonomous role exists for the gene in the development of the neural crest. If Pax3 requires a non-cell autonomous signal to enable the instruction of downstream target genes involved in neural crest development, suitable candidate factors should be considered.

Wnts are attractive signals to facilitate this role of Pax3. In Xenopus, the induction of neural crest cell fates in the dorsal neural tube has been shown to be dependent upon
$Wnt1$ and $Wnt3a$, synergising with neuralising factors such as $noggin$ (Saint-Jeannet et al., 1997). This activity could not be replaced with $Wnt5a$ or $Wnt8$, implying that the Wnt signal responsible for inducing $Pax3$ expression and that involved in neural crest specification are different (Bang et al., 1999; Saint-Jeannet et al., 1997). Similarly in Zebrafish, Wnt signalling has been shown to be essential during neural crest induction from underlying mesoderm (Lewis, 2003) (in *Xenopus* neural crest (Bang et al., 1999) this mesodermal signal has been shown to be $Wnt8$) but this signal is distinct from a second Wnt activity, however, which specifies a neural crest lineage. Both the "two signal" model of neural crest induction (where the induction of neuroectoderm by $chordin$ and $noggin$ requires a second signal to express neural crest specific markers (LaBonne and Bronner-Fraser, 1998)) and the "double gradient" model (where default anterior neural tissue is patterned combined lateralising and posteriorising gradients (Villanueva, 2002)) in *Xenopus* depend upon canonical Wnt signalling. In chick, inhibition of Wnt signalling similarly perturbs the expression of neural crest marker genes (Garcia-Castro, 2002).

Other Wnt genes that are expressed in this tissue at this stage are $Wnt1$ and $Wnt3a$. Compound knockouts of these genes (Ikeya et al., 1997) show major defects in their neural crest populations and neural crest derived structures but not in the overall dorsoventral patterning of the neural tube, suggesting these factors act after the establishment of patterning within the neural tube itself. Further experiments investigating the expression and function of Wnt receptors in Xenopus neural crest, the frizzled protein family, have revealed that the $Wnt1$ specific receptor $Frz3$ is necessary for the induction of neural crest in the developing *Xenopus* embryo in association with the Kermit co-receptor (Deardorff, 2001; Tan, 2001).
The expression of *Wntl* and *Wnt3a* in the *Pax3* mutant has been investigated, and whilst the expression of *Wnt3a* did not change, *in situ* data for *Wntl* did imply a significant downregulation in the neural tube of *Sp2H* homozygotes (Conway et al., 2000). *Wntl* therefore represents an attractive candidate gene for direct regulation by *Pax3* in this system.

**Wntl and Wnt1 Regulation**

Much work has been done on *Wntl* and the regulation of this gene, and will be briefly dealt with here (for a comprehensive review of Wnt biology in this and other systems see (Logan, 2004; Miller, 2001)). *Wntl* was the first of the *Wnt* genes to be discovered, and was originally identified as the site of a mouse mammary tumour viral insertion into the 5' and 3' regulatory regions of the gene (Van Ooyen, 1984). Later sequence alignments revealed that the gene was the mammalian homologue of the *Drosophila* Wingless (*Wg*) gene, and the ability of *Wntl* to transform mammary cells when inappropriately expressed led to its categorisation as a proto-oncogene and suggested a biological role in proliferation (Dickinson et al., 1994; Nusse et al., 1990). Knockout strategies (McMahon and Bradley, 1990; Thomas, 1990) revealed an essential role for *Wntl* in the development of midbrain and cerebellar structures. Neither of these studies illustrated defects in neural tube development in spite of the strong dynamic expression of this gene. Similarly, *Wnt3a* knockout mice (Takada, 1994) displayed defects in somitogenesis but not in the neural tube or neural crest derivatives. A redundancy in Wnt signalling in the neural tube was therefore
postulated and \textit{Wnt1} / \textit{Wnt3a} double mutant embryos exhibit a severe neural crest phenotype (Ikeya et al., 1997).

If \textit{Wnt1} represents a direct target of \textit{Pax3} regulation in the developing embryo, an elucidation of the \textit{cis}-regulatory elements driving the expression of this gene is required. Due to the insertion of the provirus into the region 5' proximal to the gene causing inappropriate \textit{Wnt1} regulation (Nusse et al., 1990; Van Ooyen, 1984), initial analysis of the \textit{Xenopus Wnt1} 5' proximal promoter was performed to delineate a functional promoter with luciferase reporter genes (Gao et al., 1994). Attempts in mouse to generate LacZ transgenics which recapitulate the expression pattern of \textit{Wnt1} using 5' proximal regions of the \textit{Wnt1} genomic locus failed to show a repeatable or faithful pattern of expression (J. Mason, unpublished data and (Echelard et al., 1994)). In the mouse, a 5.5kb 3' distal element has been shown to direct the expression of \textit{Wnt1} in the developing CNS (Echelard et al., 1994) which was later remarkably reduced to a highly conserved 110bp element sufficient for driving expression of a reporter in the midbrain and even rescuing the knockout phenotype in this region of the \textit{Wnt1} mutant embryo (Rowitch et al., 1998). This 110 bp element does not drive expression in the developing neural tube, however. The 3' distal enhancer is generally considered the region of major regulatory importance controlling the expression of \textit{Wnt1}.

It should be noted that some studies (Lagutin et al., 2003; St-Arnaud and Moir, 1993) have demonstrated the importance of the \textit{Wnt1} promoter in regulating expression of the gene in forebrain and RA differentiated P19 cells, respectively. This region of the
Wnt1 locus should not, therefore, be excluded a priori from an investigation into the regulation of Wnt1 expression.

**Pax3 and Cardiac Development**

One of the most interesting phenotypes observed in the Pax3 mutant mouse is the presence of a defect in cardiac development, see (Epstein, 1996) and (Creazzo, 1998) for a general overview. The cardiac defect in the Splotch mouse was first observed in 1954 (Auerbach, 1954) and is characterised by failure in the outflow vessels of the heart to properly septate and position over the chambers of the heart. This typically leads to anatomical conditions of the Persisting Truncus Arteriosus (PTA) family or, in less severe examples, Double Outlet Right Ventricle (DORV) (Creazzo, 1998). These defects are identical to those observed in avian models of neural crest ablation, and specifically in those where the ablation is limited to regions of cranial neural crest between mid-otic placode and somite 3 (Van den Hoff, 2000). Similarly, fate mapping neural crest cells in the mouse embryo reveals an important role for neural crest derivatives in the development of the outflow tracts of the heart (Jiang, 2000). The Splotch alleles have been shown to be mutations of the Pax3 locus (Epstein et al., 1991) and the nature of the mid-gestational lethality in Sp2H homozygous mice was convincingly shown to correlate exactly with the incidence of congenic heart malformation (Conway et al., 1997b) and not with the neural tube related phenotypes demonstrated by these mutants. The immunolocalisation of the slug protein (a marker of neural crest, downstream of Pax3 (Monsoro-Burq, 2004)) in the outflow tracts of the avian heart further supports the hypothesis that neural crest are important in the
proper development of these tissues since slug represents a downstream marker of Pax3 function in avians (Carmona, 2000).

Much work has been done to characterise the nature of this defect. Transgenic rescue of the cardiac malformations in Pax3 mutants using Pax3 driven by a promoter specific to the neural tube and not the somites (another major region of Pax3 expression) (Li et al., 1999) has shown that the cardiac defects observed in these animals is due to the expression of Pax3 in ectodermally derived tissues (i.e., the neural tube and neural crest). More recent work has precisely labelled the source of cardiac neural crest in the mouse to originate from the neural tube opposite somite two, and for migration to occur from the seven somite stage (~E9.5). The route of migration of these cells is through pharyngeal arches 3, 4 and 6 and into the developing heart (Chan et al., 2004). In the Sp2H mutant, the number of these cells migrating to the heart and ultimately present in the developing outflow tract (and other neural crest populated structures such as the dorsal root ganglia) is defective (Conway et al., 1997b). Contrasting studies have been performed to assess whether the lack of cardiac neural crest cells reaching the outflow tract of the heart is due to defects in the migratory competence of these cells (Conway et al., 1997a; Epstein et al., 2000) or their proliferative ability (Conway et al., 2000). Reciprocal grafts of cardiac neural crest cells between wild type and Sp2H/- tissues (Chan et al., 2004) have clearly demonstrated that a mutation in Pax3 affects both the migrating cells themselves and their ability to navigate in their environment. This finding is in keeping with earlier work (Mansouri et al., 2001) which found that, whilst Pax3 function seemed to be cell autonomous in the neural tube and somites, its role in neural crest was not. Candidate molecules mediating this defect in cell – cell
communication in \textit{Pax3} mutants include integrins (Bajanca and Thorsteinsdottir, 2002), NCAM (Glogarova and Buckiova, 2004), and versican (Henderson et al., 1997). Given the number of genes the overexpression of \textit{Pax3} in a cell line can alter (almost 300 (Mayanil et al., 2001)), and the range of phenotypes exhibited by \textit{Sp}^{2H} \(-/-\) mice, a complexity of cell autonomy phenotype of this nature is not necessarily surprising.

Could one of the non-cell autonomous features of this defect in \textit{Sp}^{2H} mice be due to a defect in \textit{Wnt1} signalling due to the absence of \textit{Pax3}? The establishment of \textit{Wnt1} as a direct target of \textit{Pax3} in \textit{vivo} would clearly be a significant step towards the resolution of this issue. Interestingly, \textit{in vitro}, the expression of \textit{Pax3} in Saos cells has been shown to activate the c-Jun-N-terminal kinase mediated non-canonical Wnt signalling cascade (Wiggan and Hamel, 2002), implying that the activation of Wnt signalling in response to \textit{Pax3} need not be mediated by the canonical $\beta$-catenin pathway alone.

A further line of evidence suggesting a role for Wnt signalling in the development of the outflow tract of the heart has also recently been reported (Clevers, 2002). In these studies in the mouse, a mutant component of the Wnt signalling pathway, \textit{Dishevelled 2} (\textit{Dvl2}), was shown to have an identical phenotype to the \textit{Sp}^{2H} homozygote cardiac outflow tract (Hamblet, 2002) with an similar penetrance of $\sim 50 - 60\%$ to that previously reported for outflow tract defects in \textit{Sp}^{2H} mice (Conway et al., 1997b). These mice also exhibited neural tube and somite defects, again reminiscent of the \textit{Pax3} mutation. Further analysis of this pathway revealed that the transcription factor \textit{Pitx2} (Kioussi, 2002) is a target of the \textit{Wnt/Dvl2} pathway in cardiac malformations, and that this transcription factor mediates its response through the regulation of the
G1 growth control genes (Hee Baek, 2003). Two interesting results are reported in the findings of this group. Firstly, \( \text{Pax3} \) expression appears to be unperturbed in \( \text{Pitx2} \) knockouts, suggesting the cardiac phenotype is not due to an indirect effect of \( \text{Pax3} \) misregulation in these experiments and that \( \text{Pax3} \) expression is genetically upstream of these defects. Secondly, \( \text{Wnt1} \) and \( \text{Pitx2} \) expression co-localise in migrating neural crest cells in the 3\(^{rd}\) and 4\(^{th}\) pharyngeal arches (through which the cardiac neural crest migrate (Chan et al., 2004)) (Kioussi, 2002). This expression domain mirrors the route of the cardiac neural crest in the wild type (Chan et al., 2004).

The regulation of \( \text{Wnt1} \) by \( \text{Pax3} \) in this process could therefore link the cardiac phenotypes of the \( \text{Pax3} \) and \( \text{Dvl2} / \text{Pitx2} \) mutants to a common pathway, and indicate a specific Wnt molecule involved in the developmental processes examined in these papers.

Finally, whilst the \( \text{Wnt1} \) knockout mice themselves do not show any reported signs of cardiac malformations and are largely carried to term, which would argue against the presence of a severe cardiac defect (McMahon and Bradley, 1990; Thomas, 1990), the possible redundancy between \( \text{Wnt1} \) and \( \text{Wnt3a} \) in the development of neural crest may explain this lack of cardiac phenotype in the \( \text{Wnt1} \) knockout animals. Where both \( \text{Wnt1} \) and \( \text{Wnt3a} \) are knocked out (Ikeya et al., 1997), no examination of cardiac outflow tract was reported and these embryos did not survive to term. It is also pertinent to point out that in this study the expression domain of \( \text{Pax3} \) in the hindbrain of \( \text{Wnt1} \) homozygote mutants appears normal, again arguing its position as genetically upstream. In the compound \( \text{Wnt1/Wnt3a} \) mutant it was reduced, however. Whether this reduction in the double mutant reflects the presence of \( \text{Pax3} \) in a downstream
genetic context to Wnt1 or Wnt3a, or whether this is due to the observed reduction in cell proliferation in these mutants was not determined. One group have also reported cardiac defects in mouse embryos injected with Wnt1 antisense RNA constructs (Augustine et al., 1993).

Subsequently, the following hypothesis was investigated in this thesis:

**Hypothesis One**

*Pax3 up regulates Wnt1 transcription directly, probably through the 3' distal enhancer region, in vivo during the development of the neural crest.*

Since the beginning of this project, one major paper concerning the development of the cardiac outflow tract in the mouse has been published (Kwang et al., 2002). In this paper a central role for the gene *Msx2* in the development of an outflow tract defect in mice is evidenced. The authors demonstrate that *Pax3* acts as a negative regulator of *Msx2* expression in the wild type. In the *Sp2H* homozygote an increased level and ectopic expression domain of *Msx2* is reported; the breeding of the *Sp2H* allele onto an *Msx2* -/- background to generate a compound mutant rescues the heart defect of the *Sp2H* mutant animal. Interestingly, other neural crest derived structures (dorsal root ganglia, thyroid and thymus) defective in *Sp2H* homozygotes were not rescued. Similarly, non-neural crest defects observed in *Pax3* mutants (neural tube closure, exencephaly) remained in the compound embryos. The authors conclude that *Msx2* is the principal downstream effector of *Pax3* in the development of the murine outflow tract.
To establish whether \textit{Wnt1} plays a role in the development of the outflow tract it was reasoned that, if \textit{Pax3} can be shown to regulate \textit{Wnt1}, the ectopic expression of \textit{Msx2} in the \textit{Sp}^{2H} mutant would be hypothesised to have a negative effect on the expression of \textit{Wnt1}. In support of this hypothesis, analysis of \textit{Wnt1} expression (as determined by expression of a \textit{Wnt1}-LacZ reporter gene) on a \textit{Sp}^{2H} homozygous background illustrates a loss of expression in a population of cells migrating from the neural tube, presumably the neural crest (although it should be noted that no sections are presented in this work at an appropriate level to represent cardiac neural crest) (Serbedzija and McMahon, 1997). Similarly in \textit{Sp}^{2H} \textit{-/-} embryos, \textit{in situ} hybridization for \textit{Msx2} transcript illustrates an expansion of \textit{Msx2} expression in the region of the embryo from which the cardiac neural crest would normally originate (Kwang et al., 2002). If a negative interaction between \textit{Msx2} and \textit{Wnt1} could be established, this would provide specific evidence for a role for \textit{Wnt1} in cardiac neural crest development. \textit{Msx2} and \textit{Msx1}/\textit{Msx2} compound mutants do not display any cardiac defects (Kwang et al., 2002; Satokata, 2000). No experiments examining the expression of \textit{Wnt1} in these mutants was reported.

The following hypothesis was therefore also investigated.

\textbf{Hypothesis Two}

\textit{Wnt1} transcription is directly downregulated by \textit{Msx2}, most probably by the distal 3' enhancer element, in vivo with implications for the normal development of the cardiac neural crest.
Hypothesis One and Hypothesis Two are addressed in Chapters 3 and 6 of this thesis.

**Pax3 and Somite and Limb Muscle Development**

Another major area of *Pax3* expression in the developing embryo is in the tissues that contribute to and populate the developing limb musculature. The area of tissue adjacent to the newly formed neural tube, the presomitic paraxial mesoderm (PSM) is induced to bud off into somites during vertebrate development, and this is one of the most easily observed developmental processes. The PSM buds off sequentially to form individual somites in a rostrocaudal wave of development, (see (Buckingham, 2001; Buckingham et al., 2003; Christ and Brand-Saberi, 2002) for a general overview of these processes), with more rostral somites developing first. Somites are staged numerically as they form. For example, a stage IV somite would have three somites caudal before un-segmented mesoderm, a stage XX would have nineteen and so forth. This enables both the staging of somites between different ages and allows comparative analysis of somite development between species.

*Sp*2H homozygous mutants have severe defects in their hypaxial musculature (Franz, 1993) and the *Pax3* gene has been shown to be expressed in both PSM, developing somites and in the developing limb mesenchyme (Buckingham et al., 2003; Franz, 1993). Clearly many important targets for *Pax3* regulation may exist in these mesodermal tissues.
An overview of somitogenesis is illustrated in Figure 1.6. The PSM buds off to generate the epithelial somite which expands and then, in response to signals from the notochord (Goulding et al., 1994), differentiates into the sclerotome ventrally (characterised by the expression of Pax1 and Pax9 and fated to become skeletal and cartilaginous tissues) and dermomyotome dorsally. The dermomyotome is patterned in response to factors released from the neural tube and ectoderm (Amthor, 1999). As the name would suggest, the dermomyotome constitutes the future cells of the dermis and musculature, and can be separated into the epaxial (medial) and hypaxial (lateral) components. As these structures progress further throughout their development, the myotome differentiates from the dermomyotome to lie between this tissue and the sclerotome, and ultimately generates epaxial musculature (Denetclaw, 1997). At the level of the limb buds, the ventrolateral edge of the dermomyotome generates migrating muscle precursors (which express Pax3) which then progress into the limb bud before differentiating into muscle fibres (Tremblay, 1998). It has been suggested that the paraxial mesoderm is an entirely naive tissue, and the dorsoventral and mediolateral patterning events which ultimately generate the sclerotome, myotome and dermomyotome are thought to be generated by multiple, often redundant signals arising from the neural tube, notochord, overlying ectoderm (Dietrich, 1997) and adjacent lateral plate mesoderm (Pourquie, 1995).

**Factors regulating Pax3 in the mesoderm**

Pax3 expression is initially found throughout the PSM (Goulding et al., 1994). Once the epithelial somite is formed, this expression is restricted dorsally by signals emanating from the notochord (Goulding et al., 1994), probably Shh in a manner
Figure 1.6

A. Diagram showing the development of the neural tube and somites. Key structures labeled:
- Dorsal root ganglion
- Epaxial dermomyotome
- Intercalated epaxial dermomyotome
- Neural tube
- Notochord
- Sclerotome
- Myotome
- Hypaxial dermomyotome
- Lateral mesoderm
- Presomitic paraxial mesoderm
- Epithelial somite
- Surface ectoderm

B. Images showing the differentiation of somites at different levels:
- Forelimb level
- Interlimb level
Figure 1.6:
A) Schematic representation of somitogenesis. Somites are formed and develop in a rostrocaudal gradient on both sides of the neural tube (Image taken from Buckingham, 2001)
B) Pax3 whole mount immunohistochemistry on chick embryo. Illustrating migration of muscle precursor cells into the forelimb from the ventrolateral edge of the dermomyotome (white arrow), and the high levels of Pax3 expression observed at the ventrolateral and dorsomedial edges of the somite at interlimb levels. Dorsal is up on both panels (Image taken from Venters et al, 2004)
analogous to that in the neural tube (Fan, 1997). The initial induction of Pax3 is likely to be mediated by BMP signals, specifically BMP4 from surface ectoderm and this has been shown to be important in the maintenance of Pax3 in proliferating populations of muscle precursor cells (Amthor, 1999). Sustained Pax3 expression requires constant exposure to dorsalising signals throughout somitogenesis. Later expression and maintenance of Pax3 in the somites is also likely to be Wnt dependent, since the expression patterns reported for BMP4 in the chick embryo would not enable prolonged expression of Pax3 (Dietrich, 1997). Wnt signals have been shown in vitro to induce Pax3 expression in culture (Fan, 1997), and the Wnt’s 1, 3, 3a, 4 and 6 are expressed in the appropriate patterns to induce Pax3 expression in the somite from dorsal neural tube (Wnt1, Wnt3, Wnt3a, Wnt4) and surface ectoderm (Wnt4, Wnt6) (Parr, 1993). Ectopic Wnt1 expression has been shown to up-regulate Pax3 and other dorsal somite markers (Capdevilla, 1998; Maroto et al., 1997). Furthermore, the downstream components of canonical Wnt signalling (Frz1, Lef and β-catenin) are also present in the presomitic mesoderm and are up-regulated in the dorsal somite where Pax3 is expressed (Schimdt, 2000).

Contrasting evidence has shown that neither expression of ectopically active β-catenin nor LiCl treatment (which activates the β-catenin pathway) can actually induce the expression of myogenic genes within the somite, however (Chen, 2004). Similarly, the expression of Lef1 and β-catenin in the somites is restricted to the myotomal compartment as development proceeds, and this expression is dependent upon a Shh signal from the notochord (Schimdt, 2000). Given that Pax3 is not expressed in the myotome, only in the dermomyotome, and that Shh signals are repressive to Pax3 expression in these tissues (Goulding et al., 1994), the maintenance
of this expression by Wnt signals from surrounding tissues seems unlikely to be operating via the canonical signalling pathway alone. A non-canonical Wnt responsive pathway, utilising adenylate cyclase, PKA, and CREB has been shown to both co-localise to Pax3 expressing dermomyotome and induce Pax3 expression \textit{in vitro} (Chen, 2004). The specific Wnt signal involved \textit{in vivo} in this system has yet to be identified, although both Wnt1 and Wnt7a were implicated in this study. The lack of somite or dermomyotomal defects reported in the Wnt1/Wnt3a compound mutant mice, in spite of their survival to E18.5, suggests that these Wnt’s may not be significant \textit{in vivo} inducers of Pax3 in the mesoderm (Ikeya et al., 1997). (Oddly, Wnt3a mutants, however, do demonstrate a defect in somitogenesis, however (Takada, 1994)). The closer proximity of Wnt7a in the dorsal ectoderm to the Pax3 positive dermomyotome would present this factor as a much more attractive Pax3 inductive signal (Parr, 1993). It is likely that multiple, possibly redundant, signals converge on this process of somatic differentiation, however.

As development proceeds, the expression pattern of Pax3 changes, as illustrated in figure 1.6B. The later expression of Pax3 in the dermomyotome exhibits a strong expression in the dorsomedial and ventrolateral domains, with weaker expression in the intermediate regions (Venters et al., 2004). The cells expressing Pax3 at the ventrolateral lip of the dermomyotome migrate into the developing limb bud (Figure 1.6B, white arrow), and are essential for the later expression of myogenic differentiation markers (i.e. MyoD) (Venters et al., 2004; Williams, 1994). Furthermore, the expression of Pax3 in the developing limb bud has been shown to originate from this migratory population, and is not due to \textit{de novo} Pax3 expression in limb mesenchyme by Quail – Chick Chimeras (Williams, 1994).
**Pax3 function in mesoderm**

The roles of *Pax3* in the developing somite and limb bud have been studied by several groups. Measurement of cell proliferation through the incorporation of BrdU into *Pax3* positive cells, as observed by (Amthor, 1998) led to the implication of *Pax3* as a regulator of cell cycle control and the proliferation of the dermomyotome population. Implicit with this is the converse regulation of *MyoD*, a basic helix-loop-helix transcription factor known to signal terminal differentiation of the myogenic lineage in the myotomal compartment (Amthor, 1998). Observations of somite development in the *Sp^{2H}* and *Sp^{d}* compound mutant (a *Sp^{2H} / Sp^{d}* genotype was used to enable the embryos to develop further than the cardiac defect in *Sp^{2H}* homozygotes would normally allow) are in agreement with this role (Tremblay, 1998). Mutant dermomyotome fails to elongate and organize properly at both the medial and lateral edges implying a defective proliferation or organisation of the cells in this region. This is consistent with the strong bands of *Pax3* expression recorded at these positions in the wild type (Venters et al., 2004). Similarly, the sclerotomal compartments (denoted by *paraxis* expression) and the myotomal compartments (denoted by *MyoD* and *Myf5* expression) are truncated in the *Sp^{2H}* homozygous mutant (Henderson et al., 1999), indicating a general function in overall somite proliferation for *Pax3*.

Another aspect of the *Splotch* phenotype is attributed to the inability of *Pax3* mutant cells from the dermomyotome to undertake the long-range migration required to populate the limbs at bud level. This can clearly be observed in contrasting *Pax3*-LacZ reporter gene activity at this position between wild type and *Sp^{2H}* homozygotes.
(Buckingham et al., 2003). A study of muscle precursor migration in the $Sp^{2H}$ homozygote reveals that the lack of muscle differentiation genes expressed in the developing limb bud is due to a migratory defect, as assessed by Dil labelling, and not incompetence for dermomyotome precursors to adopt myogenic cell fates since $Sp^{2H}$-/- somites transplanted into wild type limb buds are competent to express myogenin (Daston, 1996).

This migratory defect of dermomyotome derived hypaxial precursors in a Pax3 null background has been attributed to the loss of expression of the c-met transcript in these cells. The c-met gene encodes a tyrosine kinase receptor which is necessary for the migration of these precursors in the mouse (Bladt, 1995). No expression of this gene can be detected in the $Sp^{2H}$ homozygote embryo (Daston, 1996) (or detected at massively reduced levels, see (Mennerich, 1998)), and Pax3 has been shown to both bind and transactivate the human c-met promoter in culture via a defined consensus site (Epstein et al., 1996). In support of this interaction, the c-met ligand, scatter factor or hepatocyte growth factor (HGF), is expressed in the limb bud under control of signals from both the apical ectodermal ridge and zone of polarising activity and controls the motility and positioning of myogenic precursor cells (and Pax3 expressing cells in particular) in this tissue (Scaal, 1999). One group have additionally demonstrated expression patterns of specific integrin dimers corresponding to the expression domain of Pax3 in migrating muscle precursors (Bajanca and Thorsteinsdottir, 2002). This may provide these cells with the extracellular proteins necessary to complete their migration into the developing limb bud. The further observation by this group that the pattern of integrin expression changes on the surface of these cells after the induction of Myf5, another marker of terminal
myogenic differentiation, may also indicate a mechanism for changes in cell – cell contacts to reflect the morphological changes adopted by developing muscle fibres (see (Buckingham, 2001) for a review of this later process).

*Pax3* is also involved in other regulatory cascades in the developing musculature. The *Six* family of homeodomain proteins are involved in myogenesis, and mutations in the *Six1* gene generates a muscle hypoplasia phenotype which particularly affects the hypaxial muscles (Laclef et al., 2003). Dominant negative *Pax3* expression results in the inhibition of *Six1* and its regulatory co-factor *Eya2* (Ridgeway and Skerjanc, 2001), implying that *Pax3* lies functionally upstream of these factors. In contrast to these findings a recent study generating a *Six1/Six4* compound mutant mouse with an aggravated muscular phenotype has reported *Pax3* and *c-Met* expression to be downstream of *Six* gene regulation (Grifone, 2005). Direct regulatory interactions between *Pax3* and *Six* homeoproteins are not established in these studies and this contradictory evidence has yet to be resolved.

*Lbx1* is another important target of *Pax3* regulation and mice harbouring mutations in this gene also demonstrate defects in their musculature, particularly in the more distal hypaxial components (Brohmann, 2000; Gross, 2000). These mice demonstrate normal *Pax3* and *c-met* expression in the dermomyotome (Brohmann, 2000). *Sp2H* homozygous mice demonstrate no expression of *Lbx1* transcripts and a *Pax3* signal is at least partially required for the expression of this gene in the ventrolateral tips of the dermomyotome (Mennerich, 1998). These studies implicate *Pax3* as a regulator of *Lbx1* in the developing somite.
The expression of Pax3 in migrating muscle cells is essential for the later expression of myogenic determination genes, and the cells which delaminate from the dermomyotome, under the control of factors such as Pax3 and Lbx1, are thought to be committed to a myogenic fate despite not expressing myogenic markers such as MyoD or Mrf5 (Bober et al., 1994). Pax3 is essential for the expression of MyoD, since Sp2H mutants cannot express this gene in hypaxial muscles, and double Myf5/Pax3 mutants exhibit a total absence of body musculature (Tajbakhsh et al., 1997). Furthermore, ectopic expression of Pax3 alone in undifferentiated mesoderm is sufficient to cause the expression of MyF5, MyoD and myogenin markers (Maroto et al., 1997). This contrasts with the observations in the somite and migrating muscle precursors that the expression domains of Pax3 and MyoD do not overlap. The observation that in the wild type, Pax3 expression may determine myogenically fated but undifferentiated cells is of interest. The expression of Msxl in the ventrolateral tip of the dermomyotome and Pax3 positive migrating muscle precursor cells may explain this phenomenon (Bendall et al., 1999). Msxl expression in these cells was shown to maintain this population in an undifferentiated state by antagonising the transcriptional activation of MyoD by Pax3. Even more remarkably, the ectopic expression of Msxl in terminally differentiated myotubes can completely repress the expression of MyoD, Mrf4 and myogenin and enable clonal populations to be re-differentiated into non-myogenic cell types (Odelberg, 2000). One caveat to the above is the observed ability of somites obtained from a Pax3 null (Sp2H) background transplanted into a wild type chick limb bud to acquire a myogenic fate, as determined by myogenin expression (Daston, 1996). One possible explanation for this is that a second signalling pathway is able to switch on myogenin expression in this context,
such as one controlled by Mrf4, which may act at least partially independently from the Pax3/Myf5 specification of muscle cell types (Kassar-Duchossoy, 2004).

