The functional significance of multiple Pax7 isoforms

Chantel L. Burchill

Edith Cowan University

Follow this and additional works at: https://ro.ecu.edu.au/theses_hons

Part of the Genetics Commons

Recommended Citation

This Thesis is posted at Research Online.
https://ro.ecu.edu.au/theses_hons/1068
Edith Cowan University

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.

- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author’s moral rights contained in Part IX of the Copyright Act 1968 (Cth).

- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.
THE FUNCTIONAL SIGNIFICANCE OF

MULTIPLE Pax7 ISOFORMS

Chantel L. Burchill

Supervised by:
Dr. Melanie Ziman and Mr. Robert White

This Thesis Submitted in Partial Fulfilment of the Requirements for the
Award of

Bachelor of Science (Human Biology) Honours

In the Faculty of Computing, Health and Science,
Edith Cowan University, Joondalup, Western Australia.

Date: 21st November 2006
The *Pax7* gene is critical for specification of both neurons in the mid-brain and skeletal muscle satellite cells. Several alternate transcripts are transcribed from the single gene. Previous studies have shown that the resultant alternate *Pax7* isoforms differ in the structure of their paired domain (a DNA-binding domain that influences target gene selection), yet the functional significance of each isoform for specification of neurogenic and myogenic cell types remains unknown. Although previous studies have identified the presence of multiple alternate *Pax7* transcripts in both neurogenic and myogenic cell lines, more research is necessary to understand the functional significance of the alternate *Pax7* isoform that each transcript encodes. The aim of this research was to investigate DNA-binding differences of each *Pax7* isoform to reveal its specific contribution to embryonic development. A chromatin immunoprecipitation technique which purifies genomic fragments bound by individual *Pax7* isoforms was employed, for the purpose of determining DNA-binding differences. Understanding the functional specificity of each *Pax7* transcript is important as these transcripts are possible candidates for future stem cell/ gene therapy approaches aimed at developing novel treatments for Neurodegenerative diseases and Duchene Muscular Dystrophy.

Key words: *Pax7* transcripts, *Pax7* isoforms, neural sub-type specification, skeletal muscle satellite cell specification.
Copyright and Access Declaration

I certify that this thesis does not, to the best of my knowledge and belief:

(i) Incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher education;

(ii) Contain any material previously published or written by another person except where due reference is made in the text; or

(iii) Contain any defamatory material.

Student Signature: _____________________________ Date: ____________

Chantel Burchill

Supervisor Signature: ___________________________ Date: ______________

Dr. Melanie Ziman
ACKNOWLEDGMENTS

This year has been the most stimulating and challenging year out of all my years of being a student, and would not have been manageable without the support and assistance from so many remarkable people.

First and foremost, I would like to express my sincere gratitude to Dr. Melanie Ziman and the rest of the Ziman crew: Jennifer Thompson, Sandra Medic, Amira Wahdan, Dr. Meghan Thomas, Judith Blake and Robert White, for enlightening me with their knowledge and words of wisdom daily. I am most grateful to have met such an extraordinary group of people who showed me the brighter side of life when times were tough, the countless conversations and laughter shared with each and every one of you, will never be forgotten.

To Dr. Inna Sleptsova-Freidrich, thank you for accompanying me throughout endless late nights and weekend shifts in the laboratory. It was a pleasure to share working space with you and with other members of Ralph Martins group, whom always passed by with either a smile or warm greeting. These little moments made my time spent in the laboratory a joy to remember.

To my amazing honours friends, Vijesh Vaghjiani, Jasminka Murdzoska, Emily Jefferson, Shinnyi Cheah and Kirsteen Lingoh, I am thrilled to have shared the honours experience with each of you. Thank you for your friendship and support throughout the year and for reminding me that when times were tough, I was not alone. To my gorgeous circle of university and school friends, Eliza Briggs, Vishal Darji, Michelle Nguyen and Chloe Brophy, I don’t know where I would be without friends like you, your understanding and sweet messages throughout the year encouraged me to do my very best.

To Lorraine Richardson, I am so grateful to have not only a boss but a dear friend who’s endless encouragement and words of wisdom brightened up my day. Thank you for always being there when I needed a shoulder to cry on, you truly are amazing.
To my cousin Gordon Howell, thank you for taking the time to treat me with coffee breaks, I appreciate these moments. To my sweet Grandmar, Louise Anderson, thank you for always believing in me and keeping me in your thoughts daily. I don’t know where I would be without family.

And finally, I cannot end without thanking two most important people in the world, my adorable parents Philip and Virginia Burchill, whom have brought out the best in me and taught me to give one hundred percent at all times, and most importantly, to never give up. I am truly blessed to have two parents who care so much and whom believe in me, your endless encouragements and laboratory visits motivated me to keep going and made me realize that even when times are tough, when you put your mind to it, you can achieve just about anything. I love both of you unconditionally and don’t know where I would be without your advise and heartwarming presence.
TABLE OF CONTENTS

Abstract.............................................................................................................. II
Copyright and Access Declaration....................................................................... III
Acknowledgments............................................................................................. IV
Table of contents.............................................................................................. VI
List of Figures..................................................................................................... XI
List of Tables..................................................................................................... XIII
Abbreviations..................................................................................................... XIV

Chapter 1  INTRODUCTION

1.1.  The background to the study  ................................................................. 1
1.2.  Research question  ................................................................................... 1
1.3.  The significance of the study  ................................................................... 1
1.4.  The purpose of the study  ......................................................................... 2
1.5.  Aim  ........................................................................................................... 2
1.6.  Hypothesis  ............................................................................................... 2

Chapter 2  LITERATURE REVIEW

2.1.  Gene transcription  ..................................................................................... 3
2.1.1. The structure of the promoter .................................................................. 4
2.1.2. Transcription factors  ............................................................................. 5
2.1.3. Enhancers  ............................................................................................... 5
2.1.4. RNA processing  ..................................................................................... 6
2.1.5. RNA splicing  .......................................................................................... 7
2.1.6. Alternative RNA splicing  ...................................................................... 10
2.2.  Pax genes  .................................................................................................. 11
2.2.1. General function and alternate splicing of the developmental Pax genes  11
2.3.  Pax7 protein  .............................................................................................. 13
2.3.1. Structural features of Pax7: DNA binding domains 13
2.3.2. Differential splicing of the Pax7 paired box 15
2.4.  
\textit{Pax7} function 17
2.4.1. \textit{Pax7} in central nervous system development 17
2.4.2. \textit{Pax7} in muscle development 17
2.5.  
Downstream target genes 18
2.5.1. Target genes of the Pax7 protein 18
2.6.  
\textit{Chromatin immunoprecipitation} 20
2.7.  
Summary 22

Chapter 3  \textbf{METHODS}

3.1.  
Research Plan 24
3.2.  
Vector preparation 25
3.2.1. Amplification of epitope-tagged expression vectors containing \textit{Pax7(a-d)} transcripts 25
3.2.2. Isolation of alternate \textit{Pax7(a-d)} vector constructs 27
3.2.3. Agarose gel electrophoresis 27
3.2.4. Polymerase Chain Reaction (PCR) of \textit{Pax7(a-d)-pcDNA3.1D/V5-His-TOPO} cDNA constructs 28
3.2.5. Determination of DNA concentration 29
3.3.  
Cell culture 29
3.3.1. Maintenance of NIH3T3 and P19 cells in culture 29
3.3.2. Transfection and selection of stable clones 30
3.3.3. Photography of cell culture 30
3.4.  
Expression studies to confirm transfection of \textit{Pax7} transcripts in NIH3T3 cells 30
3.4.1. Western blot of transfected cells 30
3.4.1.1. Protein extraction 30
3.4.1.2. Protein concentration determination 31
3.4.1.3. Western blot 31
3.4.1.4. Polyacrylamide gel electrophoresis 31
3.4.1.5. Transfer of proteins to nitrocellulose membrane
3.4.1.6. V5 antibody membrane probing
3.4.1.7. Visualization
3.4.1.8. Stripping the nitrocellulose membrane to re-probe with the Pax7 antibody

3.4.2. **Pax7 transcript RT-PCR**

3.4.2.1. RNA extraction
3.4.2.2. Reverse transcription of RNA from NIH3T3 cells transfected with Pax7(a-d) transcripts
3.4.2.3. Polymerase Chain Reaction (PCR) to detect Pax7 transcript expression

3.4.3. **Immunocytochemistry of transfected cells to detect Pax7 expression**

3.4.3.1. Immunocytochemistry
3.4.3.2. Photography

3.5. **Chromatin immunoprecipitation**

3.5.1. DNA-shearing optimization
3.5.2. Chromatin crosslinking and lysis of nuclei
3.5.3. DNA-shearing and pre-clearing
3.5.4. Immunoprecipitation of Pax7-DNA complexes
3.5.5. Reverse crosslinking of Pax7-DNA complexes

3.5.6. **DNA purification**

3.5.6.1. Purification Kit
3.5.6.2. Phenol Chloroform isoamyl DNA purification

3.6. **ChiP-PCR to detect bound target genes**

3.7. **Regulation of target genes in Pax7d transfected cells**

3.7.1.1. RNA extraction, DNase treatment of RNA and RNA concentration
3.7.1.2. Reverse transcription of RNA from P19 cells transfected with the Pax7d transcript
3.7.1.3. Polymerase Chain Reaction (PCR) to assess target gene expression
Chapter 4  RESULTS

4.1. Introduction 43

4.2. Aim 43

4.3. *Pax7*(a-d)-vector transfection 43

4.3.1. Successful isolation of DNA containing *Pax7*(a-d)-vector constructs 43

4.3.2. Polymerase Chain Reaction (PCR) of 44

*Pax7*(a-d)-pcDNA3.1DIV5-His-TOPO cDNA constructs

4.4. Stable transfection studies 45

4.4.1. Cell morphology 45

4.5. Confirmation of *Pax7* protein expression in stably transfected cells 46

4.5.1. Western blot 46

4.6. Sonication optimization of protein-DNA complexes for ChIP 47

4.6.1. Determining DNA shearing optimization conditions 47

4.7. PCR of chromatin immunoprecipitated NIH3T3 cells 48

4.7.1. PCR of untransfected chromatin immunoprecipitated NIH3T3 cells 48

(Input chromatin)

4.7.2. PCR of stably transfected *Pax7c* chromatin immunoprecipitated cells 49

4.8. Optimization of DNA extraction for ChIP methodology 52

4.8.1. Confirming the presence and quantity of DNA throughout the chromatin 52

immunoprecipitation procedure

4.9. Confirmation of *Pax7* expression in NIH3T3 cells 54

4.9.1. mRNA expression of *Pax7* in transfected cells 54

4.9.2. Immunocytochemistry 55

4.10. Confirmation of *Pax7* expression in a stably transfected P19 clone 57

4.10.1. mRNA expression of *Pax7* in transfected cells 57

4.10.2. Immunocytochemistry 58

4.11. PCR of chromatin immunoprecipitated P19 cells 60

4.11.1. PCR of chromatin immunoprecipitations of cells stably 60

transfected with *Pax7d*-vector construct

4.12. Target gene regulation by *Pax7d* 61

4.12.1. mRNA expression of each target gene 61
Chapter 5 DISCUSSION

5.1. Introduction
5.2. Upregulation of CntfR and Rasa3 target genes by the Pax7d transcript
5.3. Appropriateness of the chromatin immunoprecipitation technique
5.4. Optimization of transfections
5.5. Implications for no target gene selection by Pax7d
5.5.1. Pax7d protein structure and DNA-binding function
5.5.2. Other factors affecting Pax7d DNA-binding
5.5.3. Utilization of chromatin immunoprecipitation for target gene selection in vivo and in vitro
5.6. Advantages and disadvantages of chromatin immunoprecipitation
5.7. Summary

Chapter 6 CONCLUDING COMMENTS

Chapter 7 REFERENCES

Chapter 8 APPENDICES

Appendix 1 Pax7 alternate transcript inserts cloned into the pcDNA3.1D/V5-His-TOPO expression vector
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure 2.1.</th>
<th>The transcription process</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.2.</td>
<td>Transcription factors binding to the promoter</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2.3.</td>
<td>The structure of a promoter region</td>
<td>4</td>
</tr>
<tr>
<td>Figure 2.4.</td>
<td>The function of the enhancer sequence</td>
<td>5</td>
</tr>
<tr>
<td>Figure 2.5.</td>
<td>Modification of the pre-mRNA transcript by capping and polyadenylation</td>
<td>6</td>
</tr>
<tr>
<td>Figure 2.6.</td>
<td>RNA splicing</td>
<td>7</td>
</tr>
<tr>
<td>Figure 2.7.</td>
<td>Exon/intron splice sites</td>
<td>8</td>
</tr>
<tr>
<td>Figure 2.8.</td>
<td>Spliceosomal removal of introns</td>
<td>9</td>
</tr>
<tr>
<td>Figure 2.9.</td>
<td>Alternative RNA splicing generating multiple transcripts</td>
<td>10</td>
</tr>
<tr>
<td>Figure 2.10.a</td>
<td>Pax7 protein structure</td>
<td>11</td>
</tr>
<tr>
<td>Figure 2.10.b</td>
<td>Pax7 protein interacting with DNA</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2.11.</td>
<td>Pax7 paired domain isoforms</td>
<td>14</td>
</tr>
<tr>
<td>Figure 2.12.</td>
<td>The Pax7 proteins crucial role in skeletal muscle development</td>
<td>15</td>
</tr>
<tr>
<td>Figure 2.13.</td>
<td>Chromatin immunoprecipitation technique</td>
<td>16</td>
</tr>
<tr>
<td>Figure 3.1.</td>
<td>A schematic representation of the ChiP protocol</td>
<td>18</td>
</tr>
<tr>
<td>Figure 3.2.</td>
<td><em>Pax7</em> alternate transcript inserted into the pcDNA3.1DIV5-His-TOPO vector containing 3' V5 and His6 sequences (Invitrogen)</td>
<td>20</td>
</tr>
<tr>
<td>Figure 3.3.</td>
<td>pcDNA3.1DIV5-His-lacZ control vector (Invitrogen)</td>
<td>21</td>
</tr>
<tr>
<td>Figure 3.4.</td>
<td>Nitrocellulose membrane sandwich</td>
<td>22</td>
</tr>
<tr>
<td>Figure 4.1.</td>
<td>Agarose gel electrophoresis of purified <em>Pax7</em>(a-d)-pcDNA3.1DIV5-His-TOPO vector DNA</td>
<td>24</td>
</tr>
<tr>
<td>Figure 4.2.</td>
<td>Agarose gel electrophoresis of isolated <em>Pax7</em>-vector DNA amplified by PCR utilizing <em>Pax7</em>-vector specific primers (F4 and BGH respectively)</td>
<td>25</td>
</tr>
<tr>
<td>Figure 4.3.</td>
<td>Micrograph images of stable <em>Pax7</em>-vector transfected clones (day 14) (20x magnification)</td>
<td>26</td>
</tr>
<tr>
<td>Figure 4.4.</td>
<td>Western Blots of Pax7-V5-His-fusion protein probed with V5-antibody (A) or Pax7 antibody (B)</td>
<td>27</td>
</tr>
</tbody>
</table>

XI
Figure 4.5. DNA shearing conditions for NIH3T3 cells

Figure 4.6. PCR products amplified from untransfected NIH3T3 cells (input sample)

Figure 4.7. ChIP-PCR to assess target gene selection by Pax7c-vector constructs immunoprecipitated with the V5 antibody

Figure 4.8. ChIP-PCR for selection of (A) CntfR and (B) Rasa3 by Pax7a identified by immunoprecipitation with the V5 antibody

Figure 4.9. Agarose gel electrophoresis images indicating the quantity of DNA at several steps during the chromatin immunoprecipitation procedure

Figure 4.10. RT-PCR products indicating Pax7 expression

Figure 4.11. Immunocytochemistry of NIH3T3 cells stably transfected with Pax7c-vector construct or vector alone

Figure 4.12. RT-PCR product indicating Pax7 expression

Figure 4.13. Immunocytochemistry of cells transfected with a Pax7d-vector construct relative to untransfected P19 cells

Figure 4.14. ChIP-PCR for target gene selection by Pax7d-vector fusion protein immunoprecipitated with the V5 antibody

Figure 4.15. RT-PCR products indicating regulation of target genes in P19 cells stably transfected with the Pax7d-vector construct