**Pax3 and Pax7**

In the developing mouse, another Pax gene, Pax7, is expressed in a highly dynamic and specific manner in the neural tube and developing musculature (Jostes et al., 1990). Pax7 is of particular interest in muscle development due to its importance in the regulation of adult muscle regeneration, which is thought to be partially due to a reprisal of embryonic developmental pathways (Zhao and Hoffman, 2004). In regenerating muscle, a group of muscle derived stem cells can be isolated, and are induced to activate Myf5 and MyoD expression in response to appropriate signals (i.e. damage) through the activation of Pax7 expression generating so-called satellite cells (Seale et al., 2004). The role of Pax7 in these adult populations has interesting parallels with Pax3 in the developing musculature. Pax7 knock-out animals display a total absence of satellite cells, but do contain a stem cell population within muscle tissues (Seale et al., 2000). In these mutants this population is competent to differentiate into a significantly greater range of cell types than those isolated from wild type mice, indicating that Pax7, like Pax3 in the migrating muscle precursors of the embryonic limb, acts to commit these cells to a myogenic lineage. Interestingly, one group have also noted that differing allelic forms of the Pax7 homeodomain correspond to differing efficiencies of muscle regeneration, implicating this portion of the protein in the regulation of the myogenic pathway (Kay et al., 1995)
An interesting aspect of this phenomenon with respect to embryonic development is that, in spite of Pax7 being involved in the regulation of myogenic targets in regenerating adult muscle, Pax7 is incompetent to replace the myogenic function of Pax3 in utero in transgenic mice (Relaix et al., 2004). Since Pax3 and Pax7 form a Pax gene subfamily, and have almost identical protein sequences across their DNA binding domains, see Figure 1.7, this difference in ability may reflect the transactivatory potential of their C-termini, rather than a difference in transcriptional targets. In keeping with this hypothesis the activation of the c-met receptor, essential in migrating muscle precursor cells and a target of Pax3 induction in the wild type (Bladt, 1995; Epstein et al., 1996), is described as "inefficient" rather than absent in these mice (Relaix et al., 2004). Pax7 knockout mice also do not demonstrate significant defects in musculature at birth, although a growth retarded post natal phenotype is demonstrated, leading to almost all (97%) of these homozygotes dying within 3 weeks after birth (Mansouri et al., 1996), possibly due to defective muscle regeneration and growth.

Pax3 has been implicated in the regulation of Pax7 through the observation of expression patterns in Sp^{2H} homozygotes (Borycki et al., 1999). This study shows a clear expansion in the expression pattern of Pax7 in the neural tube of Pax3 mutants. Normally, Pax7 is expressed in dorsal neural tube, but omitted from the roof plate (Jostes et al., 1990). In Sp^{2H} mice, the expression level is higher than in the wild type and includes the roof plate. Similarly, in the somites, the expression of Pax7 seems to extend from dorsal regions to the ventrolateral lip of the dermomyotome. This would imply a negative regulation of Pax7 by Pax3. In contrast to this, (Relaix et al., 2004) show that Pax7 does not appear to be significantly over expressed in either the neural
Figure 1.7:
Multiple alignment of the mouse Pax3 and Pax7 protein sequences (N to C-terminal). Note high level of conservation between regions corresponding to the Paired (red line, Pax3 residues 34:162) and homeodomains (purple line, Pax3 residues 228:278). Sequences diverge towards the C-terminus of the proteins (transactivatory domains); Pax7 exhibits an extended C-terminal domain containing residue sequences not found in the Pax3 protein.
tube or somites in the **Pax3** mutant mouse. These investigators attribute this difference to either genetic background or a possible truncated form of **Pax3** expressed from the **Sp2** allele interfering with **Pax7** expression. This latter explanation seems unlikely since it would imply a wild type interaction between the genes, and no truncated **Pax3** protein product has yet been reported as expressed from the **Sp2** allele. Finally, induction of P19 EC cells to a myogenic cell fate using DMSO, known to induce **Pax3** expression in this line, also induces **Pax7** (Ziman et al., 2001). Whether this effect is dependent on prior **Pax3** expression has not been established. Two very recent studies have indicated that a population of **Pax3 / Pax7** positive cells provide an important pool of skeletal muscle progenitor cells in both the embryo and the adult, and represent the common origin for satellite cells in mature muscle (Gros, 2005; Relaix, 2005). Such a cell population presents an attractive target for potential therapeutic intervention in muscular diseases and, given the requirement of **Pax3** and **Pax7** co-expression for the proliferation and expansion of this population (Relaix, 2005), any regulatory interaction between these genes may be of interest in this context.

Since no direct regulatory relationship between **Pax3** and **Pax7** has yet been explored in the literature, this thesis will also examine the possibility that **Pax7** represents a direct target of **Pax3** transcriptional regulation.

**Hypothesis Three**
Pax7 is directly downregulated by Pax3, in vitro and in vivo, via defined regulatory elements and with implications for the development of both neural and mesodermal tissues.

Hypothesis Three is addressed in Chapters 4 and 6 of this thesis.

In contrast with Wnt1, where well-defined cis-regulatory elements have been described in the mouse, at the start of this study no investigation into equivalent elements controlling transcription at the Pax7 locus has been described*. To date four studies have been performed analysing either the transcriptional start site of the human Pax7 gene (therefore providing an obvious location for a putative minimal promoter) (Schafer et al., 1994; Vorobyov et al., 1997) or the region of genomic sequence 5' proximal to this start site for cis-regulatory activity (Murmann et al., 2000; Syagailo et al., 2002). These investigators identified transcripts of varying length at the 5' end, ranging from 664bp (Syagailo et al., 2002) to 60bp (Murmann et al., 2000). This thesis will initially establish a 5' UTR for the mouse Pax7 transcript, and use this to examine the proximal region of genomic sequence for basal promoter activity. A bioinformatic approach has also been used to try and identify putative functional elements for the mouse Pax7 gene. The regulatory sequences identified were then assayed for Pax3 responsiveness and binding both in vitro and in vivo.

* Since the start of this investigation, one paper describing these elements in the mouse has been reported. These data will be considered later in detail, and related to the findings reported here. See Lang, D., Brown, C. B., Milewski, R., Jiang, Y. Q., Lu, M. M. and Epstein, J. A. (2003). Distinct enhancers regulate neural expression of Pax7. Genomics 82, 553-60.
Conclusion

The above introduction presents an overview of two major developmental pathways within which \( \text{Pax3} \) is known to play an important biological role, namely the development of neural crest and somites. The identification of two putative targets for direct \( \text{Pax3} \) regulation, and the resulting hypotheses directing much of the experimental work in the following chapters, are also included. This is not intended as an exhaustive review of all \( \text{Pax3} \) functions or previously reported target genes. Another major omission is an increasingly large body of work describing the function of \( \text{Pax3} \) (and \( \text{Pax7} \)) in the development of alveolar rhabdomyosarcomas due to a common chromosomal translocation event fusing the DNA binding domains of these factors to the transactivatory domain of the \( FKHR \) gene (see (Barr, 2001) for review). Since this reflects a pathological state only, and some groups describe a difference in the sets of target genes regulated by \( \text{Pax3} \) and \( \text{Pax3-FKHR} \) (Barber et al., 2002; Begum, 2005), this fusion protein has not been considered further in this thesis, except where directly relevant.

In addition to the hypotheses described above, the aims of this investigation were to develop a potential methodology for determining new, direct, targets of \( \text{Pax3} \) regulation \textit{in vivo}. The target genes described above were verified and then used as positive controls to verify the application of this methodology \textit{in vivo}. The determination of direct \( \text{Pax3} \) targets \textit{de novo} enables an enhanced understanding of this transcription factor’s biological role. It also helps to clarify some of the existing unresolved issues surrounding \( \text{Pax3} \) function that have been generated by the conflicting evidence presented in the past.
Chapter 2: Materials and Methods

Introduction

This chapter will describe the rationale and methodologies behind the techniques used in the following results chapters. The following is intended as a reference to enable both an appreciation of the precise methods used and a starting point for any investigator wishing to reproduce or extend the findings outlined in this thesis. General molecular biology protocols were taken from (Sambrook and Russell, 2005) and adapted where necessary. Where more specialised protocols have been utilised (i.e. QRTPCR) a discussion of the experimental rationale and data treatment has also been included. All protocols referred to in this thesis are described below, unless stated otherwise in the text. Manufacturers instructions have been referred to where appropriate and should be consulted for detailed instructions where necessary. COSHH forms and risk assessments for these methods can be obtained by contacting the University of Edinburgh. Sigma was the supplier of basic reagents and chemicals, unless otherwise stated. Supplier catalogues should be consulted to obtain information for ordering.

Organisms

Bacteria

Escherichia coli (E. coli) were used for all steps involving bacteria. Sterile conditions were used throughout, excess cultures were disposed of in 5% Virkon, and glassware washed in 5% Virkon after use. All liquid cultures were grown in 2XTY, shaken at
220 rpm for 16 hours at 37 °C, and inoculated directly from LBA plates (with the exception of Maxi-preps where a 5ml starter culture was inoculated and incubated for 8 hours before being added to the 250ml 2XTY).

<table>
<thead>
<tr>
<th>2XTY</th>
<th>1.6% Tryptone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1% Yeast Extract</td>
</tr>
<tr>
<td></td>
<td>86.2mM NaCl</td>
</tr>
</tbody>
</table>

The following volumes were used for different size cultures throughout the project, depending on the quantity of plasmid DNA required:

<table>
<thead>
<tr>
<th>Culture Size</th>
<th>Volume 2XTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mini-prep</td>
<td>1.5ml</td>
</tr>
<tr>
<td>Midi-prep</td>
<td>50ml</td>
</tr>
<tr>
<td>Maxi-prep</td>
<td>250ml (5ml starter culture)</td>
</tr>
</tbody>
</table>

*E. Coli* were plated out on LB Agar (LBA) plates. LB was made using LB tablets (Sigma) and Agar added to 1%. This was then autoclaved to sterilise and 10cm plates poured (using appropriate antibiotic) before cooling.

**Antibiotics**

The antibiotics used in this thesis were Ampicillin and Kanamycin as appropriate. Ampicillin was used at a final concentration of 1000 µg / ml. Kanamycin was used at a final concentration of 500 µg / ml.
Mammalian Cell Culture

Cell lines

Two cell lines were used in this thesis, and were ordered from ECACC. These were C2C12 (91031101) and NIH-3T3 (93061524). They were both maintained as described below.

Maintenance

All mammalian cell lines were maintained in sterile conditions, according to Class II procedures. Waste cultures and glassware were washed in 5% Virkon after use. Cells were grown at 37 °C and 5% CO₂, in a humidified incubator. All cell lines were grown to ~70% confluence (monitored by microscopy and estimated) before being routinely passaged:

1) Wash cells in 1X PBS (Phosphate Buffered Saline, Invitrogen)
2) Add 5ml 1X Trypsin-EDTA (Invitrogen)
3) Incubate at 37 °C for 5 minutes
4) Remove cells from incubator and ensure they have detached from flask
5) Add 5ml full growth medium
6) Take 1ml cell suspension and add to 30ml fresh growth medium in a new T75 Tissue Culture flask (Grenier)
7) Replace in incubator

<table>
<thead>
<tr>
<th>Growth Medium (Invitrogen)</th>
<th>440ml DMEM-N12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1X glycine</td>
</tr>
<tr>
<td></td>
<td>1X Penicillin / Streptomycin</td>
</tr>
<tr>
<td></td>
<td>10% Foetal Calf Serum (FCS)</td>
</tr>
</tbody>
</table>
Mouse Colony

CBA x C57BL/6F1 mice were maintained as according to Home Office regulations, on site, in a dedicated facility. Several breeding pairs were held over the course of the project, constituting one male heterozygote and one female wild type. The *Splotch* mutation is characterised by a white belly spot in the heterozygote. Only offspring carrying a white belly spot were kept, after we had confirmed by PCR genotyping strict concurrence of belly spot phenotype with a splotch heterozygous genotype.

Mouse Embryos

Mouse embryos were obtained by mating male and female mice, and observing vaginal plugs. Observation of a plug was taken as Embryonic Day 0.5 (E0.5), and litters were dated accordingly. Litters were collected on the day specified in the experimental protocol by anesthetising the mother, performing cervical dislocation and dissecting the embryos. Further embryo dissection was performed as described, using a dissecting microscope.

Genotyping of Mice

Genomic DNA was extracted from tails (adult) or amniotic sac (embryos) for PCR genotyping as described below.
**Nucleic Acids**

**DNA Extraction**

**Plasmid**

Plasmid DNA was extracted from bacteria using Qiagen mini, midi, or maxi-prep kits at the volumes described. DNA was then restricted to confirm the nature of the plasmid and quantify yield before any further experiments were undertaken.

**PAC (P1 Artificial Chromosome)**

Bacteria carrying PAC constructs were grown in 2XTY using 50% kanamycin, and incubated at 37 °C for 16 hours. PAC DNA was extracted as follows:

1) Spin cultures down at 6,000 rpm, 15 minutes, 4 °C  
2) Re-suspend in 7.5ml solution PAC1  
3) Add 7.5ml solution PAC2  
4) Invert gently 4-6 times  
5) Incubate at RT for 5 minutes  
6) Add 7.5ml solution PAC3, agitating gently  
7) Invert 4-6 times  
8) Incubate on ice for 5 minutes  
9) Spin at 10,000 rpm for 10 minutes at 4 °C  
10) Place tubes on ice, transfer the supernatant to a fresh tube by filtering through gauze  
11) Add 20ml isopropanol  
12) Incubate on ice for 5 minutes  
13) Spin at 10,000 rpm, 15 minutes at 4 °C  
14) Wash twice in 70% ethanol, spinning for 5 minutes to recover  
15) Air dry pellet  
16) Re-suspend in 500μl ddH₂O
Solutions:

<table>
<thead>
<tr>
<th>PAC1</th>
<th>15mM Tris-HCl (pH 8.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10mM EDTA</td>
</tr>
<tr>
<td></td>
<td>100µg/ml RNaseA</td>
</tr>
<tr>
<td>PAC2</td>
<td>0.2M NaOH</td>
</tr>
<tr>
<td></td>
<td>1% (w/v) SDS</td>
</tr>
<tr>
<td>PAC3</td>
<td>3M K.OAc (pH 5.5)</td>
</tr>
</tbody>
</table>

Genomic DNA

Genomic DNA was extracted from mouse tail tips (adults) or amniotic sac (embryos) for genotyping as follows:

1. Add proteinase K (100mg/ml) to TTLB
2. Add 500μl of this to each sample
3. Incubate at 55 °C for 16 hours
4. Vortex and spin for 15 minutes at 13,000 rpm
5. Pour off supernatant into tubes containing 500μl isopropanol
6. Mix gently and spin at 13,000 rpm for 5 minutes
7. Discard supernatant
8. Dry at 37 °C for 30 minutes
9. Add 250μl ddH2O, shake for 30 minutes at 55 °C
10. Add 250μl PCIA, vortex and incubate at RT for 5 minutes
11. Spin 13,000 rpm for 5 minutes
12. Remove upper layer into new tubes
13. Precipitate DNA as described
14. Re-suspend pellet in 100µl ddH2O

<table>
<thead>
<tr>
<th>TTLB</th>
<th>1% SDS (w/v)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10mM Tris-Base, pH7.4</td>
</tr>
<tr>
<td></td>
<td>1mM EDTA</td>
</tr>
</tbody>
</table>
PCR genotyping was then performed using primers P3GENOL/R, 1.5mM MgCl₂, on program 60°Cx30 (see PCR section for a description of this nomenclature). Products of 127bp (wild type allele) and 95bp (Sₚ²H allele) were resolved using a 2.5% agarose gel and ethidium bromide staining. Heterozygotes were identified by a doublet.

**RNA Extraction**

RNA was extracted from both cell cultures (70-80% confluence) or embryos (E9.5 – E12.5, various) using several different techniques. These are outlined below.

**Trizol**

Trizol (GibcoBRL) was used to extract total RNA from both cells and embryos according to manufacturer’s instructions. This is a Phenol-Chloroform based protocol.

**RNeasy Mini**

RNeasy Mini kits (Qiagen) were used to extract total RNA from cell cultures according to manufacturer’s instructions. This is a column based protocol for small quantities of starting tissue.
RNeasy Midi kits (Qiagen) were used to extract total RNA from embryos, according to their weight and manufacturer's instructions. This is a column based protocol for larger quantities of starting tissue (i.e. whole embryo RNA).

mRNA

A Micro-FastTrack 2.0 Kit (Invitrogen) was used to isolate mRNA specifically from both cell cultures and embryos, according to manufacturer’s instructions. This is a column based protocol, using oligo (dT) cellulose to specifically isolate mRNA from tissues.

NP-40 isolation of cytoplasmic RNA

To isolate cytoplasmic RNA, excluding hnRNA, NP-40 Lysis Buffer protocol was used on cell cultures as follows (based on one 70-80% confluent T75 flask of cells):

1) Trypsinise cells (as above)
2) Spin at 1,000 rpm for 5 minutes at RT
3) Wash in 1X PBS
4) Spin at 1,000 rpm for 5 minutes at RT
5) Re-suspend pellet in 400μl NP-40 Lysis Buffer
6) Transfer to an Eppendorf and spin at 13,000 rpm for 2 minutes at 4°C
7) Transfer supernatant to a new Eppendorf, discarding the pellet of cell nuclei
8) Add 400μl phenol, 50μl 10% (w/v) SDS
9) Vortex
10) Incubate at RT for 5 minutes
11) Spin at 13,000 rpm for 5 minutes
12) Transfer aqueous phase to a new tube
13) Precipitate RNA
14) Re-suspend in 50μl ddH2O
DNA / RNA precipitation

DNA and RNA were precipitated using ammonium acetate. The following general protocol was used for a given volume $n$ of nucleic acid:

1. Add $0.4 \times n \ 5\text{M NH}_4\text{OAc, pH5.5}$
2. Add $2.5 \times n \ 100\%$ ethanol
3. Vortex
4. Spin at 13,000 rpm for 30 minutes at $4\ ^\circ C$
5. Discard Supernatant
6. Wash in 75\% ethanol
7. Spin at 13,000 rpm for 10 minutes at $4\ ^\circ C$
8. Air dry pellet
9. Re-suspend in appropriate volume of ddH$_2$O

DNA Quantification

Plasmid DNA was quantified using agarose gel electrophoresis and comparing to a known quantity of ladder DNA. Ethidium bromide and an UV transilluminator, in conjunction with imaging software, were used. DNA would be linearised by restriction digest and the band intensity, quantified by the software, compared to that of a band on a DNA ladder of known concentration. Every six months the accuracy of the equipment would be tested using serial dilutions of ladder and linearised plasmid of a concentration specified by a supplier to ensure operating range and accuracy.
RNA Quantification

RNA was quantified using an UV spectrometer by measuring the absorbance at $\lambda = 260$ nm.

DNA Sequencing

DNA was sequenced using the MWG Sequencing Facility (http://www.mwg-biotech.com). Briefly, plasmids were precipitated as described (1µg per sequencing reaction) and posted off for sequencing. Primers used in the sequencing reactions were either supplied by MWG-Biotech or defined as required to extend sequencing runs, and synthesised by MWG-Biotech for sequencing on-site.

Agarose Electrophoresis

DNA and RNA were routinely separated using agarose electrophoresis. Low melting point agarose was dissolved by boiling in 1X TBE. Ethidium Bromide (at 10µg/ml) was added to 1µl/ml TBE-Agarose. Gels were left to set (~30 minutes) at RT before being covered in 1X TBE, samples loaded, and electrophoresed at 45mA/cm for varying amounts of time.

Polyacrylamide Electrophoresis (Denaturing)
DNA and RNA were routinely separated using denaturing polyacrylamide electrophoresis. 6% polyacrylamide gels containing 8M Urea were made to as follows:

<table>
<thead>
<tr>
<th>Acrylamide Gel Mix</th>
<th>1X TBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>15% Acrylamide / Bis-acrylamide</td>
<td></td>
</tr>
<tr>
<td>(40% mix)</td>
<td></td>
</tr>
<tr>
<td>8M Urea</td>
<td></td>
</tr>
</tbody>
</table>

20cm glass plates designed to fit in a BioRad Protean II electrophoresis system were cleaned, and the outer plate inner surface coated in RepelcoteVS (BDH). The plates were then assembled using 1mm spacers and a comb of appropriate size. A 10% (w/v) ammonium persulphate solution was then made, and 1ml added to the acrylamide gel mix to remove oxygen from solution (this is inhibitory to the polymerisation reaction). 100 μl of TEMED (N,N,N',N'-Tetramethyl-1,2-diaminomethane) was then added to initiate the radical based polymerisation. The gel mix was then poured between the plates, with care to avoid bubbles, and allowed to set. Gels were run in 1X TBE at 450 V for 2.5 hours, fixed in a solution of 5% acetic acid and 15% methanol for 15 minutes, and dried under vacuum overnight prior to exposure to X-ray film.

Cloning

Throughout the thesis several cloning strategies have been employed to manipulate DNA for various experimental purposes. The following is a general description of the techniques involved in every case. Specific cloning strategies can be found in the
results chapters pertaining to specific constructs. A full list of all DNA constructs made in this thesis and their cloning strategies can be found in Appendix 1.

**Restriction Digests**

Restriction enzymes were typically obtained from New England Biolabs (NEB), Promega or Fermentas, dependent on availability. Digests were conducted at 37 °C (unless the enzyme was active at a different temperature) for one to two hours. A general protocol for a single digest was as follows:

<table>
<thead>
<tr>
<th>Digest of 20 µl</th>
<th>2 µl 10X Reaction Buffer 2 µl 10X BSA 1 µl Enzyme (10 units) n µl DNA (required µg quantity) (20 - (5 + n)) µl ddH₂O</th>
</tr>
</thead>
</table>

Reaction buffer at 10X was supplied by the manufacturer, 100X BSA was supplied by the manufacturer and made to 10X fresh for each digest with ddH₂O. Whilst digest volumes and quantities of DNA were altered as required, all digests were scaled according to the above and the final concentration of total enzyme never exceeded 10% of the reaction volume.

**Phosphatasing to remove end phosphate groups**

Shrimp Alkaline Phosphatase (SAP, Roche) was used to phosphatase nucleic acids as according to the manufacturer's instructions. Phosphatasing was used either to
prevent open vectors from self-ligating if that were a possibility, or to remove end phosphates from a DNA ladder to enable kinase radiolabelling in Primer Extension or RPA protocols.

**Kinasing**

T4 Polynucleotide Kinase (NEB) was used to add phosphate groups to nucleic acids according to manufacturer's instructions. For radiolabelling, $^{32}$Pγ-ATP (Amersham) was added to the reaction mixture (quantities described for each protocol, below).

**Ligations**

Ligations were performed using T4 Ligase (Promega). T4 Ligase buffer was separated into 10 µl aliquots and a fresh aliquot used for each ligation reaction. Ligations were generally set up as follows:

<table>
<thead>
<tr>
<th>Ligation Reaction 20µl</th>
<th>2 µl 10X T4 Reaction Buffer</th>
<th>1 µl T4 Ligase (10 units)</th>
<th>x µl Insert (required µg)</th>
<th>y µl Vector (required µg)</th>
<th>(20 - (3 + x + y)) µl ddH$_2$O</th>
</tr>
</thead>
</table>

Quantities and ratios of vector and insert were determined empirically and were often a range within a given set of ligation experiments. Ligations were conducted at 16 °C for 16 hours before being transformed into competent *E. coli* as described and plated out on LBA with the appropriate antibiotic selection.
TOPO-cloning

A major method of creating plasmid clones of PCR products used in this thesis utilised a topoisomerase method rather than conventional ligase based cloning. These protocols used a range of vectors from Invitrogen where a topoisomerase enzyme is cross linked to an open vector designed to facilitate the rapid and efficient recombination of different PCR products before further cloning steps. Often these vectors were intermediates to enable more efficient downstream cloning steps or to enable rapid sequencing screening before further manipulation. The vectors used were as follows:

<table>
<thead>
<tr>
<th>PCR Product Type</th>
<th>TOPO vector used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taq amplified, 3' A overhang</td>
<td>pCR-II-TOPO-TA</td>
</tr>
<tr>
<td>Pfu amplified, no overhang</td>
<td>pCR-II-TOPO-Blunt</td>
</tr>
<tr>
<td>Large product (&gt; 2 kb)</td>
<td>pCR-II-XL-TOPO</td>
</tr>
</tbody>
</table>

For detailed protocols, see manufacturer’s instructions.

Heat-Shock Transformation

Plasmids and ligations were transformed into super competent \((10^9\) colonies \(\mu g^{-1}\) plasmid DNA). Top-10 \(E.\ coli\) (Invitrogen) using the heat shock protocol (at 42 °C) specified in the manufacturer’s instructions. Bacteria were then plated out on LBA with the appropriate antibiotic and incubated overnight at 37 °C.
PCR (Polymerase Chain Reaction)

PCR was used widely throughout this thesis. The following describes the general conditions used for every reaction, methods of optimisation typically used where this was required, followed by the notation used to define experimental conditions in each specific reaction. Primers were designed using MacVector 7.2 software. The names of primers used in each reaction are given in the results chapters; a table of all primers used and their sequences can be found in Appendix 2.

PCR Reaction Conditions

All PCR reactions used the following basic set of conditions, reagents were generally supplied by Promega and primers synthesised to order by MWG-Biotech.

| PCR Reaction Mix                  | 2.5 µl 10X Reaction Buffer  
x µl 25mM MgCl2  
1.25 µl 5’ Primer (10µM)  
1.25 µl 3’ Primer (10µM)  
1 µl dNTP mix (Promega)  
y µl DNA Template (required µg)  
0.5 µl Taq Polymerase (5 units)  
(25 - (x + y)) µl ddH2O |
|----------------------------------|--------------------------------------------------|
| PCR Program                      | Step 1 : 1 Minute at 95 °C  
Step 2 : 1 Minute at 95 °C  
Step 3 : 30 Seconds at Tm  
Step 4 : 1 Minute at 72 °C  
Step 5 : Go To Step 2 n Times  
Step 6 : 10 Minutes at 72 °C  
Step 7 : 4 °C Forever            |
The variations to this protocol specific to each reaction were $T_m$ (annealing temperature for Primer and template) and the number of cycles (n). These are annotated throughout as $T_m \times n$ (i.e. 55$^\circ$C to 35 cycles of amplification). All reactions were optimised to have identical reaction conditions otherwise.

**PCR Optimisation**

PCR reactions were optimised using the above basic reaction program and by varying the concentration of MgCl$_2$ between 0.5mM and 2.5mM and $T_m$ spread over the mean average predicted by the software used to design the primers. The $T_m$, cycle number, and [MgCl$_2$] used in each reaction are specified where relevant, although it is the experience of this investigator that PCR reaction conditions are generally indiosyncratic to the specific equipment, reagents and source material used. Independent optimisation of conditions should be anticipated in every case.

**TEPCR**

Template Enriched PCR (TEPCR) was used to clone from DNA where conventional, restriction digest mediated cloning was difficult. This protocol uses high concentrations of template and primers, combined with low PCR cycle number, to maximise product concentration in the final reaction mix. It was reasoned a lower cycle number would minimise the number of mutations generated by the polymerase
during the amplification cycle and therefore the number of colonies screened before an appropriate clone was found. Typically 10 – 15 cycles were used in each case.

RTPCR

Reverse Transcriptase PCR (RTPCR) was performed in a number of experiments. This technique uses the RT enzyme to generate a cDNA copy of mRNA which can then be amplified using gene specific primers. In each experiment, RNA was isolated using one of the methods described, quantified, and subjected to an RT step using SuperScript III Reverse Transcriptase (Invitrogen), as per the manufacturer’s instructions. PCR steps were then conducted on the cDNA template as previously described, product confirmed by agarose electrophoresis, and where necessary, cloning and sequencing. No RT controls were included where appropriate, and when possible primers were designed to span introns to control for contamination from genomic DNA.

QRTPCR

To enable the repeatable and rapid quantification of levels of transcript between wild type and mutant embryos, Quantified Reverse Transcriptase PCR (QRTPCR) was used (for a review see (Bustin, 2000)). In this method the incorporation of a fluorescent dye into PCR product is measured over the course of a PCR reaction. Since the dye fluoresces only when it intercalates into dsDNA (in a mechanism analogous to ethidium bromide), the level of fluorescence observed in a sample is
linearly proportional to the quantity of DNA of a given range of product sizes (>250bp). Since these samples were cDNA's synthesised from DNase treated total RNA preparations, and since all no RT controls performed during the PCR steps of this protocol were negative, the only dsDNA available for the incorporation of fluorescent dye must be PCR product derived from the amplification of cDNA. The following derivation has been adapted and simplified from (Livak and Schmittgen, 2001).

In the linear phases of a PCR reaction, the quantity of product after \( n \) cycles, \( X_n \), is related to the quantity of starting material \( X_0 \) by the relationship:

\[
X_n = X_0 \times (1 + E_x)^n
\]

Where \( E_x \) represents the efficiency of amplification of product \( X \) in the reaction (i.e. if the efficiency was 100\% \( E_x = 1 \), \( (1 + E_x)^n = 2^n \) or a doubling every cycle).

If the level of fluorescence after \( n \) cycles, \( F_n \), is directly proportional to \( X_n \) (i.e. \( F_n \propto X_n \)) then it follows:

\[
\frac{F_n}{c} \times (1 + E_x)^n = X_0
\]

Therefore, for any samples where \( E_x \) is equivalent, a measurement of \( F_n \) at cycle number \( n \) is equivalent to a measure of the quantity of starting material in that sample. Conversely, a difference of \( F_n \) after the same number of cycles in two different
samples (where $E_x$ is equivalent) is a measure of the difference in amount of starting material between those samples.