Figure 5.1.a. Pax7d isoform structure

Figure 5.1.b. Pax7d interacting with DNA
**LIST OF TABLES**

| Table 3.1. | Primer sequences, PCR annealing temperatures and expected sizes of PCR products | 40 |
| Table 3.2. | Primer sequences and expected sizes of RT-PCR products | 42 |
| Table 4.1. | Primer sequences, PCR annealing temperatures and expected sizes of PCR products analysed by RT-PCR analysis | 62 |
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>Ab</td>
<td>Antibody</td>
</tr>
<tr>
<td>AG</td>
<td>Adenine and guanine</td>
</tr>
<tr>
<td>AMP</td>
<td>Adenosine monophosphate</td>
</tr>
<tr>
<td>ATCC</td>
<td>American type culture collection</td>
</tr>
<tr>
<td>$A_{260}$</td>
<td>Absorbance reading at 260</td>
</tr>
<tr>
<td>$A_{280}$</td>
<td>Absorbance reading at 280</td>
</tr>
<tr>
<td>BGH</td>
<td>Reverse priming site</td>
</tr>
<tr>
<td>Bp</td>
<td>Base pairs</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>CAG</td>
<td>Glutamine</td>
</tr>
<tr>
<td>CCAAT</td>
<td>Cytosine, adenine and thymine</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>CH3COOK</td>
<td>Potassium acetate</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>CNTF</td>
<td>Ciliary neurotrophic factor protein</td>
</tr>
<tr>
<td>$Cnfr$</td>
<td>Ciliary neurotrophic factor receptor</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>DAB</td>
<td>3,3-diarninobenzidine</td>
</tr>
<tr>
<td>DC300</td>
<td>Digital camera 300</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>Double distilled de-ionised water</td>
</tr>
<tr>
<td>dH$_2$O</td>
<td>Distilled de-ionised water</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s modified eagle medium</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNase</td>
<td>Deoxyribonuclease</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double stranded DNA</td>
</tr>
<tr>
<td>ECL</td>
<td>Enhanced chemiluminescence</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>ECU</td>
<td>Edith Cowan University</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetra-acetate</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electrophoretic mobility shift assay</td>
</tr>
<tr>
<td>Eye4</td>
<td>Eyes absent 4 homologue</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>FCS</td>
<td>Fetal calf serum</td>
</tr>
<tr>
<td>F4</td>
<td>Forward priming site</td>
</tr>
<tr>
<td>g</td>
<td>Grams</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>GA</td>
<td>Guanine and adenine</td>
</tr>
<tr>
<td>Gbx1</td>
<td>Gastrulation brain homeobox 1</td>
</tr>
<tr>
<td>GC</td>
<td>Guanine and cytosine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>GL</td>
<td>Glycine and leucine</td>
</tr>
<tr>
<td>GL+</td>
<td>Inclusion of glycine and leucine</td>
</tr>
<tr>
<td>GL-</td>
<td>Exclusion of glycine and leucine</td>
</tr>
<tr>
<td>GTTTAG</td>
<td>Glycine and leucine</td>
</tr>
<tr>
<td>GU</td>
<td>Guanine and uracil</td>
</tr>
<tr>
<td>G418</td>
<td>Geneticin antibiotic</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>HD</td>
<td>Homeodomain</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>IL-6</td>
<td>Interleukin-6</td>
</tr>
<tr>
<td>in situ</td>
<td>In the natural place</td>
</tr>
<tr>
<td>in vitro</td>
<td>Within an artificial environment</td>
</tr>
<tr>
<td>in vivo</td>
<td>Within a living organism</td>
</tr>
<tr>
<td>IP</td>
<td>Immunoprecipitate</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobases</td>
</tr>
<tr>
<td>Kcnk2</td>
<td>Potassium channel K2</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodaltons</td>
</tr>
<tr>
<td>lacZ</td>
<td>Beta-galactosidase gene</td>
</tr>
<tr>
<td>L</td>
<td>Litre</td>
</tr>
<tr>
<td>LB agar</td>
<td>Luria-Bertani agar</td>
</tr>
<tr>
<td>LB broth</td>
<td>Luria-Bertani broth</td>
</tr>
<tr>
<td>LB medium</td>
<td>Luria-Bertani medium</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>µg</td>
<td>Microgram</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>µL</td>
<td>Microlitre</td>
</tr>
<tr>
<td>ml</td>
<td>Millilitre</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>µm</td>
<td>Micrometer</td>
</tr>
<tr>
<td>µM</td>
<td>Micromolar</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>NaCH₃CO₂</td>
<td>Sodium Acetate</td>
</tr>
<tr>
<td>NaOH</td>
<td>Sodium hydroxide</td>
</tr>
<tr>
<td>NGS</td>
<td>Normal goat serum</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometre</td>
</tr>
<tr>
<td>N3</td>
<td>Neutralization buffer</td>
</tr>
<tr>
<td>OP</td>
<td>Octapeptide</td>
</tr>
<tr>
<td>PAI</td>
<td>N-terminal paired domain subdomain</td>
</tr>
<tr>
<td>Pax</td>
<td>Paired box</td>
</tr>
<tr>
<td>PB</td>
<td>Phosphate binding buffer</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PD</td>
<td>Paired domain</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PE</td>
<td>DNA cleanup buffer</td>
</tr>
<tr>
<td>PIC</td>
<td>Protease inhibitor cocktail</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethyl Sulfonyl Fluoride</td>
</tr>
<tr>
<td>poly(A)</td>
<td>Poly A tail</td>
</tr>
<tr>
<td>Prd</td>
<td>Paired (<em>D. melanogaster</em>)</td>
</tr>
<tr>
<td>Pre-mRNA</td>
<td>Precursor messenger ribonucleic acid</td>
</tr>
<tr>
<td>Prrxl</td>
<td>Paired-related homeobox 1</td>
</tr>
<tr>
<td>P1</td>
<td>Resuspension buffer</td>
</tr>
<tr>
<td>P2</td>
<td>Lysis buffer</td>
</tr>
<tr>
<td>p15</td>
<td>Passage number 15</td>
</tr>
<tr>
<td>p20</td>
<td>Passage number 20</td>
</tr>
<tr>
<td>p21</td>
<td>Cyclin-dependent kinase inhibitor 1A</td>
</tr>
<tr>
<td>Oligo-dT</td>
<td>Homopolymeric oligo dT primers</td>
</tr>
<tr>
<td>Q</td>
<td>Glutamine</td>
</tr>
<tr>
<td>Q+</td>
<td>Inclusion of a glutamine</td>
</tr>
<tr>
<td>Q-</td>
<td>Exclusion of a glutamine</td>
</tr>
<tr>
<td>Rasa3</td>
<td>RAS p21 protein activator 3</td>
</tr>
<tr>
<td>RED</td>
<td>C-terminal paired domain subdomain</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>RPM</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RT-Buffer</td>
<td>Reverse transcriptase-buffer</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcriptase-polymerase chain reaction</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>snRNPs</td>
<td>Small nuclear ribonucleoproteins</td>
</tr>
<tr>
<td>SR</td>
<td>Serine/arginine-rich proteins</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-acetate/EDTA electrophoresis buffer</td>
</tr>
<tr>
<td>TATA</td>
<td>Thymine and Adenine</td>
</tr>
<tr>
<td>Taq</td>
<td><em>Thermus aquaticus</em></td>
</tr>
<tr>
<td>TBS</td>
<td>Tris buffered saline</td>
</tr>
<tr>
<td>TBST</td>
<td>Tris buffered saline with Tween-20</td>
</tr>
<tr>
<td>TD</td>
<td>Transactivation domain</td>
</tr>
<tr>
<td>TE</td>
<td>Tellurium</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylethylenediamine</td>
</tr>
<tr>
<td>TG</td>
<td>Thymine and guanine</td>
</tr>
<tr>
<td>TNS</td>
<td>Tris normal saline</td>
</tr>
<tr>
<td>TOPO</td>
<td>Topoisomerase I</td>
</tr>
<tr>
<td>Tris-Cl</td>
<td>Tris buffer with chloride</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>Tris buffer with hydrochloric acid</td>
</tr>
<tr>
<td>T7</td>
<td>Forward priming site</td>
</tr>
<tr>
<td>U</td>
<td>Units</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra violet</td>
</tr>
<tr>
<td>v/v</td>
<td>Volume to volume ratio</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>V5</td>
<td>Anti-V5 monoclonal antibody</td>
</tr>
</tbody>
</table>

XVI
w/v  Weight to volume ratio
(3')  Three prime)
(5')  Five prime)

Conventional notations used in this thesis:

Pax indicates the gene
Pax indicates the protein
PAX indicates the protein from Homo sapiens
PAX indicates the gene from Homo sapiens
CHAPTER ONE

INTRODUCTION
Chapter One: Introduction

INTRODUCTION

1.1. The background to the study

The Pax7 gene is one of nine members of the Pax family of developmentally important genes. During embryogenesis, Pax7 is predominantly expressed in the central nervous system (CNS), condensing somites and migrating cranial and cephalic neural crest cells (Kawakami et al., 1997; Stoykova and Gruss, 1994; Walther et al., 1991). These discrete expression patterns signify the importance of Pax7 for neuronal differentiation, skeletal muscle development, and the formation of craniofacial structures. Previous studies have identified eight alternate Pax7 transcripts which encode eight alternate Pax7 isoforms.

Pax7 isoforms differ in their paired and transactivation domains, which are regions of the Pax7 protein that are known to interact with deoxyribonucleic acids (DNA) and thus influence target gene selection (paired domain), or form protein-protein interactions and influence gene transactivation processes (transactivation domain) (Schafer et al., 1994). The purpose of this study was to investigate whether Pax7 isoforms that have structurally distinct paired domains bind to different target genes, and therefore possess functional differences that will ultimately contribute in a unique manner to embryonic development.

1.2. Research question:

"DO PAX7 ISOFORMS SELECT DISTINCT TARGET GENES?"

1.3. The significance of the study

Previous knockout and transfection studies have shown that Pax7 is critical for skeletal muscle satellite cell specification and neural sub-type specification in the mid-brain (Thompson et al., 2006; Thomas et al., 2004; Matsunaga et al., 2001; Seale et al., 1999; Jostes et al., 1990). However, the functional significance of each Pax7 isoform for cell type specification is unknown. More research is necessary to discover DNA-binding
differences of each Pax7 isoform, as this will characterize the specific role of the individual isoforms within a cell. Functionally-specific Pax7 isoforms, may provide novel treatments for future stem cell/ gene therapy approaches to muscle and neural degenerative diseases.

1.4. The purpose of the study

Structural differences in the DNA-binding domain exist between each Pax7 isoform which may induce them to select specific target genes. From identification of target specificity this study aimed to reveal the functional significance of each Pax7 isoform. Moreover, this research sought to reveal whether Pax7 isoforms are functionally distinct; are required to function in combination, or, are redundant for the overall function of the Pax7 gene.

1.5. Aim

To identify Pax7 isoform-specific differences in target gene selection, so as to characterize each Pax7 isoforms’ specific role during embryonic development.

1.6. Hypothesis

The DNA-binding domain differences of each Pax7 isoform play a distinct role in binding to specific target genes, and result in different target gene selection.
CHAPTER TWO

LITERATURE REVIEW
Chapter Two: Literature Review

2.1. GENE TRANSCRIPTION

Gene transcription is an elaborate molecular process that occurs in the nuclei of cells involving the separation of the DNA double helix and transcription of one strand by a ribonucleic acid (RNA) polymerase II enzyme (reviewed in Macfarlane, 2000). This enzyme generates a primary RNA transcript or precursor messenger RNA (pre-mRNA) molecule (Rosenthal, 1994) that is complementary to the template DNA coding strand (Figure 2.1). The pre-mRNA molecule is then spliced and modified into mature mRNA, which migrates from the nucleus to the cytoplasm of the cell (Zawel and Reinberg, 1993). Gene transcription is a critical process and is the initial step in gene expression, a dynamic process that adapts to developmental and environmental cues (Reid et al., 2005). The transcriptional event is activated and regulated by interactions between multiple, highly specialized nuclear proteins called transcription factors (Lee and Young, 2000).

Figure 2.1. The transcription process: The RNA polymerase enzyme is a complex of many proteins. This complex moves in a 3'-5' direction along the DNA strand, synthesising RNA in a 5'-3' direction (Lodish et al., 2000; Alberts et al., 1998; Rosenthal, 1994; Singer and Berg, 1991). RNA synthesis produces a precursor messenger RNA (pre-mRNA) transcript that is complementary to the template DNA strand. The DNA rewinds into the double helix behind the polymerase enzyme (Friedberg et al., 2002; Friedberg et al., 2000).
Chapter Two: Literature Review

The binding of the polymerase enzyme to the promoter region depends on the presence of bound regulatory transcription factor proteins that guide the polymerase machinery to the transcription start site (Losick and Chanberline, 1976) (Figure 2.2.).

Figure 2.2. Transcription factors binding to the promoter: Transcription factors bind to the promoter site, generally located upstream at the 5' end of the gene, facilitating the docking of RNA polymerase to the promoter start site and the activation of RNA synthesis (Lodish et al., 2000; Alberts et al., 1998; Rosenthal, 1994; Singer and Berg, 1991).

2.1.1. The structure of the promoter

The promoter region of a gene may contain a motif known as a TATA box, which is a short sequence of DNA (TATAAA or TATA) situated approximately 25 base pairs (bp) upstream from the transcription start site (Butler and Kadonga, 2002) (Figure 2.3.) which TATA-binding transcription factor proteins recognize and bind to to initiate transcription. Other promoter elements also exist in genes of eukaryotes, and include GC boxes and CCAAT boxes (Muller et al., 1988) (Figure 2.3.). The promoter may also be composed of additional DNA elements that control the rate of transcription, known as enhancers and silencers commonly situated upstream of the transcription start site (reviewed in Macfarlane, 2000).
Figure 2.3. The structure of a promoter region: The TATA box is commonly located approximately 25bp upstream of the transcription start site. The GC and CCAAT box are promoter elements often situated upstream of the TATA box and either enhance or repress gene transcription (Muller et al., 1988).

2.1.2. Transcription factors

Transcription factor proteins are located in the nucleus and interact with promoter sites of genes to enhance or repress their rate of transcription, often switching on or off the genes they control (reviewed in Macfarlane, 2000). Most transcription factors consist of at least two functional domains, a DNA-binding domain which binds to promoter or enhancer regions, and a transactivation domain that communicates via protein-protein interactions with other transcription factors/cofactors or RNA polymerase (Butler and Kadonga, 2002; Lee and Young, 2000). Certain transcription factors interact with DNA by bending and looping it in a three dimensional orientation thus influencing the rate of transcription (Brenner, 1961) (Figure 2.4.). Overall, the interactions between different transcription factors, their target gene promoter sequences and RNA polymerase, form the molecular basis of transcriptional control.

2.1.3. Enhancers

Enhancer regions are short sequences of DNA, which, when bound by their cognate transcription factors, increase the rate of gene transcription (Muller et al., 1988) (Figure 2.4.). Silencer sequences are cis-regulatory elements that inhibit the transcription of a gene (Muller et al., 1988). Repressor proteins bind to silencer regions to reduce the rate of transcription. Characterized “enhancers” may behave as “silencers” under different specific conditions (Kuras et al., 2003).
2.1.4. RNA processing

The RNA transcripts produced from most eukaryotic protein-coding genes are further processed to mature mRNA in the nuclei of cells (Janson, 2001). The modification of the pre-mRNA begins with the addition of a specialized nucleotide (7-methylguanosine triphosphate) at the 5' end of the primary transcript; this initial process is called capping (Bannerjee, 1980) (Figure 2.5.). Capping is necessary to protect the RNA transcript against nuclease degradation at the 5' end of the molecule. Polyadenylation then occurs with the addition of adenylate adenosine monophosphate (AMP) residues at the 3' end of the mRNA, forming a poly(A) tail (Figure 2.5.). The poly(A) tail, is necessary for protection against enzymatic degradation (Jackson and Standart, 1990; Darnell, 1971).
Chapter Two: Literature Review

Figure 2.5. Modification of the pre-mRNA transcript by capping and polyadenylation: The modification of pre-mRNA begins with the addition of a specialized nucleotide linkage (7-methylguanosine triphosphate) at the 5' end of the molecule in a process called capping. The capping process is then followed by the addition of adenylate (AMP) residues at the 3' end of the pre-mRNA molecule, forming a poly(A) tail in a process called polyadenylation (Jackson and Standart, 1990; Bannerjee, 1980; Darnell, 1971).

2.1.5. RNA splicing

Splicing of pre-mRNA involves splicing intronic RNA segments out of the pre-mRNA transcript, while the remaining exonic RNA segments are joined end-to-end, generating a mature mRNA (reviewed in Matlin et al., 2005; Modrek and Lee, 2002; Burge et al., 1999) (Figure 2.6.).
Figure 2.6. RNA splicing: The primary transcript generated by gene transcription contains coding segments (exons) separated by non-coding DNA sequences (introns). During RNA splicing, introns are spliced out of the primary transcript and discarded. The remaining exons join end-to-end to generate a mature mRNA transcript that is ready to migrate from the nucleus into the cytoplasm of the cell to undergo translation into a protein (Eperon et al., 1993).

A critical process in RNA splicing is the recognition of splice sites within introns consisting of dinucleotides guanine and uracil (GU) and adenine and guanine (AG) at exon/intron boundaries (Figure 2.7.). Also the intronic branch site sequence is important for correct splicing and includes an adenine nucleotide that is located towards the 3' end of the intron (Norton, 1994) (Figure 2.7.). Splice sites are recognised by the spliceosome complex. However some splice sites are degenerate (e.g. contain guanine and adenine (GA) at the 5' end of the intron or thymine and guanine (TG) at the 3' end), making spliceosome selection imperfect, resulting in alternative splice sites being selected, and leading to the generation of alternative transcripts. Ultimately, the selection of different splice sites generates mRNAs that differ in their exonic coding sequences therefore producing distinct proteins.
Figure 2.7. **Exon/intron splice sites**: RNA splicing of intronic sequences occurs when dinucleotides GU are recognised at the 5' exon/intron splice site boundary. RNA splicing of intronic sequences at the 3' splice site boundary occurs when dinucleotides AG are recognised. The branch site indicated with an adenine (A) nucleotide acts as a branch site for connecting both 5' and 3' ends generating a lariat shaped structure (Norton, 1994).

At the molecular level, RNA splicing begins with cleavage at the 5' splice site junction, followed by the formation of a lariat shape structure and cleavage at the 3' splice site junction (Figure 2.8.). This process releases the intronic RNA lariat and the remaining exonic RNA segments are spliced together (Figure 2.8.). These reactions are orchestrated by a large macromolecular enzyme complex called a spliceosome, composed of small nuclear RNA and serine/arginine-rich (SR) proteins (reviewed in Garcia-Blanco, 2003) (Figure 2.8.). This enzyme complex recognises the intronic splice sites and catalyses the removal of introns (Garcia-Blanco et al., 2001; Burge et al., 1999).
2.1.6. Alternative RNA splicing

Alternative RNA splicing is a mechanism that generates more than one mRNA transcript from a single gene (Lynch and Conery, 2003). Alternate transcripts encode multiple protein isoforms (reviewed in Matlin et al., 2005) (Figure 2.9.) all of which may possess different functions. The number of alternatively spliced genes appears to increase with the complexity of the organism; for example 10% of genes are alternatively spliced in Caenorhabditis elegans, 25-40% in Drosophila and 60-80% in Homo Sapiens (Iida et al., 2004; Kampa et al., 2004; Lee et al., 2004; Stolc et al., 2004; Reboul et al., 2003; Modrek and Lee, 2002; Beaudoing and Gautheret, 2001).
Figure 2.9. Alternative RNA splicing generating multiple transcripts: Alternative splicing is a process that generates more than one different mRNA transcript from the same precursor messenger RNA (pre-mRNA). This process may form functionally distinct proteins (Alberts et al., 1998; Stryer, 1995; Watson, 1987).

During alternative RNA splicing, it is also possible for exons and neighbouring introns to be spliced out in a process known as exon skipping (McKeown, 1992). Exon skipping also has the potential to produce different mRNA transcripts. For example, if a gene possesses 9 exons, one mRNA product may include all exons, while another may skip exon 5, resulting in production of variable mRNA transcripts. Alternatively multiple exons together with their intervening introns can also be included in the final transcript. Together these mechanisms of alternative RNA splicing produce multiple RNA transcripts, which may or may not be functional.

2.2. Pax GENES

2.2.1. General function and alternate splicing of the developmental Pax genes

Pax genes are a family of developmentally important genes encoding highly conserved transcription factors that are critical for the regulation of embryogenesis, a function first revealed for the Drosophila melanogaster Pax gene prd (Bopp et al, 1986). Nine Pax
Chapter Two: Literature Review

genes (Pax1-9) have been identified in mammals (Jostes et al., 1990). These genes have been classified into four well-defined subgroups I-IV based on similarities within their protein structure and expression patterns during embryonic development; group I (Pax1 and 9), group II (Pax2, 5 and 8), group III (Pax3 and 7) and group IV (Pax4 and 6) (Balczerak et al., 1997).

All Pax genes except Pax1 and Pax9 are involved in central nervous system (CNS) development as well as in specification of diverse organs during embryonic development. Pax1 and Pax9 play a critical role in the development of the vertebral column and in the formation of endoderm embryonic layers (Holland et al., 1995). Pax1 also plays a role in thymus development (Wallin et al., 1996). Pax2 and Pax8 are involved in the development of the brain, excretory system and the kidney (Dressler et al., 1990; Plachov et al., 1990). Pax5 is known to play a vital role in the development of the brain and immune system, specifically in the formation of B-lymphocytes (Adams et al., 1992). Pax6 is well known as a master control gene in eye and brain development (Halder et al., 1995) and together with Pax4 is involved in the development of the pancreas (Halder et al., 1995; Quiring et al., 1994). Pax3 and Pax7 are involved in the specification of the spinal cord, brain and muscular system (Jostes et al., 1990).

Diverse roles for each Pax gene may be due to the production of alternate transcripts or multiple isoforms (Ziman et al., 1997). Interestingly, alternate transcripts of some Pax genes exhibit diverse expression patterns indicating diverse roles for each transcript. For example, six alternate Pax8 transcripts have been identified; and each transcript exhibits discrete expression patterns during kidney and placental development (Poleev et al., 1995; Kozmik et al., 1993). Furthermore, Pax2 alternate transcripts are known to exhibit differential expression patterns during the formation of the eye, kidney and CNS (Dressler and Douglass, 1992); two Pax6 transcripts display unique expression patterns in both eye and brain development (Epstein et al., 1994) and alternate Pax3 transcripts display distinct expression patterns in the developing neuromuscular system (Franz et al., 1993; Epstein and Fischman, 1991).
Previous studies have identified four alternate *Pax7* transcripts *Pax7a*-*d* that possess differential expression patterns in neurogenic and myogenic cell lineages (Ziman and Kay, 1998; Ziman et al., 1997). Specifically, all four *Pax7a*-*d* transcripts are expressed in the myogenic lineage. However, only *Pax7b* and *Pax7d* transcripts are expressed in cells of neurogenic lineage (Ziman and Kay, 1998). These *in vitro* studies suggest that *Pax7* transcripts play a role in cell lineage determination, since transfection of a single *Pax7b* transcript into cultured P19 cells resulted in their differentiation into neural-like cells (Ziman et al., 2001). Interestingly however, a single *Pax7d* transcript infected into haematopoietic stem cells induced myogenic differentiation and upregulation of other *Pax7* alternate transcripts (Seale et al., 2004). The four transcripts of *Pax7* have different structural features that may alter the function of their paired DNA binding domain.

2.3. *Pax7 PROTEIN*

2.3.1. Structural features of *Pax7*: DNA binding domains

The *Pax7* protein contains several functional domains, a paired domain, an octapeptide, a homeodomain and a transactivation domain (Figure 2.10.a.). The N-terminal paired domain is a unique bipartite DNA binding structure containing approximately 128 amino acids (Mansouri et al., 1994; Gruss and Walther, 1992). This structural feature comprises two subdomains known as PAI (N-terminal) and RED (C-terminal), each containing three alpha helices (Vogan and Gros, 1997; Vogan et al., 1996) (Figure 2.10.a.). The third helix of each subdomain is primarily involved in sequence-specific DNA contacts (Ziman and White, 2006). These two subdomains are separated by a linker region and allow the *Pax7* protein to interact with DNA (Mansouri et al., 1994; Gruss and Walther, 1992) via binding to downstream target genes (Figure 2.10.b.). Both PAI and RED subdomains are capable of interacting simultaneously or independently with DNA (Jun and Desplan, 1996), as evidenced by a mutational study showing that mutations within the PAI subdomain reduced PAI DNA-binding ability, but had no effect on the DNA-binding ability of the RED subdomain (Xu et al., 1995).

The *Pax7* protein contains a second DNA binding domain, the homeodomain, which contains 60 amino acids and consists of three alpha helices (Scott et al., 1989; Frigerio et
The homeodomain functions in combination with thepaired
domain to influence target gene selection (Kozmik et al., 1997; Underhill and Gros, 1997; Vogan and Gros, 1997) (Figure 2.10.b.).

The octapeptide of Pax7 consists of eight amino acids, and is located between the paired domain and homeodomain (Noll, 1993) (Figure 2.10.a.), and possesses a protein-protein interaction function (Eberhard et al., 2000) (Figure 2.10.b.).

The Pax7 transactivation domain (Figure 2.10.a.), located at the C-terminus of the protein, is rich in amino acids proline, threonine and serine; amino acids thought to be sites of phosphorylation (Mikkola et al., 1999). The transactivation domain may interact with the general transcriptional machinery to regulate target genes to which the Pax7 protein is bound (Hadchouel et al., 2003; Schafer et al., 1994) (Figure 2.10.b.).

Figure 2.10.a. Pax7 protein structure: The Pax7 protein consists of a paired domain, comprising two subdomains (PAI) and (RED) that are each composed of three alpha helices (numbered), situated at the N-terminal of the Pax7 protein. The octapeptide is located between the paired domain and homeodomain. The homeodomain contains three alpha helices (numbered). The transactivation domain is located at the C-terminal end of the Pax7 protein and is rich in amino acids serine, proline and threonine (Ziman and White, 2006).
Figure 2.10.b. Pax7 protein interacting with DNA:
The paired domain (PD), comprising PAI (N-terminal) and RED (C-terminal) subdomains and the Homeodomain (HD) are the structures that allow the Pax7 protein to interact with DNA of downstream target genes. The homeodomain (HD) consists of three alpha helices which interact with DNA and, together with the paired domain, are assumed to influence target gene selection. The Octapeptide (OP) domain is a protein-protein interaction component within the Pax7 protein. The transactivation domain (TD) is a protein-protein interaction domain that allows the Pax7 protein to activate transcription of downstream target genes by interacting with the general transcription protein machinery. All domains together control the regulation of downstream target genes, switching them on or off (Ziman and White, 2006; Du et al., 2005; Chi and Epstein, 2002).

2.3.2. Differential splicing of the Pax7 paired box

As mentioned previously, the Pax7 gene encodes several alternate transcripts with differing sequences in the paired box gene region. Alternate Pax7 transcripts differ by inclusion or exclusion of a trinucleotide (CAG which encodes a glutamine) and/or a hexanucleotide (GTTTAG which encodes a glycine and leucine) at intron 2, exon 3 and/or intron 3, exon 4 boundaries, respectively (Ziman et al., 2001; Ziman and Kay, 1998; Vogan and Gros, 1997; Vogan et al., 1996) (Figure 2.11.). Alternate paired box Pax7 transcripts that include or exclude a tri or hexanucleotide encode four alternate Pax7 isoforms that differ by inclusion or exclusion of amino acids glutamine (Q), and/or glycine and leucine (GL) within the PAI and RED subdomain regions of the paired domain (Figure 2.11.). Isoforms are termed Pax7a (Q-, GL+), Pax7b (Q+,GL+), Pax7c (Q-,GL-) and Pax7d (Q+,GL-) (Figure 2.11.).

Isoforms containing a single glutamine residue (Q+) within the paired domain are hypothesized to use only the PAI subdomain of the paired domain to bind to DNA, whereas the exclusion of a single glutamine residue (Q-) within the paired domain is thought to result in utilization of both the PAI and RED subdomains of the paired domain for DNA binding (Du et al., 2005; Vogan and Gros, 1997; Vogan et al., 1996). Thus,
alternate Pax7 isoforms that differ in their paired domain structures may bind to different downstream target genes, and possess differential functions.

![Diagram of Pax7 paired domain isoforms]

**Figure 2.11. Pax7 paired domain isoforms:** The alternate Pax7 transcripts differ by inclusion or exclusion of a trinucleotide (CAG) and/or a hexanucleotide (GTTTAG) at exon/intron boundaries (Intron 2 / Exon 3 and Intron 3 / Exon 4, respectively) which encode alternate Pax7 isoforms which differ by inclusion or exclusion of amino acids glutamine (Q) and/or glycine and leucine (GL) within the paired domain producing four alternate Pax7 isoforms Pax7a (Q-,GL+), Pax7b (Q+,GL+), Pax7c (Q-,GL-) and Pax7d (Q+,GL-) (Ziman et al., 2001; Ziman and Kay, 1998; Vogan and Gros, 1997; Vogan et al., 1996).

The Pax7 paired box transcripts Pax7a-d are tissue-specific and thought to be important for cell differentiation along either myogenic and/or neurogenic cell lineages (Kay and Ziman, 1999). Furthermore, encoded alternate Pax7 isoforms may differentially affect transcription rate of target genes (i.e. either activation or repression) and thus have the potential to possess distinctive functions in the developing embryo (Vogan et al., 1996). This complexity of Pax7 isoform structure and DNA binding efficiency and specificity may be responsible for the observed complexity of function. In this project we aim to shed light on some of the diverse functions of the Pax7 protein.
2.4. **Pax7 Function**

2.4.1. **Pax7 in central nervous system development**

From expression patterns it is evident that the Pax7 protein plays a crucial role in the development of the central nervous system. The Pax7 protein is expressed in the dorsal neural tube prior to formation of the brain and spinal cord and has been observed in the telencephalon and mesencephalon of mouse and chick brain (Nomura et al., 1998; Kawakami et al., 1997; Stoykova and Gruss, 1994). Restricted expression has also been observed in anterior and posterior regions of the caudal mesencephalon and rostral hindbrain in zebrafish (Seo et al., 1998). Later in the development of mouse, chick and zebrafish, Pax7 expression is confined to the dorsal midbrain, particularly the tectum/superior colliculus and epiphysis (Picker et al., 1999; Kawakami et al., 1997; Krauss et al., 1991). Other studies have shown that Pax7 is important for neural sub-type specification in the chick and mouse mid-brain (Thomas et al., 2006; Thompson et al., 2006; Thomas et al., 2004); Pax7 null mice lack superior colliculus neurons and have severe facial deformities (Thompson et al., in preparation) (Mansouri et al., 1994). Moreover, Pax7 induces neural differentiation in vitro (Ziman et al., 2001).

2.4.2. **Pax7 in muscle development**

*In situ* hybridisation studies in mouse have shown that Pax7 is expressed in both epaxial and hypaxial domains of condensing somites (Figure 2.12.), which form the dorsal body wall, back muscles and specific precursor cells or satellite cells that migrate to the limb buds (Lamey et al., 2004). These expression patterns define the Pax7 protein's crucial role in skeletal muscle development. Satellite cells are important for repair and regeneration of skeletal muscle tissue and growth by differentiation into myocytes. Seale and colleagues (1999) have shown that Pax7 is critical for skeletal muscle satellite cell specification. Significantly, Pax7 null mice completely lack skeletal muscle satellite cells (muscle progenitor cells) hence muscle regeneration is severely impaired (Seale et al., 1999).
2.5. DOWNSTREAM TARGET GENES

2.5.1. Target genes of the Pax7 protein

In order to identify the target genes bound by Pax7 and thus identify the gene pathways regulated by Pax7 a chromatin immunoprecipitation (ChiP) technique (White and Ziman, 2006) was utilized. Using this technique several genomic fragments bound by the Pax7 protein in developing mouse embryos were isolated, allowing identification of a definitive set of target genes bound by Pax7 in vivo. These target genes are predominantly development genes involved in cellular proliferation and cell signalling important for development (White and Ziman, 2006).

A cohort of these identified target genes have been chosen for further study in this project based on their critical cellular and developmental functions associated with neurogenesis and myogenesis. These genes are as follows:

The Pax7 target gene, CntfR (ciliary neurotrophic factor receptor), encodes a receptor for the cytokine, Interleukin-6 (IL-6), known to play a critical role in both peripheral and central nervous systems (Ip et al., 1993). CntfR also plays a functional role in muscle
repair and regeneration; since application of CNTF to rejuvenating muscles in vivo stimulates myotube formation (Marques and Neto, 1997).

The target gene, *Eya4* (*eyes absent 4 homologue*) encodes a transcriptional regulatory factor that has been shown to play a critical role during dorsal somite development in early myogenesis (Borsani et al., 1999; Davis et al., 1999; Esteve and Bovolenta, 1999; Relaix and Buckingham, 1999; Mishima and Tomarev, 1998; Xu et al., 1997; Williams and Ordahl, 1994; Jostes et al., 1990). During mouse embryonic development, *Eya4* is also expressed in the craniofacial mesenchyme (Borsani et al., 1999). Interestingly, *Eya4* expression in the somite and the dermomyotome as well as in cells migrating to the developing limb buds strongly resembles *Pax7* expression patterns during development (Christ and Ordahl, 1995; Bober et al., 1994).

The *Gbxl* target gene (*gastrulation brain homeobox 1*) encodes a homeodomain transcription factor that controls the development of neuronal cell types in the developing dorsal spinal cord and its expression domain establishes the midbrain-hindbrain boundary (Rhinn et al., 2004). *Gbxl* is also expressed later in the developing hindbrain, in the ventral telencephalon, optic vesicles and spinal cord (Rhinn et al., 2004).

The *Kcnk2* target gene (*potassium channel K2*) commonly known as *TREK1* encodes a potassium channel protein critical for the physiological properties and modulatory potential of neurons (Bockenhauer et al., 2001; Talley et al., 2001). The gene has shown differential patterns in the CNS, being expressed in dorsal root ganglion and the superior colliculus (Talley et al., 2001), the midbrain region specified by *Pax7* (Thomas et al., 2006; Thomas et al., 2004; Jostes et al., 1990).

The *Rasa3* target gene (*RAS p21 protein activator 3*) is associated with cell signalling and cellular transformation. It has previously been demonstrated that cyclin-dependent kinase inhibitor 1A (p21) Ras protein plays a functional role in stimulating cell survival, signalled by nerve growth factors in rat sympathetic neurons (Nobes et al., 1996).
Chapter Two: Literature Review

The *Prrxl* target gene (*paired-related homeobox 1*) is a paired-class homeobox gene that is expressed in the craniofacial mesenchyme and limb buds. *Prrxl* mutants show skeletal deficiencies in the skull, vertebral column and limbs (Martin et al., 1995). Further evidence has been obtained by gene targeting experiments demonstrating that the homeobox gene *Prrxl* is critical for craniofacial and limb development as well as vascular regulation (Chesterman et al., 2001).

The binding of alternate Pax7 isoforms individually to these previously identified target genes formed the basis of the project.

2.6. CHROMATIN IMMUNOPRECIPITATION

Investigating protein-DNA interactions by identifying the target genes of transcription factors is crucial for understanding the mechanisms of gene expression (reviewed in Taverner et al., 2004). Chromatin immunoprecipitation (ChiP) is a new and valuable methodology offering a snapshot of transcription factor binding to DNA promoter sites *in vivo* and has revolutionized our knowledge of gene transcription (Weinmann, 2004; Weinmann and Farnham, 2002; Wells and Farnham, 2002; Kuo and Allis, 1999).

Before ChiP was developed, deoxyribonuclease (DNase) I footprinting and Electrophoretic Mobility Shift Assays (EMSA) were employed to investigate protein-DNA interactions (Peng and Chen, 2005; Chen and Zack, 2000). DNase I footprinting is/was used to investigate the binding of transcription factors to promoters or enhancer sites over large regions of DNA (Peng and Chen, 2005; Chen and Zack, 2000). In this technique approximately 200-400bp DNA fragments are labeled with radioisotopes and digested with DNase I in the presence or absence of a protein. The region of DNA that is bound by transcription factor proteins are protected from DNase I digestion, and are identifiable as a gap in DNA sequence on a denaturing polyacrylamide DNA sequencing gel (Peng and Chen, 2005; Manuel et al., 2002; Roulet et al., 2002). In contrast, EMSA is/was used to confirm the binding of specific transcription factors to short DNA sequences or oligonucleotide (Peng and Chen, 2005; Chen and Zack, 2000). The binding
Chapter Two: Literature Review

is detected by high molecular weight band shifts in a nondenaturing polyacrylamide gel (Peng and Chen, 2005; Chen and Zack, 2000; Chen et al., 1997).

Although both assays investigate protein-DNA binding *in vitro* they are unable to adequately identify specific target genes of transcription factors (Peng and Chen, 2005). To overcome the limitations of these techniques, the chromatin immunoprecipitation methodology was developed, specifically to study the binding of endogenous transcription factors to the promoters of specific DNA sequences *in vivo* (Kirmizis and Farnham, 2004; Wells and Farnham, 2002).

When using ChIP methodology, transcription factors associated with chromatin DNA at any one time are cross-linked to the DNA by formaldehyde. This step is followed by sonication to shear chromosomal DNA, the protein-DNA complexes are then immunoprecipitated using an antibody against the crosslinked protein. Immunoprecipitated protein-DNA complexes are extensively washed, eluted, decrosslinked and the DNA is purified and analyzed for the presence or absence of specific DNA sequences by polymerase chain reaction (PCR) for known DNA target sequences utilizing gene-specific primers (Wells et al., 2002; Wells and Farnham, 2002) (Figure 2.13.) or by cloning and sequencing to identify targets (White and Ziman, 2006).

The technique has been utilized in tissues from a variety of species, including mouse embryos (White and Ziman, 2006), protozoa (Dedon et al., 1991a; Dedon et al., 1991b), sea urchin embryos (Jasinskiene et al., 1995), and *Drosophila* (Orlando et al., 1997; Orlando and Paro, 1993). Very few studies have utilized this technique *in vitro* (Agbottah et al., 2006; Esumi et al., 2006; Gong et al., 2006; Valineva et al., 2006).
Chapter Two: Literature Review

Figure 2.13. Chromatin immunoprecipitation technique: The initial step of the technique involves formaldehyde crosslinking of protein expressed in either cultured cells or tissues to DNA. Protein-DNA complexes are then sonicated, immunoprecipitated with an antibody against the protein of interest. The immunoprecipitated complex is washed, eluted, decrosslinked and the DNA (target gene) purified and analyzed via PCR utilizing gene-specific primers for known DNA target sequences or by cloning for unknown targets. A sample that is not subjected to immunoprecipitation (no antibody) acts as a negative control (Peng and Chen, 2005).

2.7. SUMMARY

Previous studies have identified eight structurally distinct Pax7 isoforms. This DNA-binding study, focuses on identification of DNA binding specificities of four of the Pax7 paired domain isoforms. These isoforms are Pax7a (Q-, GL+), Pax7b (Q+, GL+), Pax7c (Q-, GL-) and Pax7d (Q+, GL-). Given that Pax7 isoforms differ in their paired domain structure, it is probable that DNA-binding domain differences will result in selection of specific downstream target genes. Previous studies have identified alternate Pax7
transcripts present in both neurogenic and myogenic cell lines, however more research is necessary to understand the functional significance of the encoded alternate Pax7 isoforms.

From identification of target specificity this study sought to reveal the functional significance of each Pax7 isoform. This study therefore aimed to investigate DNA-binding differences utilizing the chromatin immunoprecipitation technique. It was felt that by using ChiP, we would identify specific downstream target genes of each Pax7 isoform. We used a specific subset of target genes CntfR, Eya4, Gbx1, Rasa3, Kcnk2 and Prrxl previously identified as Pax7 targets, and assessed binding of Pax7 isoforms to these genes in a cell culture system. It was anticipated that these experiments would provide new insights into the DNA-binding differences of each Pax7 isoform and thus identify their roles during embryogenesis. Recent experiments have used specific Pax7 transcripts to induce neural and muscle cell differentiation from stem cells. The Pax7 transcripts are therefore candidates for advances in future stem cell/ gene therapy approaches aimed at developing novel treatments for Neurodegenerative diseases and Duchene Muscular Dystrophy. It is important therefore to understand the pathways regulated by each transcript.
CHAPTER THREE

METHODS
3.1. Research Plan

The purpose of this study was to identify Pax7 isoform specific downstream target genes, to provide insight into the functional significance of each isoform. To do this cells were transfected with each Pax7 alternate transcript and then subjected to Chromatin Immunoprecipitation (ChIP) (Figure 3.1.).

Alternate transcripts of Pax7 cloned into a V5-epitope tagged expression vector pcDNA3.1D/V5-His-TOP6 were transfected into a mouse NIH3T3 fibroblast cell line. Stably transfected cells were selected in geneticin antibiotic (G418) and assessed for Pax7-vector expression by reverse transcriptase-polymerase chain reaction (RT-PCR) and for translated protein by immunocytochemistry and western blot. Stably transfected cells were treated with formaldehyde to crosslink the Pax7 protein to DNA, then protein-DNA complexes were extracted and sonicated. Individual Pax7 isoforms and bound DNA were immunoprecipitated, and the purified genomic fragments were analysed by PCR with primers selecting for known Pax7 target genes (White and Ziman, 2006).

Target genes that were analysed in this study were: CntfR (ciliary neurotrophic factor receptor), Eya4 (eyes absent 4 homologue), Gbx1 (gastrulation brain homeobox 1), Kcnk2 (potassium channel K2), Rasa3 (RAS p21 protein activator 3) and Prrx1 (paired-related homeobox 1).

Cells transfected with the pcDNA3.1D/V5-His-lacZ expression vector and untransfected cells were included as controls. All transfection experiments and controls were performed in triplicate. Experimental procedures were conducted utilizing NIH3T3 mouse embryonic fibroblasts and P19 mouse embryonic carcinoma cells.
3.2. VECTOR PREPARATION

3.2.1. Amplification of epitope-tagged expression vectors containing Pax7(a-d) transcripts

Glycerol stocks of pcDNA3.1D/V5-His vectors containing individual Pax7 transcripts inserted in frame so as to produce a Pax7-V5His6 fusion protein (Figure 3.2.) were kindly donated by Robert White. Clones were streaked on Luria-Bertani (LB) agar containing 100μg/ml ampicillin and incubated overnight at 37°C. Individual colonies were randomly picked and inoculated into 5ml of Luria-Bertani (LB) (10g/L NaCl, 10g/L tryptone, 5g/L yeast) broth containing 100μg/ml of ampicillin. Cultures were grown overnight on a rocking platform at 37°C. E. coli colonies containing control pcDNA3.1D/V5-His-lacZ
vector alone were also cultured (Figure 3.3.). A small quantity of each culture was amplified by PCR (Section 3.2.4.) to validate the presence and correct orientation of alternate Pax7 inserts in the pcDNA3.1D/V5-His-TOPO expression vector.