To ensure that $E_x$ is invariant between reactions, a plot of the difference of the number of cycles required to reach a given level of fluorescence against the difference in concentration of starting product must be linear. If

$$\frac{F_1}{c.(1+E_1)^n} = X_1$$ and $$\frac{F_2}{c.(1+E_2)^n} = X_2$$

Then

$$X_1 - X_2 = \frac{F_1}{c.(1+E_1)^n} - \frac{F_2}{c.(1+E_2)^n}$$

If

$$E_1 = E_2$$

Then

$$k.\Delta X = \Delta F$$

Where $k$ = amplification constant $c.(1+E_2)^n$.

Experimentally this can be tested by a dilution series of cDNA ($X$) against fluorescence ($F$) after $n$ cycles. If this plot is linear then the efficiency of PCR
amplification is identical across that range of cDNA concentrations, and then the measurement of fluorescence can be interpreted as an indicator of quantity of starting material.

The purpose of the QRTPCR experiments conducted in this thesis was to determine the difference in quantity of mRNA of each product between wild type and mutant embryo samples. This was performed in the following manner.

Firstly, two $Sp^{2H} +/-$ were mated to generate embryos which were genotyped using PCR as described, and $Sp^{2H} +/+$ and $Sp^{2H} +/-$ E9.5 sibling pairs used in the experiments. Total RNA was extracted and quantified as described, and equal quantities of RNA were used in three separate RT experiments for each embryo (with one –RT for each embryo). These replicates control for variation in cDNA synthesis from each sample.

For each of the genes analysed (i.e. for each PCR), a series of dilutions for the wild type samples were performed, from 1.0 µl of cDNA to 0.01 µl of cDNA. This generates a plot of concentration of cDNA vs. Fluorescence observed after $n$ cycles (cycle threshold). In every case this was linear, showing that amplification efficiencies were identical over this concentration range.

This also generates an standard curve of arbitrary units where the wild type samples are set at 1 (undiluted). Each of the three +RT reactions from the wild type and mutant embryos are then interpolated off of this graph, at this cycle threshold, to show the relative amounts of starting material present in the cDNA.
Each sibling pair generates 18 readings. Three sets of 6 readings for each of these genes analysed, namely GAPDH, Wnt1 and Pax7. For each gene, there are 3 wild type and 3 mutant readings. Variation between these replicates is a measure of the variability of the production of cDNA from the RNA samples, samples where the variation was high were excluded from the final data pool.

Four sibling pairs were analysed in this way and contributed to the final data pool. Mean averages of the level of expression of each transcript in the Sp<sup>2H</sup> mutants were expressed as % of wild type expression.

i.e. wild type = ~1.0 or each gene (since we are quantifying relative expression from the standard curve generated by serial dilution of wild type cDNA), and mutant = < 1.0 if the transcript is less represented than the wild type or > 1.0 if more abundant in the mutant.

This method does not simply normalise by the value obtained for GAPDH but tests to ensure GAPDH transcript levels remain unchanged. If this is the case, any changes in transcript levels observed have been assumed to be real.

Total RNA was extracted from littermate wild type and Sp<sup>2H</sup>-/- embryos as described, and quantified. Identical µg quantities of RNA were then subjected to an RT step (above) and the cDNAs then used with QuantiTect SYBR Green reagent (Qiagen) in PCR reactions on an Opticon single wavelength PCR machine and data analysed using the Opticon Monitor software. Dilution series were performed for all genes
analysed to ensure linearity of response across a concentration gradient, and to enable relative quantisation of each transcript. All reactions were run on the program 55°140, and [MgCl₂] 1.5mM. Amplicons were all < 250bp to ensure linearity of SYBR Green incorporation (i.e. c is identical between all reactions, Qiagen), and melting point curves were examined to ensure the specificity of each reaction. Levels of Sp^{H^{}}^{1}/- transcript expression were expressed as % of wild type, and GAPDH was used a normalising control.

**Site Directed Mutagenesis**

Mutation of the core Pax3 binding consensus was performed using the QuikChange Site Directed Mutagenesis Kit (Stratagene) according to manufacturer’s instructions. Briefly, this protocol uses primers carrying the desired mutations to anneal to the unmutated sequence on the plasmid of interest. A Pfu polymerase then extends the primers around the plasmids on both strands, generating nicked, non methylated complementary strands. The parental DNA (methylated) is then digested away with the DpnI enzyme leaving only the new, mutated strands. These are then annealed and transformed into E. coli as described above.

Primers used were W1_sub_FWD/REV in conjunction with the pW1P(1.2)Luc plasmid and P7_sub_FWD/REV in conjunction with the p1.5I1RLuc plasmid. Colonies were grown up and plasmid DNA extracted for sequencing, as described, to confirm the substitution of the Pax3 core consensus site.
Primer extension

Primer extension was used to size the 5’ Untranslated Region (UTR) of the Pax7 mRNA, a diagrammatic representation of this method can be seen in Chapter 4 (Figure 4.3A). Briefly this method employs a radiolabelled primer which anneals to a known sequence within the transcript of interest (i.e. Pax7 coding exons) and is then used by the RT enzyme to prime a cDNA extension to the 5’ end of the mRNA. Since the primer is first kinase labelled with $^{32}$P at the 5’ end, this can be visualised on a denaturing polyacrylamide gel against an ssDNA ladder (also radiolabeled). This enables the 5’UTR of the transcript to be sized.

Total and messenger RNA was extracted from C2C12 (Pax7 positive) cells and embryos as described above. This was quantified and 8µg RNA used in each extension step. Primers were ordered from MWG-Biotech (detailed in Chapter 4 and Appendix 2) and 100 pmol kinase labelled with $^{32}$P γ-ATP at 3000 Ci/mMole using the T4 polynucleotide kinase protocol. Labelled primer was then separated from unincorporated $^{32}$P γ-ATP by application of the reaction mix to a G-25 size exclusion column (Amersham). Purified, labelled primer were then frozen and equal quantities used in each extension experiment.

8 µg of RNA was then placed in the following reaction mix:

<table>
<thead>
<tr>
<th>Extension Mix (uses SuperScript RT (Invitrogen) reagents)</th>
<th>8 µl MgCl$_2$ (25mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 µl 10X Reaction Buffer</td>
</tr>
<tr>
<td></td>
<td>4 µl dNTP mix (Promega)</td>
</tr>
<tr>
<td></td>
<td>1 µl RNase OUT (Ambion)</td>
</tr>
<tr>
<td></td>
<td>n µl RNA (8 µg)</td>
</tr>
</tbody>
</table>
This was then heated to 70 °C for 10 minutes to denature RNA secondary structure and then snap cooled on ice for 5 minutes. 4 µl Labelled Primer was then added and the mixture incubated at RT for 5 minutes to allow primer – RNA complexes to form. 1 µl SuperScript III RT (Invitrogen, final reaction volume 40 µl) was then added and the mixture incubated at 55 °C for 45 minutes. Reactions were then stopped by heating to 100 °C for 5 minutes. Samples were then boiled in Formamide Loading Buffer II (Ambion) and analysed by running on a 1mm thick 6% polyacrylamide 8M Urea denaturing gel (see above). A DNA ladder of pBR322 (MspI, NEB) was phosphatased using SAP (Roche) and then kinase labelled using 32P γ-ATP as described. This was used to size the primer extension fragments.

The gels were ran for 2.5 hours at 450 V and then fixed in a solution of 5% acetic acid and 15% methanol in ddH2O for 20 minutes. The gels were then dried overnight under vacuum, and exposed to X-ray film (Kodak) for 12, 24, 48 and 72 hours at −80 °C in an imaging cassette.

5'RACE

To further examine the Pax7 transcript, 5' RACE (Rapid Amplification of cDNA Ends) was used to amplify, clone and sequence the 5'UTR. This methodology uses a set of gene specific primers and a set of primers complementary to sequences annealed to the 5' end of the transcript to perform nested PCR and amplify the 5'UTR of the gene. Oligonucleotides of known sequence are generally annealed to the very
5' end of the mRNA or cDNA. The precise method by which this occurs differs from protocol to protocol, and is described below. Once these linker sequences have been annealed, PCR or RTPCR (depending on whether the linkers anneal to the mRNA or cDNA population) can then be used to specifically amplify the gene of interest.

The SMART RACE cDNA amplification kit was used in the first instance (Clontech). In this protocol total RNA is extracted from the source of interest (in this instance C2C12 or wild type E11.5 embryos) and an RT step performed with an enzyme blend designed to leave an oligo-C overhang at the very 3' end of the cDNA. This is then used to anneal a primer containing an oligo-G 5' end with a 3' known sequence. The RT enzyme then switches templates and extends along this new primer, generating a cDNA with a known 5' sequence. This population of cDNA's can then be used in a conventional nested PCR reaction using gene specific primers to isolate and sequence the 5'UTR required.

Due to the seemingly foreshortened nature of many of the 5'RACE clones isolated with the above method when compared to the primer extension data (see Chapter 4), a second method of performing 5'RACE was then utilised to try and exclusively select for full length transcripts. A FirstChoice RACE-Ready cDNA kit (Ambion) was used. This method selects for full length 5'UTR's by exploiting the 5' terminal 7-methylguanine cap present only on full length mRNA's. Firstly, a phosphatase activity is used to remove phosphates from any RNA species without a 5' cap (i.e. those partially degraded). Next the 5' cap guanine is removed from full length mRNA molecules. This generates a population of RNA's in which only those which were full length will have a 5' phosphate moiety capable of accepting the ligation of
an RNA primer of known sequence to the 5' end of the transcript. Thus a population of mRNA's with a 5' end of known sequence are generated. This is then used in an RTPCR reaction with primers designed against this sequence and within the gene of interest to amplify and clone the 5'UTR.

Cytoplasmic RNA was isolated from C2C12 cells of wild type embryos, and quantified, as described. Gene specific primers were designed using MacVector, and the 5' RACE followed according to the manufacturer's instructions. PCR products were TOPO cloned and sequenced using MWG-Biotech. Sequences were aligned to the genomic using a CustalW local alignment.

**Ribonuclease Protection Assay (RPA)**

RPA analysis was employed to further analyse the *Pax7* 5'UTR. RPA utilises a radio labelled antisense RNA probe against the sequence of interest (see figure 4.5A) to anneal to the mRNA. This forms an RNA-RNA duplex where the sequences are complementary. RNase is then used to digest away un-annealed RNA and probe sequences, leaving a probe – target complex intact. These are then purified and run out on a denaturing polyacrylamide gel. Since the probe sequence is known (and overlaps with the coding sequence of *Pax7* at one end), this can be used to size the 5’UTR accurately. Since the probe was designed to cover the UTR’s predicted by both primer extension and 5’RACE it was hoped that this method might help resolve differences between the 5’UTR’s predicted by PEXT and 5’RACE for *Pax7*. 
A probe was cloned into a vector containing an SP6 promoter in the appropriate orientation (see Chapter 4). This was linearised at a final concentration of 0.5µg/µl and in vitro transcription performed with the MAXIscript in vitro transcription kit (Ambion) according to manufacturer’s instructions and using $^{32}$P α-UTP at 800Ci/mmol as the radiolabel. Probes were then subjected to DNasel digestion and a G-25 size exclusion column used to purify as described previously. Probes were then run out on a 6% acrylamide gel and exposed to X-ray film for 1 hour to position within the gel. Full length probes were then extracted overnight into buffer.

RNA was then extracted from C2C12 cells (cytoplasmic RNA) or E11.5 wild type embryos (total RNA) and quantified. 5µg of RNA was used in each RPA reaction and this was made up to 50µg RNA using yeast tRNA (a negative control reaction using Pax7 negative NIH-3T3 was also performed). Hybridisation was conducted using the HybSpeed RPA kit (Ambion) according to manufacturer’s instructions. Samples were then treated with RNase and protected fragments separated on a denaturing polyacrylamide gel, as described. Fragments were sized using pBR322(MspI) kinase – labelled ladder as before, fixed, dried and either exposed to X-ray film for two weeks or to a phosphorimager.

**PAC Library Screen**

A mouse PAC genomic library was utilised to clone the mouse Pax7 genomic locus. PAC clones can carry up to 150kb of insert and, given the size of the Pax7 genomic region (figure 4.1A), these were deemed an appropriate source from which to clone
this region of DNA. The PAC library, gridded on to nylon filters, was obtained from the HGMP. A probe was synthesised from plasmid pP7HYB (520bp EcoRI fragment, human sequence, see Chapter 4) using $^{32}$P $\alpha$-CTP and the HighPrime kit (Roche). Unincorporated radioactivity was removed by applying the reaction mix to a G-25 size exclusion column (Amersham). Filters were probed using the following protocol:

1) Soak filters in 2X SSC for 1 hour at RT
2) Remove from 2X SSC
3) Stack filters one on top of another, DNA facing upwards, with a nylon membrane placed between the sheets
4) Roll filters up and place into a Hybridisation Tube
5) Add 20ml Church Hybridisation Mix
6) Incubate at 65 °C for 2 hours (pre-hybridisation step)
7) Pre-warm 15ml Church Hybridisation Mix to 65 °C
8) Pour Church pre-hybridisation solution off filters and discard
9) Add labelled probe to 15ml pre-warmed Church Hybridisation Mix
10) Place in tube with filters, rotate for 16 hours at 65 °C, ensuring even coverage of filters with probe solution
11) Remove probe solution and discard
12) Wash filters three times in Church Wash solution, pre-warmed to 65 °C
13) Monitor filters for background
14) Starting with 2X SSC, wash filters with increasing stringency of SSC until filter background is acceptable and specific hot spots can be detected
15) Wrap films in cling film
16) Expose to X-ray film in imaging cassette at −80 °C for 24 hours to one week, depending on signal strength
17) Identify positive clones and order from HGMP collection

The solutions used were as follows:

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>20X SSC</td>
<td>3M NaCl</td>
</tr>
<tr>
<td></td>
<td>300mM Na-citrate</td>
</tr>
<tr>
<td>1M NaPO$_4$ pH 7.2</td>
<td>1M Na$_2$HPO$_4$·7H$_2$O</td>
</tr>
<tr>
<td></td>
<td>0.4% H$_3$PO$_4$</td>
</tr>
<tr>
<td>Church Hybridisation Mix</td>
<td>250 mM NaPO$_4$ pH 7.2</td>
</tr>
<tr>
<td></td>
<td>7% SDS (w/v)</td>
</tr>
<tr>
<td></td>
<td>1mM EDTA</td>
</tr>
<tr>
<td></td>
<td>1% BSA (w/v)</td>
</tr>
<tr>
<td>Church Wash</td>
<td>1% SDS</td>
</tr>
</tbody>
</table>
Southern Blot

To map and confirm the PAC clones identified by the PAC library screen, southern blotting was performed using a 770bp PCR probe amplified from mouse genomic DNA, see Chapter 4. Probes were labelled as described for the PAC library screen. PAC clones were grown up and PAC DNA extracted as described. The DNA obtained from each clone was then digested with 7 different enzymes for 16 hours at 37 °C. The digests were then run out on a large (200ml) 1.2% agarose gel for 16 hours at 30 V. The gels were then stained and photographed to ensure digestion and separation. Southern blotting was then conducted as follows:

1) Wash gel in ddH₂O for 30 minutes, shaking gently on a rotating platform
2) Wash gel in ddH₂O with 12.5ml concentrated HCl for 30 minutes
3) Wash gel in 2 x excess of Denaturing Solution for 30 minutes each
4) Wash gel in 2 x excess of Neutralising Solution for 30 minutes each
5) Place gel on blotting set-up containing 20X SSC
6) Cover with cling film and cut a hole around the gel, removing this
7) Cover gel surface with 2X SSC
8) Cut nitrocellulose filter to the size of the gel and place on top, taking care to remove any air bubbles
9) Wet filter with 2X SSC
10) Place two pieces of Whatmann paper, soaked in 2X SSC, cut to the appropriate size on top of the filter
11) Stack absorbent paper towels on top of the Whatmann papers
12) Apply heavy weight to the top of the blotting set up
13) Leave for 16 hours
14) Replace Whatmann paper and towels, re-apply weight
15) Leave for a further 8 hours
16) Remove filter, wrap in cling film
17) UV cross link
18) Air dry, then bake at 120 °C for 15 minutes
19) Place filter in hybridisation tube with nylon membranes
20) Add 20ml Church Hybridisation Mix, pre-warmed to 65 °C, containing 20 μl denatured salmon sperm DNA
21) Pre-hyb for 30 minutes
22) Remove pre-hyb solution
23) To 10ml pre-warmed Church Wash add 10 µl denatured Salmon Sperm DNA and radiolabeled probe
24) Rotate in a hybridisation oven for 16 hours at 65 °C
25) Remove probe wash
26) Wash filter 4 x 30 minutes in 20 ml Church Wash
27) Wrap moist filter in cling film
28) Expose to X-ray film in an imaging cassette for 24, 48 and 72 hours, as required

The solutions used were as follows:

<table>
<thead>
<tr>
<th>Revelation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>20X SSC</td>
<td>As above</td>
</tr>
<tr>
<td>1M NaPO₄ pH 7.2</td>
<td>As above</td>
</tr>
<tr>
<td>Church Hybridisation Mix</td>
<td>As above</td>
</tr>
<tr>
<td>Church Wash</td>
<td>As above</td>
</tr>
<tr>
<td>Denaturing Solution</td>
<td>0.5M NaOH</td>
</tr>
<tr>
<td></td>
<td>1.5M NaCl</td>
</tr>
<tr>
<td>Neutralising Solution pH 7.5</td>
<td>0.5M Tris Base</td>
</tr>
<tr>
<td></td>
<td>1.5M NaCl</td>
</tr>
<tr>
<td></td>
<td>pH with HCl or NaOH</td>
</tr>
</tbody>
</table>

**Transfections**

**Luciferase Assay Transfections (protocol for adherent NIH-3T3 and C2C12 cells)**

Cultures of NIH-3T3 and C2C12 cells were maintained in class II conditions as described. One day prior to transfection, cells were trypsinised and re-suspended in antibiotic free DMEM N-12. Cells were counted using a haemocytometer and 2 x 10⁶ cells were added to each well of a sterile 24 well plate containing 500µl antibiotic free DMEM N-12. These cultures were left incubated for 16 hours to reach 80-90% confluence prior to transfection. 2 hours prior to transfection the growth medium was
aspirated, the cells washed with PBS twice and covered in 500µl Opti-MEM transfection medium (Invitrogen) per well. Cultures were then replaced in the incubator for two hours prior to transfection.

Plasmid DNA was grown and purified as described before being linearised and quantified. Equal µg quantities of DNA were added to the cells in each transfection. Also, equal µg quantities of CMV and SV40 strong and weak promoters were added to the cells where appropriate. This was to prevent variations in luciferase output and Renilla / luciferase ratios due to promoter – promoter interactions and competition for transcription factors between experiments. For example:

<table>
<thead>
<tr>
<th>µg luciferase plasmid DNA</th>
<th>µg Renilla plasmid DNA</th>
<th>µg pCMV-PAX3 plasmid DNA</th>
<th>µg pCMV-Script plasmid DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well A</td>
<td>0.5</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Well B</td>
<td>0.5</td>
<td>0.1</td>
<td>0.05</td>
</tr>
<tr>
<td>Well C</td>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Here, all three wells have a total of 1µg plasmid DNA added, varying the amount of Pax3 expression plasmid, enabling the effect of several concentrations of Pax3 expression to be investigated on the luciferase reporter plasmid, the total amount of CMV promoter is constant. (NB. Since the sizes of pCMV-Script and pCMV-Pax3 vary this is not strictly speaking true. pCMV-Pax3 is 25% larger and so, weight for weight, will carry 25% fewer ‘molecules’ of CMV promoter into the transfection. Since an excess of the smaller plasmid is used in each experiment, the actual difference in number of CMV promoters is minimal, i.e. 0.4 in Well A compared with
(0.3 + (0.75 x 0.1)) = 0.375 = 6% difference in Well C). Since the observed luciferase induction was always much greater than these small differences in CMV promoter composition, and DNA-Lipofectamine ratios thought to have a much greater effect on transfection ratios, (Invitrogen, product insert), this was considered acceptable). For 24 well plates the optimal total amount of DNA was found to be 1.5µg per well for both cell lines.

<table>
<thead>
<tr>
<th>µg luciferase plasmid DNA</th>
<th>µg Renilla plasmid DNA</th>
<th>µg pCMV-PAX3 plasmid DNA</th>
<th>µg pCMV-Script plasmid DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>x</td>
<td>0.4 + y</td>
</tr>
</tbody>
</table>

Where x + y = 0.5µg DNA

Lipofectamine-2000 reagent was then used to transfect the plasmid DNA into each well as follows:

1) Add all DNA to 50µl Opti-MEM
2) Mix and incubate at RT
3) Add Lipofectamine-2000 to a second 50µl Opti-MEM. Add nµl Lipofectamine-2000 where n = (µg DNA x 2). For example, 1µg total DNA = 2ul Lipofectamine-2000
4) Mix and incubate at RT for 5 minutes
5) Mix DNA - Opti-MEM and Lipofectamine-2000 - Opti-MEM together
6) Incubate at RT for 30 minutes
7) Add to well containing 500µl Opti-MEM
8) Mix gently and incubate at 37 °C, 5% CO₂ for 24 hours

After 24 hours, cells were washed in PBS and harvested using the Passive Lysis Buffer component of the Dual Luciferase Reporter Assay System (Promega). Luciferase and Renilla Luciferase expression was then assayed according to manufacturers instructions and using a TD-20/20 Luminometer (Turner Biosystems).
Relative Light Units (RLU) were measured for each luciferase, and recorded for every sample. RLU was measured cumulatively over 10 seconds and performed twice for each sample to ensure accuracy of measurement. Individual transfections were conducted in triplicate in each run to ensure each set of conditions gave a reliable result within a run of transfections. These runs were repeated between 3 and 12 times, depending on the specific experiment, and this is then recorded as the n value for each experiment.

Appendix 4 contains a detailed description of the Luciferase rationale and data treatment used in this thesis.

Transfections for Western Blotting

To confirm the expression of Pax3 and Msx2 full length protein product from the plasmids pCMV-Pax3 and pCMV-Msx2 Western blotting was also used. These plasmids were transfected into both NIH-3T3 and C2C12 and the proteins harvested for Western blotting. Transfections containing varying quantities of expression vector and pCMV-Script, or pCMV-Script alone (negative control), were conducted as described above but in 6 well plates. All quantities of cells and reagents were multiplied by 5 to account for the larger size of the wells and transfections conducted using Lipofectamine as normal.

After incubating for 24 hours, cells were washed twice in PBS, trypsinised and re-suspended in DMEM N-12 medium. These were then spun down (1000 rpm, 5 minutes), washed in cold PBS, re-spun and the re-suspended in 100μl cold Western
Lysis Buffer. These were transferred to ice for 15 minutes, passed through a small gauge syringe 10 times, spun down at 13,000 rpm and the protein concentration of the supernatant assayed as described. Protein samples were stored at -80 °C, and Western blotting performed as described below.

**Proteins**

**Choice of peptides used to raise α-Pax3 antibody**

The amino acid sequence of the Pax3 protein was derived from the cDNA sequence, and used to design immunogenic peptides to raise against the Pax3 protein. Peptide sequences and modification specifications were emailed to Genosphere Biotechnologies Ltd. who synthesised, modified and finally conjugated the peptides to Keyhole Limpet Hemocyanin. This was then injected into Rabbits over several weeks, and the serum checked for immunogenicity using an ELISA with the original peptides as primary substrate. Once good titres had been obtained the un-purified sera was shipped back from the company for analysis and purification in our lab.

**Protein Quantification**

Protein concentration was determined using the BCA Reagent Kit (Pierce), against a known range of BSA concentrations, according to the manufacturer's instructions. Assays were conducted in a 96 well plate and absorbance at λ562 was measured on a
plate reading spectrophotometer. A standard curve was drawn, linear regression performed and the resulting equation used to calculate the concentration of total protein of samples.

**Western Blotting**

Western blotting was used to confirm the specificity of the α-Pax3 antibody, expression of Pax3, Msx2 and Pax7 in cell culture and embryonic tissue. The following general protocol was used in all cases; the specific antibody concentrations used were determined empirically for each serum.

**SDS-PAGE**

10% Tris-Glycine polyacrylamide Novex pre-cast gels (Invitrogen) were used to separate proteins in one dimension. 10µg protein was loaded on each gel after being reduced by boiling in the following mixture:

<table>
<thead>
<tr>
<th>Protein Loading Mixture</th>
<th>n µl protein sample (10µg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 µl β-Mercaptoethanol</td>
</tr>
<tr>
<td></td>
<td>5 µl SDS 4X Sample Loading Buffer</td>
</tr>
<tr>
<td></td>
<td>(n + 6) – 20 µl Western Lysis Buffer</td>
</tr>
</tbody>
</table>

Samples were boiled for at least 10 minutes at 100 °C and then loaded into a lane filled with 1X SDS Running Buffer. The gels were set up in the Xcell SureLock Mini-Cell (Invitrogen) and run for 1 hour and 40 minutes at 150V. Magic Mark (Invitrogen) was used as a ladder to enable sizing of bands.
Protein Transfer

Immediately after electrophoresis proteins were transferred onto nitrocellulose membrane. The gel was removed from the SDS-PAGE apparatus, rinsed in ddH₂O and Transfer Buffer before being placed in a sandwich between two sheets of Whatmann paper and two thick nylon sponges on the outer surface. Next to the gel inside the sandwich a piece of nitrocellulose membrane (Biorad) was placed on the side which would face the positive electrode in the transfer tank. This was then placed into a transfer tank and covered in Transfer Buffer. Transfer was conducted at 4 °C at 50mA for 16 hours. The transfer apparatus was then dismantled and the nitrocellulose stained with Amido Black to demonstrate transfer and examine equal loading of protein samples. Excess Amido Black was then washed off in ddH₂O and an image of the membrane recorded. Membranes were chopped into separate pieces at this stage if necessary.

Immunodetection

The following protocol was used during immunodetection.

1) Block membrane in Blocking Buffer for 1 hour
2) Add primary antibody at appropriate concentration, diluted in Blocking Buffer
3) Incubate with shaking at 4 °C for 16 hours
4) Wash three times in PBS-T
5) Add secondary antibody at appropriate concentration, diluted in Blocking Buffer
6) Incubate with shaking at RT for 2 hours
7) Wash three times in PBS-T
8) Detect using the ECL+ detection system (Amersham) according to manufacturer’s instructions and exposing to X-Ray film

The solutions used were as follows:

<table>
<thead>
<tr>
<th>Western Lysis Buffer</th>
<th>150mM NaCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25mM Tris-HCl (ph 7.5)</td>
</tr>
<tr>
<td></td>
<td>1% Triton X-100</td>
</tr>
<tr>
<td></td>
<td>0.1% SDS</td>
</tr>
<tr>
<td></td>
<td>+ Complete Protease inhibitor tablets</td>
</tr>
<tr>
<td></td>
<td>(Roche)</td>
</tr>
<tr>
<td>4X Sample Loading Buffer</td>
<td>0.5M Tris-HCl (pH 6.8)</td>
</tr>
<tr>
<td></td>
<td>10% SDS (w/v)</td>
</tr>
<tr>
<td></td>
<td>10% Glycerol</td>
</tr>
<tr>
<td></td>
<td>0.1% Bromophenol Blue</td>
</tr>
<tr>
<td>5X SDS Running Buffer</td>
<td>0.125M Tris-Base (pH 8.3)</td>
</tr>
<tr>
<td></td>
<td>0.96M glycine</td>
</tr>
<tr>
<td></td>
<td>0.5% SDS</td>
</tr>
<tr>
<td>Western Transfer Buffer (make frc)</td>
<td>100mM glycine</td>
</tr>
<tr>
<td></td>
<td>120mM Tris-Base</td>
</tr>
<tr>
<td></td>
<td>200mL methanol</td>
</tr>
<tr>
<td>Amido Black</td>
<td>0.1% Amido Black</td>
</tr>
<tr>
<td></td>
<td>25% isopropanol</td>
</tr>
<tr>
<td></td>
<td>10% acetic acid</td>
</tr>
<tr>
<td>PBS-T</td>
<td>1X PBS</td>
</tr>
<tr>
<td></td>
<td>0.1% Tween-20</td>
</tr>
<tr>
<td>Blocking Buffer</td>
<td>5% powdered milk solution in PBS-T</td>
</tr>
</tbody>
</table>

**Immunohistochemistry**

Mouse embryos were obtained and dissected at specific ages as described. Embryos were then fixed in 10% paraformaldehyde for 16 hours at 4 °C and placed in 70% ethanol before being embedded in paraffin. Waxed embryos were then sectioned at 10 μm thickness and placed on glass slides.

The following protocol was used for immunohistochemistry.
1) Dewax sections in xylene (twice for 10 - 20 minutes)
2) Rehydrate sections in series of descending ethanols (100% - 50%)
3) Wash with PBS (3 times for 5 minutes)
4) Block endogenous peroxidases with a solution of 3% H₂O₂ and 10% methanol in PBS for 15 minutes in the dark
5) Wash with PBS-TX (3 times for 5 minutes)
6) Wash with (3 times for 5 minutes)
7) Wash in 10mM Sodium Citrate Buffer for 15 minutes
8) Boil slides in microwave for a total of 20' (5'-20') as follows:
9) 5 minutes at full power
10) top up with dH₂O
11) repeat up to 3 more times
12) Rest slides for 20 minutes to cool down
13) Incubate with Blocking Solution for 2 hours at RT
14) Wash with PBS-TX for 5 minutes
15) Incubate with primary antibody diluted in Blocking Solution for 16 hours at 4 °C
16) Wash with PBS-TX 5 times for 5 minutes
17) Incubate with biotinilated secondary antibody (1:200) Blocking Solution for 1 hour at RT
18) Wash with PBS-TX 5 times for 5 minutes
19) Incubate with the Streptavidine-HRP complex (1:400, ABC Kit) Blocking Solution for 1 hour at RT
20) Wash with PBS-TX twice for 5 minutes
21) Wash with PBS 3 times for 10 minutes
22) Develop with a solution of 0.03% DAB and 0.003% H₂O₂ in PBS
23) Wash slides 3-5 times in PBS
24) Counter stain with Crystal Violet if necessary
25) Dehydrate sections in a series of ascending ethanols and xylene
26) Mount the tissue with DPX

Sections were examined under an inverted light phase contrast microscope and images taken using a digital camera.

The solutions used were as follows:

<table>
<thead>
<tr>
<th>PBS-TX</th>
<th>0.2% solution of Triton X-100 in PBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium Citrate Buffer (pH 6.0)</td>
<td>100mM sodium citrate</td>
</tr>
<tr>
<td>Blocking Solution</td>
<td>0.2% gelatine</td>
</tr>
<tr>
<td></td>
<td>10% goat serum</td>
</tr>
<tr>
<td></td>
<td>in PBS-TX</td>
</tr>
</tbody>
</table>
Antibody Purification

After the immunogenicity of the α-Pax3 serum was confirmed, column purification was performed using Protein-G sepharose (Amersham). The sepharose was degassed under vacuum for 30 minutes at RT before being packed onto a 1cm diameter glass EconoColumn (Biorad). Flow rate was adjusted to 1ml per minute, and the column washed with 10 bed volumes of ice cold TBS. All purification was conducted in the cold room unless otherwise stated.

The volume of serum purified was calculated as follows:

Binding capacity of column (5ml) = 100mg IgG

Assume 90% binding efficiency = 90mg, and that rabbit IgG concentration is 5mg per ml, therefore 90/5 = 18ml serum applied to column.

Any debris from the serum was spun out at 15,000g for 5 minutes at 4 °C, and then applied to the column at a flow rate of 1 ml per minute. The serum was passed through the column twice. The column was then washed with 180ml ice cold TBS until no protein could be detected from the eluate using the BCA Kit (Pierce).

100µl Neutralisation Buffer was added to 20 x 1.5ml Eppendorfs as all the TBS was allowed to drain from the column. The column was transferred to RT and 15ml RT
pH 2.7 Elution Buffer was applied. 10 x 1ml fractions were collected. 15 ml of RT pH 1.9 Elution Buffer was the applied and 1ml fractions were collected. Samples were mixed immediately to ensure neutralization and placed on ice. The BCA Kit (Pierce) was then used to quantify the protein concentration and the λ.562 was measured. λ.562 vs. fraction number was then plotted and two peaks were identified and separated into two fractions for later testing with Western blot and immunohistochemistry. For long term storage, the fractions had BSA added to 5%, aliquoted and stored at −80 °C.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TBS</td>
<td>50mM Tris-HCl (pH 7.4)</td>
</tr>
<tr>
<td></td>
<td>150mM NaCl</td>
</tr>
<tr>
<td>Neutralization Buffer</td>
<td>1M Tris-HCl (pH 8.0)</td>
</tr>
<tr>
<td></td>
<td>1.5M NaCl</td>
</tr>
<tr>
<td></td>
<td>1mM EDTA</td>
</tr>
<tr>
<td>Elution Buffer pH 2.7</td>
<td>50mM glycine (pH 2.7)</td>
</tr>
<tr>
<td>Elution Buffer pH 1.9</td>
<td>50mM glycine (pH 1.9)</td>
</tr>
</tbody>
</table>

**Chromatin Immunoprecipitation (ChIP)**

Chapter 6 is dedicated to ChIP experiments, and so a detailed description of the rationale behind this method is not given here.

PCR primers were designed to amplify a region of either the promoter or downstream cis-regulatory elements of both the Pax7 and Wnt1 gene loci. PCRs were optimised for use on genomic DNA. The position of these primers and the reasoning behind their design is included in Chapter 6. ChIP primer sequences are recorded in Appendix 2.
Mouse embryos were obtained as described and sacrificed at E10.5. Embryos were dissected out and the telencephalon removed. These regions were deemed as being likely to contain tissues where the hypothesised interactions suggested from Chapters 3 and 4 (i.e. neural tube and somite / mesenchyme) might be found.

An in vivo ChIP methodology was designed specifically for use in this thesis as an amalgamation of in vitro methods used by other groups, and published on the internet. Several approaches were tested, and the following is a description of the method which proved successful. One of the major considerations in using an embryological source is the rapid and uniform fixation of the source material. Over exposure to formaldehyde would result in excessive crosslinking making subsequent enrichment steps impossible. To enable embryonic tissues to be manipulated in a manner closer to that of the tissue culture cells usually used in this type of analysis, a gentle disassociation protocol was utilised after dissection to generate a suspension of single cells. This also enabled cell counting and a uniform quantity of starting material to be used in each repeat experiment.

Dissected embryos were then immediately disassociated using the Papain Dissociation System (Worthington Biochemical Corporation) to obtain a suspension of cells in PBS. Cells were counted and equal numbers of cells used in every ChIP experiment (1 x 10⁷). Cell suspensions were then fixed by adding formaldehyde to 1%, and incubating with gentle agitation at 37 °C for 10 minutes. The fixing process was then quenched by the addition of glycine to 125mM. The cells were then pelleted, washed in PBS, and then re-suspended in 200μl Lysis Buffer and incubated on ice for 10 minutes.
Sonication conditions were previously optimised on disassociated mouse embryos to generate fragments of 500bp to 1kb, see Figure 6.3B. This should be optimised for every individual sonicator and probe before attempting these experiments since large variation between individual equipment and probes is generally observed.

Suspensions were then sonicated to generate fragments of an appropriate length. Samples were then centrifuged at 13,000 rpm for 5 minutes at 4 °C to remove debris. Supernatants were then transferred to new tubes. 300μl fresh Lysis Buffer and 50μl protein G sepharose (Amersham) with Salmon sperm DNA (2.5 μg/μl) was added to pre-clear the lysates. The samples were then gently rotated at 4 °C for 2 hours. The beads were spun down, 6,000 rpm for 1 minute, and the supernatants transferred to fresh tubes. 20 μg of purified α-Pax3 antibody was then added and the samples rotated at 4 °C for 16 hours.

50μl Protein G Sepharose was then added to each sample, and rotated for a further 4 hours at 4 °C. Immunoprecipitated complexes were then spun down at 6,000 rpm for 1 minute and the supernatant discarded. Complexes were then washed in 500 μl of the following, spinning down to recover the beads each time (all steps performed at 4 °C):

1) Lysis Buffer, 10 minutes
2) Lysis Buffer, 5 minutes
3) Low Salt Buffer, 5 minutes (low stringency)
4) High Salt Buffer, 5 minutes (high stringency)
5) LiCl Buffer, 5 minutes (removal of non-specific chromatin – agarose interactions)
6) TE Buffer (pH 8.1), 30 minutes
7) TE Buffer (pH 8.1), 5 minutes
Finally the complexes were re-suspended in fresh Elution Buffer, and vortexed at RT for 15 minutes. The sepharose was then spun out and the supernatant transferred to a fresh tube. Formaldehyde cross linking was then reversed by adding 5µl 8M NaCl, 1µl RNaseH (Qiagen), 5µl Protinase K (Promega) and incubating at 65 °C for 16 hours.

DNA was then purified using phenol chloroform and precipitated using ammonium acetate as described. The seeDNA nucleic acid reagent (Amersham) was added during precipitation to visualise the pellet. Pellets were re-suspended in 50µl ddH₂O and 5µl ran out on an agarose gel against serial dilutions of 'input' sonicated chromatin (the source of positive control, see Chapter 6) to ensure that similar quantities of template DNA were added to each PCR reaction. 5µl of ChIP and input chromatin were then used in each PCR reaction. PCR conditions were as shown in Chapter 6 and PCR products visualised on a 2% agarose gel.

The solutions used were as follows:

| **Lysis Buffer**          | 150 mM NaCl  
|                          | 25 mM Tris-HCl pH 7.5 
|                          | 1% Triton X-100 
|                          | 0.1% SDS 
|                          | 0.5 % Deoxycholate 
| **Low Salt Buffer**       | 0.1% SDS, 
|                          | 1% Triton X-100 
|                          | 1mM EDTA 
|                          | 20mM Tris-HCl, pH 8.0 
|                          | 150mM NaCl 
| **High Salt Buffer**      | 50 mM Tris-HCl, pH 8.0 
|                          | 500 mM NaCl 
|                          | 0.1 % SDS 
|                          | 1% Triton X-100 
<p>|                          | 1 mM EDTA |</p>
<table>
<thead>
<tr>
<th>Buffer</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>LiCl Buffer</td>
<td>50 mM Tris, pH 8.0</td>
</tr>
<tr>
<td></td>
<td>1 mM EDTA</td>
</tr>
<tr>
<td></td>
<td>250 mM LiCl</td>
</tr>
<tr>
<td></td>
<td>1% NP-40</td>
</tr>
<tr>
<td></td>
<td>0.5% Deoxiholate</td>
</tr>
<tr>
<td>TE Buffer</td>
<td>10 mM Tris-HCl, pH 8.0</td>
</tr>
<tr>
<td></td>
<td>1 mM EDTA</td>
</tr>
<tr>
<td>Elution Buffer</td>
<td>1% SDS</td>
</tr>
<tr>
<td></td>
<td>0.1M NaHCO₃</td>
</tr>
</tbody>
</table>

Every solution had protease inhibitors added fresh before use in each experiment, and were kept at 4 °C. Complete Protease Inhibitor Cocktail tablets (Roche) were used.

**Bioinformatics**

DNA sequences were obtained, where available, from [www.ensembl.org](http://www.ensembl.org) and annotated manually. Global alignments were performed using PipMaker (Schwartz, 2000), Jemboss, and ClustalW software, all available from [www.hgmp.ac.uk](http://www.hgmp.ac.uk). Transcription factor binding sites were predicted using the MatInspector program, available from Genomatix; data outputs were then dumped into Excel and analysed manually with this software. Cloning and DNA analysis was performed using the MacVector 7.2 software from Accelrys.
Chapter Three: Wnt1 regulation

Introduction

In this chapter, the interaction between Pax3 and the regulatory elements of the Wnt1 gene are investigated. As described in Chapter 1, Wnt1 provides an attractive potential target for regulation by Pax3 due to both its expression pattern, timing and its apparent position as genetically downstream of the Pax3 signal (see (Conway et al., 2000; Deardorff, 2001; Meulemans and Bronner-Fraser, 2004; Monsoro-Burq, 2004) for representative evidence). This chapter will show that, whilst Pax3 cannot be demonstrated to have any significant regulatory effect on the 3’ distal enhancer region, (Echelard et al., 1994), the identification of a putative Pax3 consensus site 5’ proximal to the transcriptional start site of the Wnt1 gene enables a regulatory interaction between this region and the Pax3 transcription factor to be established in vitro. Hypothesis One will be addressed first.

Hypothesis One

Pax3 regulates Wnt1 directly, most probably through the 3’ distal enhancer region, in vivo during the development of the neural crest.
Wnt1 expression is decreased in Sp\textsuperscript{2H} homozygotes

Before establishing \textit{Wnt1} as a direct target of regulation by \textit{Pax3}, it was important to unequivocally demonstrate the misregulation of this gene \textit{in vivo}. Whilst other groups have shown \textit{Pax3} to be unperturbed in \textit{Wnt1} mutants (Ikeya et al., 1997) and later studies have shown evidence from \textit{in situ} hybridisations which suggest \textit{Wnt1} is significantly downregulated in the \textit{Sp}\textsuperscript{2H} homozygous mutant (Conway et al., 2000), no quantitative proof of this misregulation in the \textit{Splotch} mouse has been reported. Demonstrating a change in \textit{Wnt1} expression on a \textit{Pax3} null background was perceived as an essential step before an investigation into the potential regulatory relationship could be taken further.

To achieve this, Quantified Reverse Transcriptase PCR (QRTPCR) was performed on RNA extracted from wild type and mutant sibling pair embryos at E9.5 as described (see Chapter 2 for a detailed description of the QRTPCRT technique). Primers pairs used were GAPLITEF/R and W1LITEF/R, (all primers are listed in Appendix Two). The assay was performed on four pairs of embryos, genotyped by PCR (primers P3GENOF/R), which were then homogenised and RNA extracted using a RNeasy Midi kit (Qiagen). RNA was then quantified using a UV spectrophotometer, and a fixed quantity added to a reverse transcriptase reaction. QRTPCR, using the SYBR Green dye (Qiagen) and a Opticon Monitor system (MJ Research), was then performed using the 55\(^{\circ}\)35 program as described. Expression levels for mutant embryos were then reported as a \% of wild type. A wild type standard curve was used to interpolate and the results averaged across the four sib-pairs. Blank and no RT
controls were negative, and blank fluorescence automatically subtracted from the data set by the Opticon software. These data can be seen in Figure 3.1.

This figure clearly shows a significant reduction of Wntl expression of around 40% (p = 0.003). In contrast no detectable difference in GAPDH expression was observed in the same embryo pairs. This figure is in keeping with previously published estimates of 50% (Conway et al., 2000) and was interpreted as meaning that the levels of Wntl expression in the Sp2H mutant are significantly reduced in comparison to the wild type. This result does not indicate why this difference is observed. Wntl may be entirely independent of Pax3 regulation, but expressed within a cell population that is reduced in Sp2H homozygous mutants. To establish a regulatory role for Pax3, a more detailed analysis of the regulatory elements known to control Wntl expression was performed.

**Pax3 does not interact with the Wntl 3’ enhancer**

The expression of Wntl has been faithfully recaptured by the expression of a LacZ reporter cassette driven by a 5.5kb enhancer fragment 3’ distal to the transcriptional start site (Echelard et al., 1994). This region has been investigated in detail, and a number of potential transcription factor binding sites have been mapped to this region (Iler et al., 1995; Rowitch et al., 1998). To enhance this investigation a bioinformatic analysis of this 5.5kb sequence was performed using MatInspeetor software (MatInspeetor). This searches genomic sequence for significant matches to known transcription factor binding sites based on known targets and sequence affinities. Whilst over 600 sites were predicted, with a core similarity of 75% or over, only three
Figure 3.1: QRTPCR of Wnt1 and GAPDH transcript levels in Pax3 -/- embryos. Levels expressed as % of wild type expression. GAPDH 94.01% of wild type (no difference), Wnt1 59.4% of wild type expression (40.6% reduction). n = 4 wild type / homozygous mutant sibling pairs. Error bars are SEM. p = 0.003 (two sample t test)
potential Pax binding sites were identified. These were all predicted as Pax5 binding sites. Since Pax3 and Pax5 are not from the same Pax gene subfamily, and Pax5 is thought to be primarily involved in B-lymphocyte development (Urbanek, 1994), these seemed unlikely to be potential Pax3 binding sites.

A lack of predicted binding sites does not preclude a regulatory interaction between this fragment and the Pax3 protein via a novel site. To investigate this possibility the 5.5kb enhancer fragment spanning this region was cloned from the pWEXP3 plasmid (A. McMahon, kind gift) as BglII fragment into the pGL3promoter Luciferase reporter construct (Promega) BamHI site and clones screened to ensure genomic orientation relative to the Luciferase ORF. The pGL3promoter construct carries a Firefly Luciferase reporter gene under the control of a weak SV40 promoter, and is designed to enable the dissection of enhancer elements. This construct was then co-transfected into NIH-3T3 cells (Pax3 and Wnt1 negative by RTPCR, data not shown) with a Pax3 expression construct, pCMV-Pax3. The expression construct was generated by cloning full length Pax3 cDNA from wild type E10.5 RNA by RTPCR using the primers 3/5 P3IRES to amplify full length Pax3 cDNA. This was then TOPO cloned into pCR-II-TOPO-Blunt and sequenced using the Pax3 sequencing primers shown in Appendix Two. Full length Pax3 with no mutations was then sub-cloned into pCMV-Script (Stratagene) as a SacII – BamHI fragment. This generated a plasmid with the full length Pax3 cDNA under the control of the CMV promoter. These constructs are shown in Figure 3.2A.

pCMV-Pax3 was then co-transfected into NIH-3T3 cells with pW1ELuc, and the fold induction over pGL3promoter baseline calculated for 6 triplicate runs, see Figure
Figure 3.2:
A) Illustrating Wnt1 enhancer Luciferase reporter construct and pCMV-Pax3 expression construct, B = BgIII - BamHI cloning sites.
B) Luciferase data illustrating normalised luminescence for pW1ELuc with and without Pax3 expression (100ng pCMV-Pax3 or 100ng empty vector). Data is expressed as fold induction over pGL3promoter (pW1ELuc without Wnt1 enhancer fragment). n = 6, error bars are SEM, induction is non-significant (p = 0.24, two sample t test)
3.2B. A detailed description of the Luciferase protocol and the data treatment used in this thesis can be found in Appendix 4 and is not given here. Whilst a very slight induction was observed with the addition of pCMV-Pax3 versus empty vector, pCMV-Script, this was not significant. The reverse enhancer orientation was also examined in this assay, but again no change in Luciferase activity could be detected on co-transfection with pCMV-Pax3 (data not shown).

To ensure the pCMV-Pax3 expression vector was producing full length Pax3 protein in these cells, Western blots on protein extracts from samples transfected with pCMV-Pax3 and those transfected with pCMV-Script were performed. The primary antibody was derived as described in Chapter 5. Un-purified serum was used in these experiments since non-specific bands enable an easy control of loading equivalence between lanes. A representative blot can be seen in Figure 3.3. This clearly shows the expression of full length Pax3 protein in these samples (red arrow). It was concluded that the lack of Luciferase induction in these samples could not be attributed to a problem with Pax3 expression from the pCMV-Pax3 construct.

In this *in vitro* assay, Pax3 could not be shown to interact with the 3' distal enhancer element of Wnt1.

**The Wnt1 5' promoter is Pax3 responsive**

Since the 3' distal enhancer of Wnt1 did not show any reaction to Pax3 co-expression, the 5' promoter region of the locus was then examined. The examination of this region, in spite of the control of Wnt1 expression being generally attributed to the
Figure 3.3:
Western blot confirming expression of Pax3 in Luciferase assay system (NB. samples taken from actual luciferase experiment and unpurified α-Pax3) Pax3 ~53KDa
+ = transfected lane, - = un-transfected lane.
enhancer element investigated above, was seen as important for several reasons. Firstly, (Serbedzija and McMahon, 1997) reported that the expression of a Wnt1-LacZ transgene was altered on a Pax3 null background. This transgenic was constructed using the WZT9B construct, and contains both the 5.5kb 3' enhancer region and a portion of the Wnt1 genomic 5' proximal promoter (Echelard et al., 1994). Further, studies in *Xenopus* (Gao et al., 1994) have implicated this region as important in inducing reporter gene expression concomitant with neural cell specification. Also, the differentiation of P19 embryonal carcinoma cells along a neuroectodermal lineage using retinoic acid treatment has been illustrated to induce both Pax3 (Pruitt, 1992) and the expression of Wnt1 via an element in the 5' promoter region (St-Arnaud and Moir, 1993). Finally, the Wnt1 promoter is essential in Six3 mediated repression of Wntl expression in the developing telencephalon (Lagutin et al., 2003), establishing a regulatory precedent for this region in anterior neural tissues.

The Wnt1 5' proximal promoter was then subjected to bioinformatic binding site analysis (MatInspector). This identified 178 potential transcription factor binding sites in the proximal 1.5kb upstream of the Wnt1 gene. Several Pax gene consensus sites were found, including Pax1, Pax4 and Pax9. Most importantly, one Pax3 consensus was discovered in extreme proximity to the Wnt1 transcriptional start site (46bp upstream). This is shown in Figure 3.4A. ClustalW local alignments were then performed on this region between human and mouse genomes, and illustrated a high degree of conservation (Figure 3.4B) suggesting functional significance. Finally, analysis of the human, mouse, *Fugu* and *Xenopus* Wnt1 promoters (from (ensembl)) revealed putative Pax3 binding sites in similar positions in all 4 species (Figure 3.4C). This site is also found
Figure 3.4:
A) Identification of Pax3 binding consensus in the Wnt1 5' promoter.
B) Local alignment of human and mouse promoters showing high overall level of conservation. Red bar highlights Pax3 consensus site.
C) Alignments showing conservation of consensus in Xenopus, Fugu, mouse and human Wnt1 promoter regions
on the murine \textit{Wntl} promoter fragments used in the constructs examined in (Serbedzija and McMahon, 1997) and (St-Arnaud and Moir, 1993).

To investigate the possible functional significance of this conserved element, Luciferase constructs were made to test this in \textit{Pax3} co-transfection assays. The \textit{Wntl} proximal promoter was obtained from pMT86 (R. Nusse, kind gift) as a 4.9kb BamHI / NcoI fragment and cloned into the BglII / NcoI sites of the pGL3basic (Promega) Luciferase reporter vector. The NcoI site in the \textit{Wntl} sequence contains the ATG start codon of the \textit{Wntl} gene. Cloning into the NcoI site of pGL3basic (also the ATG of the Luciferase ORF) positions the \textit{Wntl} promoter in an identical position to the Luciferase ORF as it is to the translational start of \textit{Wntl} \textit{in vivo}. pGL3basic contains a Firefly Luciferase ORF in a promoterless construct to enable the investigation of putative promoter elements. This construct was named pW1PLuc and is illustrated in Figure 3.5A. This was co-transfected into NIH-3T3 cells with either the pCMV-\textit{Pax3} expression construct or empty vector. Four triplicate runs were completed for these assays, and the normalised luminescence baselined to pGL3basic background to generate fold induction. A clear induction of Luciferase activity concomitant with \textit{Pax3} expression can be seen in Figure 3.5A.

To compare this induction with the results obtained with the pW1ELuc construct, the data was rebased to uninduced samples (i.e. co-transfections with empty vector) and illustrated in Figure 3.5B. It was concluded from this analysis that \textit{Pax3} is able to cause induction of a reporter gene via the proximal promoter of \textit{Wntl}, and not via the distal enhancer element, \textit{in vitro}.
Figure 3.5:
A) Induction of luciferase activity from pW1PLuc on co-transfection with pCMV-Pax3. Data is normalised luminescence expressed as fold induction over pG3basic (pW1PLuc without the 5.2kb Wnt1 promoter fragment). n = 4, error bars are SEM.
Figure 3.5:
B) Comparison of Pax3 induction of pW1PLuc and pW1ELuc. Data is normalised luminescence, fold induction over baseline (respective luciferase vectors without Wnt1 regulatory sequence), indexed to un-induced to allow direct comparison. n = 4 and 6, respectively, error bars = SEM.
To further investigate the relationship between the Wnt1 promoter and the Pax3 protein, a series of deletion constructs were made using restriction sites native to the element. pW1PLuc was digested with Asp718 to remove 3.7kb of sequence distal to the Luciferase ATG. The backbone was then re-ligated to generate pW1P(1.2)Luc. Similarly, twin ApaI sites were used to remove the proximal 3.1kb of the Wnt1 promoter to generate pW1P(-1.8)Luc. In this latter construct the Pax3 consensus site is also removed. These constructs are illustrated in Figure 3.6A. Each of these constructs was then co-transfected into NIH-3T3 cells with either pCMV-Pax3 or empty vector, as previously described, to assay any transcriptional response to the Pax3 protein. These results can be seen in Figure 3.6B. Clearly, whilst both pW1PLuc and pW1P(1.2)Luc retain their response to Pax3 expression, the pW1P(-1.8)Luc construct does not. Since this latter plasmid has lost the Pax3 consensus site identified in Figure 3.4, this data further argues for a functional role for this site in binding Pax3, at least in vitro. Finally, the pW1PLuc and pW1P(1.2)Luc were exhaustively tested in pCMV-Pax3 co-transfection assays using a range of expression vector quantities (this was also shown to generate a proportional range of Pax3 expression levels within the cells transfected, see Figure 4.15A). This was to ensure both specificity of Pax3 response by illustrating a concentration dependence and to test the result reported by (Chalepakis et al., 1994) which demonstrates a biphasic transcriptional response of increasing Pax3 concentrations when binding to reporter gene constructs carrying the NCAM promoter. This can be seen in Figure 3.6C. Whilst these data clearly demonstrate a concentration specific response to Pax3 gene expression over the concentration range examined, no biphasic response (i.e. loss of transcriptional inductive ability at higher Pax3 concentrations) was observed. Whilst it remains possible that in these experiments insufficient Pax3 was co-expressed to
Figure 3.6

A

Wnt1 4.9kb promoter fragment

1K 2K 3K 4K 5K 6K

Pax3

Luciferase

pW1PLuc

pW1P(1.2)Luc

pW1P(-1.8)Luc

Asp

Deleted

1K 2K 3K 4K 5K 6K

Apa

Deleted

1K 2K 3K 4K 5K 6K

B

Indexed Fold Induction

pW1Luc pW1Luc + 100ng pCMV-Pax3 p(1.2)W1Luc p(1.2)W1Luc + 100ng pCMV-Pax3 p(-1.8)W1Luc p(-1.8)W1Luc + 100ng + 100ng pCMV-Pax3

C

Indexed Fold Induction

ng pCMV-Pax3
Figure 3.6:
A) Schematic of pW1PLuc and deletion constructs pW1P(1.2)Luc and pW1P(-1.8)Luc, Asp = Asp718, Apa = Apal sites
B) Luciferase data comparing the induction of the Wnt1 promoter deletion constructs with 100ng pCMV-Pax3 vs. empty vector. Data is normalised luminescence expressed as fold induction over baseline (pG3basic) and indexed (empty vector = 1) to enable comparison between constructs. n > 3 for each data point.
C) Variation in Luciferase induction from pW1PLuc and pW1P(1.2)Luc over a range of Pax3 concentrations (expressed as ng pCMV-Pax3 plasmid used in each transfection). Data is normalised luminescence expressed as fold induction over baseline (pGL3basic) and indexed to enable comparison between experiments n = 5 - 11 (depending on data point), error bars represent SEM.
generate this biphasic response, the apparent plateau of transcriptional induction from 250ng of pCMV-Pax3 per well would suggest this is not the case. Perhaps the mode of Pax3 transcriptional induction on the Wnt1 and NCAM promoters differs, explaining this discrepancy.

The data presented above strongly argue that Pax3 can bind to and transactivate the Wnt1 promoter in vitro, and that this interaction occurs via the Pax3 consensus site identified. To show that the Pax3 consensus site is necessary and sufficient for this observed Pax3 induction, site directed mutagenesis was performed to ablate the core binding residues, and replace them with a low affinity sequence. Primers were designed to mutate the TCGC core Pax3 binding consensus (as defined by MatInspector and (Chalepakis and Gruss, 1995; Phelan and Loeken, 1998; Underhill, 1997), etc) to the sequence ATAA. The actual primers held a mutation in an additional base 5' to the TCGC sequence, altering the native sequence CTCGC to GATAA. The reason for this additional change is that the mouse promoter sequence contains a tandem repeat of the core consensus; TCGCTCGC. It was feared that the preceding TCGC might be able to substitute for the mutated core consensus, so the final base of this was also mutated. An alignment of these sequences can be seen in Figure 3.7A. Site directed mutagenesis was performed on pW1P(1.2)Luc using a QuikChange Site Directed Mutagenesis kit (Stratagene), and the primers W1_sub_FWD and W1_sub_REV (see Appendix 2). Clones were screened by sequencing (from the GL3 primer within the Luciferase ORF (Promega)) and aligned to the wild type sequence to confirm the specificity of the mutagenesis. The resulting plasmid was named pW1PALuc.
Figure 3.7:
A) Sequence alignment of wild type and mutated Pax3 core consensus sequence on the plasmids pW1P(1.2)Luc and pW1P<delta>, respectively
B) Luciferase data contrasting luciferase induction on co-transfection with pCMV-Pax3 between pW1P(1.2)Luc and pW1P<delta>. Mutant Wnt1 promoter shows a complete loss of Pax3 induction; co-transfection experiments with wild type promoter were conducted in tandem to ensure any loss of luciferase induction in the mutant could not be attributed to other factors. n = 3, error bars are SEM.
Co-transfection experiments, using pCMV-$Pax3$ at levels previously shown to elicit a strong induction of Luciferase activity from the un-mutated plasmid (100ng) were then performed using pW1PALuc as the reporter construct. A total loss of $Pax3$ induction was observed on co-transfection with the pCMV-$Pax3$ plasmid. To ensure this loss of induction was due to the consensus site mutation and not a general failure of the Luciferase assay system, experiments using the un-mutated pW1PLuc reporter construct were run in tandem. No loss of $Pax3$ induction was observed. These data can be seen in Figure 3.7B.

It was concluded from these experiments that the transcriptional induction driven from the $Wnt1$ promoter by $Pax3$ was due to the interaction of the protein with this consensus binding site.

Electrophoretic mobility shift analysis (EMSA) was not performed on this sequence. This was because the mutation incorporated into the pW1PALuc plasmid was based on one made by (Kwang et al., 2002). In this study, the core consensus site for $Pax3$ binding to the $Msx2$ promoter, TCAC (TCGC here) was mutated to GATA (GATAA here). This mutation was used to show the specificity of $Pax3$ binding in a series of EMSA experiments where this mutation causes a complete loss of $Pax3$ binding. Given the similarities of the core consensus sites in question, and the nature of an EMSA experiment (i.e. purified protein binding to radiolabelled oligonucleotide), an EMSA performed using the $Wnt1$ promoter $Pax3$ consensus described above would be an almost exact copy of the experiment reported by this group. It was decided that an EMSA could not generate a sufficiently informative result in this context and was therefore omitted.
Discussion (Hypothesis One)

The above studies address several of the issues raised by Hypothesis One. Pax3 does appear to be able to up-regulate the expression of a reporter gene, at least in vitro, using the Wnt1 regulatory elements and via a defined consensus site. This interaction does not occur via the 3’ distal enhancer element but by the 5’ proximal promoter, as shown here. These experiments do not confirm the occupation or biological relevance of this site in vivo. The experimental data reported in Chapters 5 and 6 will complete the investigation of this hypothesis by addressing this final issue.