![Pax7 transcript inserted into vector](image)

**Figure 3.2.** *Pax7* alternate transcript inserted into the pcDNA3.1D/V5-His-TOPO vector containing 3' V5 and His6 sequences (Invitrogen).

![pcDNA3.1D/V5-His-lacZ control vector](image)

**Figure 3.3.** *pcDNA3.1D/V5-His-lacZ* control vector (Invitrogen).
3.2.2. Isolation of alternate Pax7(a-d) vector constructs

Once the E. coli colonies were identified as having the correct Pax7 insert by PCR, they were grown in 5ml LB medium containing 100μg/ml of ampicillin. A miniprep plasmid purification kit (Qiagen) was used to purify plasmid DNA from each E. coli culture as per manufacturer’s instructions. Briefly, 5ml cultures were pelleted at 4500rpm for 1.5 minutes, resuspended in 250μl of resuspension Buffer (P1) (50mM Tris·Cl, pH 8.0; 10 mM ethylenediaminetetra-acetate (EDTA); 100 μg/ml ribonuclease (RNase) A), followed by the addition of lysis Buffer (P2) (250μl) (200 mM Sodium hydroxide (NaOH), 1% Sodium dodecyl sulfate (SDS). Neutralization Buffer (N3) (350μl) (3.0 M potassium acetate (CH3COOK), pH 5.5) was added to each sample, mixed immediately and samples were centrifuged at 13 000rpm for 10 minutes. Supernatants were applied to QIAprep Spin Columns and centrifuged at 13 000rpm for 30 seconds. Plasmid DNA was then washed with phosphate binding buffer (PB) (500μl), and centrifuged at 13 000rpm for 30 seconds, then DNA cleanup buffer (PE) (600μl), was added and the DNA preparation was centrifuged at 13 000rpm for 30 seconds, then for 1 minute to remove residual wash buffer. DNA was eluted by applying distilled deionised water (50μl) to the column for 1 minute at room temperature, then centrifuged for 1 minute at 13 000rpm.

3.2.3. Agarose gel electrophoresis

Agarose gel electrophoresis was performed to assess the quantity and quality of Pax7(a-d)-pcDNA3.1D/V5-His-TOPO vector DNA.

Isolated DNA (5μl) from plasmid preparations was diluted in 6x loading buffer (50mM EDTA, 40%v/v glycerol, 0.24%w/v Bromophenol blue) applied to a 1%w/v agarose gel and electrophoresed in TAE buffer (40mM Tris-acetate pH 8, 1mM EDTA) containing 5μg/ml ethidium bromide (Sigma) at a constant 105V for 1 hour. Sizes of products were estimated by comparison to a DNA size ladder (1·Kb Plus, Invitrogen). The gel was viewed on a UV Biorad Gel Doc System and photographed.
3.2.4. Polymerase Chain Reaction (PCR) of Pax7(a-d)-pcDNA3.1D/V5-His-TOPO cDNA constructs

To confirm that alternate Pax7 transcripts were correctly incorporated into the vector in the expected orientation PCR reactions were performed on Pax7a-d-vector constructs. The forward primer (F4) 5'-CACCGACTCTGGATTCGTCTCCAGCGTG-3' was utilized since the first four bases are vector specific (underlined bases) and the rest of the primer spans the 5' untranslated region of Pax7. This primer was utilized in combination with a reverse primer (BGH) 5' -TAGAAGGCACAGTCGAGG-3', that spans a region in the vector that is 3' of the insert. PCR reactions were conducted utilizing an MJ Research PTC-200 Peltier thermal cycler (Geneworks) with a heated lid.

PCR reactions (20µl) contained DNA template (1µl/ml), primers (1µM each), deoxynucleotide triphosphate (dNTP) (200µM each) (Finnzymes), PCR buffer (1x Qiagen), Q solution (1x Qiagen), Taq DNA polymerase (2.5 Units (U), Qiagen). Optimal cycling conditions utilising the forward primer F4 5'-CACCGACTCTGGATTCGTCTCCAGCGTG-3' in combination with the reverse primer BGH 5'-TAGAAGGCACAGTCGAGG-3', for Pax7-pcDNA3.1D/V5-His-TOPO complementary deoxyribonucleic acid (cDNA) constructs were as follows; an initial denaturation step of 94°C for 3 minutes, followed by 36 cycles of 94°C for 30 seconds (denaturation), 52°C for 2 minutes (annealing), 72°C for 30 seconds (extension) and final extension at 72°C for 10 minutes.

A forward primer (T7) 5'-TAATACGACTCCTATAGGG-3' was utilized in combination with the reverse primer (BGH) 5' -TAGAAGGCACAGTCGAGG-3', to confirm the correct control vector sequence.

Optimal cycling conditions utilizing the forward primer T7 5'-TAATACGACTCCTATAGGG-3' in combination with the reverse primer BGH 5'-TAGAAGGCACAGTCGAGG-3', for pcDNA3.1D/V5-His-lacZ vector alone were as follows; an initial denaturation step of 94°C for 3 minutes, followed by 36 cycles of 94°C
for 30 seconds (denaturation), 52°C for 2 minutes (annealing), 72°C for 30 seconds (extension), then a final extension at 72°C for 10 minutes.

Agarose gel electrophoresis was performed to determine the size of PCR products for Pax7(a-d)-pcDNA3.1/DV5-His-TOPO vector DNA. PCR products (5μl) were analysed as in Section 3.2.3.

3.2.5. Determination of DNA concentration

Purified plasmid DNA, was diluted 50 fold in distilled deionised water and the concentration was measured using a biophotometer (Eppendorf). Absorbance at 260nm (A$_{260}$) was measured in triplicate for each sample. The following formula were used to calculate concentration:

\[
\text{[Double stranded DNA (dsDNA)]: } 1 \times \text{A$_{260}$ unit x 50 = 50μg/ml}
\]

3.3. CELL CULTURE

3.3.1 Maintenance of NIH3T3 and P19 cells in culture

NIH3T3 mouse embryonic fibroblast cells (ATCC) at fifteenth passage (P$^{15}$) were grown in 5ml Dulbecco’s Modified Eagles Medium (DMEM) containing 10% Fetal Bovine Serum (FBS) (Invitrogen) and 1% L-glutamine (Invitrogen) in 25cm$^2$ cell culture flasks (T25). Cells were incubated at 37°C in a 5% carbon dioxide (CO$_2$) atmosphere. The medium was replaced every 48 hours until cells became 90-95% confluent. Confluent cells were passaged by adding 2ml accutase (Sigma) to the flask to lift the cells, the cells were transferred to a 15ml falcon tube and centrifuged at 500rpm for 5 minutes at room temperature. The supernatant was discarded, new medium (1ml) was added and cells were resuspended and distributed into 25cm$^2$ flasks or 10cm$^2$ 6 well culture plates at a concentration of 1×10$^5$cells/ml. Cells were grown to confluency, at which time individual Pax7(a-d) transcripts were transfected into each 10cm$^2$ well. Cells not subjected to transfection were grown in 25cm$^2$ cell culture flasks and subjected to chromatin immunoprecipitation (ChIP) studies as input samples (untransfected, NIH3T3 chromatin samples not subjected to immunoprecipitation) (Section 3.5.4.). A P19 mouse
embryonic clone (ATCC) P^20 already transfected with the \( Pax7d \) transcript, and untransfected P19 cells were also cultured in conditions as stated for NIH3T3 cells.

### 3.3.2. Transfection and selection of stable clones

Once cells were 90-95% confluent, DNA containing purified \( Pax7(a-d)-pcDNA3.1/V5-His-TOPO \) cDNA constructs (4.0\( \mu \)g) or \( pcDNA3.1/V5-His-lacZ \) vector (control) were transfected into NIH3T3 cells in 10cm\(^2\) culture wells, using lipofectamine 2000 reagent (10\( \mu \)l) (Invitrogen). Since the \( pcDNA3.1D/V5-His-TOPO \) vector contains a neomycin resistant gene, constructs that had stably integrated into the NIH3T3 genome were selectively grown in cell culture medium containing G418 antibiotic (Invitrogen). Medium, containing 0.5mg/ml G418 was replaced every 72 hours for the next 21 days for selection. Thereafter, medium containing 0.25mg/ml G418 was replaced every 72 hours for 7 days to maintain selection.

After the 28-day selection period, NIH3T3 cells that were transfected with \( Pax7(a, b, c \) and \( d) \)-vector constructs were transferred into individual flasks (25cm\(^2\)) and allowed to propagate for subsequent use in expression and Chromatin Immunoprecipitation (ChIP) studies.

### 3.3.3. Photography of cell culture

Alternate \( Pax7(a-d) \) clones resistant to G418 were viewed with a Leica microscope and photographed using an Olympus Digital camera at 20x magnification.

### 3.4. Expression studies to confirm transfection of \( Pax7 \) transcripts in NIH3T3 cells

#### 3.4.1 Western blot of transfected cells

##### 3.4.1.1. Protein extraction

NIH3T3 cells transfected with \( Pax7(a-d)-pcDNA3.1D/V5-His-TOPO \) were washed with 1ml phosphate buffered saline (PBS), scraped into individual eppendorf tubes, and centrifuged at 1500rpm for 5 minutes. Cells were resuspended in 50\( \mu \)l cell lysis buffer (50mM Tris-HCL, pH 7.8, 150mM NaCl, 1% Nonidet P-40) containing 1x proteinase
inhibitor cocktail (Roche), briefly vortexed and incubated at room temperature for 15 minutes. Debris was then centrifuged at 10000 rpm for 10 minutes at 4°C, and supernatants were collected and stored at -80°C.

3.4.1.2. Protein concentration determination

Cell lysates were diluted (20 fold, 40 fold and 50 fold) in distilled deionised water and protein concentration was measured using a biophotometer. Absorbance at 280 nm ($A_{280}$) was measured in triplicate for each sample. The following formulae were used to calculate protein concentration:

$$\text{[Protein mg/ml]}: (1.55 \times A_{280} \text{ unit}) - 0.76 \times A_{260} \text{ unit}$$

3.4.1.3. Western blot

Western Blot was employed to confirm the correct expression and translation of Pax7 isoforms. The presence of the individual Pax7 (a-d)-vector proteins in transfected cells were identified by either the Anti-V5 monoclonal antibody or the Pax7 monoclonal antibody.

3.4.1.4. Polyacrylamide gel electrophoresis

A 8% separating acrylamide gel was prepared (Tris-HCl 1.5M pH 8.8, 8% Acrylamide/Bis and 10% SDS) the gel was degassed for 10 minutes at room temperature, followed by the addition of catalysts ammonium persulfate (10%) and $2 \times 10^{-8}$ M TEMED. The gel was loaded into the gel apparatus and left to polymerize. The stacking gel was prepared (Tris-HCl 0.5M pH 6.8, 4% Acrylamide/Bis and 10% SDS) degassed for 10 minutes at room temperature, followed by the addition of catalysts ammonium persulfate (10%) and $4 \times 10^{-8}$ M Tetramethylethylenediamine (TEMED), loaded into the gel apparatus on top of the stacking gel and left to polymerize.

Cell lysates were diluted in 2x SDS sample buffer (0.5M Tris-HCl, 10% Glycerol, $5 \times 10^{-6}$ M $\beta$-mercaptoethanol, 0.001M SDS, $3 \times 10^{-5}$ M Bromophenol blue), vortexed and boiled for 5 minutes at 100°C. The electrophoresis tank (Biorad) was filled with 1x SDS
Chapter Three: Methods

running buffer (3.03g Tris, 14.4g Glycine, 1g SDS). Protein samples were loaded onto the gel (20μl) and the gel was run at constant 200V (Laemmli, 1970).

3.4.1.5. Transfer of proteins to nitrocellulose membrane

Proteins were transferred from the gel to a nitrocellulose membrane as follows: A transfer sandwich was assembled under transfer buffer (0.02M Tris, 0.08M Glycine, 0.001M SDS, 20% methanol (MeOH) consisting of 1 fibre sheet, 2 filter papers, polyacrylamide gel electrophoresis (PAGE), nitrocellulose paper, another filter paper and a final fibre sheet (Figure 3.4.). The sandwich was placed into the plastic transfer cassette and into the transfer tank, and internal and external sections were filled with transfer buffer (0.02M Tris, 0.08M Glycine, 0.001M SDS, 20% MeOH) and run at a constant 20V overnight at 4°C. After disassembly, the nitrocellulose membrane was cut for orientation of protein samples and the ladder (Precision Plus standard protein ladder, BioRad) marked on the membrane.

Figure 3.4. Nitrocellulose membrane sandwich: The sandwich was assembled as follows; 1 fibre sheet, 2 filter papers, PAGE gel, nitrocellulose paper, filter paper, fibre sheet (Laemmli, 1970).

3.4.1.6. V5 antibody membrane probing

The nitrocellulose membrane was immersed in blocking buffer containing 1x Tris buffered saline (TBS) pH 7.4, Tween-20 (0.1%) and dry non-fat skim milk (5%), for 2 hours on an orbital shaker at a constant medium rotation. Anti-V5 antibody (1200μg/ml) (Invitrogen) was diluted 1:5000 in blocking buffer (1x Tris buffered saline with Tween-20 (TBST), dry non-fat skim milk 2.5%). The Anti-V5 antibody was added to the
nitrocellulose membrane and incubated for 2 hours at room temperature on an orbital shaker at a constant slow rotation. The membrane was washed sixteen times in 1x Tris buffered saline with Tween-20 (TBST) for 15 minutes on an orbital shaker at room temperature.

IgG-linked to horseradish peroxidise diluted 1:2000 in 1x TBST, was applied to the nitrocellulose membrane incubated for 1 hour at room temperature on an orbital shaker at a constant slow rotation. The membrane was washed 8 times in 1x TBST for 15 minutes on an orbital shaker at room temperature.

3.4.1.7. Visualization

Probed membranes were visualized by chemiluminesance utilizing the ECL Plus Western Blot Detection Kit (Amersham). ECL kit detection reagents were equilibrated at room temperature, and mixed together (solution A 98% solution B 2%) and added to the nitrocellulose membrane, then incubated at room temperature for 5 minutes. The nitrocellulose membrane was sandwiched between transparent overhead sheets and placed into a cassette. The following steps were performed in the dark room; medical x-ray film (AGFA) was placed onto the nitrocellulose membrane, and exposed for various times (15 seconds, 1 minute, 3 minutes, 15 minutes and overnight). After exposure, the medical x-ray film was placed in developer solution (solution A 80%, solution B 20%) for 2 minutes and agitated, washed by water and then placed in fixer solution (fixer 100ml and double distilled de-ionised water (ddH₂O) 300ml) for 2 minutes, washed (water) and air dried.

3.4.1.8. Stripping the nitrocellulose membrane to re-probe with the Pax7 antibody

The nitrocellulose membrane was immersed in stripping buffer (100mM β-mercaptoethanol, 2% SDS, 62.5Mm Tris-HCl pH 6.7), incubated at 50°C for 30 minutes with agitation. The membrane was washed 2x in TBS for 10 minutes at room temperature on an orbital shaker, blocking buffer (1x TBS pH 7.4, Tween-20 0.1% and 5% dry non-fat skim milk) was applied to the membrane, incubated for 2 hours at room temperature on an orbital shaker. Pax7 monoclonal antibody (DSHB) was diluted 1:20 in 1x TBST
and added to the nitrocellulose membrane, incubated overnight at 4°C on a rotating platform. The membrane was washed sixteen times in 1x TBST for 15 minutes on an orbital shaker at room temperature.

The secondary antibody Immunoglobulin G-horseradish peroxidase (IgG-HRP) (diluted 1:2000 in 1x TBST) was applied and visualised using the ECL kit as described earlier (Section 3.4.1.7.). The medical X-Ray film was exposed to the membrane and developed and air dried as before.

Once confirmation of protein expression for each Pax7 transcript was achieved, NIH3T3 cells transfected with alternate Pax7 transcripts as well as cells transfected with the vector alone and untransfected cells were subjected to chromatin immunoprecipitation (ChiP) studies.

3.4.2. Pax7 transcript RT-PCR

3.4.2.1. RNA extraction

RNA was isolated from transfected and untransfected control NIH3T3 mouse embryonic fibroblast cells utilising an Aurum RNA extraction Kit (BioRad) as per manufacturer’s instructions. RNA samples were then treated with DNase 1 utilizing the DNA-free™ Kit (Ambion) to eliminate DNA from each RNA preparation. RNA concentration was determined by measuring absorbance at 260nm [A_260] using a biophotometer (eppendorf). Extracted, DNase treated RNA isolated from cells transfected with each alternate Pax7(a-d) transcript or vector transfected or untransfected cells were used in RT-PCR studies to confirm Pax7 expression in transfected cells.

3.4.2.2. Reverse transcription of RNA from NIH3T3 cells transfected with Pax7(a-d) transcripts

RNA was converted to cDNA utilizing Omniscript Reverse Transcriptase (Qiagen). Conditions were; reaction volumes of 20μl containing final concentrations of 1μg total RNA template, 1μM Oligo-dT primer, 0.5mM of each dNTP (Qiagen), 1x reverse transcriptase-buffer (RT-Buffer) (Qiagen), RNase out (10 U, Qiagen), Omniscript reverse
transcriptase (4 U, Qiagen). Each reaction was incubated for 60 minutes at 37°C, in the PCR machine.

3.4.2.3. Polymerase Chain Reaction (PCR) to detect Pax7 transcript expression

A region of Pax7 was amplified utilizing forward primer F4: 5’-CACCGACTCTGGATTCTCCACGCGTG-3’ and reverse primer Pax7R: 5’-ACAACCATGAGAACACGGAGC-3’, that span the 5’ paired box region of Pax7. RT-PCR reactions were conducted using an MJ Research PTC-200 Peltier thermal cycler with a heated lid. Samples were set up on ice and loaded into a pre-heated thermal cycler (95°C).

Conditions for PCR reactions were; 20μl reaction volumes containing 1/10 of the final concentration of cDNA (0.1μg), 1μM forward primer F4: 5’-CACCGACTCTGGATTCTCCACGCGTG-3’ and reverse primer Pax7R: 5’-ACAACCATGAGAACACGGAGC-3’, 1x PCR buffer, 1x Q solution, 200mM of each dNTP and Taq DNA polymerase (2.5 U/reaction). Optimal cycling conditions were; an initial denaturation of 94°C for 3 minutes, followed by 45 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 30 seconds, then a final extension at 72°C for 10 minutes. RT-PCR products for each Pax7 transcript were then subjected to agarose gel electrophoresis. Products were visualized on ethidium bromide stained gels relative to DNA size ladders and photographed using a UV Biorad Gel Doc System.