Regulation of Wnt1 by Msx2

To investigate the potential role for Wnt1 in the regulation of the cardiac neural crest, as implicated by both the Sp2H mutant phenotype (Conway et al., 2000) and Dvl2 and Pitx2 knockout phenotypes (Hamblet, 2002) and (Kioussi, 2002), the possibility of a further regulatory interaction between Msx2 and Wnt1 was examined, given the data presented in (Kwang et al., 2002).

Hypothesis Two
Wnt1 transcription is directly downregulated by Msx2, most probably by the distal 3' enhancer element, in vivo with implications for the normal development of the cardiac neural crest.

To investigate this possibility in vitro, a series of luciferase co-transfection experiments were conducted in a manner analogous to that described above. Firstly, an Msx2 expression vector was required. A search of the HGMP clone database was conducted and a potential full length Msx2 cDNA was identified. This was ordered and sequenced to confirm this was indeed a full length and mutation free Msx2. The vector within which the cDNA was supplied was a CMV expression vector (pCMV-Sport6 (Invitrogen)) analogous to the one used to generate the pCMV-Pax3 construct. Sequencing also confirmed that the Msx2 cDNA was in the correct orientation and position to enable its expression to be driven by the CMV promoter on transfection into a mammalian cell line. This construct is illustrated in Figure 3.8.

NIH-3T3 cells were transfected with pCMV-Msx2, and a Western blot using an anti-Msx1/2 monoclonal obtained from the Developmental Studies Hybridoma Bank (DSHB) was then conducted on the protein extracts to confirm expression. This can be seen in Figure 3.8; the predicted molecular weight of Msx2 is ~28kDa (ensembl). This confirms the expression of Msx2 in cells transfected with the pCMV-Msx2 plasmid.

The MatInspector analyses of putative transcription factor binding sites in the Wnt1 promoter and enhancer elements were then re-examined to identify any Msx2 consensus in these regions. This analysis revealed no Msx2 sites in the Wnt1 promoter.
Figure 3.8:
Schematic of pCMV-Msx2 and Western blot confirming expression in transfected NIH-3T3 cells using anti-Msx1/2 from DSHB resource.

![Diagram showing schematic and Western blot results for pCMV-Msx2 expression in NIH-3T3 cells.](image-url)
region, but an Msx2 consensus site in the distal 3' enhancer. Furthermore, this site was found to lie within the conserved 110bp element shown by (Rowitch et al., 1998) to control Wnt1 expression in the early neural plate and developing brain. This site is illustrated in Figure 3.9B.

To investigate this further, co-transfections of the pCMV-Msx2 vector and the pW1PLuc and pW1P(1.2)Luc reporter plasmids were performed. No consistent response to Msx2 co-expression could be observed over a range of pCMV-Msx2 concentrations with either plasmid, see Figure 3.9A. The pCMV-Msx2 construct was then tested with the pW1ELuc plasmid, see Figure 3.9B. This experiment illustrated a potentially interesting and highly significant repression of around 50% at 250ng pCMV-Msx2, n = 3.

To analyse this interaction further, a series of deletion constructs were made of the Wnt1 enhancer. In these constructs the 5.5kb Wnt1 enhancer fragments present in pW1ELuc was restricted further by BamHI and SpeI digests to extract internal fragments. The backbone was then re-ligated to generate the serial deletion plasmids pW1E(bam)Luc and pW1E(spe)Luc. These are illustrated in Figure 3.10A. Neither of these constructs carry a deletion for the Msx2 consensus site predicted by MatInspector analysis. Deleting a small region containing this site in the plasmid by restriction digest proved impossible. Two MbiI sites were observed flanking a small region containing the Msx2 consensus and the highly conserved 110bp element within the larger enhancer region. Since the presence of other MbiI sites elsewhere within the backbone prevented a conventional deletion, the inverse experiment was performed. In this case, the 816bp MbiI fragment was then extracted from a digest of
Figure 3.9

A. Wnt1 4.9 kb promoter fragment with Pax3 and Luciferase reporter constructs.

B. Graph showing fold induction of luciferase activity with varying ng of pCMV-Msx2.
Figure 3.9:
A) Wnt1 promoter constructs used to test the response of the Wnt1 promoter to co-expression with Msx2 (from pCMV-Msx2). Luciferase response to co-expression with Msx2. Data expressed as fold induction over baseline (pGL3basic). n = 3, error bars represent SEM.
B) Wnt1 enhancer construct used to test the response of the Wnt1 enhancer to co-expression with Msx2. Luciferase response expressed as fold induction over baseline (pGL3promoter). n = 3, error bars represent SEM.
the pW1ELuc plasmid and then shuttled into an empty pGL3Promoter vector restricted with Smal. This not only extracts the Msx2 containing fragment from the Wnt1 enhancer, but also places it within a different context within the Luciferase reporter vector. Since this region has been shown to behave as a canonical enhancer element by other groups (Echelard et al., 1994; Rowitch et al., 1998) and is therefore thought to be position and orientation independent, this was hoped to have no adverse effect on the ability of this element to respond to Msx2. This construct is also illustrated in Figure 3.10A.

Co-transfection with pCMV-Msx2 expression plasmid revealed that neither the pW1E(bam)Luc nor the pW1E(spe)Luc reporter plasmids demonstrated any significant loss of Msx2 inhibition. Both of these plasmids retain the Msx2 consensus site. When Msx2 was co-expressed with the pW1E(mbi)Luc plasmid with the Msx2 containing fragment shuttled into a new vector, an almost identical level of repression was observed. These data can be seen in Figure 3.10B and 3.10C. (N.B. Previous controls (see the Luciferase Assay subsection in Chapter 2) were performed to ensure that the SV40 weak promoter carried on the pGL3promoter vector did not respond to Msx2 co-expression).

From the above it was concluded that the Wnt1 enhancer, and not the Wnt1 promoter, was able to respond to Msx2 to repress transcription of a reporter gene. Furthermore, this repressive response could be mapped to an 816bp region containing an Msx2 consensus binding site within an area of the Wnt1 enhancer demonstrated by previous studies to be both highly evolutionarily conserved and of great regulatory
Figure 3.10

A

SV40 Luciferase

Wnt1 enhancer 110bp region

Msx2

pW1ELuc

B

SV40 Luciferase

Deleted 110bp region

Msx2

pW1E(Bam)Luc

C

SV40 Luciferase

Deleted 110bp region

Msx2

pW1E(Spe)Luc

Msx2

pW1E(mbi)Luc

B

Fold Induction

0.7

0.6

0.5

0.4

0.3

0.2

0.1

0

pw1E(bam)Luc

pw1E(bam)Luc + 100ng Msx2

pw1E(Spe)Luc

pw1E(Spe)Luc + 100ng Msx2

pw1E(mbi)Luc

pw1E(mbi)Luc + 100ng Msx2
Figure 3.10:
A) Schematic of Wnt1 enhancer deletion constructs used to dissect the response of the Wnt1 enhancer to co-expression with Msx2. B = BamHI, S = SpeI, M = MboI sites used in cloning
B) Luciferase response of deletion constructs to co-expression with 100ng pCMV-Msx2.Data expressed as fold induction over baseline (pGL3promoter), n = 3,error bars represent SEM
C) Luciferase response of pW1E(Mbi)Luc construct to co expression with 100ng pCMV-Msx2.Data expressed as fold induction over baseline (pGL3promoter), n = 3,error bars represent SEM
significance. Unfortunately, attempts to mutate this site with site directed mutagenesis were not be performed due to time constraints within this project.

**Pax3 and Msx2: a regulatory relationship at the Wnt1 locus?**

Any possible role of *Wnt1* regulation in the development of the cardiac neural crest is likely to be extremely complex. The above study, however, has demonstrated that, *in vitro*, there is the possibility that *Pax3* may act to induce the expression of *Wnt1* via its promoter and that *Msx2* may conversely inhibit *Wnt1* expression via its 3' enhancer. The basis for much of this hypothesis in terms of the cardiac neural crest is derived from a study by (Kwang et al., 2002) where they demonstrate that *Msx2* is a downstream target for negative regulation by *Pax3*, and its expression domain is expanded in the *Sp2H* mutant background (See Chapter 1 for a full discussion of these findings). It would be interesting to examine the possibility that the inhibition of *Wnt1* expression by *Msx2* can further be overridden by the induction of *Pax3* binding to the *Wnt1* promoter. This would give *Pax3* two mechanisms through which to ensure the correct level of *Wnt1* expression in the wild type; both inhibition of *Msx2* expression, as established by (Kwang et al., 2002), and also through the alleviating the repression of *Wnt1* by *Msx2* via an interaction on the regulatory elements of this gene.

To begin to examine this possibility *in vitro*, a construct was made carrying both the *Pax3* responsive *Wnt1* promoter and the *Msx2* responsive *Wnt1* enhancer. This construct was named pWIEPLuc and was made by shuttling the *Wnt1* promoter fragment from pW1P(1.2)Luc as an Asp718 – NcoI fragment into the same sites on the pWIE(bam)Luc reporter construct. This construct is illustrated in Figure 3.11A.
Figure 3.11:
A) Schematic of construct pW1EPLuc, A = Asp718 N = Ncol sites used in cloning
B) Luciferase response of pW1EPLuc on co-transfection with 250ng pCMV-Script (blue), 250ng pCMV-Msx2 (orange) and 250ng of both pCMV-Msx2 and pCMV-Pax3. n = 9, 3 and 3, respectively. * p = 0.0072, ** = 0.0001 (two sample t test)
pW1EPLuc was co-transfected with pCMV-Msx2 to show that Msx2 was still able to significantly inhibit the transcription of the Luciferase reporter in this context. Next pCMV-Pax3 was also co-transfected with pCMV-Msx2 and it was demonstrated that the induction by Pax3 via the Wnt1 promoter was clearly able to override any inhibitory effect mediated by Msx2 in this experimental system, see Figure 3.11B.

This enables the very tentative conclusion that Pax3 may be able to alleviate any repression of Wnt1 expression generated by Msx2. Obviously, there are numerous substantial flaws with this final piece of analysis, and these results should be interpreted with great caution. Firstly, the relative levels of expression of Msx2 and Pax3 from their respective expression constructs could not be meaningfully assessed; even Western blotting could not control for differing affinities of the two antibodies. It is impossible in this system to control for the levels of protein produced. An alleviation of repression by Pax3 could simply be a reflection of the pCMV-Pax3 plasmid producing significantly more protein than pCMV-Msx2. Secondly, the observable level of repression from Msx2 co-expression with the pW1EPLuc reporter is inherently much lower than with the pW1ELuc family of reporters. This is because the endogenous activity of the Wnt1 promoter in the NIH-3T3 line is very low compared to that driven from the SV40 weak promoter found on all pGL3promoter derived vectors used to generate the pW1ELuc vectors (no cell line naturally expressing Wnt1 has been reported). Any reduction of expression in this context is necessarily going to be difficult to observe. Finally, this experiment contains four variables (two regulatory elements and two transcription factors) and is therefore almost impossible to meaningfully control. This data should be treated with extreme
caution and is presented here merely as a final observation of a potential relationship, rather than as proof of any interaction.

**Discussion (Hypothesis Two)**

The above studies address the issues raised by hypothesis two in a number of ways. It was demonstrated that *Msx2* can repress reporter gene expression *in vitro* and that this occurs via the 3' enhancer element as predicted. The responsive region was mapped to an 816bp MbiI fragment containing an *Msx2* consensus, but mutagenesis of this fragment was not conducted. It is concluded that this repression is likely to be mediated by this binding site, in light of the other evidence presented. Finally, evidence is provided that the inductive ability of *Pax3* mediated by the *Wnt1* promoter may be able to override the inhibitory effects of *Msx2* on the *Wnt1* enhancer, although this is a tentative conclusion. None of the above experiments demonstrate occupancy of these sites *in vivo*. This aspect of hypothesis one and hypothesis two is addressed in Chapter 6.
Chapter Four: Pax7 regulation

Introduction

In this chapter regulatory elements for the Pax7 gene are described, and potential direct regulatory interaction between these and the Pax3 protein tested and mapped *in vitro*. Pax7 represents an interesting potential target of Pax3 regulation not only due to its developmentally dynamic expression pattern (Jostes et al., 1990) or role in the development of regenerating muscle from an adult stem cell population (Seale et al., 2000), but also due to the existing controversy over the relationship between these two genes in the mouse (Borycki et al., 1999; Relaix et al., 2004). A range of both wet and dry approaches were used to characterise the mouse Pax7 genomic locus, identify a number of interesting features and map a candidate region for both a 5' proximal promoter and a regulatory feature within intron 1. These were then verified experimentally in a Pax7 expressing cell line (C2C12) and a Pax3 response determined.

Pax7 regulatory regions: Bioinformatics

Since no information on the mouse Pax7 genomic locus had been reported, and no regulatory regions had been experimentally verified, a bioinformatic analysis of the Pax7 gene was performed. Using the mouse genome sequence database 100kb on either side of the predicted mouse Pax7 gene was downloaded and used in subsequent
analyses (ensembl). Firstly, the exons and introns were mapped using the published mouse cDNA sequence (Ziman et al., 2001). The mouse Pax7 locus is shown in Figure 4.1A, and shows a similar genomic organisation to that reported for human PAX7 (Vorobyov et al., 1997).

In the search for gene regulatory elements, the obvious starting point is the 5' end of the gene which is likely to contain a minimal promoter driving transcription. To this end, ~10kb around exon 1 (N.B. since the 5' UTR for the mouse Pax7 gene had not been determined experimentally, ‘exon 1’ refers to the coding sequence from the ATG start site through to the first intron boundary) was subjected to a NIX analysis (HGMP). This uses a range of sequence analysis software to identify putative features of genomic sequence, taking into account species of origin and size of fragment. This analysis identified a CpG island, a putative promoter region and a potential RNA polymerase II start site, all 5' proximal to exon 1 (within 1kb upstream). These are illustrated in Figure 4.1B.

To take this analysis a step further, the mouse Pax7 locus was then analysed in tandem with the human Pax7 genomic region. Comparative sequence analysis can be used to identify likely regions of regulatory significance in non-coding DNA on the basis of sequence conservation. The assumption is that sequence changes in important genomic elements (i.e. coding or regulatory sequences) will be strongly selected against in evolution. Similarly, sequences encoding no genetic information can be mutated without any adverse effect on the fitness of the organism and are therefore not selected against. If a region of non-coding DNA is conserved at the sequence level between two species which shared a common ancestor millions of years ago, this
Figure 4.1

A

1 3

2 4

80K 100K 120K 140K 160K 180K 200K

B (inset)

60K

51K

50K

49K

48K

47K

48K

NIX PolII Start

NIX Promoter

DotplotID Regions

NIX CpG1

PiPMaker CpG Island

1 3

2 4

14K 45K 46K 48K 49K 50K 51K 52K

140K 160K 180K 200K

1

III
Figure 4.1

CpG >60%

Window Size = 100

D

Dotmatcher: mouse vs human

ATG...
Figure 4.1:
N.B. Exon 1 in all figures starts at ATG of coding sequence.
A) Annotated Pax7 locus, spanning 200kb. Red boxes are exons predicted from cDNA sequences, shaded blue boxes represent areas of significant PIPMaker identified homology of >60% between human and mouse genomic sequences (see Figure 4.1E, and appendix 3 for a complete output).
B) Inset of 5’ end of Pax7 locus shown in A. Shows relative positions of exons 1 - 4, CpG islands predicted by NIX and PIPMaker analyses, RNA PolII and promoter sites identified by NIX, and regions of homology identified by Jemboss Dotmatcher software (see Figure 4.1D)
C) %C+G plot for 3kb around Pax7 exon 1 (upper) and corresponding genomic locus (lower) showing NIX identified features. %C+G calculated as moving average from window of 100 bp. Note increase in CG content proximal to the 5’ end of exon 1 (red line).
D) Dotmatcher dotplot of human (y axis) vs mouse genomic (x axis) Pax7 locus upstream of ATG. Sequences of high position and sequence identity represented by a diagonal line of dots between the two axis. Results summarised graphically in Figure 4.1B. (Window size = 10, Threshold = 23, Blossum matrix)
E) PIPMaker output for the first 40kb of the Pax7 locus. Showing CpG islands, repetitive sequences, exons and percentage identity to human sequence (dots in vertical dimension running beneath genomic locus in horizontal plane represent % identity at that point)
homology can be interpreted as a potentially important site of regulation (see (Cooper and Sidow, 2003; Wasserman and Sandelin, 2004) for review).

Firstly a simple alignment program, representing regions of homology between two sequences as a dot plot was performed. Regions of significant homology are summarised in Figure 4.1B and the Dotmatcher output (HGMP) shown in Figure 4.1D for the 5kb immediately upstream of exon 1. The presence of a CpG island in this region also potentially infers a regulatory significance (Cross and Bird, 1995). The MacVector program was then used to independently analyse the CG content regions upstream of exon 1, and also identified a CpG island in the upstream 1kb (Figure 4.1C). Finally, the PIPMaker suite was used to compare the entire genomic Pax7 locus between human and mouse (Schwartz, 2000). The sample output can be seen in Figure 4.1E, and complete output in Appendix 3. This program found clear regions of homology in the sequences upstream of exon 1, and also identified a CpG island across the 2.5kb surrounding exon 1 (Figure 4.1B). The PIPMaker output can be seen in summary in Figure 4.1A as blue boxes demarcating regions of high human - mouse homology (>60%) next to the corresponding genomic region. Since a number of different bioinformatics programs had identified significant homologies between human and mouse sequences, and several other features of potential regulatory importance, it was concluded that the 1-1.5kb upstream of exon 1 provided an attractive candidate location for a Pax7 promoter.

**Pax7 regulatory regions: 5'UTR mapping**
Bioinformatics can only identify regions of potential regulatory interest. To clone and test the Pax7 promoter, it was necessary to first characterise the 5' untranslated region (5'UTR) of the Pax7 transcript. This is because the regions of homology identified upstream of the coding exon 1 may represent novel upstream non-coding exons, rather than a promoter element, therefore positioning the transcriptional start site for the gene elsewhere. A schematic of the Pax7 transcript, with a hypothetical 1kb 5'UTR drawn for illustrative purposes, can be seen in figure 4.2A. Several groups have studied the Pax7 5'UTR in humans, using two different techniques, and these results are summarised in figure 4.2B. In all cases, the 5'UTR reported represented an extension of the first coding exon, and not a novel upstream non-coding exon.

To investigate the 5'UTR of the mouse Pax7 gene a number of different approaches were employed. Firstly, primer extension (PEXT) was used to size the 5'UTR of the mouse Pax7 transcript. A schematic of this protocol can be seen in Figure 4.3A. Briefly, an antisense primer (kinase labelled at the 5' end with $^{32}$P) anneals to the Pax7 mRNA coding sequence, and the reverse transcriptase (RT) enzyme is then used to extend a cDNA from this point to the end of the 5' end of the 5'UTR. This is a quick method of approximately sizing the 5'UTR of Pax7 from a number of different sources. These results are shown in Figure 4.3B. Several gels are shown illustrating results from using different Pax7 mRNA sources and extending with different primers. Two 5'UTR sizes were repeatably (7 PEXT experiments, 4 different RNA sources / preparations) obtained from this method, corresponding to a 5'UTR of 300 – 500bp and another of around 100bp (after correction for the positioning of the antisense primers within the coding region of Pax7). These are referred to as High $M_r$ and Low $M_r$, respectively. The Low $M_r$ band indicates a 5'UTR corresponding to one
Figure 4.2:
N.B. Pax7 5'UTR shown is for illustrative purposes only
A) Mouse Pax7 transcript illustrating positions of all 9 exons, 5'UTR, start (ATG) and stop (TGA) codons and 5'Cap (5' terminal 7-methylguanine) at the very distal end of the mRNA
B) Table of published data to date highlighting source material, technique used and 5'UTR length obtained

<table>
<thead>
<tr>
<th>Reference</th>
<th>Source RNA</th>
<th>Experimental Method</th>
<th>5'UTR size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schafer, B. et al</td>
<td>Human A673 cells</td>
<td>5' RACE</td>
<td>103 bp</td>
</tr>
<tr>
<td>1994</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vorobyov, E. et al</td>
<td>Human ARMS tumor</td>
<td>5'RACE</td>
<td>599 bp</td>
</tr>
<tr>
<td>1997</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murmann, O. et al</td>
<td>Human A673 cells</td>
<td>5'RACE</td>
<td>95bp to 60bp</td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td>Primer Extension</td>
<td>70bp</td>
</tr>
<tr>
<td>Syagailo, Y. et al</td>
<td>Human skeletal</td>
<td>5'RACE</td>
<td>655bp</td>
</tr>
<tr>
<td>2002</td>
<td>muscle</td>
<td>Primer Extension</td>
<td>640bp and 664bp</td>
</tr>
</tbody>
</table>
Figure 4.3:
N.B. Pax7 5'UTR shown is for illustrative purposes only
A) Illustration of 5' end of Pax7 mRNA, showing positions of primers EX2, B2 and B4 used in primer extension experiments shown in 4.3B and others. Primers are complementary to coding sequence and anneal 78bp (EX2), 171bp (B2) and 183bp (B4) downstream of the ATG at their 5' ends. Primers are radiolabelled at the 5' end, annealed to Pax7 mRNA and RT activity used to extend along the 5'UTR to the 5'CAP. ssDNA products are then ran out on acrylamide and sized according to kinase labeled ladder. Size of fragment - distance from 5' end of primer to ATG = size of 5'UTR
B) Representative images of primer extension gels, showing banding pattern obtained (pBR322(Mspl) ladder used in every case)
Bi and Bii:
Primer EX2 used. Lane 1 = C2C12 total RNA, 2 = C2C12 mRNA, 3 = C2C12 total RNA, 4 = E11.5 total RNA, 5 = E11.5 total RNA from a Pax3 -/- embryo. Illustrating PEXT products at 350 - 600 bp (high Mr) and 160 - 180 (low Mr)
Biii)
All lanes are E11.5 mRNA, Primers EX2, B2 and B4 used as marked. Illustrating PEXT products for EX2 at high and low M, as for Bi and ii (red lines), and corresponding bands for primers B2 and B4 (blue lines), taking into account differing positions of primer annealing sequences within the Pax7 transcript (green arrow designates a non-specific band caused by gel drying)
Figure 4.3

A gel electrophoresis diagram showing two sets of lanes labeled Bi and Bii. Each lane contains protein bands at molecular weights indicated by markers. Lanes 1 and 2 are labeled Bi, with bands at 622, 404, 307, 240, 180, 160, 147, 127, and 110. Lanes 3, 4, and 5 are labeled Bii, with bands at 622, 404, 307, 240, 180, 160, 147, 127, and 110. The labels 'high M_' and 'low M_' indicate the molecular weight markers for high and low molecular weight proteins, respectively.
Figure 4.3

- High Mr
- Low Mr

ns
previously published for human Pax7, see Figure 4.2B, but the High M, bands are novel.

Since the 5'UTR sizes obtained in the PEXT experiments represented fragment sizes which would be easily amplifiable by PCR, a 5'RACE (Rapid Amplification of cDNA Ends) methodology was then used to try and clone and sequence these fragments from wild type embryo RNA. Two different kits were used, see Chapter Two, to minimise experimental artefacts due to specific methodologies. Nested PCR was used from the Pax7 coding region, see Figure 4.4A, and the product TOPO-cloned for sequencing. Over 40 clones were sequenced and 18 of these were found to be carrying an insert. The results can be seen in the alignment in Figure 4.4B. A range of 5'UTR sequences were cloned, ranging from 93bp (which would correspond to the Low M, bands observed in the PEXT experiments) to 11bp. All of these sequences align exactly to the genomic sequence upstream of the Pax7 coding region.

Interestingly, no clone which would correspond to the High M, bands observed in the PEXT gels could be identified by 5'RACE. Bands of ~500bp were sometimes seen when running out amplification products, but these proved to be unclonable in spite of several attempts.

Finally, in an attempt to characterise the High M, bands from the PEXT experiments and confirm the results from the 5'RACE, a Ribonuclease Protection Assay (RPA) was performed. A riboprobe was made by amplifying 770bp of the Pax7 upstream sequence by PCR, see Figure 4.5A, using primers P7RT1L/R. This was then sub-cloned into pBSII(KS) as a 686bp Asp718–FspI fragment to generate pP7RiboProbe,
Figure 4.4:
N.B. Pax7 5'UTR shown is for illustrative purposes only
A) Schematic of 5'RACE reaction: nested PCR using two gene specific primers annealing at the 5' end of the Pax7 transcript were used to amplify the Pax7 5'UTR with primers annealed to the distal end after an appropriate RT step
B) Alignment of the 5'RACE products obtained. Reference line (bottom) represents mouse genomic sequence from ensembl and positions ATG of coding sequence. Sequences color coded to enable visualization of alignment.
and confirmed by sequencing. Figure 4.5A illustrates the position of this riboprobe and the predicted sizes of protected fragments which would be obtained from the High M_r, Low M_r, and 5'RACE 5'UTR's (assuming the High M_r bands are not obtained from novel exons upstream of this probe binding site). The RPA gel is shown in Figure 4.5B for a C2C12 total RNA sample and an E11.5 total RNA sample. Protected fragments present in both RPA's are at 74, 105 and 135bp, in agreement with the PEXT and 5'RACE data. No larger protected fragment which would correspond to the High M_r bands observed in the PEXT gels could be detected.

To summarise the above data, a plot of 5'UTR length and frequency of observation has been drawn in Figure 4.6. This shows that all 3 methods employed agree on a 5'UTR length of around 100bp. Both 5'RACE and RPA experiments indicate this 5'UTR represents an upstream extension of the coding region and not a novel 5' exon for Pax7. This is in agreement with some data published for the human Pax7 locus (Murmann et al., 2000; Schafer et al., 1994). This UTR also contains a stop TGA codon on the + strand 40bp upstream of the Pax7 ATG. The existence of several shorter 5'UTR's determined by 5'RACE, and several repeatable larger PEXT fragments is puzzling, however. It is worthy to note that this region of genomic DNA is extremely CG and repeat rich. These sequences are reportedly often difficult for RT enzymes and polymerases to metabolise properly in vitro, and may explain the foreshortened 5'RACE products observed. The reliance of this method on PCR and cloning, which may preferentially select for shorter products, could explain why these fragments are observed in RACE clones but not by either PEXT or RPA. The repeated observation of a larger set of PEXT products is more difficult to explain.

Whist several groups have reported longer 5'UTR's for the human Pax7 gene
Figure 4.5

Lanes:
1 = Positive control, undigested probe
2 = Negative control, (NIH-3T3 RNA)
3 = pBR322(Mspl) ladder
4 = pBR322(Mspl) ladder
5 = RPA, C2C12 RNA
6 = RPA E11.5 embryo RNA

- ~135
- ~105
- ~74
Figure 4.5: N.B. Pax7 5'UTR shown is for illustrative purposes only.
A) Illustration of Pax7 mRNA, showing 5'UTR's predicted so far by 5'RACE and primer extension (PEXT). The riboprobe cloned and used in the RPA shown in B is positioned in blue (685bp). The predicted fragments are shown below (to scale). A 5'UTR in agreement with RACE and the smaller primer extension products would generate protected fragments of around 100bp. 5'UTR's of the size predicted by the larger primer extension product would generate protected fragments of around 350 - 500 bp (only 500bp fragment shown here for simplicity) from the same riboprobe. Populations of transcripts containing both should generate fragments of both sizes.
B) RPA gel, visualized on a phosphorimager. Band and ladder sizes as marked, lane key as shown. Note no evidence of any bands corresponding to the larger primer extension products shown in Figure 4.3.
Figure 4.6:
Summary graph of all experiments performed to determine the length of the mouse *Pax7* 5'UTR. The graph is a plot of length of 5'UTR (x axis) against frequency of experimental observation (y axis). Primer extension is represented by the blue trace (n = 40), RPA is pink (n = 6) and 5'RACE is yellow (n = 18). All methods cluster a predicted 5'UTR at around 100 bp (in the region reported by Schafer et al 1994 and Lang et al 2003) but primer extension repeatedly determines a longer extension products from ~300 to 500 bp (in the region reported by Vorobyov et al 1997, Murmann et al 2000 and Syagailo et al 2002) which has not been detected by either 5'RACE or RPA.
(Syagailo et al., 2002; Vorobyov et al., 1997), these are all at least 150bp bigger than the bands observed in these PEXT experiments, and since these were also 5' extensions of the Pax7 coding sequence and not novel upstream exons, these should also have been detected by RPA. It is possible that these PEXT bands represent 5'UTR's of genes other than Pax7. Since three different primers all show extension products in this region, the most likely candidate for this would be Pax3, given the level of homology of between Pax3 and Pax7. C2C12 cells do not express Pax3 (data not shown and (Borycki et al., 1999)) and give PEXT products of this size. It would therefore seem unlikely that this PEXT product represents an accidental extension of the Pax3 5'UTR. Furthermore, all primers were subjected to BLAST analysis to ensure specificity during the design process. The data presented here does not adequately explain this observation, and further investigation is needed. Given the convergence of all three methodologies on a transcriptional start site of around 100bp upstream of the Pax3 coding sequence, it was decided for the purposes of this thesis to take this as the major transcriptional start site for mouse Pax7.