3.4.3. Immunocytochemistry of transfected cells to detect Pax7 expression

3.4.3.1. Immunocytochemistry

Cells were transferred to Lab-Tek® II Chamber Slides™ (Nalge Nunc International) at a concentration of 1x10⁵ cells/ml and grown overnight. Cell culture medium was removed from transfected NIH3T3 cells, and cells were washed 3 times with sterile 1x TBS. Freshly prepared 4% paraformaldehyde in PBS was added to the cells and incubated for 10 minutes at room temperature. The paraformaldehyde was then removed. Cells were rinsed three times in 1x TBS and quenched for 10 minutes in 3% hydrogen peroxide for depletion of endogenous hydrogen peroxide. Cells were rinsed in 1x TBS. 1x TBST
supplemented with 1% normal goat serum (NGS) and 1% fetal calf serum (FCS) was used for blocking non-specific reactions for 1 hour at room temperature. Primary antibodies, Anti-V5 (1:500 Invitrogen) or Pax7 (1:10 DSHB) diluted in TBST supplemented with 1% NGS and 1% FBS were applied to cells and incubated overnight at 4°C. Cells were washed with 1x TBS and the secondary antibody, IgG-linked to biotin (Lsab, Biorad) diluted in 1x TBST, was applied to cells and incubated for 1 hour at room temperature. Three 1x TBS wash steps followed for 5 minutes and then streptavidin linked to horse-radish peroxidise (Lsab, Biorad) diluted in 1x TBS was applied to the cells and incubated for 1 hour at room temperature. A colorimetric reaction was obtained by application of 3,3-diaminobenzidine (DAB, Dako) to the cells for 1-2 minutes. Cells were washed in 1x TBS and 1x Tris normal saline (TNS), mounted using glycin (gelatine 10g/ml, glycerine 70ml, phenol 1ml), coverslipped and viewed with an Olympus Bx41 microscope.

3.4.3.2. Photography
Immunocytochemistry results were photographed using a Leica DC300 digital camera mounted on an Olympus Bx41 microscope at 40x magnification.

3.5. CHROMATIN IMMUNOPRECIPITATION
3.5.1. DNA-shearing optimization
T25 flasks containing untransfected NIH3T3 cells (1x10^6/ml) were utilized to determine optimal sonication conditions. The complete experimental procedure was performed on ice. Crosslinked protein-DNA cell lysates extracted from untransfected NIH3T3 cells by the addition of 1ml SDS lysis buffer containing 10mM EDTA, 50mM Tris-HCl, pH 8.1 and 1% SDS, were sheared utilising an Ultrasonicator (Sonics Vibra cell™) for the following times; 60 seconds, 120 seconds, 160 seconds, 200 seconds and 300 seconds. An unsonicated sample was included as a control to assess successful sonication procedures. Sonicated samples were centrifuged at 14 000rpm for 10 minutes at 4°C, the supernatant was collected and 5M NaCl (600 μl) was added to each sample, vortexed briefly and incubated overnight at 65°C (to de-crosslink protein-DNA complexes). The de-crosslinked protein-DNA complexes were then pelleted at
14 000rpm for 30 seconds at room temperature, protein was digested with 10μl of 0.5M EDTA, 20μl of 1M Tris-HCl, pH6.5 and 30μl of 10mg/ml Proteinase K (Qiagen), by incubation for 1.5 hours at 45°C. The digested protein samples were then subjected to DNA purification utilizing the QIAquick PCR Purification kit (Qiagen) as per manufacturer’s instructions (Section 3.5.6.1.). Purified DNA products were resolved by agarose gel electrophoresis to visualise the optimal sonication time required to shear crosslinked protein-DNA complexes to fragments of approximately 200-1000 bp. Sonicated, purified DNA (5μl) were diluted in 6x loading buffer and applied to a 1%w/v agarose gel and electrophoresed in TAE buffer containing 5μg/ml ethidium bromide and run at a constant 105V for 1 hour. Sizes of products were estimated by comparison to a DNA size ladder (1 Kb Plus). The gel was viewed on a UV Biorad Gel Doc System and photographed.

The sonication time of 120 seconds showed adequate genomic shearing evidenced by DNA fragments between 200-1000 bp and was selected as the optimal sonication time for future ChIP experiments.

3.5.2. Chromatin crosslinking and lysis of nuclei

Transfected cells (1x10^6 cells/ml) were treated with 1% formaldehyde in 10mls of PBS and incubated for 10 minutes at room temperature. Glycine at a final concentration of 0.125M, was applied to the cells to prevent Pax7-isoforms from over-fixing to the DNA. Cells were then washed twice with 10ml ice cold PBS. Transfected cells were resuspended in SDS lysis buffer (10mM EDTA, 50mM Tris-HCl, pH 8.1 and 1% SDS) supplemented with proteinase inhibitors (Protease inhibitor cocktail (PIC) 20x, Phenylmethyl Sulfonyl Fluoride, PMSF 200x) and incubated on ice for 10 minutes to lyse nuclei.

3.5.3. DNA-shearing and pre-clearing

Cell lysates were sonicated for 120 seconds (according to optimisation procedures described in Section 3.5.1.), centrifuged at 14 000rpm for 10 minutes at 4°C, then supernatant was collected and diluted with 600μl ChIP dilution buffer (0.001% SDS,
1.1% Triton X 100, 1.2mM EDTA, 16.7mM Tris-Cl pH 8.1, 167mM NaCl), containing 1x PIC and pre-cleared (to reduce non-specific immunoprecipitation) by the addition of 30μl DNA/Protein A Agarose beads (Santa Cruz Biotechnology). The samples were incubated on a rotating platform for 30 minutes at 4°C, centrifuged at 1200rpm for 1 minute and the supernatant was collected.

3.5.4. Immunoprecipitation of Pax7-DNA complexes

Supernatants from pre-clearing were divided between experimental samples and control (no antibody) samples. The V5 antibody (2μg) was added to each experimental sample. All samples were incubated overnight on a rocking platform at 4°C. DNA/Protein A Agarose beads (60μl) were added to each sample followed by a 2 hour incubation (rocking platform at 4°C) and then centrifuged at 1200rpm for 1 minute at 4°C. The beads were washed with 1ml ice cold low salt buffer (0.1% SDS, 1% Triton, 2mM EDTA, 20mM Tris, 150mM NaCl), then with 1ml ice cold high salt buffer (0.1% SDS, 1% Triton, 2mM EDTA, 20mM Tris, 500mM NaCl), and then three times with 1ml ice cold TE buffer (10mM Tris, 1mM EDTA). Each wash was incubated for 5 minutes on a rotating platform at 4°C followed by centrifugation at 1200rpm for 1 minute at 4°C; the supernatant was discarded each time and the DNA/Protein A Agarose beads retained. Antibody/protein/DNA complexes were then eluted from beads by addition of 2x 250μl elution buffer (0.1M sodium hydrogen carbonate, 1% SDS), the mixture was shaken gently and incubated for 15 minutes at room temperature, followed by centrifugation at 1200rpm for 1 minute. The supernatant was collected and transferred into a new tube.

3.5.5. Reverse crosslinking of Pax7-DNA complexes

To reverse formaldehyde crosslinks, NaCl (final concentration 1x10^{-5}M) was added, followed by overnight incubation at 65°C. Samples were then centrifuged (14 000rpm, for 1 minute), protein was digested with 1nM Proteinase K, 30nM EDTA and 0.2nM Tris-HCl and incubated at 45°C for 1.5 hours.
3.5.6. DNA purification

DNA was purified using either the QIAquick spin column (Qiagen) or by Phenol Chloroform extraction.

3.5.6.1. Purification kit

Digested protein samples were resuspended in 800μl PB buffer and applied to QIAquick® spin columns (Qiagen), centrifuged at 13 000rpm for 60 seconds, washed with 750μl DNA cleanup buffer (PE), centrifuged at 13 000rpm for 60 seconds, then for an additional 1 minute (to remove residual ethanol from the PE buffer). Samples were then eluted with ddH₂O (50μl), incubated for 1 minute at room temperature and centrifuged (13 000rpm, for 1 minute). Purified DNA samples were subjected to PCR utilising primers specific for the representative target genes chosen for investigation (Table 3.1.).

3.5.6.2. Phenol chloroform isoamyl DNA purification

Following proteinase K digestion, samples were resuspended in Tellurium (TE)-saturated phenol: chloroform: isoamyl alcohol (25:24:1) (500μl), vortexed for 30 seconds and centrifuged (14 000rpm, for 5 minutes). The upper aqueous layer of each sample was placed into a new tube, followed by the addition of 100% ethanol (1000μl) supplemented with 3M Sodium Acetate (NaCH₃CO₂) (50μl). Samples were inverted 6 times and incubated overnight at -20°C. Samples were then centrifuged (14 000rpm, 4°C, 15 minutes), supernatant was removed and samples were air dried for 10 minutes then resuspended in 50μl TE buffer. Samples were subjected to PCR utilising primers specific for the representative target genes (Table 3.1.).

3.6. ChIP-PCR to detect bound target genes

PCR was performed using specific primers as indicated in (Table 3.1.) Optimal conditions for ChIP-PCR were; reaction volumes of 20μl containing final concentrations of 1μg/ml DNA template, 10μM each primer, 200μM each dNTP, 1x PCR buffer, 1x Q solution, Taq DNA polymerase (2.5 U). Optimal cycling conditions were; an initial denaturation step of 94°C for 3 minutes, followed by 35 cycles of 94°C for 1 minute.
(denaturation), primer-specific annealing temperature (Table 3.1.) for 1 minute, 72°C for 1 minute (extension), then a final extension at 72°C for 10 minutes. The PCR products were separated by agarose gel electrophoresis.

A positive control was also amplified using input samples (untransfected NIH3T3 cell DNA) that were crosslinked, lysed, sonicated for 120 seconds, de-crosslinked, protein digested and DNA purified to determine if the primers designed for each target gene chosen for investigation (Table 3.1.) successfully amplified the target genes (produce PCR products of the correct size). Amplified input samples were subjected to agarose gel electrophoresis to confirm the presence and sizes of the amplified target genes.

Table 3.1. Primer sequences, PCR annealing temperatures and expected sizes of PCR products.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequences</th>
<th>Amplicon size (bp)</th>
<th>PCR annealing Temperature</th>
</tr>
</thead>
</table>
| CntfR | F:5’-TTCCCAGAGCCCAGTGGAG-3’  
R:5’-CGGCAAGCAGCAGCACACAGCAG-3’ | 593 | 58°C |
| Eya4  | F:5’-TTGGAGTCAGGAGCAGGAC-3’  
R:5’-ACACAAACTCGTACACCCAC-3’ | 775 | 56°C |
| Gbxl  | F:5’-TGCTACACTACCTCTCCACAG-3’  
R:5’-TTCTTAGAGCAGCCTTGAGG-3’ | 378 | 56°C |
| Rasa3 | F:5’-TGCTACACTACCTCTCCACAG-3’  
R:5’-TTCTTAGAGCAGCCTTGAGG-3’ | 601 | 56°C |
| Kcnk2 | F:5’-TGCAATCAAGCTGAGTGC-3’  
R:5’-AAATAGGAGGAACACTGGAG-3’ | 725 | 56°C |
| Prrx1 | F:5’-AAATACGTTCCAGACCACCTCC-3’  
R:5’-ACCTATAACAGAGTGACTG-3’ | 470 | 50°C |
3.7. REGULATION OF TARGET GENES IN Pax7d TRANSFECTED CELLS

3.7.1. RNA extraction, DNase treatment of RNA and RNA concentration

RNA was isolated from P19 mouse embryonic cells transfected with Pax7d utilising the RNA extraction Aurum Kit (Qiagen) as per manufacturer's instructions. RNA samples were then treated with DNase 1 utilizing the DNA-free™ Kit (Qiagen) and RNA concentration was determined by measuring absorbance at 260 nm [\(A_{260}\)] using a biophotometer (eppendorf). Extracted, DNase treated RNA was used in RT-PCR reaction studies to determine regulation of target genes in transfected cells. Untransfected P19 cells were utilized as a negative control.

3.7.1.2. Reverse transcription of RNA from P19 cells transfected with the Pax7d transcript

RNA was first converted to cDNA utilizing Omniscript Reverse Transcriptase kit (Qiagen). Optimal conditions for converting RNA to cDNA were; reaction volumes of 20 \(\mu\)l containing final concentrations of 1 \(\mu\)g total RNA template, 1 \(\mu\)M Oligo-dT primer, 0.5 mM of each dNTP (Qiagen), 1x RT-Buffer (Qiagen), RNase out (10 U, Qiagen), Omniscript reverse transcriptase (4 U, Qiagen). Each reaction was incubated for 60 minutes at 37°C.

3.7.1.3. Polymerase Chain Reaction (PCR) to assess target gene expression

The cDNA of target genes was amplified utilizing primer pairs corresponding to the 5' flanking region or sequences beyond the 3' end of target genes (Table 3.2.). Optimal conditions for PCR were; reaction volumes of 20 \(\mu\)l containing 1/10 of the final concentration of cDNA (0.1 \(\mu\)g), 1 \(\mu\)M primers, 1x PCR buffer, 1x Q solution, 200 mM of each dNTP and Taq DNA polymerase (2.5 U/reaction). Optimal cycling conditions were; 94°C for 3 minutes, followed by 45 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 30 seconds, then a final extension at 72°C for 10 minutes and a cool down step at 10°C for 10 minutes. RT-PCR products for each target gene were resolved by agarose gel electrophoresis as described in (Section 3.2.3.).
P19 cells transfected with the *Pax7d* transcript were subjected to RT-PCR for mRNA expression levels as described in (Section 3.4.2.) and immunocytochemistry as described in (Section 3.4.3.). Once expression of *Pax7* and *V5* were confirmed, cells were treated for ChIP as described in (Sections 3.5.2.-3.5.5., 3.5.6.2) and finally amplified by PCR (Section 3.6.). Due to time constraints a western blot analysis of the *Pax7d* transcript was not performed.

**Table 3.2. Primer sequences and expected sizes of RT-PCR products.**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequences</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
</table>
| **CntfR** | F: 5’-ACACGCAGAAACACAGTCCAC-3’  
R: 5’-TTGGAGCGGCGAGCTGACAC-3’ | 274 |
| **Eya4** | F: 5’-GTAAAGAAAAACGTGGCCAG-3’  
R: 5’-AAAAGTGTCAAGCGCTCC-3’ | 258 |
| **Gbx1** | F: 5’-ACGCTGAGGAGCTGCTGC-3’  
R: 5’-TCGCTGCCTGCTGGGTCTC-3’ | 163 |
| **Rasa3** | F: 5’-AGGGACACATGTTCCAGC-3’  
R: 5’-GTTCACAATGGGGAGACC-3’ | 162 |
| **Kcnk2** | F: 5’-TCAAGTTAGTCATGGGAC-3’  
R: 5’-ATGTGTCTTCCACTTGGC-3’ | 221 |
CHAPTER FOUR

RESULTS
Chapter 4  RESULTS

4.1  Introduction

To determine the target genes that each Pax7 isoform selects, chromatin immunoprecipitation (ChIP) was performed on Pax7-isoform-DNA (target gene) complexes isolated from mouse NIH3T3 embryonic fibroblast cells transfected with individual Pax7(a-d)-pcDNA3.1D/V5-His vector transcripts.

Transfected cells were stably selected with G418 for 28 days, assessed for Pax7-vector and Pax7 expression by western blot, then bound target genes of each isoform were analysed by chromatin immunoprecipitation. Pax7-vector-target-gene complexes were immunoprecipitated with the anti-V5 antibody, to ensure only Pax7 isoforms expressed from the Pax7-vector constructs (and not endogenous isoforms) were selected. Previously identified downstream target genes chosen for investigation of Pax7 isoform binding were: CntfR, Eya4, Gbx1, Prrx1, Rasa3 and Kcnk2.

4.2.  AIM: To identify Pax7 isoform-specific differences in target gene selection.

4.3.  Pax7(a-d)-vector transfection

4.3.1. Successful isolation of DNA containing Pax7(a-d)-vector constructs

In previous studies, full length Pax7(a-d) transcripts were cloned into the pcDNA3.1D/V5-His-TOPO expression vector and sequenced to confirm correct orientation and sequence. Here clones were amplified in E. coli, and then individual clones were picked at random, grown in LB broth and vectors were isolated utilizing a miniprep plasmid purification kit (Qiagen). Agarose gel electrophoresis was performed to assess the quantity and quality of Pax7(a-d)-pcDNA3.1D/V5-His-TOPO vector DNA (Figure 4.1.). Bands were observed at 12000 bp, 7000 bp and 5000 bp indicating circular plasmid forms of the cloned DNA.
Chapter Four: Results

Figure 4.1. Agarose gel electrophoresis of purified \textit{Pax7(a-d)-pcDNA3.1D/V5-His-TOPO} vector DNA: (Left to right) Lane 1 (Pax7d-), 2 (Pax7c-), 3 (Pax7b-) and 4 (Pax7a-) vector constructs respectively. DNA bands were observed at approximately 12000 bp, 7000 bp and 5000 bp. Lane M, displays the 1Kb Plus standard DNA ladder (Invitrogen).

Results indicate successful isolation of plasmid vectors, given that the set of three large DNA bands at around 12000 bp, 7000 bp and 5000 bp are indicative of circular plasmid DNA in various conformations.

4.3.2. Polymerase Chain Reaction (PCR) of \textit{Pax7(a-d)-pcDNA3.1D/V5-His-TOPO} cDNA constructs

PCR was performed on isolated purified plasmid DNA using vector specific primers that span the vector and \textit{Pax7} insert at the 5' end and the vector at the 3' end to determine the size and orientation of alternate \textit{Pax7} transcript vector inserts. PCR products of the correct size of 1650 bp were obtained for each cloned transcript (Figure 4.2.).

Figure 4.2. Agarose gel electrophoresis of isolated \textit{Pax7}-vector DNA amplified by PCR utilizing \textit{Pax7}-vector specific primers (F4 and BHG respectively): Lanes 1, 2, 3, and 4 display alternate \textit{Pax7} transcripts inserted correctly into the \textit{pcDNA3.1D/V5-His-TOPO} vector. (Left to right) Lane 1 (Pax7d), Lane 2 (Pax7c), Lane 3 (Pax7b), and Lane 4 (Pax7a). Lane M, the 1 Kb Plus standard DNA ladder. Lane B, blank sample consisting of all PCR reaction constituents excluding template DNA (negative control). All transcripts display the predicted size of 1650 bp.
4.4. Stable transfection studies

4.4.1. Cell morphology

In order to assess the DNA-binding differences between each Pax7 isoform, single Pax7(a-d) transcripts cloned into the pcDNA3.1D/V5-His-TOPO expression vector were transfected into NIH3T3 mouse embryonic fibroblast cells. Constructs that had stably integrated into the mouse NIH3T3 genome were selectively grown in cell culture medium containing G418.

NIH3T3 cells transfected with the pcDNA3.1D/V5-His-lacZ vector alone, was used as a negative control. Photographs of clones 14 days after stable transfection of each Pax7 transcript and vector alone were taken utilizing a Leica DC300 digital camera (Figure 4.3.). No difference in morphology was observed between cells transfected with each transcript versus vector transfected cells (Figure 4.3.).

Figure 4.3. Micrograph images of stable Pax7-vector transfected clones (day 14) (20x magnification): (A) Cells transfected with vector alone. NIH3T3 cells transfected with (B) Pax7a transcript; (C) Pax7b transcript; (D) Pax7c transcript; (E) Pax7d transcript. Scale bar = 40μm.
4.5. Confirmation of Pax7 protein expression in stably transfected cells

After selection of transfected cells, Pax7 protein expression was assessed by western blot methodology.

4.5.1. Western blot

Western blot was performed to determine the protein expression levels in cells stably transfected with each Pax7 transcript or with vector alone. Initially the nitrocellulose membrane was probed with the anti-V5 antibody, to ensure that cells express full-length Pax7 isoforms fused to the V5 epitope tag transcribed from the vector (Figure 4.4. A). The membrane was then stripped and re-probed with the Pax7 antibody to further confirm protein identity (Figure 4.4. B). A band corresponding to a Pax7-V5-His-fusion protein was observed on the western blot at the correct size of 59kDa (Figure 4.4. A, Lane 3). A similar band was observed when the membrane was probed with the Pax7 antibody (Figure 4.4. B, Lane 3). Positive Pax7 expression at the correct size of 59kDa was obtained for the Pax7c isoform only (Figure 4.4. A and B, Lane 3).

![Western Blots of Pax7-V5-His-fusion protein probed with V5-antibody (A) or Pax7 antibody (B):](image)

Figure 4.4. Western Blots of Pax7-V5-His-fusion protein probed with V5-antibody (A) or Pax7 antibody (B):(Left-right) Lane 1, Pax7a; Lane 2, Pax7b; Lane 3, Pax7c; Lane 4, Pax7d; Lane 5, Control vector; and Lane 6, Untransfected NIH3T3 cells. (A) Positive expression observed for the Pax7c isoform (Lane 3) at the correct size of 59 kDa when probed with the V5 antibody, or (B) with the Pax7 antibody. Lane M, standard protein size ladder.
The size of 59kDa is anticipated since the Pax7 protein is 30kDa in size and when fused to the V5 and His epitope tags, a fusion protein of 59kDa is produced. From western blot it was apparent that specific Pax7-isoform-epitope-tag expression was achieved for the Pax7c transfectant (Figure 4.4. A. Lane, 3) and possibly Pax7 expression for the Pax7a transfectant (Figure 4.4, B. Lane, 1). No endogenous Pax7 appeared to be consistently present in any of the cells (Figure 4.4. A and B Lanes 1-6). Since the Pax7c isoform demonstrated positive Pax7-V5-His protein expression, the isoform was subjected to chromatin immunoprecipitation and PCR (Section 4.7.2.) to observe target gene selection.

4.6. Sonication optimization of protein-DNA complexes for ChIP

4.6.1. Determining DNA shearing optimization conditions

To perform chromatin immunoprecipitation it was necessary to optimize the shearing of genomic DNA into 200-1000 bp fragments to ensure that analysed protein-DNA interactions are confined to distinct chromosomal regions. The optimal sonication time was chosen on the basis of obtaining small fragments of sonicated DNA between 200-1000 bp. Sonication optimization was performed on untransfected NIH3T3 cells (Figure 4.5.). As this had not been performed on cultured cell lysates in our laboratory before, but only on embryonic tissue, it was necessary to attempt sonication at a number of different times and determine the time at which optimal DNA shearing was obtained.

![Figure 4.5. DNA shearing conditions for NIH3T3 cells: A volume of 1ml was pulsed at 30% power for 60 seconds (Lane 6), 120 seconds (Lane 5), 160 seconds (Lane 4), 200 seconds (Lane 3) and 300 seconds (Lane 2) on ice with 30 seconds rest between each pulse. An unsonicated sample was used as a control (Lane 1). Samples were purified then subjected to agarose gel electrophoresis (1%), and visualized by a gel doc UV illuminator system following ethidium bromide staining of the gel. Lane M, 1Kb Plus standard DNA ladder (Invitrogen).](image)
Results indicate that the maximum amount of DNA at the required size between 200-1000 bp was obtained after 120 second sonication time (Figure 4.5. Lane, 5).

4.7.  PCR of chromatin immunoprecipitated NIH3T3 cells

4.7.1. PCR of untransfected chromatin immunoprecipitated NIH3T3 cells (Input chromatin)

Primer pairs of target genes CntfR, Eya4, Gbxl, Prrxl, Rasa3 and KcnK2 (Table 3.1.), were first trialled against a positive control (input sample) containing chromatin isolated from untransfected NIH3T3 cells subjected to formaldehyde crosslinking (Methods section 3.5.2.), sonication (Methods section 3.5.3.), de-crosslinking (Methods section 3.5.5.) and DNA purification (Methods section 3.5.6.2.) without immunoprecipitation. PCR primer pairs correspond to the 5'- flanking region or sequences beyond the 3' end of target genes for target gene amplification (Table 3.1.). The PCR reactions generated products of the expected size for most target genes (Figure 4.6.).

Figure 4.6. PCR products amplified from untransfected NIH3T3 cells (input sample): (Left-right) Lanes 1, 601 bp (Rasa3); 2, 775 bp (Eya4); 3, 725 bp (KcnK2); 4, 378 bp (Gbxl); 5, no product is visible (Prrxl) and 6, 593 bp (CntfR). Lanes labelled B are negative controls consisting of all PCR reaction constituents minus the template DNA. No amplification was observed for these samples. Lane M, 1Kb Plus standard DNA ladder. Further optimisation for the Prrxl gene was conducted until specifically designed primers successfully amplified the target gene region.
Chapter Four: Results

Results indicate amplification of target genes chosen for investigation, demonstrating that primers specifically designed for most target genes successfully amplified target gene regions and are therefore appropriate for later stages of the investigation when the presence of these gene regions bound by Pax7 isoforms were to be assessed.

4.7.2. PCR of stably transfected Pax7c chromatin immunoprecipitated cells

Since positive Pax7c expression was obtained by western blot, the cell line transfected with the Pax7c-vector construct was subjected to chromatin immunoprecipitation to observe target gene selection (Figure 4.7.). Chromatin immunoprecipitation of transfected cells was performed with the V5 antibody which will bind to the V5 epitope tag on the Pax7c-vector fusion protein. The immunoprecipitated chromatin DNA was analyzed by PCR for the presence of regulatory sequences of selected target genes. This was achieved utilizing primer pairs corresponding to the 5'-flanking region or sequences beyond the 3' end of target genes as described above (Table 3.1.).

Upon completion of this process, negative results were obtained. Results show no target genes were amplified from DNA isolated from chromatin immunoprecipitates when this process was performed on Pax7c-vector transfected cells using the V5 antibody (Figure 4.7. Lane, +). Furthermore, no PCR products were visible even when DNA from cells transfected with Pax7c-vector constructs but not immunoprecipitated with V5 (no antibody control) were used (Figure 4.7. Lane, -). In contrast, products of expected sizes were visible for input chromatin (untransfected NIH3T3 cells) not precleared, not immunoprecipitated with V5, washed and eluted (Figure 4.7. Lane, I). No products were visible for blank samples that consisted of all PCR reaction components minus template DNA (Figure 4.7. Lane, B).
Figure 4.7. ChIP-PCR to assess target gene selection by \textit{Pax7c}-vector constructs immunoprecipitated with the V5 antibody: (Top-bottom) CntfR, Eya4, Gbx1, Rasa3, Kcnk2 and Prrxl (Left-right) (B) blank PCR samples consisting of all PCR reaction constituents minus template DNA, no products are visible. (I) Input samples showing visible products of target genes of expected size (Top-bottom) CntfR, 593 bp; Eya4, 775 bp; Gbx1, 378 bp; Rasa3, 601 bp; Kcnk2, 725 bp and Prrxl, 470 bp. (–) No antibody controls amplified with target gene primer pairs. For all target genes, no PCR products are visible. (+) Pax7c isoform immunoprecipitated with V5 antibody and amplified with target gene primer pairs (Top-bottom) CntfR, Eya4, Gbx1, Rasa3, Kcnk2 and Prrxl. Negative results show no target genes were selected by Pax7c.

Results indicate that no DNA was amplified when isolated from immunoprecipitated samples, since there are no visible PCR products for the set of target genes chosen for investigation. In contrast, input chromatin (untransfected NIH3T3 cells not immunoprecipitated) successfully amplified each target gene (Figure 4.7. Lane, I).

Second ChIP attempt

Given that there appeared to be Pax7a-vector protein expression by western blot probed with the Pax7 antibody (Figure 4.4. B, Lane 1), the cells transfected with \textit{Pax7a}-vector construct were subjected to chromatin immunoprecipitation and a PCR was performed with specific primers (Table 3.1.) designed to amplify the promoter region of CntfR (Figure 4.8. A) and Rasa3 (Figure 4.8. B) genes to observe whether these genes were selected by Pax7a. Results show that these target genes were not selected by Pax7a (Figures 4.8. A and B, Lane +). In contrast, input chromatin (untransfected NIH3T3 cells) amplified CntfR and Rasa3 target genes and produced a PCR product of the correct size (Figure 4.8. A and B, Lane I respectively).
Figure 4.8. ChIP-PCR for selection of (A) \(CnffR\) and (B) \(Rasa3\) by Pax7a identified by immunoprecipitation with the V5 antibody: (A) and (B) Lane (B) blank PCR sample consisting of all PCR reaction constituents minus template DNA, no products are visible. (A) (Left-right) (I) Input sample showing a visible product of expected size 593 bp (\(CnffR\)). (-) No antibody control amplified with \(CnffR\) primer pairs generated no PCR product. (+) cells transfected with Pax7a-vector construct immunoprecipitated with V5 antibody and amplified with \(CnffR\) primer pairs, generated no PCR product. (B) (Left-right) (I) Input sample showing a visible product of expected size 601 bp (\(Rasa3\)). (-) No antibody control amplified with \(Rasa3\) primer pairs generated no PCR product. (+) cells transfected with Pax7a-vector construct immunoprecipitated with V5 antibody and amplified with \(Rasa3\) primer pairs generated no PCR product. No other target genes were attempted for the Pax7a-transfected cells.

Summary

Although results show that ChIP methodology was not successful in isolating target genes bound by Pax7, several reasons for no target gene amplification were possible. Firstly, it was possible that not all cells were expressing high levels of Pax7. Therefore, to test this the Pax7-vector transfected cells were subjected to mRNA analysis. The Pax7c-vector transfected cells as well as control vector transfected cells were also subjected to immunocytochemistry to determine expression levels and the ratio of cells stably transfected and expressing the Pax7c-vector fusion protein containing the V5 epitope.

Moreover, secondly, it was possible that negative results were obtained due to losses of DNA throughout the ChIP procedure.
4.8. Optimization of DNA extraction for ChIP methodology

4.8.1. Confirming the presence and quantity of DNA throughout the chromatin immunoprecipitation procedure

Because no PCR products were evident from ChIP isolated DNA, the amount of DNA material was then assessed at each step to determine whether DNA was lost during the multi-step procedure (Methods section 3.5.). When DNA analysis was performed on untransfected sonication trials, loss of DNA was apparent (Figure 4.5.). Thus to try to correct this, DNA concentration was analysed at each step prior to immunoprecipitation using NIH3T3 control cells.

The amount of DNA was analyzed after cell collection from cell culture flasks by either accutase treatment or by applying nuclear lysis buffer directly to cells (Figure 4.9. A-C.). DNA was then analyzed after optimizing protein-DNA sonication shearing conditions for each method of cell collection (Figure 4.9. B and C respectively). The quantity of DNA was greater when nuclear lysis buffer was applied directly to the cells (Figure 4.9. B) compared to DNA levels after cells were removed by accutase (Figure 4.9. C).

The presence of DNA was again analyzed after DNA was purified (Methods section 3.5.6.) by either the PCR DNA purification kit (Qiagen) or by phenol chloroform isoamyl extraction (Figure 4.9. D and E respectively). The quantity of DNA was greater when phenol chloroform isoamyl extraction was used (Figure 4.9. Image D) compared to DNA levels obtained as a result of DNA purification by the DNA purification kit (Figure 4.9. E).
Figure 4.9. Agarose gel electrophoresis images indicating the quantity of DNA at several steps during the chromatin immunoprecipitation procedure: (Left-right) Lanes labeled M, 1kb Plus standard DNA ladder (Invitrogen). (A) displays the presence of DNA after cells were lysed directly with nuclear lysis buffer (Lane 1). (B) Lane 1 displays cells directly lysed, sonicated for 120 seconds followed by a centrifugation concentration step. Sheared chromatin was between the expected sizes of 200-1000 bp. The presence of DNA after cells directly lysed with nuclear lysis buffer were subjected to shearing at the following times, 60 seconds (Lane 3), 120 seconds (Lane 2). (C) displays the presence of DNA after cells treated with accutase were subjected to shearing for 120 seconds (Lane 1), the smear was between the expected size of 200-1000 bp. (B) and (C) indicate the presence and quantity of DNA is better when cells were lysed by direct application of nuclear lysis buffer.

(D) displays the presence of DNA after DNA purification with the PCR purification Kit (Qiagen). (Left-right) Lanes 1, 2, 3, 4, and 5 display cells to which nuclear lysis buffer was directly applied followed by sonication for 300 seconds, 200 seconds, 160 seconds, 120 seconds and 60 seconds respectively. DNA smears were between expected sizes of 200-1000 bp. (E) displays the presence of DNA after DNA purification with phenol chloroform isoamyl alcohol extraction. (Left-right) Lanes 1, 2, 3, 4 and 5 display DNA from cells to which lysis buffer was directly applied followed by sonication for 300 seconds, 200 seconds, 160 seconds, 120 seconds and 60 seconds respectively. DNA smears were larger in size and DNA concentration was greater.
Chapter Four: Results

Results indicate that optimal DNA levels were obtained when cells were treated directly with nuclear lysis buffer (Figure 4.9. B.), prior to sonication for 120 seconds. When this step was followed by centrifugation, the DNA levels were concentrated allowing more DNA to be used in later steps (Figure 4.9. B, Lane 1). Decrosslinked cells were finally DNA purified by phenol chloroform isoamyl extraction procedure rather than by purification kit as this extraction procedure yielded higher DNA concentrations (Figure 4.9. E) prior to proceeding with PCR analysis on chromatin immunoprecipitated cells.

Because DNA of target genes was not found bound by Pax7 in Pax7-vector transfected cells, the levels of Pax7 expression and the number of Pax7 expressing cells were assessed in NIH3T3 cells transfected with Pax7(a-d)-vector. It was felt that lack of ChIP results might also be due to poor transfection efficiencies as individual clones had not been selected during the selection process in G418 but rather all surviving cells had been used.

4.9. Confirmation of Pax7 expression in NIH3T3 cells

4.9.1. mRNA expression of Pax7 in transfected cells

Pax7(a-d) expression in NIH3T3 cells was assessed by RT-PCR after treatment with DNase 1 to remove the DNA of the Pax7a-d-vector constructs from transfected cells. Using primers F4 and Pax7R, a region of Pax7 was amplified that spans the 5′ paired box region. A PCR product was obtained of the expected size of 300 bp, in all cells transfected with Pax7(a-d)-vector constructs (Figure 4.10. Lanes 1, 2, 3 and 4). In contrast, no product was visible in untransfected NIH3T3 cells (negative control) (Figure 4.10. Lane, N) and in the blank sample containing all PCR reaction constituents minus template DNA (Figure 4.10. Lane, B).
Chapter Four: Results

Figure 4.10. RT-PCR products indicating Pax7 expression: RNA was extracted from NIH3T3 cells treated with DNase 1 and converted into cDNA. Primers F4 and Pax7R were employed to amplify a region of Pax7. (Left to right) Lanes 1, Pax7d; 2, Pax7c; 3, Pax7b; and 4, Pax7a demonstrate a PCR product of 300 bp. Results from Pax7 transfected NIH3T3 cells compared to untransfected NIH3T3 cells where no product (Lane N) is visible. No product is visible in the blank sample (Lane B). Lane M, 1kb Plus standard DNA ladder.

RT-PCR results confirm that Pax7-vector constructs were successfully expressed in stably transfected NIH3T3 cells.

4.9.2. Immunocytochemistry

Positive Pax7c-vector expression evident from western blot and RT-PCR analysis of transfected cells suggested that sufficient Pax7-vector expression was present in transfected cells for chromatin immunoprecipitation to succeed. However, results obtained from ChIP studies were negative. Therefore the expression of Pax7-vector constructs in transfected cells was analysed by immunocytochemistry to assess the number of cells expressing Pax7 as well as the level of Pax7 expression. Immunocytochemistry was performed with mouse monoclonal Pax7 and V5 (IgG2a) antibodies respectively. Since the cells transfected with Pax7c-vector construct appeared to have the highest levels of Pax7-vector expression when assessed by western blot, only cells transfected with Pax7c-vector construct were analysed by immunocytochemistry. All cells that survive selection in G418 should contain the Pax7c-vector construct stably inserted into their genomic DNA and should therefore express the Pax7c-vector construct via the cytomegalovirus (CMV) promoter present in the vector.

Since Pax7 is a transcription factor, the Pax7-vector fusion protein is likely to be predominantly present in the nuclei of transfected cells when probed with either Pax7 or V5 antibodies (Figure 4.11. C and D respectively). Since the pcDNA3.1D/V5- His-lacZ vector alone would express lacZ with a V5-epitope-tag, the cells transfected with vector
alone were expected to display predominant cytoplasmic epitope expression when probed with the V5 antibody (Figure 4.11. B).

![Figure 4.11. Immunocytochemistry of NIH3T3 cells stably transfected with Pax7c-vector construct or vector alone: Micrographs of immunocytochemistry showing mouse monoclonal Pax7 or V5 antibodies, 40x magnification. (Left to right) (A) and (B) NIH3T3 cells transfected with vector alone. (C) and (D) NIH3T3 cells transfected with Pax7c-vector construct. (A) Vector alone visualized with Pax7 antibody, (B) Vector alone visualized with V5 antibody, showing V5 expression in the nucleus and cytoplasm of the cell; (C) Pax7c-vector transfected cells visualized with Pax7 antibody, showing Pax7 expression in the nucleus of the cell. (D) Pax7c-vector expression visualized with V5, showing positive V5 expression in the nucleus of the cell. Scale bar = 20μm.]

Results clearly indicate nuclear Pax7c-vector expression in NIH3T3 cells transfected with the Pax7c-vector construct when reacted with either V5 or Pax7 antibodies (Figure 4.11. C and D respectively). NIH3T3 cells transfected with the pcDNA3.1D/V5-His-lacz
expression vector alone demonstrate cytoplasmic V5-epitope tag expression (Figure 4.11. B). Cells not transfected with the Pax7c-vector construct showed no Pax7c expression (Figure 4.11. A). In all cases, cells transfected with either Pax7c-vector construct or with vector alone demonstrate low levels of V5 expression in only 1 in 100 cells (Figure 4.11. B, C and D).

These results provide a valid reason for the negative ChIP results. From these results it is clear that not enough of the Pax7c isoform was present in sufficient numbers of transfected cells to be found bound to DNA during the formaldehyde crosslinking procedure (Methods section 3.5.2.). As a consequence of this firstly, cells were retransfected.

Re-transfection and selection of each Pax7 transcript was conducted utilizing an antibiotic concentration curve to assess the optimal concentration at which untransfected cells would be lost. However, after four weeks of selection all cells remained completely confluent. Immunocytochemistry was conducted on cells with the highest antibiotic concentration applied. Unfortunately, re-transfected cells did not express any of the isoforms. Therefore, all re-transfected cells were discarded since they were resistant to the antibiotic. A third attempt to re-transfect was not performed due to time constraints.

4.10. Confirmation of Pax7 expression in a stably transfected P19 clone

A P19 embryonic carcinoma stem cell line previously transfected with Pax7d and kindly donated by Vijesh Vaghjiani was utilized for further experiments.

4.10.1. mRNA expression of Pax7 in transfected cells

Pax7d expression in P19 cells was assessed by RT-PCR after treatment with DNase 1 to remove the DNA of the Pax7d-vector constructs from transfected cells. Using primers F4 and Pax7R, a region of Pax7 was amplified that spans the 5’ paired box region. A PCR product was obtained of the expected size of 300 bp, in cells transfected with Pax7d-vector construct (Figure 4.12. Lane, 1). In contrast, no product was visible in
untransfected P19 cells (negative control) (Figure 4.12. Lane, N) and in the blank sample containing all PCR reaction constituents minus template DNA (Figure 4.12. Lane, B).

![Figure 4.12. RT-PCR product indicating Pax7 expression](image)

**Figure 4.12. RT-PCR product indicating Pax7 expression:** RNA was extracted from P19 cells treated with DNase 1 and converted into cDNA. Primers F4 and Pax7R were employed to amplify a region of Pax7. (Left to right) Lane 1 demonstrates a PCR product of 300 bp. Results from Pax7 transfected P19 cells compared to untransfected P19 cells where no product (Lane N) is visible. No product is visible in the blank sample (Lane B). Lane M, PUC DNA ladder (Fermentas).

RT-PCR results confirm that the Pax7d-vector construct was successfully expressed in stably transfected P19 cells.