**Pax7 5' Promoter: Cloning**

Since the 5' extent of the Pax7 transcript was taken to include around 100bp of sequence upstream of the coding region, it was reasoned that the genomic sequence upstream of this, containing a CpG island and a high level of homology to the equivalent region in the human genome, might contain a minimal promoter for the mouse Pax7 gene. In the human, a similar homologous region has been shown to confer Pax7 promoter activity in cell culture (Murmann et al., 2000; Syagailo et al., 2002).
To investigate this experimentally, this region of the Pax7 locus was cloned. Initially, a mouse genomic PAC library was screened using a 520bp probe (made by amplifying human genomic DNA (primers P7HYB1L/R) and TOPO-cloning into pCR-II-TOPO-TA to generate pP7HYB). The predicted hybridisation region is illustrated in Figure 4.7. Several positive clones were identified and these were then confirmed by Southern blot. Figure 4.8A and B illustrate the position of the probe used (a 770bp PCR product amplified from mouse genomic DNA using primers P7RT1L/R) and fragment sizes predicted from the mouse genomic sequence. PAC clones were digested using the restriction enzymes indicated and Southern blotting performed, see Figure 4.8C. Clone 14 gave positive signals of the correct size in each case and was used to clone the Pax7 promoter region.

Since this region proved resistant to cloning via conventional methods, TEPCR was used to TOPO clone a 3.5kb fragment into pCR-XL-TOPO using primers PAX7PROML/R to generate p3.5TOPO. The insert was then sequenced completely, and a 3kb (BamHI – FspI) putative promoter fragment was then shuttled into pGL3basic, (via an intermediate construct, see appendix one), generating p3.0Luc, see Figure 4.9A and B. To test the proximal 1kb region of human – mouse homology initially identified by bioinformatics, two deletion constructs were made using tandem Asp718 or BglII sites to delete the distal and proximal regions of this putative promoter region, respectively (Figure 4.9B). These were named p1.5Luc (distal deletion) and p-1.8Luc (proximal deletion).
Figure 4.7: 
Pax7 genomic locus showing predicted hybridization of 520bp fragment (cloned into pP7HYB, human sequence) used to probe a mouse genomic PAC library.
Figure 4.8:
A) Illustration of Pax7 locus and predicted restriction sites used in Southern blotting to verify validity of PAC clones
B) Table of enzymes and predicted restriction fragments
C) Representative Southern blots of PAC clone I4 showing predicted fragments
To test the ability of these regions of proximal genomic sequence to drive Luciferase reporter gene expression, these plasmids were transfected into the C2C12 mouse myoblast cell line which is Pax7 positive. Figure 4.9C illustrates these results. A detailed description of the Luciferase protocol and the data treatment used in this thesis can be found in Appendix 4 and is not dealt with here. In the C2C12 cell line, the constructs p3.0Luc and p1.5Luc were clearly able to drive reporter gene expression over baseline (pGL3basic, or empty vector). In the construct p-1.8Luc, where the sequence proximal to the transcriptional start site has been deleted, no induction could be observed. This enabled the proximal 1.5kb of mouse genomic sequence to be established as a Pax7 minimal promoter in the C2C12 cell line.

During the bioinformatic analysis of this region of the mouse genome, a potential regulatory element within the first intron of the Pax7 gene was also identified. This region was highlighted by the PIPMaker program as having >91% homology between human and mouse sequences. Given that the coding sequences for Pax7 generally only demonstrate 75 – 80% identity, this represented a potentially interesting feature. This is illustrated in Figure 4.10A. To investigate experimentally, the region was cloned using TEPCR from the I4 PAC clone, with primers P7INT1REGF/B (containing BamHI sites) and directly into the BamHI site (3’ to the Luciferase ORF) of both pGL3promoter and p1.5Luc. Both orientations of the element were cloned and screened with restriction digests, before sequencing to confirm insertion. These constructs are represented in Figure 4.10B, and were named pl1RLuc and p1.5I1RLuc.
Figure 4.9

3.5kb TEPCR fragment

Promoter fragment cloned into p3.0Luc

BamHI

44K 45K 46K 47K 48K 49K 50K

FspI

A

3.0kb Pax7 Promoter

Luciferase

B

1K 2K 3K 4K

Asp718

Asp718

1K 2K 3K 4K

BglII

B

Fragment deleted in p1.5Luc

Luciferase

Asp718

Asp718

1K 2K 3K 4K

BglII

C

Fragment deleted in p-1.8Luc

Luciferase

BglII

BglII

1K 2K 3K 4K

p3.0Luc

p1.5Luc

p-1.8Luc
Figure 4.9:
A) Genomic locus of Pax7 showing region cloned by TEPCR into p3.5TOPO, and the BamHI and FspI sites then used to clone a 3kb fragment into pGL3basic via shuttling through pIntermediate1 (pBSII backbone)
B) p3.0Luc, p1.5Luc and p-1.8Luc constructs showing regions deleted and restriction sites used.
C) Luciferase data of Pax7 promoter luciferase plasmids on transfection into the Pax7 expressing C2C12 cell line. Data is normalised luminescence expressed as fold induction over baseline (pGL3basic). n = 3, error bars are SEM
Figure 4.10

A

3.5kb TEPCR fragment

>91% homology

Fwd  Rev

45K  46K  47K  48K  49K  50K  51K  52K

B

SV40

luciferase

>98%

pl1RLuc

p1.5pl1RLuc

1.5kb Pax7 promoter

luciferase

>98%

I1R element
Figure 4.10

C

Fold Induction

<table>
<thead>
<tr>
<th>pGL3promoter</th>
<th>p1.5RLuc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Fold Induction

<table>
<thead>
<tr>
<th>p1.5Luc</th>
<th>p1.5I1RLuc</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>35</td>
</tr>
</tbody>
</table>

Fold Induction

<table>
<thead>
<tr>
<th>p1.5Luc</th>
<th>p1.5I1RLuc +</th>
<th>p1.5I1RLuc -</th>
</tr>
</thead>
<tbody>
<tr>
<td>140</td>
<td>60</td>
<td>40</td>
</tr>
</tbody>
</table>
Figure 4.10:
A) Schematic of Pax7 genomic locus showing region of very high sequence conservation between human and mouse genomes (>91%). Since this homology is greater than that demonstrated for the exon sequences, primers were designed to amplify and clone this sequence into Luciferase reporter vectors (shown). B = BamHI sites engineered into PCR product and used for cloning into pGL3Promoter and p1.5I1RLuc
B) p1RLuc and p1.5I1RLuc reporter vectors used to test the effects of the highly conserved sequence found within intron one. The I1R fragment was cloned behind the Luciferase ORF in plasmids containing the SV40 weak and Pax7 endogenous promoters, respectively.
C) Luciferase data showing the silencing effects of the I1R element on reporter gene activity driven by both the SV40 and Pax7 endogenous promoters (top panels). The silencing effect is also orientation independent (bottom panel). Data is normalised luminescence, expressed as fold induction over empty vector (pGL3promoter and pGL3basic). n = 3, error bars are SEM.
To test the effects of this element on transcription, these plasmids were transfected into the C2C12 cell line as before, and levels of reporter gene expression examined with and without the I1R insert. These data are shown in Figure 4.10C. The I1R element demonstrated a robust silencing activity which was both promoter and orientation independent. This silencing activity was also observed in NIH-3T3 cells (data not shown). It was therefore hypothesised that this element may also function to regulate *Pax7* expression *in vivo* through the binding of inhibitory elements.

**Regulation of Pax7 by Pax3**

The identification of some potential regulatory regions within the *Pax7* locus now enables the final hypothesis described in Chapter One to be investigated.

**Hypothesis Three**

*Pax7 is directly downregulated by Pax3, in vitro and in vivo, via defined regulatory elements and with implications for the development of both neural and mesodermal tissues.*

To test this hypothesis, the regulatory regions identified above were initially subjected to a bioinformatic analysis to identify potential transcription factor binding sites using the MatInspector program (MatInspector). Both the minimal promoter and I1R element were analysed in this way. One *Pax3* binding consensus was identified in the proximal promoter region. Local alignment using the ClustalW program illustrated
that this region of the *Pax7* promoter is highly conserved between human and mouse, and this consensus is present in both species, see Figure 4.11A. Figure 4.11B positions this site in both p3.0Luc and p1.5Luc.

To test the *Pax3* responsiveness of this site in *vitro*, these constructs were co-transfected into C2C12 with and without the pCMV-*Pax3* expression construct, as described previously (see Chapter Three). These data are shown in Figure 4.11C. Despite relatively high variation in normalised luminescence, no significant change in reporter gene expression could be obtained for either construct. This indicates that *Pax3* cannot interact with this promoter consensus in *vitro*. To ensure this result was not an artefact of cell line activity, these experiments were also conducted in the NIH-3T3 cell line; and an identical result was obtained (data not shown).

Since the hypothesis under investigation predicts that *Pax3* down regulates *Pax7* activity, and a silencing element was identified at the *Pax7* locus, it was decided to investigate any potential interaction between *Pax3* and this element in spite of a failure of MacVector to identify any consensus binding sites. A similar approach was adopted and pI1RLuc (Figure 4.12A) was co-transfected with and without the pCMV-*Pax3* expression vector, and the changes in normalised luminescence recorded. Again, no change in fold induction was observed, Figure 4.12B, indicating this element does not respond to *Pax3* co-expression in *vitro*.

Finally, the p1.5I1RLuc construct was tested. This construct contains both the *Pax7* proximal promoter and the I1R element around the Luciferase ORF, Figure 4.13A. This was co-transfected into C2C12 cells with and without pCMV-*Pax3*, in a repeat
A

mouse  
human  

---TGGGGTTGGAGTGTTTGTTTGTTTGAACTTCCTTGTCGCCACCTTCCCTCCCCCCAA  3619
TGGGGTTGGAGTGTTTGTTTGTTTGAACTTCCTCGTCGTCGCCACCTTCCCTCCCCCCAA  3619

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

mouse human

 mouse human
CCCACCTCACCCCCCAGCTTCTGGGAAGGGGCGCGTCTGCAACCGGGTG  3856
CCTCCAACCACCTCAACCCTTCCCCCTCCCCAGCTTCTGGAGCGGT-TGACTGCAAGCCAGGG  3678

** *** * **** * *****  ************ * *  ***** ** ** **
Figure 4.11:
A) CustalW alignment of mouse and human Pax7 promoter regions showing high level of homology and position of Pax3 consensus sequence (highlighted). Consensus site has been gapped in mouse by alignment software but is clearly present in both species.
B) Illustration of p3.0Luc and p1.5Luc showing position of Pax3 consensus site.
C) Luciferase response to co-transfection with the pCMV-Pax3 expression plasmid. No significant change in reporter gene expression observed, n = 3 error bars are SEM. Data expressed as fold induction over baseline (pGL3basic).
Figure 4.12:
A) Plasmid pl1RLuc
B) No change in reporter gene activity on co transfection with pCMV Pax3, n = 3 error bars are SEM.
Figure 4.13

A

1.5kb Pax7 promoter

B

Fold Induction

C

Indexed Fold Induction

ng pCMV-Pax3
Figure 4.13:
A) Plasmid p1.5I1RLuc
B) Pax3 response to co-transfection of p1.5I1RLuc with 100ng pCMV-Pax3. A clear induction in Luciferase activity is observed, n = 3 error bars are SEM.
C) To test the observations in Figures 3.11, 3.12 and B, a range of ng quantities of pCMV-Pax3 were co-transfected with each of p1.5Luc, pI1RLuc and p1.5I1RLuc. An induction of reporter gene activity is observed only on co-transfection of the pCMV-Pax3 expression construct with p1.5I1RLuc (containing both the Pax7 promoter and Pax7 intron 1 silencer element). Blue trace is p1.5I1RLuc, pink is p1.5Luc, yellow is pI1RLuc. Data is expressed as normalised luminescence fold induction over baseline and indexed to 0ng pCMV-Pax3 to enable direct comparison between data sets. n = 3 - 5 (dependant on data point) error bars are SEM.
of the above, and this time a clear response to Pax3 co-expression was obtained, Figure 4.13B. To test this response further, and to ensure a concentration dependent response to Pax3, a range of quantities of pCMV-Pax3 were co-transfected for each of the plasmids p1.5Luc, pI1RLuc, and p1.5I1RLuc. Only the latter construct responded to pCMV-Pax3 co-transfection over the range investigated, see Figure 4.13C.

Interestingly, this result implies that Pax3 can interact with the Pax7 regulatory elements defined here, at least in vitro, but that this interaction would appear to ‘de-repress’ the silencing effect of the I1R element, rather than being a simple activating or inhibitory relationship.

Since the reporter gene assays described above do not indicate whether Pax3 is interacting with the promoter or silencer element, it remained possible that the Pax3 consensus identified by the MatInspector analysis mediates this response. To investigate this possibility, the core consensus Pax3 binding site within the Pax7 promoter was mutated using site directed mutagenesis. The p1.5I1RLuc plasmid was used with the QuikChange Site Directed Mutagenesis kit (Stratagene) and primers P7_sub_FWD/REV. Clones were screened by direct sequencing to confirm substitution of the core consensus TCGC to ATAA, Figure 4.14A. Reporter gene co-transfections were repeated using this mutated construct (pP7P<delta>) and the original p1.5I1RLuc plasmid. No change to the Pax3 response was observed, see Figure 4.14B, indicating that Pax3 does not bind to this consensus sequence to mediate its effect in this system.
Figure 4.14:
A) Site directed mutagenesis of the Pax3 consensus site in the Pax7 promoter sequence alignment
B) Luciferase response of p1.5I1RLuc and pP7P<delta> illustrating no change in Pax3 response on co-transfection with 250ng pCMV-Pax3. Data is fold induction over baseline (pGL3basic), n = 4, error bars are SEM

| Wild Type | C T T C C T T | G T C C C C C C C C | C T T C C T T A T A A C C A C C C C C C C C |
| Mutant    | C T T C C T T | G T C C C C C C C C | C T T C C T T A T A A C C A C C C C C C C C |

<table>
<thead>
<tr>
<th>p1.5I1RLuc</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1.5I1RLuc + 250ng pCMV-Pax3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pP7P&lt;delta&gt;</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>pP7P&lt;delta&gt; + 250ng pCMV-Pax3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Finally, since this ‘de-repression’ of Pax7 regulatory elements by Pax3 was not predicted in the literature, and is in fact in direct contradiction to data presented by (Borycki et al., 1999) for the expression of Pax3 in the C2C12 cell line, Western blotting analysis of Pax3 and Pax7 was performed to see if the implied changes in Pax7 regulation could be observed at the protein level in C2C12 cells. The Pax3 antiserum (rabbit) described in Chapter 5 and the DSHB α-Pax7 (mouse) were used for all pCMV-Pax3 quantities used in the above reporter gene assays. A β-actin antibody (goat) was also used to control for loading differences on the Pax7 blot, after stripping. This result can be seen in Figure 4.15A.

These data show that, as more pCMV-Pax3 is transfected into the C2C12 cells, more Pax3 protein is made. They also indicate that the level of Pax7 protein expressed by these cells is initially undetectable in this assay, but is induced by increasing quantities of Pax3. The β-actin control indicates that this change is not an artefact of loading. To contrast this result, a northern blot published by (Borycki et al., 1999) is included in Figure 4.15B, clearly showing a loss of Pax7 activity in C2C12 cells stabling expressing Pax3.

**Discussion (Hypothesis Three)**

The aims of this chapter were to first try and establish experimentally some regions of the mouse Pax7 genomic locus that might control expression of this gene in vivo and then to use these elements to test the third hypothesis outlined in Chapter One. Since no regulatory elements had been defined for the mouse Pax7 gene at the start of this
Figure 4.15:
A) Western blotting for Pax3, Pax7 and β-Actin from C2C12 cells transfected with 0 - 500ng pCMV-Pax3, showing increase in Pax3 protein and Pax7 protein (β-Actin shown as lane loading control)
B) Contrasting evidence from Borycki, A. et al 1999, of a Northern blot illustrating Pax7 RNA levels from two C2C12 cell lines stably transfected with a Pax3 expression construct
project, the minimal promoter was perceived as good starting point. To this end a comprehensive bioinformatic analysis was performed over the entire Pax7 region, and highly conserved CpG island was identified in the 1 – 2kb upstream of the Pax7 coding exon. To ensure this region could act as a minimal promoter, the 5'UTR of the Pax7 gene was also characterised in the mouse to see if upstream, non-coding exons could be detected. Since the coding sequence of the Pax7 gene covers around 200kb of genomic DNA the potential existence of upstream exons, which would position any minimal promoter away from the area identified by bioinformatics, had to be considered. The results from PEXT, 5'RACE and RPA indicated a definite transcriptional start site around 100bp upstream of the start of the Pax7 coding sequence. This is in agreement with some reports published concerning the human Pax7 gene (Murnann et al., 2000; Schafer et al., 1994). PEXT repeatably demonstrated a second population of 5'UTR’s of between 300 – 500bp. This product was not detected in 5'RACE or RPA experiments. It was concluded that, given the range of sources of RNA used in the PEXT experiments, this seemed unlikely to be an artefact of PEXT methodology. Previous reports have indicated Pax7 may have a longer (600 – 650bp) 5'UTR, which would be closer to the size reported here (Syagailo et al., 2002; Vorobyov et al., 1997). None of these researchers report a novel non-coding exon, however. Since the start of this project, one paper analysing the Pax7 promoter in the mouse has been published (Lang et al., 2003). This group conclude that the Pax7 5'UTR is 73bp, and encodes no novel exons. Whilst this is in agreement with my data, the evidence presented by this group does not rule out the possibility of larger 5'UTR’s. Given the heterogeneity of actual start sites reported for this gene, and the extremely high CG content of the proximal sequence, it possible a discrete single start site for Pax7 does not exist. In other systems, such conditions
have been attributed as causative factors for a ‘stuttery’ transcriptional start sites, (Ayoubi and Van den Ven, 1996), where transcription starts at numerous places within a defined region. This may explain the differences observed between both the data obtained from different techniques here, and that reported in the literature. Importantly, no novel upstream exons have been reported for Pax7, allowing a putative promoter region to be positioned here.

This region was cloned and used in reporter gene assays in Pax7 positive cell lines, and promoter activity mapped to the proximal 1.5kb. Initial bioinformatic analyses also indicated a highly conserved element within intron one. This was cloned and found to act as a canonical silencer element in vitro. This element was also identified by (Lang et al., 2003) and shown in transgenic assays to act as a Pons-specific enhancer, in contrast to the role indicated here.

To address hypothesis three, the minimal promoter was assayed for an in vitro response to Pax3 in co-transfection assays, since a Pax3 consensus had been identified in this region. This did not demonstrate a response, nor did the I1R silencer. When both elements were carried on the same reporter vector, a Pax3 dependant ‘de-repression’ of reporter gene activity was observed. Mutagenesis indicated this activity was not mediated by the Pax3 binding consensus identified in the promoter.

Finally, the effect of Pax3 expression on endogenous C2C12 Pax7 expression was conducted by Western blotting. Interestingly, a clear induction of Pax7 was observed with increasing quantities of Pax3. This was in contrast to previously published reports
It was concluded from the above that \textit{Pax3} is able to effect the transcription of \textit{Pax7} \textit{in vitro}, but that this regulation may be from either the proximal promoter or the I1R element identified above. Therefore hypothesis three can be partially confirmed; \textit{Pax3} does seem to have an effect on the regulatory elements defined here \textit{in vitro}, but the mechanism and \textit{in vivo} significance of that effect is unclear. Chapter Six will attempt to investigate which of these elements, if any, are occupied by the \textit{Pax3} protein \textit{in vivo}, and attempt to establish a role for \textit{Pax3} in the regulation of \textit{Pax7} in an embryonic context.
Chapter Five: Generation of a novel α-Pax3 antibody

Introduction

Chromatin Immunoprecipitation (ChIP) requires relatively large (µg) quantities of specific antibody, to reduce background and increase the sensitivity and specificity of detection. At the start of this project, only one anti-Pax3 antibody had been described (Stark, 1997). This was a rabbit polyclonal derived by another group, and therefore not reliably available in large quantities. The decision was made to raise and design a new anti-Pax3 antibody, using specific peptide sequences, to enable multiple ChIP experiments to be performed and optimised. This chapter describes the methods used to design and raise a rabbit polyclonal antibody to specific peptide sequences within the Pax3 protein. The anti-sera obtained were characterised using both Western blotting and immunohistochemistry to confirm detection of the Pax3 epitope at the correct molecular weight and expression pattern. One of the sera was then selected and purified to obtain mono-specificity on Western blot and clean expression patterns with immunohistochemistry.

Peptide Design and Antibody Production

The Pax3 sequence and predicted structure were considered in the design of the peptides used to generate anti-Pax3 sera. From X-ray analyses of both Paired domains (Xu, 1995) and structural dissections of Pax3 (Chalepakis et al., 1994) it has been
shown that the N-terminal regions contain the DNA binding domains, whilst the C-terminus of the protein is involved in recruiting transcriptional complexes. The C-terminus of the Pax3 protein sequence was examined to identify possible immunogenic peptide sequences. It was reasoned that the N-terminal DNA binding domains may contain less structural epitopes accessible to an antibody when bound to DNA due to conformational restrictions and the possible presence of associated cofactors or the RNA polymerase machinery. The C-terminal portion of the related Pax6 protein has been hypothesised to be less structurally ordered on contact with DNA than the paired domain (Xu, 1999). Structural disorder has been cited as a possible factor increasing the immunogenicity of an epitope experimentally due to its increased availability for antibody binding. Furthermore, since ChIP depends upon transcription factor – DNA contact peptides in the N-terminal portion of the protein would be undesirable since they may disrupt protein – DNA interactions.

Figure 5.1 shows the full peptide sequence of Pax3, and highlights the peptide sequences initially considered for immunisation. Peptides A and B were chosen. Peptide A represents the natural C-terminus of Pax3, contains ~60% polar residues (green) one charged residue (lysine at position 12 - red) and the negatively charged carboxyl group at position 16. Peptide A was chemically modified to include an cysteine residue at the very N-terminus; this was to enable chemical linkage to the carrier protein (Keyhole Limpet Haemocyanin) via the sulphydryl group. Peptide B is a shorter peptide, has an overall negative charge, and is also composed of ~60% polar residues. This peptide was chemically modified to amidate the C-terminus; since this sequence is internal in the protein this would better represent the chemical environment of the N-terminal region of this sequence. Analysis of the C-terminus of
Figure 5.1:
A) Scale drawing of the Pax3 polypeptide. Major features illustrated; yellow boxes show the position of peptide sequences A, B, and C.
B) Peptide sequences (green = polar, red = charged side chains).
C) Antigenicity and flexibility plots for the C-terminus of Pax3; Peptide A and B sequences highlighted in red.
the protein using MacVector antigenicity index and flexibility plot prediction software further illustrated that these peptide sequences represented regions of the protein which had both high flexibility and antigenicity (Figure 5.1C). Peptide C was initially identified as potentially antigenic sequence, but was ultimately not used as this epitope was considered too far into the protein, risking the epitope being occluded by tertiary structure in the intact protein. Also bioinformatic analysis suggested a low antigenicity index (Figure 5.1C). Anti-sera were raised in rabbits as described in Chapter 2; two rabbits were immunised for each peptide, generating four batches of anti-sera.

**Verification of α-Pax3 immunogenicity**

To evaluate the sera obtained from immunisation with peptides A and B, Western blotting was performed on NIH-3T3 cells transiently transfected with pCMV-Pax3 plasmid, as described in Chapter Two. The pCMV-Pax3 plasmid contains a full length Pax3 cDNA driven by a strong (CMV) promoter, and was used in these experiments as a source of in vitro expressed Pax3 protein. NIH-3T3 were previously shown not to express Pax3, as detectable by RTPCR (data not shown). Mock transfected NIH-3T3 cells (transfected with the pCMV-Script plasmid) were therefore used as a source of negative control protein. All four sera were tested by comparing pre-immune with post immune sera (to ensure any bands observed were the result of immunisation against the Pax3 peptides), on protein extracts from transfected and mock-transfected NIH-3T3 cultures. Serum generated from peptide A (rabbit 1) gave the best potential-band at this stage and was therefore used in all later experiments.
Peptide A (rabbit 2) and one of the rabbits immunised with peptide B did also generate Pax3 bands, but these were weaker and therefore not used (data not shown).

Figure 5.2A shows the initial Western blot for the peptide A serum, comparing pre and post-immune sera on protein samples containing Pax3 and matched negative controls. The Pax3 band, at 53 kDa is highlighted in the post-immune western. Western blotting conditions were further optimised, Figure 5.2B, to reduce background. The antibody was then used to detect Pax3 protein from NIH-3T3 (± pCMV-Pax3), C2C12 (±pCMV-Pax3) and E11.5 wild type and Sp^{2H} homozygous mutant (Pax3 null) mice (Figure 5.2C). As Figure 5.1A illustrates, the Sp^{2H} mutations would truncate any protein product from the sploitch allele before the peptide sequences used to raise the anti-sera. Therefore no Pax3 protein should be detectable by Western blot in extracts from Sp^{2H} homozygous mutant embryos.

To further confirm the specificity of the serum to Pax3, immunohistochemistry was then attempted using the un-purified serum. Unfortunately the background levels on all sections examined was too high to enable specific expression patterns to be reliably determined (data not shown). This is perhaps unsurprising given the number of non-specific bands observed on Western blots using this serum.

**Antibody Purification**

ChIP requires a source of anti-serum as specific as possible. To prepare the anti-Pax3 serum for ChIP, and to enable its use in immunohistochemistry, a protein-G column purification, followed by a two step acid elution, protocol was used. After washing,
Figure 5.2:
A) Western blot using serum from rabbits immunised with peptide A. Comparison of pre-immune serum with post-immune serum on NIH-3T3 transfected with pCMV-Pax3 expression plasmid or empty vector (±). Pax3 pro-band highlighted.
B) Western blot optimised using serum from rabbits immunised with peptide A. Pax3 clearly visible at ~53kDa, and specific to cell transfected with pCMV-Pax3.
C) Western blot showing presence of Pax3 in wild-type E11.5 embryos (lane 1), cells transfected with pCMV-Pax3 (lanes 3 and 5). Homozygous Splotch embryo protein was loaded in lane 2. Lanes 4 and 6 represent protein extracts from cells lines transfected with empty vector.
elute was collected in 1 minute intervals (1ml) and the protein concentration determined by Bradford assay. A plot of protein concentration (expressed as OD at λ562) vs. elution volume can be seen in Figure 5.3.

A clear elution peak can clearly be seen between elution volumes pH2.7(5) and pH2.7(9). These volumes were combined and referred to as α-Pax3(HT) (High Titre). The slow tail-off in protein concentration between volumes pH2.7(10) and pH2.7(14) was interpreted as a potential second peak, poorly resolved due to the relative crudity of the protein concentration assay. These elutes were also combined and referred to as α-Pax3(LT) (Low Titre). The mean (OD = 0.15 n = 6, SD = 0.004) blank (background) OD λ562 is shown on the plot for reference. No protein over background could be detected in any of the samples eluted with the pH1.9 buffer. These were discarded.

α-Pax3(HT) and α-Pax3(LT) were therefore tested using the Western blot parameters outlined above (Figure 5.4A). This clearly shows that, whilst both fractions contain an anti-Pax3 immunogenicity, the α-Pax3(HT) still retains many of the non-specific bands previously described for the unpurified serum. This is possibly due to the use of a protein-G column, which binds IgG heavy chain irrespective of immunospecificity, rather than a peptide A specific column. The α-Pax3(LT), however, gave an extremely clear signal at the appropriate Mr for Pax3. During the process of deriving and purifying this antibody, a monoclonal anti-Pax3 resource became available from the DSHB (Developmental Studies Hybridoma Bank, Iowa). To compare specificities, Western blots were performed comparing the two antibodies (DSHB and α-Pax3(LT)) under identical conditions (Figure 5.4B). The DSHB antibody gave a
Figure 5.3: IgG acid elution curve to generate monospecific anti-Pax3 antisera. Time (volume) of sample eluted from column is plotted against OD562 (protein concentration). HT = samples collated to generate α-Pax3(HT), LT = samples collated to generate α-Pax3(LT). Solid red line represents background protein concentration in eluate.
Figure 5.4:
A) Western blot on NIH-3T3 cells transfected with either pCMV-Pax3 (+) or empty vector (-). Comparing the LT and HT α-Pax3 fractions collected
B) Western blot on NIH-3T3 cells transfected with either pCMV-Pax3 (+) or empty vector (-). Comparing α-Pax3(LT) with the monoclonal antibody from the DSHB
non-specific band of about 60kDa which was absent from the α-Pax3(LT) westerns. This correspond with Western blots obtained using a sample of the source polyclonal used to derive this monoclonal (A kind gift from M. Goulding, data not shown) when verifying the production of Pax3 from transfected NIH-3T3 cells. Since ChIP requires as mono-specific an antibody as possible, and the fact that the DSHB antibody was produced from the C-terminal 183 residues, rather than a specific peptide sequence (Venters et al., 2004), the α-Pax3(LT) was chosen for the ChIP experiments. α-Pax3(LT) is henceforth referred to as α-Pax3.