4.10.2. Immunocytochemistry

The expression of the Pax7d-vector fusion protein was visualized in transfected cells, by immunocytochemistry utilizing a mouse monoclonal Pax7 antibody or a V5 (IgG2a) antibody respectively (Figure 4.13.).

Since the transfected cells were selected using G418 and an individual clone was selected and assessed for Pax7 expression by RT-PCR, all cells that survived the selection process should contain Pax7d-vector construct stably inserted into their genomic DNA, and thus should express the Pax7d-vector construct via the CMV promoter present in the vector in the nuclei of transfected cells.

58
Figure 4.13. Immunocytochemistry of cells transfected with a Pax7d-vector construct relative to untransfected P19 cells: Micrographs of immunocytochemical staining utilizing mouse monoclonal Pax7 or V5 antibodies, 100x and 40x magnification, respectively. (Left to Right) (A) and (B) Untransfected P19 cells. (C) and (D) P19 cells transfected with Pax7d-vector construct. (A) Untransfected P19 cells visualized with Pax7 antibody, (B) untransfected cells with V5 antibody (C) Pax7d-vector fusion protein visualized with Pax7 antibody, showing positive Pax7 expression in the nuclei of cells. (D) Pax7d-vector fusion protein visualized with V5, showing positive V5 expression in the nuclei of cells. Scale bar = 20µm.

Results clearly indicate nuclear Pax7d-vector expression in P19 cells transfected with the Pax7d-vector construct visualized with Pax7 or V5 antibodies (Figure 4.13. C and D). Untransfected P19 cells showed no V5 or Pax7 expression (Figure 4.13. A and B). Pax7d-vector transfected cells visualized with Pax7 or V5 show that approximately 1:2 cells express the fusion protein (Figure 4.13. C, D). P19 cells transfected with vector alone did not survive and no further transfections were conducted due to time constraints.
Chapter Four: Results

Since the Pax7d-vector fusion protein was present in approximately half the cells, the isoform was subjected to ChIP and PCR to observe target gene selection.

4.11. PCR of chromatin immunoprecipitated P19 cells

4.11.1. PCR of chromatin immunoprecipitations of cells stably transfected with Pax7d-vector construct

Immunoprecipitations were performed with the V5 antibody against the V5-epitope-tag present in the Pax7-pcDNA3.1D/V5-His fusion protein. The immunoprecipitated DNA was isolated and purified using optimized procedures and analyzed by PCR for the presence of regulatory sequences of selected target genes. This was achieved utilizing primer pairs corresponding to the 5’-flanking region or sequences beyond the 3’ end of the target genes (Table 3.1.).

Results were negative yet again. No target genes were selected by the Pax7d-vector fusion protein (Figure 4.14. Lane, +). Moreover, as expected no PCR products were visible for samples isolated from transfected cells but without immunoprecipitation (no antibody control) (Figure 4.14. Lane, -). In contrast, products of expected sizes were visible for input chromatin (untransfected P19 cells) that were not precleared, immunoprecipitated, washed or eluted (Figure 4.14. Lane, I). No products were visible for blank samples that consisted of all PCR reaction components minus template DNA (Figure 4.14. Lane, B).
Figure 4.14. ChiP-PCR for target gene selection by Pax7d-vector fusion protein immunoprecipitated with the V5 antibody: (Left-right) (B) blank PCR samples consisting of all PCR reaction constituents minus template DNA, no products are visible. (I) Input samples showing visible products of target genes of expected sizes (Top-bottom) CntfR, 593 bp; Eya4, 775 bp; Gbx1, 378 bp; Rasa3, 601 bp; Kcnk2, 725 bp, and Prrxl, 470 bp. (-) No antibody controls amplified with target gene primer pairs (Table 3.1.) (Top-bottom) CntfR, Eya4, Gbx1, Rasa3, Kcnk2 and Prrxl. For all target genes, no PCR products are visible. (+) Pax7d isoform immunoprecipitated with V5 antibody and amplified with target gene primer pairs (Top-bottom) CntfR, Eya4, Gbx1, Rasa3, Kcnk2 and Prrxl, no products are visible.

Results indicate that using the ChIP procedure performed here, no DNA from the genes was present in the immunoprecipitated samples, since there are no visible PCR products for target genes chosen for investigation. In contrast, input chromatin (untransfected P19 cells not immunoprecipitated) amplified each target gene (Figure 4.14. Lane, I).

Attempts to immunoprecipitate Pax7-vector-DNA complexes with the Pax7 antibody were not possible without agarose G beads which were not available in our laboratory at the time.

4.12. Target gene regulation by Pax7d

Since we were not able to identify target genes bound by Pax7 using the ChIP procedure, Pax7d transfected cells were analysed for differential regulation of target genes relative to untransfected cells.

4.12.1. mRNA expression of each target gene

P19 cells stably transfected with the Pax7d-vector construct were propagated to analyze the regulation of target genes CntfR, Gbx1, Rasa3, Eya4 and Kcnk2 by RT-PCR. The regulation of target genes by the Pax7d isoform in stably transfected P19 cells was
assessed by RT-PCR after treatment with DNase 1 (to remove transfected DNA). Target genes were amplified utilizing primer pairs corresponding to the exonic sequences in each of the genes (Table 4.1). RT-PCR products at expected sizes of 274 bp and 162 bp were obtained for CntfR and Rasa3 target genes respectively (Figure 4.15. CntfR and Rasa3, Lanes +). In contrast, no products were visible for all other target genes for cells either transfected or untransfected with the Pax7d-vector construct (Figure 4.15. Eya4, Gbxl and Kcnk2 Lanes, + and C for each target gene respectively). No product was obtained for the blank sample containing all PCR reaction constituents minus template DNA (Figure 4.15. Lanes labelled B).

Table 4.1. Primer sequences, PCR annealing temperatures and expected sizes of PCR products analysed by RT-PCR analysis.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequences</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CntfR</td>
<td>F: 5′-ACACGCAGAAACACACAGTCCAC-3′</td>
<td>274</td>
</tr>
<tr>
<td></td>
<td>R: 5′-TTGGAGCCGCACAGCTGACAC-3′</td>
<td></td>
</tr>
<tr>
<td>Eya4</td>
<td>F: 5′-GTAAAGAAGAAACGTGCCCAG-3′</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>R: 5′-AAAAGTGTCAAGCGCTCC-3′</td>
<td></td>
</tr>
<tr>
<td>Gbxl</td>
<td>F: 5′-ACGCTGAGGAGCTGCTGC-3′</td>
<td>163</td>
</tr>
<tr>
<td></td>
<td>R: 5′-TCGCTGCTGCTGAGCCTTC-3′</td>
<td></td>
</tr>
<tr>
<td>Rasa3</td>
<td>F: 5′-AGGGACACATGTTCCAGGC-3′</td>
<td>162</td>
</tr>
<tr>
<td></td>
<td>R: 5′-GTTCACAATGCGGAGGACCC-3′</td>
<td></td>
</tr>
<tr>
<td>Kcnk2</td>
<td>F: 5′-TCAAGTTAGTCACCTGGGACC-3′</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td>R: 5′-ATGTGTCTTCCACTTTGAC-3′</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.15. RT-PCR products indicating regulation of target genes in P19 cells stably transfected with the *Pax7d*-vector construct: (Left to right) + Lanes (*Rasa3,*) and (*CntfR*) generated PCR products of expected sizes 162 bp and 274 bp respectively. Results from *Pax7*-vector transfected P19 cells were compared to untransfected P19 cells where no PCR products are visible for target genes *CntfR, Eya4, Gbx1, Rasa3* and *Kcnk2* (Lanes labeled C). No products are visible in the blank samples (Lanes labelled B), consisting of all PCR reaction constituents minus the template DNA. Lane M, 1kb plus DNA ladder.

RT-PCR results confirm that only the target genes *CntfR* and *Rasa3* were successfully up-regulated in P19 cells stably transfected with the *Pax7d*-vector construct.
CHAPTER FIVE

DISCUSSION
Chapter 5  **DISCUSSION**

5.1 **Introduction**

Previous knockout and transfection studies have revealed the importance of *Pax7* for nervous and muscular system development (Seale et al., 1999; Mansouri et al., 1994; Jostes et al., 1990). The role of *Pax7* within both myogenic and neurogenic lineages signifies the diverse biological functions of *Pax7*, which may be explained by the presence of structurally distinct Pax7 isoforms. These isoforms differ in the structure of their DNA-binding domain (paired domain) known to influence target gene selection (Schafer et al., 1994), and are hypothesized to account for the diverse functions of Pax7. Although the transcripts encoding these isoforms are present in both neurogenic and myogenic cell lines (Ziman and Kay, 1998; Ziman et al., 1997), the functional significance of each Pax7 isoform for specification of these cell types remains unknown.

It was envisaged that this present study would be one of the first to identify the DNA-binding differences between each Pax7a-d isoform in terms of selecting specific target genes thereby characterizing each Pax7 isoform's specific role within a cell. This research is necessary since functionally-specific Pax7 isoforms are possible candidates for use in future stem cell/gene therapy approaches.

5.2 **Upregulation of CntfR and Rasa3 target genes by the Pax7d transcript**

*Pax7* is known to be involved in neural sub-type specification in the midbrain and in specification of skeletal muscle satellite cells (Seale et al., 1999; Jostes et al., 1990). A previous study has delineated the role of *Pax7* further by identification of several downstream target genes (White and Ziman, 2006). Here we provide the first evidence confirming a link between *Pax7* and upregulation of *CntfR* and *Rasa3*. The present study shows that transfection of the *Pax7d* transcript into P19 cells, causes the upregulation of *CntfR* and *Rasa3* target genes.

*CntfR* plays a critical role in the developing and mature central nervous system (Ip et al., 1993) and has a functional role in the degeneration and regeneration of myofibers during muscle repair and regeneration (Marques and Neto, 1997). Research has shown that the
ligand for CntfR, CNTF is predominantly expressed by glial cells (Stockli et al., 1989) and schwann cells of peripheral motor neurons and in astrocytes in the central nervous system (Stockli et al., 1989). CNTF exerts functional effects from schwann cells (Ip et al., 1993; Sendtner et al., 1992) and has been shown to cause neuronal growth and cholinergic formation, as well as stimulate choline acetyltransferase and the production of acetylcholine in specific neuroblastoma cell lines (Kirsch et al., 1997; Kirsch et al., 1996; Fuhrmann et al., 1995; Meyer-Franke et al., 1995; LaVail et al., 1992; Ernsberger et al., 1989; Lehwalder et al., 1989; Hofmann, 1988).

Furthermore, CNTF has been implicated in determination of the cell fate of neural stem cells, proliferation of neuronal precursor cells and the survival and differentiation of a variety of neurons in the peripheral and central nervous systems (Bonni et al., 1997; Ip and Yancopoulos, 1996; Ip et al., 1994; Patterson, 1994; Sendtner et al., 1994). The protein acts in response to spinal cord and striatal injuries (Lee et al., 1998; Oyesiku et al., 1996; Asada et al., 1995), hippocampal deafferentation (Guthrie et al., 1997) and ischemia (Park et al., 2000).

An in vivo study has shown that exogenous application of CNTF to rejuvenating muscle increases myotube formation (Marques and Neto, 1997) and that CNTF expression increases in response to denervation of muscles (Helgren et al., 1994; Ip et al., 1993). Furthermore, research has found that a mutation in the CntfR gene in mice induced severe motor neuron defects (DeChiara et al., 1995), confirming that CNTFR plays a critical role in nervous system development. In view of the importance of CntfR and its ligand in muscle and nerve development, specification and regeneration and the diverse functional roles ascribed to Pax7, the emerging scenario suggests that the Pax7d transcript may induce neural and myogenic specification via regulation of CntfR.

The suggestion that the Pax7d transcript is important for both neural and myogenic development is further supported by previous studies on alternate Pax7 transcripts which detail dominant Pax7d expression in the neurogenic cell line (Ziman and Kay, 1998; Ziman et al., 1997) but when transfected into haematopoietic stem cells Pax7d causes
myogenic differentiation (Seale et al., 2004). The results presented here confirm a role for Pax7 in these processes, possibly via upregulation of CntfR.

Additionally, the target gene Rasa3 plays an important role in cell signalling and cellular transformation. Previous studies have suggested a crucial role for the gene in cell survival signalled by nerve growth factors in sympathetic neurons (Nobes et al., 1996). In view of the functional roles ascribed to both Rasa3 and Pax7, the results from the present study indicate that the role of Pax7d in specification of cells during neural development may be mediated by Rasa3.

The Pax7d transcript excludes the hexanucleotide sequence (GTTTAG) within its mRNA structure (Ziman et al., 2001). Interestingly, the finding that Pax7 transcripts excluding the hexanucleotide sequence in ascidians (Halocynthia roretzi), resulted in the formation of neural structures (Wada et al., 1997) adding further support for a neural function for Pax7d.

Due to time constraints, further studies on cells transfected with Pax7a, Pax7b and Pax7c were not completed and thus determination of target genes selected by individual transcripts could not be determined. However, several important advances on ChIP methodology were made as a result of this research project.

5.3. Appropriateness of the chromatin immunoprecipitation technique

Immunoprecipitation of epitope-tagged Pax7-isoform-DNA complexes is a highly specific and complex procedure and was a suitable technique with which to investigate the DNA-binding differences between each Pax7 isoform. By using immunoprecipitation of transfected epitope-tagged isoforms bound to target genes it should have been possible to determine the selected target genes bound by each Pax7 isoform (separately from these bound by multiple endogenous Pax7 isoforms). This was an important consideration since, a single Pax7d transcript infected into haematopoetic stem cells upregulated endogenous Pax7 transcripts (Seale et al., 2004). However in transfection studies
conducted for this project, endogenous \textit{Pax7} did not appear to be upregulated, (as judged by a single band obtained by western blot utilizing a Pax7-specific antibody).

Chromatin immunoprecipitation encompasses four experimental components including an initial fixation step, followed by sonication, immunoprecipitation and finally an analysis of the immunoprecipitated DNA. Das and colleagues (2004) reinforce that it is a complicated procedure, requiring optimization of reaction conditions, an obstacle faced in this present study.

In this study, NIH3T3 cells stably transfected with each \textit{Pax7} transcript were subjected to the first few stages of chromatin immunoprecipitation involving formaldehyde crosslinking, sonication and lysing of nuclei. Western blot procedures were then conducted to confirm the level of Pax7 isoform expression before proceeding with ChiP. Since positive Pax7 fusion protein was observed for the Pax7c isoform (Figure 4.4.A.) these cells were subjected to the remainder of ChiP. However, whilst ChiP was performed on the Pax7c isoform, trials were simultaneously conducted on untransfected NIH3T3 cells to obtain optimal DNA shearing conditions, only to discover that DNA was being lost somewhere within the procedure.

The maintenance of DNA quantity throughout the ChiP procedure is vital since the loss of DNA minimizes the possibility of obtaining meaningful results from the overall investigation. In the experiments conducted here, it is possible that no DNA was present when Pax7-isoform-DNA complexes were immunoprecipitated from the NIH3T3 genome due to DNA losses and as a result the final PCR product revealed a negative result for the selection of target genes.

A large amount of time was spent minimizing the amount of DNA lost during the procedure. Method optimization was conducted with untransfected NIH3T3 cells, and several different methods of cell collection were assessed. Results suggested that direct application of nuclear lysis buffer to cells resulted in greater DNA recovery compared to other methods of cell collection after formaldehyde crosslinking. It is therefore possible
to conclude that in the first ChIP experiment, loss of DNA due to accutased cell collection, was responsible for unsuccessful collection of the maximum number of cells, subsequently decreasing immunoprecipitation efficiency during ChIP. DNA purification after immunoprecipitation by DNA purification kit rather than by phenol chloroform isoamyl extraction caused further DNA losses.

5.4 Optimization of transfections

While ChIP studies conducted on Pax7(a-d) transfected cells were not successful, it would not be credible to conclude that the Pax7c isoform does not bind to these target genes since numerous possibilities can account for the negative result. The loss of DNA during the ChIP procedure is a major factor contributing to lack of positive results. Another major contributing factor to the negative outcome was the low transfection efficiency. Not enough cells expressed the Pax7c isoform to give results; immunocytochemistry revealed that only 1:100 cells were relatively expressing the Pax7 isoform.

It is critical for all stably transfected cells to express Pax7 isoforms to ensure that enough protein is present in cells to be bound at any one time to DNA during formaldehyde crosslinking procedures. Crosslinking of Pax7 to DNA targets would essentially freeze 'a moment in time' when a number of Pax7 molecules would be expected to be bound to different target genes. To identify these target genes bound at any one moment requires that a large amount of Pax7 be bound to the target genes at that moment for positive identification by ChIP. Therefore, insufficient Pax7 levels at any one time may have resulted in the levels of target genes selected being too low to be detected by ChIP.

Most ChIP studies are conducted on whole tissue (Antonini et al., 2006; Jelinic et al., 2006; Matsukuma et al., 2006; Plateroti et al., 2006; Pedchenko et al., 2005; Peng and Chen, 2005; Pierreux et al., 2004) or embryos (Terranova et al., 2006; White and Ziman, 2006; Magness et al., 2004; Jasinskiene et al., 1995) where there are large numbers of cells containing transcription factors bound to specific target genes at any one moment in time.
However, other studies have subjected transiently transfected cells to chromatin immunoprecipitation and have obtained conclusive results (Agbottah et al., 2006; Esumi et al., 2006; Gong et al., 2006; Valineva et al., 2006). In these previous studies stable transfections were used and the multi-step procedures of ChIP were modified relative to our own study. For example one such study conducted formaldehyde crosslinking for a longer period of time (Gong et al., 2006). In effect, this may result in greater preservation of already bound proteins for subsequent coarse steps of the ChIP procedure for positive identification by ChIP. These results suggest that perhaps the chromatin immunoprecipitation technique was not sufficiently optimized to produce a conclusive result in our own study.

While the limited amount of cells expressing the isoform in the nucleus was a major reason for negative results in the first ChIP experiment, in the second experiment when a Pax7d stably transfected cell line was utilized and more cells were expressing the isoform in the nucleus, ChIP-PCR results still displayed no target gene selection after DNA purification. Since many of the methodological errors from the initial experimental attempt were resolved, it may be possible to conclude that the Pax7d isoform does not bind to Eya4, Gbx1, KcnK2 and Prrx1 target genes in vitro. However, since RT-PCR results of the Pax7d transfected P19 cells show that the target genes CntfR and Rasa3 were up-regulated by Pax7d, it is not credible to conclude that the isoform does not bind to these target genes. The only other possible explanations are that either Pax7d does not bind directly to target genes CntfR and Rasa3 to upregulate their expression or that the chromatin immunoprecipitation procedure was inadequate for analyzing Pax7 isoform DNA-binding differences in transfected cell lines. Moreover, since Pax7 has diverse functional roles in both neurogenic and myogenic cell lines (Ziman and Kay, 1998; Ziman et al., 1997), it is possible that the Pax7 protein binds to different target genes in different cell lines, or that the Pax7 protein is unable to bind to different target genes simultaneously.

Alternately, perhaps the wrong set of genes were chosen for investigation. The choice of target genes for another testing is an important consideration. The immunoprecipitated
DNA is possible to assess by promoter-specific PCR for a genuine target (Wang et al., 2002; Boyd and Farnham, 1999). However, while we checked several of the identified target genes of Pax7 (White and Ziman, 2006), it was not possible to check them all in the short time available.

5.5. Implications for no target gene selection by Pax7d

5.5.1. Pax7d protein structure and DNA-binding function

It is known that all Pax7 isoforms differ in the structure of their DNA-binding domain (paired domain) which influences the selection of downstream target genes. The Pax7d isoform includes a trinucleotide sequence (CAG) that encodes the amino acid residue glutamine (Q+) within the linker region between the PAI and RED subdomains of the paired domain (Figure 5.1.a.).