**α-Pax3 Immunohistochemistry**

Given the specificity of the α-Pax3 polyclonal on Western blot, immunohistochemistry on wax embedded, transverse sectioned, wild-type E12.5 mouse embryos was performed. Figure 5.5A illustrates a transverse section showing clear specific staining in dorsal neural tube (NT), dorsal root ganglion (DRG) and facial mesenchyme. Figure 5.5B is a high-power image of the same section, showing nuclear staining in both NT and DRG cells, as expected for a transcription factor, and 5.5C a high power image from another section, illustrating Pax3 positive cells delaminating from the dorsal edge of the NT, possibly neural crest. Also illustrated are comparative sections from published work (Dottori et al., 2001) and (Terzie and Saraga-Babic, 1999). Since the purpose of this exercise was to confirm the Pax3 specificity of the α-Pax3 polyclonal produced, further expression analysis was not performed. The lack of a comprehensive publication detailing Pax3 protein expression
Figure 5.5:
A) Transverse section through E12.5 WT embryo. Pax3 expression in dorsal neural tube (NT), facial mesenchyme (FM) and dorsal root ganglia (DRG).
B) High power image of inset highlighted in A. Note Pax3 expression in migrating neural crest cells (NC).
C) Comparison of similar published images illustrating Pax3 expression. Human 9 week fetus Pax3 immuno (Terzic et al 1999), Mouse E12.5 transverse section Pax3 immuno (Pax3 = red and yellow signal) (Dottori et al 2001), another high power transverse section showing nuclear Pax3 staining.
across a developmental series would make this anti-serum a potentially attractive tool for future investigators, however.

**Discussion**

The α-Pax3 polyclonal described here was then aliquoted and stored appropriately for ChIP, described in Chapter 6. All experiments described in this thesis using an α-Pax3 antibody were performed using the above polyclonal. One interesting aspect of all Western blots shown here, whichever antibody was used, was the presence of a doublet (see Figure 5.4B). This was specific to Pax3 positive lanes only, and therefore could not be attributed to background. There are several possible explanations for this. Firstly, the doublet could represent one of the glutamine or exon isoforms of Pax3 (see Chapter 1 for an overview of relevant literature). Since many of the Pax3 bands detected on these westerns were obtained from protein driven from a Pax3 cDNA expression construct such isoforms must represent either novel splicing events within the mRNA produced from the plasmid or protein produced from the cells on genomic copies of Pax3 via an auto-regulation mechanism. Pax3 may have different phosphorylation states and this would clearly explain a band migrating a different rate on an SDS-PAGE gel. A more mundane explanation would be that the doublet simply represents a common degradation product which is detected in transfected cells due to the high quantities of Pax3 produced in these transfections. Further investigation is required to differentiate between these possibilities and determine the cause of this doublet.
Chapter Six: *in vivo* Chromatin immunoprecipitation

**Introduction**

In this chapter the occupancy of the transcription factor binding sites previously investigated on the *cis*-regulatory target sequences identified will be tested *in vivo*. This will bring together the *in vitro* studies that have attempted to establish a possible role for the regulation of *Wnt1* and *Pax7* by *Pax3*. To this end the anti-*Pax3* antibody, designed and purified as described in Chapter 5, and the anti-*Msx1/2* antibody obtained from the DSHB were used to develop specific Chromatin Immunoprecipitation (ChIP) assays on mouse embryonic tissue. This technique allows the localisation of transcription factor binding sites to specific regions of genomic DNA. The data described below confirms the *in vitro* observations for the *Pax3* regulation of *Wnt1* via the 5' proximal promoter region, raises some interesting possibilities for the interaction of *Msx2* with these regulatory elements, and resolves the uncertainty over the location of *Pax3* interaction within the *Pax7* genomic locus.

**ChIP: an overview**

The ChIP technique isolates and enriches DNA sequences which are involved in protein-DNA interactions. The technique was originally designed for use with chromatin remodelling proteins (such as DNA methyltransferases and histone deacetylases) but has since been adapted for a number of other assays to assess any
kind of protein-DNA interaction (Ren, 2000b). An increasingly widely used application of this technology is the characterisation of transcription factors binding to target sequences within the genome. Generally this technique has been used to examine these interactions in cell culture, due to the relative ease of obtaining large amounts of source material, but the demand for in vivo ChIP to examine the validity of findings (such as those outlined in Chapters 3 and 4) in the tissues and organisms of interest is growing.

A schematic of the ChIP process is shown in Figure 6.1. In the first stage, the transcription factor under investigation is assumed to bind to and activate / repress transcription of its target gene by binding to appropriate sequences within the genome (stage I - II). At this point the embryo or tissue is dissected and disassociated into single cells as rapidly as possible, and the cell suspensions treated with formaldehyde. The formaldehyde acts as a chemical cross linking agent between proteins (generally between the ε-amino group of lysine residues and adjacent amide bonds) and between proteins and DNA (between lysine residues and the NH₂ exocyclic moiety of adenine, guanine or cytosine (Siomin, 1973) –interestingly this requires the DNA to be at least partially denatured, or for the protein to be making close contact with the bases through the major or minor grooves of the DNA molecule. Both of these are thought to occur during the interaction of a transcription factor with its target sequence), see stage III. The transcription factor is then effectively ‘trapped’ on its target genomic sequence. The cells are then placed in detergent-based lysis buffer and subjected to sonication to disrupt the cell and nuclear membranes and to break the chromatin into fragments of a desired size, stage IV. The tissue sample has now been broken into a pool of chromatin fragments with transcription factors, histones, and other DNA
Figure 6.1

ChIP Overview

Stage I

Stage II

Stage III

Stage IV

Stage V

Stage VI

Stage VII

PCR or cloning and sequencing = verification of target site occupancy
Figure 6.1:
Schematic of *in vivo* ChIP protocol.
Stage I: Transcription factor (circle) of interest binds target element (box) on genomic DNA (line) in appropriate tissue / developmental stage
Stage II: Transcription factor drives / alters expression of target gene
Stage III: Cells / tissue extracted and treated with weak formaldehyde treatment to crosslink proteins and DNA *in situ* (curved lines)
Stage IV: Crosslinking quenched, and cells lysed. Crosslinked protein - DNA complexes sonicated to break up genomic DNA into smaller fragments
Stage V: Antibody specific to transcription factor of interest the added to pool of sonicated protein-DNA complexes
Stage VI: Immunoprecipitation performed to isolate Antibody - transcription factor - DNA complexes
Stage VII: Protein - DNA complexes uncrosslinked, proteins and antibodies degraded with proteases, and DNA isolated by phenol - chloroform. Primers specific to the regulatory regions being tested for transcription factor occupancy then used to screen pool of DNA fragments and amplify sequences if present.
associated proteins bound co-valently. To enrich specifically for those complexes containing the transcription factor under investigation, an immunoprecipitation is performed using an antibody raised against the protein of interest (stage VI). Once all unwanted species have been washed away, the antibody – transcription factor – DNA complexes (stage VI) are then eluted from the IP and the formaldehyde cross links reversed using heat and halide (typically Cl). The proteins (antibody and transcription factor) and any co-purifying RNA are then enzymatically digested. This should leave a population of DNA fragments, of broadly uniform size due to the sonication process employed, which were bound to the transcription factor under investigation. These are then phenol chloroform purified and re-suspended in clean ddH₂O.

At this stage the occupancy of a specific site can be tested by designing flanking PCR primers and testing the pool of precipitated fragments for the presence of this site (stage VII). Enrichment can be meaningfully assessed by also attempting to amplify a genomic sequence to which the factor is known / thought not to bind to. It is important to ensure that the second sequence is not too close to the putative binding sequence so as not to co-purify on the same fragment. This makes the sonication step essential; care must be taken to ensure all chromatin is fragmented to a size sufficiently small that these two amplicons cannot occupy the same fragment, but large enough to ensure that both test primers can anneal to the same target fragment and enable PCR amplification to occur. A positive control PCR of ‘input’ chromatin (i.e. sonicated chromatin from the source sample which is phenol chloroform purified without any intermediate immunoprecipitation) and a negative control ‘no antibody’ PCR (i.e. a sonicated chromatin sample where water instead of antibody was added at the IP stage).
Alternatively the populations of enriched fragments can be cloned and sequenced, or applied to genome coverage microarrays. Both of these endpoints are generally used to identify novel target sequences and, whilst not being experimentally utilised in this study, will be discussed in more detail later.

**Thesis findings: an overview**

Figure 6.2 presents a diagrammatic review of the findings so far described in this thesis and the interactions that the following ChIP experiments were designed to investigate. Figure 6.2A summarises the *in vitro* interactions suggested between *Pax3* and the *Wnt1* regulatory elements in an *in vivo* context. Illustrated is the mouse genomic *Wnt1* region, highlighting the four exons, the distal 110bp enhancer (blue box) and the *Pax3* consensus site contained within the 1.2kb proximal promoter region (green arrow). Deletion and site directed mutagenesis studies have strongly suggested that this *Pax3* consensus may mediate an active binding site for the *Pax3* protein *in vivo*. ChIP was performed on these regions using the purified anti-*Pax3* polyclonal antibody described in Chapter 5 to confirm *in vivo* occupation of this site, and not the *Wnt1* distal enhancer.

Figure 6.2B summarises the putative relationship of the *Msx2* protein with the *Msx2* consensus binding site in the *Wnt1* distal enhancer, in the context of the mouse genomic *Wnt1* locus. This interaction was supported by deletion and transposition studies using luciferase reporter vectors *in vitro*. No mutagenesis experiments were conducted on this site during this investigation. ChIP using the anti-*Msx1/2* antibody
Figure 6.2

A

1.2kb proximal promoter

Wnt1 mouse genomic locus

B

110bp region

Wnt1 mouse genomic locus

C

1.5kb proximal promoter

Pax7 mouse genomic locus
Figure 6.2:
A) Summary of model proposed from *in vitro* data described in Chapter 3 and in investigation of hypothesis one. Pax3 binds to and transactivates the Wnt1 proximal promoter via a defined Pax3 consensus site, confirmed by mutagenesis and deletion studies.

B) Model proposed from *in vitro* data and investigation of hypothesis two described in Chapter 3. Msx2 binds to the Wnt1 distal enhancer (within the conserved 110bp element, see Echelard et al 1994) to repress transcription, confirmed by deletion and translocation studies.

C) Model proposed by *in vitro* data described in Chapter 4 and in investigation of hypothesis three. Pax3 binds to either the 1.5kb proximal promoter or the highly conserved I1R element within intron one of the Pax7 locus, but not via the Pax3 consensus site present in the promoter region. Confirmed by mutagenesis and deletion studies.
was used to try and show the \textit{in vivo} occupation of \textit{Msx2} at this site, and not the \textit{Wnt1} proximal promoter, and is discussed below.

Figure 6.2C summarises the \textit{in vitro} data obtained during an investigation into the relationship between \textit{Pax3} and the \textit{Pax7} regulatory elements \textit{in vitro}. The \textit{Pax7} genomic locus is shown, illustrating the first three coding exons, the 1.5kb proximal promoter and the I1R silencer within intron 1. Luciferase experiments demonstrated that both the proximal promoter and the I1R element were necessary to observe a \textit{Pax3} driven transcriptional effect, although this was a de-repression (contrary to the relationship previously described by one group), and was not mediated by the \textit{Pax3} consensus in the 5' proximal promoter (as determined by site directed mutagenesis). To try and see if either of these sites were occupied by \textit{Pax3} \textit{in vivo}, and support the findings of Chapter 4 as potentially biologically significant (rather than an artefact of the Luciferase assay system), ChIP using the \textit{Pax3} antibody and primers specific to these regions was also performed.

\textbf{ChIP: technical considerations}

To examine the role of \textit{Pax3} or \textit{Msx2} in the transcriptional contexts described above and in previous chapters, it was necessary to utilise the correct embryonic tissues. The role of \textit{Pax3} in regulating \textit{Wnt1} transcription has been hypothesised to occur during the development of the murine neural tube, when both \textit{Pax3} and \textit{Wnt1} are being expressed, see Chapter 1. Similarly, any interaction between \textit{Msx2} and the \textit{Wnt1} locus is hypothesised to occur in the developing cardiac neural crest, also found within and around the dorsal neural tube. Finally the interaction between \textit{Pax3} and the \textit{Pax7}
locus has been hypothesised to be in both the somitic mesoderm, limb bud and neural tube. Embryos were collected at E10.5, and the neural tube, forelimb bud and mesoderm dissected as shown in figure 6.3A. Care was taken to remove telencephalon and any tissues caudal to the forelimb bud. E10.5 is a little late to isolate tissues containing migrating neural crest cells. However, one of the major technical difficulties in performing in vivo ChIP is obtaining enough starting material to generate an amplifiable pool of DNA fragments at the end of the procedure. Trials with embryos from ages ranging from E8.5 to E12.5 were initially performed. The most reliable ChIP fractions (generating consistently strong positive control PCR products for all amplicons described below) were generated by embryos at E10.5. Since it was felt that a compromise had to be made experimentally between embryo age and reproducibility of data, E10.5 embryos, dissected as shown, were used in the following ChIP assays.

The sonication of genomic DNA to the correct size is an essential stage in the design and implementation of ChIP experiments. Figure 6.3B shows the optimisation of sonication conditions. The left panel represents DNA extracted from E10.5 dissected, disassociated and fixed embryos which were then sonicated under a range of conditions. Conditions which favoured the generation of fragment sizes between 500bp and 1kb were used in all following experiments and example sonications can be seen from four embryo preparations in the right panel. This average size of DNA fragment length is large enough to enable all of the ChIP PCR reactions to amplify (product range between 282 and 595bp) but small enough to ensure that the two amplicons cannot fall entirely on the same fragment (minimum distance between
Figure 6.3:
A) Image of an E10.5 embryo showing locations of dissection when preparing sample for ChIP analysis (red lines). Mid / Forebrain and regions below the forelimb bud were discarded.
B) Agarose gels to illustrate the optimisation of sonication conditions to generate genomic DNA fragments of 0.5 - 1kb. Left panel shows initial optimisation of conditions for fixed, embryonic chromatin. Bar indicates increasing sonication power, U denotes un-sonicated chromatin. Right panel illustrates sample sonicated chromatin showing required fragmentation in four samples. Red line denotes 1kb level.
amplicons is 2.127kb on the Pax7 locus) thus preventing a false positive signal due to fragment co-purification.

**Pax3 interacts with the Wnt1 proximal promoter in vivo**

Chapter 3 provides strong evidence from both QRTPCR *in vivo* and Luciferase co-transfection assays *in vitro* that Wnt1 represents a direct target of Pax3 regulation. Furthermore this interaction appears to be mediated via a Pax3 consensus binding site, identified by bioinformatic analysis and confirmed with mutagenesis, in the Wnt1 proximal promoter region. To test occupancy of this site *in vivo* ChIP assays were performed as described. Primers were designed to amplify the Wnt1 genomic locus at either the 5' proximal promoter region (a 541bp product flanking the Pax3 consensus, primers W1_PROM_CHIP_FWD/REV) or the 3' distal enhancer (a 282bp product containing the 110bp conserved element and the Msx2 consensus binding site, primers W1_ENHAN_CHIP_FWD/REV). These sites are separated by 8.789kb of genomic DNA and therefore are unlikely to be contained on the same sonicated fragment. An illustration of the Wnt1 genomic locus, showing exons, Pax3 consensus and primer binding sites (labelled W1P and W1E respectively) is illustrated in Figure 6.4A.

The ChIP output for this experiment can be seen in Figure 6.4B. Input control was positive, indicating both fragments are present and competent to be amplified from the pre-IP pool of DNA and no antibody control negative, indicating that any bands observed are due to specific antibody binding during the IP. A clear positive signal is observed for three different wild type E10.5 dissected embryos, showing Pax3 is
Figure 6.4: A Genomic Wnt1 locus showing ChIP primer binding sites, and table of predicted mplicon sizes.

B) PCR products of anti-Pax3 ChIP using primers specific to the Wnt1 locus. Wnt1 promoter positive in all three experiments, confirming occupancy of this site in vivo by Pax3.
bound to the *Wnt1* promoter region and not the distal enhancer at this stage in development in these tissues.

This enables the conclusion that *Pax3* mediates the transcriptional upregulation of *Wnt1* *in vivo* via the 5' proximal promoter and the *Pax3* consensus site described and in agreement with hypothesis one.

**Msx2 may associate with the Wnt1 locus in vivo**

Chapter 3 also provides evidence that the *Msx2* transcription factor may mediate the repression of *Wnt1* transcription through the binding of a consensus site in the 3' distal enhancer. These experiments were conducted in response to a study performed by (Kwang et al., 2002) where the misregulation of *Msx2* in the *Pax3* homozygous mutant was shown to be responsible for the cardiac phenotype observed in these animals. Given the importance of the *Wnt – Dvl2* pathway in the development of the cardiac outflow tract in a second series of studies (Hamblet, 2002), it was postulated here that the Wnt signal involved could be *Wnt1*, and that the expansion of *Msx2* expression in the *Splotch* mutant was able to repress *Wnt1* expression in these migratory cardiac neural crest. To test this hypothesis, the *cis*-regulatory elements of *Wnt1* were tested in co-transfection assays with an *Msx2* expression construct. A repressive activity of *Msx2* was observed in association with the *Wnt1* 3' distal enhancer. This was mapped with deletion and transposition constructs to a region of the enhancer containing a highly conserved 110bp element and an *Msx2* consensus site.
To test occupancy of this site \textit{in vivo}, ChIP assays were performed on three wild type E10.5 embryos, dissected as described. Identical conditions and PCR primers to those used to test the occupancy of Pax3 in the \textit{Wnt1} promoter were employed here. These, and the position of the Msx2 consensus site, are illustrated in Figure 6.5A. The Msx1/2 antibody obtained from the DSHB was used in the IP stage of these assays.

Figure 6.5B illustrates the ChIP output for these experiments. Unfortunately, whilst the positive and negative controls were correct, the ChIP assays themselves did not give a consistent or easily interpretable result. Two of the embryos assayed gave positive signals for both the promoter and the enhancer, whereas the third gave no signal at all. There are several ways in which to interpret this data. Firstly it is possible that the experiment simply did not work properly and has simply given random signals. More repeats are necessary to determine if this is the case. Given that the Pax3 ChIP worked consistently under identical conditions for this locus other possible causative factors are considered. Assuming that the third embryo examined simply did not work the presence of bands in both the promoter and enhancer lanes may suggest a number of other outcomes. It is possible that Msx2 is binding to one or other region of the \textit{Wnt1} locus, but the sonication did not separate the two fragments efficiently or these regions were in close physical proximity and were bound together during the formaldehyde treatment. Since this was not the case for the Pax3 ChIP illustrated in Figure 6.4 using identical source tissues and cross linking conditions, this would seem improbable. It is possible that Msx2 was bound to both the promoter and enhancer regions in these tissues, although the lack of \textit{in vitro} response from the \textit{Wnt1} promoter fragments to the Msx2 protein would argue against this. One major fault with this experiment was the use of a bi-specific anti Msx1/2 antibody. Clearly
Figure 6.5: A) Genomic Wnt1 locus showing ChIP primer binding sites, and table of predicted amplicon sizes. Wnt1 promoter and enhancer positive in two experiments, negative in one.

B) PCR products of anti-Msx1/2 ChIP using primers specific to the Wnt1 locus. Wnt1 promoter and enhancer positive in two experiments, negative in one.

<table>
<thead>
<tr>
<th>Fragment</th>
<th>PCR</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wnt1 Promoter (W1P)</td>
<td>55^30</td>
<td>541 bp</td>
</tr>
<tr>
<td>Wnt1 Enhancer (W1E)</td>
<td>56^30</td>
<td>282 bp</td>
</tr>
</tbody>
</table>
this cannot differentiate between the binding of either protein to the \textit{Wnt1} regulatory elements, and may account for the observed result. Further investigation is obviously required, perhaps using a mono-specific antibody or site directed mutagenesis of the \textit{Msx2} site within the \textit{Wnt1} enhancer. It is tentatively concluded that \textit{Msx2} is found in association with the \textit{Wnt1} locus \textit{in vivo}, although the precise position of this interaction is unclear. This is therefore in partial agreement with second hypothesis presented in this thesis, although much further work is required.

\textbf{Pax3 binds to the Pax7 I1R element \textit{in vivo}}

Chapter 4 presents data which both establishes two regulatory elements potentially controlling \textit{Pax7} expression in the mouse genome, and presents data that \textit{Pax3} may interact with one or both of these elements to regulate the transcription of \textit{Pax7 \textit{in vivo}}. Luciferase assay experiments investigating the interaction of \textit{Pax3} with either the \textit{Pax7} proximal promoter (as established by activity in the C2C12 cell line and by (Lang et al., 2003)) or the I1R silencer revealed that both elements were required to generate a response \textit{in vitro}. The mutagenesis of a \textit{Pax3} binding site consensus in the \textit{Pax7} proximal promoter region did not abrogate this effect (in contrast to the \textit{Pax3} binding site in the \textit{Wnt1} promoter) indicating that the interaction was not mediated through that site.

To examine if \textit{Pax3} does indeed interact with the \textit{Pax7} locus \textit{in vivo}, and whether the interaction is through the promoter or I1R silencer element, ChIP was performed on the \textit{Pax7} locus on dissected E10.5 mouse embryos. Primers were designed to amplify a 595bp region of the \textit{Pax7} promoter (primers P7\_PROM\_CHIP\_FWD/REV) or a
546bp region of the Pax7 intron 1 silencer (P7_INTRON_CHIP_FWD/REV). The Pax7 genomic locus with these primers positioned and showing the first three exons, Pax3 consensus binding site and I1R element can be seen in Figure 6.6A.

ChIP assays were then performed using the Pax3 antibody described in Chapter 5 on three wild type embryos; the positive and negative controls were correct in each case. Figure 6.6B illustrates that the ChIP assay repeatably enriched for fragments carrying the I1R element and not the Pax7 proximal promoter or Pax3 consensus site. This confirms the mutagenesis result, and suggests that Pax3 is interacting with the Pax7 locus via the I1R silencer element in vivo.

 Whilst this result enables hypothesis three to be partially accepted this, and the work described in this thesis, does not fully resolve whether Pax3 is acting to repress as suggested by (Borycki et al., 1999), de-repress as suggested by Chapter 4’s Luciferase data, or do nothing at all as suggested by (Relaix et al., 2004) the Pax7 gene. The occupancy of a site within a highly conserved element of apparent regulatory function would argue against Pax3 paying no role whatsoever in the transcription of Pax3. It is possible that the de-repression observed in Chapter 4 reflects an artefact of the Luciferase assay whereby the absence of Pax3 co-factors, usually present in vivo, enables the Pax3 transcription factor to mediate an unusual effect. Since Pax3 carries domains for both transcriptional activation and repression of target genes (Chalepakis and Gruss, 1995; Lang, 2005)) an observation of this nature in vitro would not be entirely unsurprising. This would not be in keeping with the clear up-regulation of Pax7 protein observed from C2C12 cells transfected with increasing quantities of Pax3, shown in Figure 4.15. Clearly further work, to resolve this issue in vivo and to
Figure 6.6:
A) Genomic Pax7 locus showing ChIP primer binding sites, and table of predicted amplicon sizes
B) PCR products of anti-Pax3 ChIP using primers specific to the Pax7 locus. The Pax7 silencer within intron is specifically amplified in each experiment, confirming Pax3 binding to this region in vivo.
map the binding of $Pax3$ to the IIR element precisely, is needed. Given the role $Pax7$ plays in both embryonic development and adult stem cell derived muscle regeneration in the adult, a greater understanding of $Pax7$ regulation may also be of therapeutic value.

**Conclusions**

In the above chapter a series of ChIP experiments, bringing together and testing the observations made in the Chapters 3 and 4 with the antibody designed and described in Chapter 5, are described. $Pax3$ has been shown to associate with the $Wnt1$ promoter and the $Pax7$ IIR silencer *in vivo*, via a defined and as yet undefined binding site, respectively. $Msx2$ may have some association with the $Wnt1$ locus *in vivo*, although the data presented here is inconclusive and much work remains to be done to complete this analysis. Importantly, the above data validates the use of the $Pax3$ antibody with *in vivo* ChIP and paves the way for the use of these tools for a high level analysis of $Pax3$ function in development.
Chapter Seven: Discussion

In this thesis a series of direct transcriptional interactions were postulated from published studies on Pax3. Methodologies to test these interactions were developed for in vitro and in vivo analyses. In the post-genomic era it is increasingly important to dissect precisely the genetic interactions involved in biological processes. Transcription is one of the first lines of regulation between the information encoded within the genome and the translation of that information into biochemically active material. It is essential that this process is studied in detail and that gene regulatory networks and relationships are reliably established.

The process of transcription transfers coding sequences in DNA to an RNA format that is then translated into protein by the cellular ribosomal machinery. The RNA polymerases and factors constituting the general transcriptional holoenzyme are the same irrespective of gene, cell type, tissue, or even organism (Ayoubi and Van den Ven, 1996). The proteins that recruit the polymerases to the correct position on the genome are myriad and form one of the major regulatory points in this process. These transcription factors bind to specific sequences in the cis-regulatory regions of genes and mediate their upregulation (i.e. recruitment of the polymerase machinery and/or chromatin remodelling factors) or inhibition depending on their precise biological context. One of the problems faced in experimental biology is that the techniques used to examine such protein – DNA interactions precisely in vitro are often difficult to confirm in an in vivo context. A relationship observed with in vitro techniques will
not necessarily be borne out in an \textit{in vivo} setting, nor is an \textit{in vivo} approach always practical when examining relationships of this kind experimentally.

In this thesis, the developing mouse embryo was used as a model system to examine transcriptional interactions between \textit{Pax3} and its targets. An \textit{in vitro} approach was used in the first instance to map these interactions precisely. This was then followed by ChIP to confirm these interactions \textit{in vivo}. It is hoped that this technique could then be applied to other \textit{Pax3} – target interactions, or even be scaled up to scan the mouse genome for equivalent sites demonstrating \textit{Pax3} binding \textit{in vivo} to enable a fuller understanding of this important gene’s function in development.

\textit{Pax3 and Wnt1}

Chapter 1 outlines some of the developmental roles \textit{Pax3} is known to have in the developing mouse embryo. One of the earliest post-gastrulation events is the induction of cells fated to become the nervous system from the ectoderm, known as neurulation (see (Copp, 2003) for review). \textit{Pax3} is intimately associated with this process, and the expression of this gene in the developing neuroectoderm is one of the earliest observable markers of the future spinal cord (Bang, 1997; Goulding et al., 1991). The roles of \textit{Pax3} in the development of these tissues are varied, and much work has been done to characterise the effects of both \textit{Pax3} overexpression (Tremblay et al., 1996) and mutation (see (Epstein et al., 1991; Epstein, 1996) and references within Chapter 1 for overview). Assaying the direct targets of \textit{Pax3} is important when considering the role of this factor in development. This will also
enable the separation of direct and indirect effects of Pax3 when manipulating this factor in these systems.

Wnt signalling is also fundamental to the patterning of the developing embryo, and several Wnt genes are expressed in the developing nervous system (Miller, 2001; Parr, 1993). Whilst Wnt8 has been shown to be one of the inductive factors establishing Pax3 expression in the developing neural folds (Bang et al., 1999) the expression of Wnt1 in the neural tube has been shown to be follow Pax3 temporally (Deardorff, 2001) and the expression of neural crest cell marker genes, such as slug, has been shown to depend upon a Pax3 mediated Wnt signal (Monsoro-Burq, 2004; Sato, 2005). The observation that Wnt1 expression levels are reduced on a Pax3 null background (Conway et al., 2000) but that Pax3 expression levels remain unchanged on a Wnt1 mutant background (Ikeya et al., 1997) suggest that Pax3 is genetically upstream of Wnt1 and led to the investigation of the following hypothesis:

**Hypothesis One**

*Pax3 up regulates Wnt1 transcription directly, probably through the 3' distal enhancer region, in vivo during the development of the neural crest.*

This hypothesis was tested in Chapters 3 and 6. In Chapter 3 this problem was initially addressed by sub-cloning the 5.5kb 3’ distal Wnt1 enhancer, first defined as necessary and sufficient to recreate the Wnt1 expression pattern in the developing mouse CNS by (Echelard et al., 1994) into a series of Luciferase reporter constructs. An interaction between the 3’ distal Wnt1 enhancer and Pax3 could not be detected.
Bioinformatic analysis had identified a *Pax3* consensus within the *Wnt1* 5’ proximal promoter region. Despite the inability of the *Wnt1* proximal promoter to drive reporter gene expression in a faithful reproduction of the *Wnt1* expression pattern in the mouse, this region of DNA has been shown to be important in a regulatory context (Lagutin et al., 2003). The 5’ proximal promoter was then also cloned into a Luciferase reporter construct and shown to be *Pax3* responsive. Deletion studies confirmed that this interaction was mediated by a region of the *Wnt1* promoter carrying the *Pax3* consensus site. The specificity of this interaction was finally confirmed by site directed mutagenesis.

These experiments illustrated that *Pax3* was competent to up-regulate the transcription of a reporter gene under the control of the *Wnt1* 5’ proximal promoter. To confirm this interaction *in vivo*, ChIP was conducted using an antibody designed specifically for this purpose. This experiment confirmed the occupancy of *Pax3* specifically on this region of mouse genomic DNA, and not on the distal 3’ enhancer.

This enables hypothesis one to be concluded, with the modification that *Pax3* is binding to the proximal promoter and not the distal enhancer element of this gene. This result sits well with previous studies examining the roles of these genes in the development of the neural tube and induction of neural crest. A model of the proposed interactions can be seen in Figure 7.1 (Brault et al., 2001; Dickinson et al., 1994; Ikeya et al., 1997; Monsoro-Burq, 2004; Sato, 2005; Serbedzija and McMahon, 2004). Whilst other factors must clearly regulate the expression of *Wnt1* in addition to *Pax3* (i.e. *Msx1* (Lallemand, 2003; Monsoro-Burq, 2004), or *WiF* (St-Arnaud and Moir, 1993)), the above data in conjunction with the observed reduction of *Wnt1*
Figure 7.1:
A summary of the genetic network involved in neural crest induction, some of the factors involved, and the proposed position of the Pax3-Wnt1 relationship determined by this thesis.
expression by around 40 (Chapter 3) – 50% (Conway et al., 2000) on a Pax3 null background clearly implicates Pax3 as an important regulator of Wntl expression in vivo.

Only one study to date, (Burstyn-Cohen, 2004), has cited evidence that Wntl may lie upstream of Pax3 in the chick. The data presented by this group does not differentiate between specific Wnt signals, however, and has only demonstrated that Pax3 is downstream of a canonical Wnt signal. Since Pax3 has been shown to be induced in the neural plate by Wnt8 in previous work in Xenopus (Bang et al., 1999), and given that the rostrocaudal wave of Pax3 expression is widely held to precede that of Wntl (see (Deardorff, 2001) for a time course analysis of dorsal neural tube markers in Xenopus), the role of Wntl as an inducer of Pax3 expression in the neural tube seems unlikely.

When a Wntl-LacZ transgene was crossed onto a Pax3 null background a clear loss of migrating LacZ positive neural crest cells was observed (Serbedzija and McMahon, 1997). The possibility that a Pax3 – Wntl interaction may play a role in events after the induction of the neural crest population was also considered. Mice carrying a mutation in the Dvl2 gene (Dvl2 is a downstream component of both canonical and PCP Wnt signalling pathways) share an identical cardiac neural crest phenotype to Pax3 mutant mice (Briata, 2003; Clevers, 2002; Hamblet, 2002; Kioussi, 2002). This places Wntl in an attractive position to act as the Wnt signal in this process. This would also link the Pax3 and Dvl2 cardiac outflow tract phenotypes with a common mechanism. A mutation in Pax3 would reduce the expression of Wntl in cardiac neural crest, thus inhibiting signalling through the Wnt/Dvl2/Pitx2 pathway and
generating a cardiac outflow tract phenotype. Another interesting feature of this pathway is the partially non-cell autonomous nature of the Sp$^{\text{m}}$ cardiac phenotype (Chan et al., 2004; Conway et al., 1997a; Mansouri et al., 2001), again implicating a secreted signal such as Wnt1 in the development of these tissues. Finally, one group has noted in cardiac neural crest ablation experiments that defects in the developing chick heart actually begin to manifest whilst the migrating cells would still be populating the pharyngeal arches (i.e. before reaching the heart) (Waldo, 1999). One of these defects was in Ca$^{2+}$ transients, a measure of excitation – contraction coupling, and is a phenotype also observed in Pax3 mutant embryos (Creazzo, 1998). Wnt1, signalling via $\beta$-catenin, has also been shown to regulate the expression of the connexin43 gene (critical in the formation of gap junctions in the heart which mediate the electronic coupling in cardiac tissues) (Zhaowei, 2000). Wnt1-cre targeted inactivation of $\beta$-catenin also generates defects in neural crest derived structures and cell survival (Brault et al., 2001). It is tempting to speculate that Wnt1, as a target of Pax3 and signalling via $\beta$-catenin, mediates these effects in vivo.

Msx2 and Wnt1

To establish a role for Wnt1 as a possible effector of Pax3 function during cardiac development the findings of (Kwang et al., 2002) needed to be addressed. In this study Pax3 was shown to negatively regulate the Msx2 gene in the neural crest. In the Sp$^{\text{m}}$ -/- mouse, the Pax3 null phenotype generates an expanded expression domain of Msx2 at the level of the post-otic neural tube (the area from which the cardiac neural crest originate (Chan et al., 2004)). This upregulation of Msx2 on a Pax3 null
background was shown to be causative in the development of the *Splotch* cardiac phenotype, since a double *Msx2/Pax3* homozygous mutant genotype rescues this defect specifically. It was reasoned that if *Wntl* is also an effector of *Pax3* signalling in the development of the cardiac neural crest, *Msx2* should act to downregulate *Wntl* expression.

Therefore hypothesis two was also investigated:

**Hypothesis Two**

*Wntl* transcription is directly downregulated by *Msx2*, most probably by the distal 3' enhancer element, in vivo with implications for the normal development of the cardiac neural crest.

To investigate this hypothesis the *Msx2* protein was co-expressed in a series of Luciferase transfection experiments, and a responsive element mapped to a region of the *Wntl* enhancer containing a highly conserved 110bp sequence (first defined as necessary and sufficient for LacZ transgene expression in the *Wntl* domain of the mouse midbrain (Rowitch et al., 1998)) and an *Msx2* binding site (Chapter 3). Unfortunately, mutational analysis of this site was not performed and the information obtained from a series of ChIP experiments using the DSHB Msx1/2 antibody was inconclusive (Chapter 6). It was tentatively concluded that *Msx2* might associate with the *Wntl* locus in vivo, but more experiments are required to confirm this association. Also, ChIP experiments using both a specific *Msx2* antibody and possibly using isolated neural crest cells would help to clarify this result.
The hypotheses outlined above were designed to identify and test possible direct interactions between *Pax3* and target gene cis-regulatory sequences both *in vitro* and *in vivo*. The developmental systems outlined above and in Chapter 1 are presented as a relevant context for these interactions. To understand what the relationship between *Pax3*, *Wntl* and *Msx2* may be functionally during development, the data described here can now be used to design specific transgenic experiments (for example the mutation of the *Pax3* binding site in the *Wntl* promoter in a transgenic line, or assaying for a rescue phenotype if the *Pax3* gene is expressed under the control of the *Wntl* regulatory elements on a *Sp<sup>−/−</sup>* background) to test these relationships further.

**Pax3 and Pax7**

Another developmental system was chosen for the second candidate target of *Pax3* regulation. The controversy over the role which *Pax3* may play in the regulation of *Pax7* is described in Chapter 1 (see (Borycki et al., 1999; Relaix et al., 2004)). To test any possible direct regulatory interaction between *Pax3* and its sister *Pax* family member a similar approach to that outlined for *Wntl* was adopted, see Chapter 4. Because regulatory regions for the *Pax7* gene had not been described for the mouse the identification of these regions formed the initial part of the investigation into this candidate target.

Three techniques were used to try and position putative *Pax7* promoter regions on the basis of the size and sequence of the *Pax7* 5'UTR. All techniques converged on a
5'UTR of around 100bp of upstream genomic sequence, which was consistent with a paper published subsequently on the mouse Pax7 promoter. This study evidenced a 5'UTR of 73bp although the PEXT gel presented by this group does indicate faint bands at higher and lower molecular weights (Lang et al., 2003). One interesting exception in my data was the presence of consistently larger PEXT products, indicating a second transcriptional start site at a different location for the mouse Pax7 gene. Studies in human (Syagailo et al., 2002) have also reported a larger 5'UTR (~650bp), and the methods presented in (Lang et al., 2003) may have missed a larger PEXT product for technical reasons (running the PEXT products on a sequencing gel may mean larger (>500bp) fragments do not resolve). 5'RACE may not have amplified these larger fragments due to technical issues with the PCR and cloning basis of this protocol and the CG and repeat rich nature of the genomic sequence in question. The lack of confirmation of these larger fragments by RPA is puzzling; this may indicate the presence of an upstream exon in the mouse Pax7 gene outside of the region covered by the RPA probe. It is interesting to note that, of the six studies to date characterising the mouse and human Pax7 5'UTR's, no definitive consensus on the transcriptional start has been reached. It is possible that a defined start site may not exist for this gene. The existence of promoters with multiple transcriptional start sites has long been accepted and may even be of regulatory importance (see (Ayoubi and Van den Ven, 1996) for review). Of the LacZ reporter transgenics reported in (Lang et al., 2003), even those containing 10kb of genomic sequence upstream of the Pax7 transcriptional start did not fully recreate the Pax7 binding pattern. This implies that additional regulatory elements to those reported here and in (Lang et al., 2003) must be involved in the regulation of this gene. It is possible the larger PEXT product observed here can be explained in this manner. Perhaps the PIPMaker alignment
output, shown in Appendix 3, could be used as a starting point to identify regions of putative regulatory importance.

The I1R silencer within intron 1 was also reported by (Lang et al., 2003) although this element was not isolated and tested. Interestingly, they report that, in transgenic mice carrying both the I1R element and 4kb of proximal Pax7 genomic sequence, the I1R silencer seems to act as a pons-specific enhancer with respect to LacZ staining.

Once candidate regulatory regions for the Pax7 gene had been described the following hypothesis was investigated:

**Hypothesis Three**

\[ \text{Pax7 is directly downregulated by Pax3, in vitro and in vivo, via defined regulatory elements and with implications for the development of both neural and mesodermal tissues.} \]

In this study we cloned a minimal Pax7 promoter, restricted to 1.5kb by activity in C2C12 cells, and the I1R element both individually and together in a series of Luciferase reporter constructs. These were co-transfected with Pax3 to test the above hypothesis. It was discovered that both elements were required on a reporter construct to elicit a response to Pax3. This response was a de-repression of the I1R silencing activity of the Pax7 promoter observed for these constructs \textit{in vitro}. This relationship was unexpected since, if any regulatory interaction existed between Pax3 and Pax7 at all, it had been proposed to be negative (Borycki et al., 1999). Mutagenesis confirmed
that the *Pax3* response was not via a *Pax3* consensus in the *Pax7* promoter region, and ChIP indicated that *Pax3* bound the I1R element *in vivo*. These data confirm that *Pax7* is a direct target of *Pax3*, but imply a positive relationship which is at odds with one published account of *Pax7* expression on a *Pax3* null genotype (Borycki et al., 1999). Future work should map this *Pax7* binding site on the I1R element precisely through mutagenesis, and then a transgenic approach should be adopted to interpret the developmental significance of this. QRTPCR was attempted for *Pax7* between wild type and *Sp2H*/-/- embryos, but a consistent result could not be obtained despite several attempts (data not shown). Recent work identifying a *Pax3* and *Pax7* dependant population of skeletal muscle progenitor cells highlights the importance these genes have in adult muscle regeneration (Relaix, 2005). Establishing a regulatory interaction between these factors may therefore be significant in this context.

**Pax3 ChIP: future directions**

The establishment of direct vs. indirect targets for the action of *Pax3* and other developmentally important transcription factors is essential in the interpretation and investigation of the function of these genes *in vivo*. Previous attempts have been made using a microarray approach (Mayanil et al., 2001) to screen for candidate targets of *Pax3*. The inability of these techniques to differentiate between primary and secondary effects of *Pax3* function make the meaningful interpretation of these data difficult.
It is intended that the ChIP technique designed in this thesis may be used as a way of rapidly testing direct functional relationships between \(Pax3\) and potential target genes \textit{in vivo}. In addition to testing specific DNA-protein interactions, ChIP can be used in combination with genome wide arrays or CpG island arrays to screen for sites of direct transcription factor occupation (Buck, 2004; Hanlon, 2004; Weinmann, 2002). This approach, whilst relatively new in mammals due to limitations in the microarrays and bioinformatics available, has been used with great success in yeast (Lee, 2002; Lieb, 2001; Ren, 2000a). It is hoped that the validation of ChIP in mouse embryonic tissues described here will be used in future to screen for direct transcriptional targets of \(Pax3\) during development.
### Appendix 1: Constructs and cloning strategies

<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Cloning Strategy Outline</th>
</tr>
</thead>
<tbody>
<tr>
<td>pP3IRESEGFP / pP3IRES</td>
<td>RTPCR from E10 RNA using Primers 3/5 P3IRES, topo cloned and then shuttled using the Sacll-BamHI sites form the primers into pIRES-2-EGFP (Clontech)</td>
</tr>
<tr>
<td>pW1ELuc</td>
<td>Wnt-1 enhancer plasmid (McM -pWEXP3) BglII fragment of 5.4kb using a BglII / Sall double digest and then cloned into pGL3Promoter (Promega) BamHI site.</td>
</tr>
<tr>
<td>pW1PLuc</td>
<td>pMT86 (JOM) digested with BamHI/NcoI to generate 4.89kb promoter fragment. This cloned into pGL3Basic BglII/NcoI sites.</td>
</tr>
<tr>
<td>pP7HYB</td>
<td>PAC Library Pax-7 promoter screen. Human DNA amplified using P7HYB1L / 2R primers, 520bp product topoisomerased into vector.</td>
</tr>
<tr>
<td>pP7Probe</td>
<td>PCR topo clone using P7RT1I/R primers to generate 770bp product (across the CpG island of Pax-7 5' proximal region)</td>
</tr>
<tr>
<td>pCMV-Pax-3</td>
<td>pP3IRES (Sacll - BamHI) into pCMV-Script (stratagene) (Sacll - BamHI).</td>
</tr>
<tr>
<td>p3.5TOPO</td>
<td>Pax-7 promoter fragment TEPCR cloned from I4 PAC and crystal violet extracted before topoisomerasing into pCR-XL-TOPO (invitrogen). HindIII screened to check orientation.</td>
</tr>
<tr>
<td>pP7P-EGFP</td>
<td>See Step 1, above.</td>
</tr>
<tr>
<td>pP7RiboProbe</td>
<td>Take pP7Probe and clone a 686bp Pax-7 promoter fragment using Asp718 - Fspl and insert into pBSII(KS) with Asp718 - EcoRV. BamHI or HindIII can be used to linearise at either end (or double to confirm fragment) to synth. riboprobe from with T3 or T7, depending on sense or antisense being required.</td>
</tr>
<tr>
<td>p3.5TOPO</td>
<td>p3.5TOPO (BamHI - Fspl) -take 2.8kb frag into pBSII(KS) (BamHI-EcoRV) to create plintermediate1 (below) and then digest plintermediate1 with Sacl - Xhol - Fspl triple, extract the 2.9kb frag and insert into pGL3basic (promega) via Sacl - Xhol. NB. p3.0Luc called p3.5Luc in earlier notes, renamed to be more accurate.</td>
</tr>
<tr>
<td>p3.0Luc</td>
<td>See above</td>
</tr>
<tr>
<td>p1.5Luc</td>
<td>Take p3.0Luc, digest with Asp718, excise 1.404kb frag of upstream promoter, re-ligate backbone to generate p1.5Luc with proximal 1.5kb of Pax-7 promoter.</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>p-1.8Luc</td>
<td>Take p3.0Luc, digest with BglIII, excise 1.755kb of Pax-7 promoter proximal to ATG of Luciferase. Re-ligate backbone to generate Luciferase construct with distal 1.8kb of Pax-7 promoter (deleted in p1.5Luc) controlling Luciferase.</td>
</tr>
<tr>
<td>pl1RLuc</td>
<td>TEPCR PAC DNA using P7INT1REGF/B, generate 546bp, digest frag with BamHI (sites in primers), clone into BamHI site of pGL3Promoter (both orientations, orientation screen performed by SacI / Apal / Xbal triple digest)</td>
</tr>
<tr>
<td>p1.5I1RLuc</td>
<td>TEPCR PAC DNA using P7INT1REGF/B, generate 546bp, digest frag with BamHI (sites in primers), clone into BamHI site of p1.5Luc (both orientations, orientation screen performed by SacI / Apal / Xbal triple digest).</td>
</tr>
<tr>
<td>pW1P(1.2)Luc</td>
<td>pW1PLuc, Asp718 digest to delete the distal 3.5kb of Wnt-1 promoter</td>
</tr>
<tr>
<td>pW1P(-1.8)Luc</td>
<td>pW1PLuc, Apal digest to delete the proximal 3.0kb of Wnt-1 promoter (mis-named) KEEPS 250bp upstream of ATG so a better plasmid than p(-1.8)Luc which only keeps 50bp</td>
</tr>
<tr>
<td>p7SP6RiboProbe</td>
<td>pP7RiboProbe, BamHI - HindIII extraction of 686bp frag and then cloning into pGEM-3ZF using BamHI - HindIII. Linearise with BamHI later for SP6 transcription of riboprobe for RPA</td>
</tr>
<tr>
<td>pCMV-MSx2</td>
<td>Ordered from HGMP, seq to ensure full length cDNA and correct orientation wrt CMV prom (originally pCMV-Sport (Invitrogen?)</td>
</tr>
<tr>
<td>pW1E(Bam)Luc</td>
<td>pW1ELucBF (BamHI) -fragment excised and vector re-ligated to make enhancer deletion construct</td>
</tr>
<tr>
<td>pW1E(Spe)Luc</td>
<td>pw1ELucBF (SpeI) -fragment excised and vector re-ligated to make enhancer deletion construct</td>
</tr>
<tr>
<td>Plasmid</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>pW1E(Mbi)Luc</td>
<td>pw1ELucBF (MbiI), 816bp fragment carrying putative Msx2 response element excised and re-ligated into pGL3Promoter (Promega) SmaI site and then screened for orientation using</td>
</tr>
<tr>
<td>pW1EPLuc</td>
<td>pW1P(1.2)Luc (Asp718 - Ncol), take 1.236kb fragment, ligate into pW1E(Bam)Luc (Asp718 - Ncol) 7.937 backbone fragment (loose 277bp containing the SV40 minimal promoter used in pGL3Promoter based vectors).</td>
</tr>
<tr>
<td>pW1PΔLuc</td>
<td>Site directed mutagenesis of Pax-3 core consensus in the Wnt-1 promoter sequence on this plasmid. CTCGC - GATAA.</td>
</tr>
<tr>
<td>pP7PΔLuc</td>
<td>Site directed mutagenesis of Pax-3 core consensus in the Pax-7 promoter sequence on this plasmid. TCGC - ATAA.</td>
</tr>
</tbody>
</table>
## Appendix 2: Primer List

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence (5' - 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3GENOF</td>
<td>AAGCAGCGCAGGAGCAGAACC</td>
</tr>
<tr>
<td>P3GENOR</td>
<td>CCTCGGTAAGCTTCGCCCCCTCT</td>
</tr>
<tr>
<td>W1LITEF</td>
<td>CGCTCTCTCCAGTTTCAGACAC</td>
</tr>
<tr>
<td>W1LITER</td>
<td>CAGGATGGCAAAGGGGTTCG</td>
</tr>
<tr>
<td>GAPLITEF</td>
<td>GGGTGTGAACCACGAGAAAT</td>
</tr>
<tr>
<td>GAPLITER</td>
<td>CTTCCACAATGCCAAAGTT</td>
</tr>
<tr>
<td>3P3IRES</td>
<td>GATACAGGATCCTAGAACCTCCAACAGGCTTACTTTG</td>
</tr>
<tr>
<td>5P3IRES</td>
<td>TGTATCGCAGGAAGACGCCGCAAGTCCAACAGGCTTACTTTG</td>
</tr>
<tr>
<td>NT - reverse</td>
<td>TGTCACCTTGGGTTTGCTG</td>
</tr>
<tr>
<td>E1L</td>
<td>GCTGTGCCCCAGGATGATG</td>
</tr>
<tr>
<td>E4L</td>
<td>GAGTTCTATCAGCCGCATCC</td>
</tr>
<tr>
<td>E4R</td>
<td>CCATCGATGCTGTGTTTAGC</td>
</tr>
<tr>
<td>E6L</td>
<td>GCAAGATGGAGAAACAA</td>
</tr>
<tr>
<td>E6R</td>
<td>GGTAAGAGTGCTCCGACAGC</td>
</tr>
<tr>
<td>E7L</td>
<td>CCGTGTCAATCCCCAGTAGC</td>
</tr>
<tr>
<td>E7R</td>
<td>GATGGAGGCAAAAGCTGTC</td>
</tr>
<tr>
<td>E8R</td>
<td>GCCCATACTGCTAGGCTGTG</td>
</tr>
<tr>
<td>CT - forward</td>
<td>TTTTCACATCTCTCTTG</td>
</tr>
<tr>
<td>W1_sub_FWD</td>
<td>CGGCAAGAGCCACAGCTTGCTTAACACTC</td>
</tr>
<tr>
<td>W1_sub_REV</td>
<td>AAGCTGTGCTTCCTTGGCC</td>
</tr>
<tr>
<td>EX2</td>
<td>AACTACCCCGGCACCAGGCTTCACACACACACCGT</td>
</tr>
<tr>
<td>B2</td>
<td>CGGAACCACATCCGTCAAA</td>
</tr>
<tr>
<td>B4</td>
<td>CACATCCGTCACAAGGATGTTAAG</td>
</tr>
<tr>
<td>5'RACE (outer)</td>
<td>CTGCCGAACCACATCCGTCACAAG</td>
</tr>
<tr>
<td>5' RACE (inner)</td>
<td>TCAATCAGCTTTGCTGCCGGT</td>
</tr>
<tr>
<td>P7RT1L</td>
<td>CCAAGAGGTTTATCCAGCCGA</td>
</tr>
<tr>
<td>P7RT1R</td>
<td>TTGATGAAGAGCCACACAAAGCC</td>
</tr>
<tr>
<td>P7HYB1L</td>
<td>GTAGAGTAAGAACCAGGACACCG</td>
</tr>
<tr>
<td>P7HYB2R</td>
<td>CAACTGAATGATCGAGGAGTCAGG</td>
</tr>
<tr>
<td>P7PROML</td>
<td>CCCTGCTGGATGTAACATCG</td>
</tr>
<tr>
<td>P7PROMR</td>
<td>CGGTCTGCTGAGGAGGGTAGG</td>
</tr>
<tr>
<td>P7INT1REGF</td>
<td>GTAGATGGATCTTCTAGGGCCGGTTAGGTCATTG</td>
</tr>
<tr>
<td>P7INT1REGB</td>
<td>AGTTTGAATCCTACCGGCCCTTTCCTTACCTT</td>
</tr>
<tr>
<td>P7_sub_FWD</td>
<td>AGTGGTTTGGTTTGAACCTCTTGCTGATATACCTT</td>
</tr>
<tr>
<td>P7_sub_REV</td>
<td>CACTCCTCCCGCCCC</td>
</tr>
<tr>
<td>P7_sub_FWD</td>
<td>GGGGCCGGGAGGTGAAGGTATATCCAAAGGA</td>
</tr>
<tr>
<td>P7_sub_REV</td>
<td>GTTCAAAACACACAAACACT</td>
</tr>
</tbody>
</table>
P7_PROM_CHIP_FWD   CCCTCGCTTTTCCTCTTGTTCC
P7_PROM_CHIP_REV   GCCAAGAGGTTTATCCAGCCGA
P7_INTRON_CHIP_FWD TTAGCCGCCCTCTTGTCATTG
P7_INTRON_CHIP_REV GAAAGTAAAGAGGGGGCGGTAG
W1_PROM_CHIP_FWD   GGGACAGAGGAGACGGACTTC
W1_PROM_CHIP_REV   CATCACTGCCCTCACCCT
W1_ENHAN_CHIP_FWD  CGTCAGCCTGGATTAATCTTCG
W1_ENHAN_CHIP_REV  CGTTCACGTAGTGCTCCCCAA
Appendix 3: PIPMaker Output
Appendix 4: Luciferase Assay Data Treatment

Luciferase transfections were performed as above using the Promega DLR Assay system and a TD-20/20 Luminometer (Turner Biosystems). There were two distinct types of Luciferase assay performed in this thesis. The nature and purpose of these, and the way in which the data has been processed in each case, are described here.

Luciferase assays are transient transfection experiments where the expression of the Luciferase reporter gene can be used to infer the properties of various cis-regulatory elements in vitro. Whilst these methods generally involve plasmid DNA in secondary cell lines and are therefore not subject to all of the levels of gene regulation thought to occur in an in vivo context (i.e. chromatin remodelling, post-transcriptional modification, etc), the benefit of these methods is their rapid reproducibility and the chance to carefully dissect functional elements of non-coding DNA in much greater detail than would otherwise be practical.

In this thesis all Luciferase assays were conducted as transfections in 24 well plates, using a final total of 1.5µg DNA in each well. This was to ensure no variability in transfection efficiency arose from different quantities of DNA in each experiment. Between two and four plasmids were simultaneously transfected into cells in any one transfection. In all cases a reporter vector, carrying the Firefly Luciferase ORF under the control of various pieces of DNA, was transfected simultaneously with an identical quantity of a second plasmid carrying a Renilla Luciferase ORF whose transcription is controlled by a known element (either the SV40 or Thymidine Kinase promoters). Both gene products are enzymes capable of metabolising specific
substrates with the concomitant production of light. In these assays the metabolite specific for the Firefly luciferase gene is added first, and the quantity of light produced over a fixed period of time (10 seconds) is measured by the Luminometer. The quantity of light produced is directly proportional to the quantity of enzyme produced, as predicted by Michaelis-Menten kinetics where rate, \( v \), is related to concentration of enzyme by the relationship:

\[
 v = \frac{V_{max} [S]}{K_m + [S]}
\]

where \( V_{max} = k_2 [E_{total}] \), \([S]\) = concentration of enzyme substrate, and \( k_2 \) and \( K_m \) represent kinetic constants. Where substrate is constant and in great excess (i.e. so \( \Delta[S] \to 0 \) over the time course observed), increasing the concentration of enzyme \( ([E_{total}] \) will increase the rate of reaction and therefore the quantity of light produced over a fixed period. A detailed derivation of this relationship can be found in (Price, 2001).

After the first measurement of light produced by the Firefly Luciferase has been made, the reaction in the sample is quenched chemically and a second substrate is added which is specific to the Renilla Luciferase enzyme. This activity is then measured in the same way. Since the first reaction is quenched before the second reading, two independent and specific Luciferase activity readings can be taken from the one sample. This is important since the second measurement of the Renilla Luciferase activity is made to control for the differences in transfection efficiency, cell lysis, etc which may vary. If the first reading increases by 50% between a treated and untreated sample, for example, and the second reading remains constant it is
assumed that the treatment has increased the transcription of the Firefly Luciferase gene by 50% through interacting with whichever DNA sequences are controlling its transcription. If both readings increase by 50%, it is assumed that this change is due to differences in experimental conditions.

The first Luciferase reading is therefore divided by the second reading to correct for this type of experimental variation and the resulting ratio, called normalised luminescence throughout, is taken as the real measure of variation between samples.

Data is expressed as the mean value of several transfections. Each experimental run was performed in triplicate, to ensure each transfection was repeatable within a run, and then experimental runs repeated at least three times to generate data. Each reading on the Luminometer was also taken twice to ensure no variability in luminescence could be observed within the same sample over the time period assayed (i.e. that $\Delta[S] \to 0$ is true).

In the first type of assay, the measurement of the basal activity of a putative promoter element is examined. In these assays normalised luminescence is measured for transfections where the Firefly Luciferase gene is placed under the control of a specific piece of DNA (such as the Wnt1 promoter). Two plasmids are co-transfected in each case (a test firefly Luciferase plasmid and a control Renilla Luciferase plasmid). Data is gathered for these transfections and expressed as fold induction. This is the number of times greater (or smaller) the normalised luminescence for these samples is over a baseline. In this case the baseline is defined as the normalised luminescence obtained for sample transfected with the Firefly Luciferase gene under
the control of the empty vector sequences (specifically the pGL3basic or pGL3promoter plasmids).

In the second type of experiment the effect of co-transfecting an expression plasmid containing the ORF of a transcription factor of interest with plasmids containing the Firefly Luciferase gene under the control of specific regulatory element is performed. The normalised luminescence of samples transfected with differing ratios of expression construct to empty expression construct are compared. Three or four plasmids are transfected in each case; Firefly reporter, Renilla control, expression construct and/or empty expression construct to ensure the total quantity of DNA added remains constant (this is also to ensure a consistent amount of CMV promoter, driving expression of the transcription factor of interest, is added to prevent 'squelching' or an expression data artefact due to competition of strong promoters for RNA polymerase machinery in the cell). In these experiments fold induction of normalised luminescence over baseline is generally reported.

Occasionally, it was necessary to represent data as 'indexed fold induction'. This was used to compare between samples where a direct comparison of fold induction was not possible. This was always for one of two reasons. Firstly, when considering the interaction of a transcription factor with either a promoter element or an enhancer element a meaningful comparison is impossible. This is because the two vectors used as baseline (pGL3basic or pGL3promoter, respectively) have very different inherent Luciferase activities. As a result, the actual numbers generated for fold induction are different. Secondly, when comparing effects of co-expression across a large data set or between different reporter plasmids of different sizes, the differences in absolute
value of fold induction was observed to vary, although the relative changes were similar. For example, one data set might generate a change in fold induction of 0.1 to 10 whilst another might generate 10 to 1000. Clearly, the relative effect of treatment is identical, but an efficient comparison of the absolute values is impossible. Since all these experiments are concerned with the changes induced by co-expression of a particular cDNA, not the relationship to background observed in each separate experimental run, these data were indexed by setting the fold induction observed for the untreated sample (i.e., 0ng pCMV-Pax3) at 1. In these cases, relative or indexed fold induction is reported (essentially the % change induced by a specific quantity of expression plasmid) to enable direct comparison.

Finally, control experiments where pGL3basic and pGL3promoter empty reporter constructs were co-transfected with a range of quantities of pCMV-script, pCMV-Pax3 and pCMV-Msx2 expression constructs, using both pRL-TK and pRLSV40 Renilla transfection control plasmids were performed. No change in baseline Luminescence was ever observed for any of these controls (i.e. < 3-5%, data not shown).
Bibliography


outflow tract development, somite segmentation and neural tube closure. 

*Development* **129**, 5827 - 5838.


**HGMP.** http://www.hgmp.mrc.ac.uk, (ed.


Kay, P. H. and Ziman, M. R. (1999). Alternate Pax7 paired box transcripts which include a trinucleotide or a hexanucleotide are generated by use of alternate 3’ intronic splice sites which are not utilized in the ancestral homologue. *Gene* 230, 55-60.


MatInspector, G.-. Transcription factor binding site prediction software. 


Relaix, F., Polimeni, M., Rocancourt, D., Ponzetto, C., Schafer, B. W. and Buckingham, M. (2003). The transcriptional activator PAX3-FKHR rescues the
defects of Pax3 mutant mice but induces a myogenic gain-of-function phenotype with ligand-independent activation of Met signaling in vivo. *Genes Dev* 17, 2950-65.


systems reveal that the PAX3-FKHR fusion oncoprotein regulates CXCR4, PAX3, and PAX7 expression. *Lab Invest* **84**, 1060-70.