Previous studies have noted that a single glutamine residue within the paired domain causes only the PAI subdomain to bind to DNA (Du et al., 2005; Vogan and Gros, 1997; Vogan et al., 1996) (Figure 5.1.b.). Although the third alpha helix present within each subdomain is the principle component allowing each subdomain to interact with DNA (Figure 5.1.b.) (Kay and Ziman, 1999; Ziman and Kay, 1998; Balczarek et al., 1997; Vogan et al., 1996; Treisman et al., 1991), research has shown that a single glutamine (Q+) within the linker region extends the length of this linker region, which inhibits binding of the RED domain to target genes (Xu et al., 1999; Fortin et al., 1998; Jun et al., 1998; Vogan and Gros, 1997; Vogan et al., 1996). In effect, the structural alteration may restrict the DNA-binding potential of the Pax7 protein to target genes. It is possible that such a structural alteration resulted in no target gene selection by Pax7d.

Furthermore, research has shown that Pax proteins are capable of utilizing the paired domain in combination with the homeodomain to recognize specific DNA sequences, and thus influence target gene selection (Xu et al, 1999; Jun et al, 1998; Kozmik et al, 1997; Underhill and Gros, 1997; Vogan and Gros, 1997; Jun and Desplan, 1996; Miskiewicz et al, 1996; Xu et al, 1995) (Figure 5.1.b.). Therefore, it is possible that interactions between
the PAI subdomain region and the homeodomain may have restricted Pax7d DNA-binding ability resulting in limited target gene selection.

**Chapter Five: Discussion**

**Figure 5.1.a. Pax7d isoform structure:** The Pax7d isoform consists of a paired domain, comprising two subdomains (PAI) and (RED) that are each composed of three alpha helices (numbered). The isoform Pax7d includes the amino acid residue glutamine (Q+) between the PAI and RED subdomains but excludes amino acids glycine and leucine (GL-) within the RED subdomain. The octapeptide is located between the paired domain and homeodomain. The homeodomain contains three alpha helices (numbered) (Ziman and White, 2006).

**Figure 5.1.b. Pax7d interacting with DNA:** The Pax7d isoform includes the amino acid residue glutamine (Q+). This isoform only utilizes the (PAI) subdomain to bind to DNA influencing target gene selection. The third alpha helices (red) allow subdomain-DNA interaction. The Homeodomain (HD) is known to function in combination with the paired domain to influence target gene selection (Du et al., 2005; Chi and Epstein, 2002).
Chapter Five: Discussion

Additionally, previous studies have shown that the transactivation domain interacts with other transcription factors and the proteins of the general transcriptional machinery to activate transcription of target genes (Ziman and White, 2006; Hadchouel et al., 2003; Schafer et al., 1994). Interestingly, it is possible that the epitope tag fused to the Pax7 protein may have blocked interactions of the transactivation domain with other proteins, restricting the activation of target genes, and thus limiting target gene selection.

5.5.2. Other factors affecting Pax7d DNA-binding

Interestingly, multiple transcription factors and interacting proteins present in particular cell types have the potential to significantly affect the DNA-binding capacity of transcription factors (Chakrabarti et al., 2002; Qiu et al., 2002; Dutta et al., 2001; Glick et al., 2000; Ohneda et al., 2000; Peers et al., 1994). Results from this thesis show that the Pax7d isoform did not select target genes *Eya4, Gbx1, Kcnk2* and *Prrxl*, in vitro, whereas these genes were identified as Pax7 target genes in vivo, suggesting that several other transcription factor proteins not present in P19 cells are required to interact with Pax7d to facilitate binding to the target gene.

Additionally, previous studies have shown that the nuclear chromatin environment affects the binding of transcription factors to promoter sites (Chakrabarti et al., 2002; Spencer et al., 1999; Grunstein, 1997). It is possible that in this study the use of cell lines NIH3T3 and P19 with different chromatin structure to cells in vivo may have affected Pax7d DNA-binding ability in vitro relative to binding in vivo (Chakrabarti et al., 2002; Qiu et al., 2002; Dutta et al., 2001; Glick et al., 2000; Ohneda et al., 2000; Peers et al., 1994).

Furthermore, it is possible that all isoforms of Pax7 are required to function in combination to select target genes. Evidence on the combined function of Pax7 isoforms remains to be determined. Alternatively, the possibility exists that Pax7d only binds to *CntfR* and *Rasa3*, a result thought possible since only *CntfR* and *Rasa3* were noticeably upregulated in Pax7d transfected cells. It is not possible to comment on the ability of Pax7 to bind to other genes since no change in their gene expression levels was noted in
transfected cells. It remains possible that Pax7 bound to selected target genes to inhibit their expression \textit{in vitro} as these genes were not upregulated in P19 cells and thus inhibition of their expression after Pax7 transfection could not be assessed.

5.5.3. Utilization of chromatin immunoprecipitation for target gene selection \textit{in vivo} and \textit{in vitro}

Many studies have obtained remarkable findings that have expanded knowledge and understanding of protein-DNA interactions within mammalian cells or tissues by utilization of the chromatin immunoprecipitation technique. However, in the majority of studies, the technique is frequently applied to embryonic tissue as opposed to cell lines (Antonini et al., 2006; Jelinic et al., 2006; Matsukuma et al., 2006; Mayanil et al., 2006; Plateroti et al., 2006; Terranova et al., 2006; White and Ziman, 2006; Pedchenko et al., 2005; Peng and Chen, 2005; Magness et al., 2004; Pierreux et al., 2004; Jasinskiene et al., 1995). For example, Mayanil and colleagues (2006), applied the chromatin immunoprecipitation technique to E10.0 $\text{Pax3}^{+/+}$ mouse embryos utilizing rabbit polyclonal Pax3 antibody to investigate the regulation of a specific target gene by Pax3 during early embryonic development. Furthermore, Peng and Chen (2005), utilized photoreceptor transcription factors Crx, Otx2, Nrl and Nr2e3 to investigate binding of these proteins to promoter sites of photoreceptor genes in the retina of wild type and mutant mice. Both investigations performed PCR analysis with designed primer sets, and obtained conclusive results. These studies suggest that chromatin immunoprecipitation is better suited for \textit{in vivo} studies in embryonic or adult tissue.

A previous chromatin immunoprecipitation study showed that Pax7 target genes $\text{CntfR}$, $\text{Eya4}$, $\text{Gbx1}$, $\text{Rasa3}$, $\text{KcnK2}$ and $\text{Prrx1}$ are bound by the Pax7 protein in developing mouse embryos (White and Ziman, 2006). In this project, while both mouse NIH3T3 and P19 embryonal cells were utilized to observe Pax7 isoform DNA-binding \textit{in vitro}, no conclusive results were obtained. Therefore, the possibility exists that target genes bound by a transcription factor \textit{in vivo} may not be bound by the same genes \textit{in vitro}. The suggestion is supported by a study conducted with several \textit{Drosophila} transcription
factors known to bind to target genes \textit{in vivo} whereas these same genes were not bound \textit{in vitro} (Carr and Biggin, 1999; Walter and Biggin, 1996).

\subsection{Advantages and disadvantages of chromatin immunoprecipitation}

According to Peng and Chen (2005), chromatin immunoprecipitation technology has numerous advantages over other methodologies for studying protein-DNA interactions \textit{in vivo} and \textit{in vitro}, however there are limitations to utilization of the method.

Chromatin immunoprecipitation is a novel methodology used to investigate specific protein-DNA interactions \textit{in vivo} (Peng and Chen, 2005). The technique provides binding profiles for an array of transcription factors. However, only DNA sequences directly bound by a given transcription factor at any one time within a living cell can be identified (Weinmann and Farnham, 2002). This strengthens the possibility that the Pax7d isoform interacts with other factors to upregulate \textit{Cntfr} and \textit{Rasa3} but does not bind directly to the target genes under investigation and thus may explain why no target genes were selected by ChiP.

\subsection{SUMMARY}

\textit{Pax7}, \textit{Cntfr} and \textit{Rasa3} all perform important functional roles in nervous system and or skeletal muscle development. The direct or indirect upregulation of \textit{Cntfr} and \textit{Rasa3} by \textit{Pax7d} remains to be determined. However, from results presented in this thesis, the role of \textit{Pax7d} in neural and myogenic functions is strengthened, and adds weight to previous experiments which merely detail expression of \textit{Pax7d} in neurogenic and myogenic lineages.

The Pax7d isoform did not appear to bind directly to target genes \textit{Cntfr}, \textit{Eya4}, \textit{Gbx1}, \textit{Rasa3}, \textit{Kcnk2} and \textit{Prrx1} \textit{in vitro}. Several reasons for this have been identified; it is possible that the Pax7d isoform structure and DNA-binding ability do not allow direct interaction with the selected target genes, whereas it may bind to many other genes not identified here. However, the possibility that all Pax7 isoforms are necessary for target gene selection requires further investigation.
Other possibilities for the negative results are associated with the methodology and the use of *in vitro* cell lines. Perhaps the chromatin immunoprecipitation technique is not effective when applied to a stably transfected cell line, as previous studies have applied the technique to embryonic tissues and transiently transfected cells. DNA-binding abilities of a given transcription factor *in vitro* are significantly affected by the chromatin structure of the cell type utilized and the presence of interacting proteins. Further optimization of the method may provide results in the future.

Studies on the binding affinities and specificity of each alternate isoform are important and would provide a clearer understanding of the specific roles of each isoform within a cell and would expand the diverse functions ascribed to the *Pax7* gene.
CHAPTER SIX

CONCLUDING COMMENTS
From the results presented in this thesis it can be concluded that the *Pax7d* transcript is important for both neural and myogenic function. This is supported by research that has shown predominant *Pax7d* expression in a neurogenic lineage and that transfection of the transcript into haematopoietic stem cells induced myogenic differentiation. The *Pax7d* transcript encoding the Pax7d isoform was not shown to be bound directly to target genes *Cntfr*, *Eya4*, *Gbx1*, *Rasa3*, *Kcnk2* and *Prrx1 in vitro*. No target gene selection may be attributed to the incorrect choice of target genes, the Pax7d isoform structure and DNA-binding function, interactions between the paired domain, homeodomain and epitope tag interfering with Pax7d DNA-binding, and the ineffective application of the chromatin immunoprecipitation technique in cell lines (given that DNA-binding is affected by chromatin structure environment and the presence of interacting proteins).

Moreover, the possibility that Pax7 isoforms function in combination to bind to DNA requires further investigation. Further investigation on the DNA-binding differences between each Pax7 isoform are necessary to provide much needed answers for the overall function of the *Pax7* gene.
CHAPTER SEVEN

REFERENCES
Chapter Seven: References

Chapter 7  REFERENCES


Chapter Seven: References


Chapter Seven: References


81


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.


Chapter Seven: References


CHAPTER EIGHT

APPENDICES
Appendix 1: *Pax7* alternate transcript inserts cloned into the pcDNA3.1D/V5-His-TOPO expression vector

**Pax7a** -/+ cloned region:

```
GACTCTGGATTTCTCTCCAGCTGGTGCAGAAATGGCGGCGCTGCCCGGCGCGGTCCCCAGGATGATGAGACCCGCCGCCCGGGGCACTGAGCTGGTGGGGTCTTCATCAATGGTCGACCCCTGCCGAACCACATCCGTCACAAGATAGTTGAAATGGCCCACCATGGCATCCGGCCCTGCGTCATCTCCGTACAGCTCCGTGTTTCTCATGGTTGTC
TCCAAAGATCTCTGCGCCGATATCAGGAGACTGGGTCCATCCGGCCCGGGGCTATCGGAGGCAGCAAGCCGAGTGGCGACTCCGGATGTGGAGAAAAAGATTGAGGAGTATAAGAGAGAGAACCCCGGGATGTTCAGCTGGAAATCCGGGACCGGCTGCTGAAGGACGGTCACTGCGACCGAAGCACGGTGCCCTCAGGTTAGTGATGTCGATTAGCCGAGTGCTCAGAATCAAGTTCGGGAAGAAAGAGGACGACGAGGAAGGAGACAAGAAAGAAGATGGCGAGAAGAAAGCCAAACACAGCATCGATGGCATCCTGGGCGACAAAGGGAACCGTCTGGATGCGCTCAGATGTGGAATCAGAACCCGACCTCCCCCTGAAGCGCAAGCAGCGCCGCAGTCGGACCACGTTCACAGCCGAGCAGCTGGAGGAGCTAGAGAAGGCCTTTGAGAGGACCCACTACCCGGACATCTACACCCGGGAGGAGCTGGCACAGAGGACCAAGCTCACGGAGGCACGCGTCCAGGTCTGGTTCAGTAACCGGCGTGCCCGCTGGCGCAAGCAGGCAGGAGCTAACCAGCTGGCCGCCTTCAACCACCTTCTCCCCAAGCAGATGCCCACGCTGCCACCCTACCAGCTGCCGGACTCTACCTACCCCACCACCACCATCTCCCAAGATGGGGGCAGCACAGTACACAGGCCCCAGCCCCTCCGCCATCAACCATGCATCAGGGTGGGCTGGCTGCAGGGCTGCAGCAGCAGACACCAGCTCTGCCTACGGAGCCCGCCACAGCTTCTCCAGCTACTCTGACAGCCTCATGAACCCTGGGGCTCCCTCCAACCACATGAACCCTGTCAGCAATGGCCTGTCTCCTCAGGTCATGAGCATCCTTAGCAACCCGAGTGCCGTGCCTCCACAGCCCCAGGCCGACTTCTCCATCTCCCCGCTGCATGGAGGCCTGGACTCGGCTTCCTCCATCTCAGCCAGCTGCAGCCAACGGGCCGACTCCATCAAGCCAGGAGACAGCTTGCCCACGTCCCAGTCTTACTGCCCACCCACCTACAGCACCACTGGCTACAGTGTGGACCCTGTGGCTGGCTACCAGTACAGCCAGTATGGCCAAACTGCTGTTGATTACCTGGCCAAAAACGTGAGCCTGTCCACAGCGCCGTATGAAGCTTGGGGAACACTCCGCTGTGCTGGGACTTCTTCCTGTGGAAACGGGACAAGCTAC
```

93
Chapter Eight: Appendices

Pax7b +/- cloned region:

GACCTCTGGATPTCGTTCACCGCGCGTTCAAGAAATGGGCGGCTGCGGCAGCGCGGGATGATGGACCCCGGCCCGGGGCAGAACTACCCGCGCACCGGCTTCCCCTGGAAGTGTCCACCCCTTGGCCAAGGCCGGGTCAATCAGCTTGGTGGGGTCTTCATCAATGGTCGACCCCTGCCGAACCACATCCGTCACAAGATAGTGGAAATGGCCCACCATGGCATCCGGCCCTGCGTCATCTCCCGTCAGCTCCGTGTTTCTCATGGTTGTGTCCTCCAAGATTCTGTGCCGATATCAGGAGACTGGGTCCATCCGGCCCGGGGCTATCGGAGGCAGCAAGCCAGACAGGTGGCGACTCCGGATGTGGAGAAAAAGATTGAGGAGTATAAGAGAGAGAACCCCGGGATGTTCAGCTGGGAAATCCGGGACCGGCTGCTGAAGGACGGTCACTGCGACCGAAGCACGGTGCCCTCAGGTTTAGTGAGTTCGATTAGCCGAGTGCTCAGAATCAAGTTCGGGAAGAAAGAGGACGACGAGGAAGGAGACAAGAAAGAAGAAGATGGCGAGAAGAAAGCCAAACACAGCATCGATGGCATCCTGGGCGACAAAGGGAACCGTCTGGATAGGGCTCAGATGTGGAATCAGAACCCGACCTCCCCCTGAAGCGCAAGCAGCGCCGCAGTCGGACCACGTTCACAGCCGAGCAGCTGGAGGAGCTAGAGAAGGCCTTTGAGAGGACCCACTACCCGGACATCTACACCCGGGAGGAGCTGGCACAGAGGACCAAGCTCACGGAGGACA CGCGTCCAGGTCTGGTTCAGTAACCGGCGTGCCGCTGGCGCAAGCAGGCAGGAGCTAACCAGCTGGCCGCCTTCAACCCTCTGCCGGGAGGTTTCCCACCCACCGGCATGCCCACGCTGCCACCCTACCAGCTGCCGGACTCTACCTACCCCACCACCATCTCCC

AAGATGGGGCGAGCAAGACGACAGCAGCGCCCAGCGCCCCTCTCGCCCGCATGACCCACATCAGACATCAGCGTTGCGGCTGC

TGGGCGCGCTGCGAGCCGACAGCAGCGCCCAGCGCCCCTCTCGCCCGCATGACCCACATCAGACATCAGCGTTGCGGCTGC

AGGTCATGGAAACCCTGGGCTCCTCCCTCAAACCACATGAAACCTGCTACAGATGGGTGGCTGCTGCA

TGCGGTCCATTGACACCGGAGGCGGCGGTCCTCAGAGCCCCCCAGGCGGCTCTTTCTCGCCCGCATGACCCACATCAGACATCAGCGTTGCGGCTGC

TGGGCGCGCTGCGAGCCGACAGCAGCGCCCAGCGCCCCTCTCGCCCGCATGACCCACATCAGACATCAGCGTTGCGGCTGC

GACTCATGGGCTCAGTACATGCTGCTGCTGCTGCTGCTGCCTTCTCGCCCGCATGACCCACATCAGACATCAGCGTTGCGGCTGC

GCCTAC
Chapter Eight: Appendices

Pax7c -/- cloned region:

```
GACTCTGGATTGTCTCACAAGCTTGCTGCAAGAAAAGCTCCGGGGGCGCGCTCCACAGGATGATGAGACCCCGGCCCGGGCAGAACTACCCGCGCACCGGCTTCCCCCTGGAAGTGTCCACCCCTCTTGGCCAAGGCCGGGTCAATCAGCTTGGTGGGGTCTTCATCAATGGTCGACCCCTGCCGAACCACATCCGTCACAAGATAGGGGAAATGGCCCACCATGGCATCCGGCCCTGCGTCATCTCCCGTCAGCTCCGTGTTTCTCATGGTTGTGTC
```

95
Pax7d +/- cloned region:

GACTCTGGATTCGTCCTCCAGCGTGTGCAGAAATGGCGGCGCTGCCCGGCGCGGTCCCCAGGATGATGAGACCCTGCCCGGGCGAGAACTACCCGCGCACCGGCTTCCCCCTGGAAGTGTCCACCCCTCTTGGCCAAGGCCGGGTCAATCAGCTTGGTGGGGTCTTCATCAATGGTCGACCCCTGCCGAACCACATCCGTCACAAGATAGTGAAATGGCCCACCATGGCATCCGGCCCTGCGTCATCTCCCGTCAGCTCCGTGTTTCTCATGGTTGTGTCCTGACATTCCTGTGCAGATATCAGGAGACTGGGTCCATCCGGCCCGGGGCTATCGGAGGCAGCAAGCACAAGGTGGCGACTCCGGATGTGGAGAAAAAGATTGAGGAGTATAAGAGAGAGAACCCCGGGATGTTCAGCTGGGAAATCCGGGACCGGCTGCTGAAGGACGGTCACTGCGACCGAAGCACGGTGCCCTCAGTGAGTTGATTAGCCGAGTGCTCAGAATCAAGTTCGGGAAGAAAGAGGACGACGAGGAAGGAGACAAGAAAGAAGATGGCGAGAAGAAAGCCAAACACAGCATCGATGGCATCCTGGGCGACAAAGGGAACCGTCTGGATGAGGGCTCAGATGTGGAATCAGAACCCGACCTCCCCCTGAAGCGCAAGCAGCGCCGCAGTCGGACCACGTTCACAAGCAGCTGGAGGAGCTAGAGAAGGCCTTTGAGAGGACCCACTACCCGGACATCTACACCCGGGAGGAGCTGGCACAGAGGACCAAGCTCACGGAGGCACGCGTCCAGGTCTGGTTCAGTAACCGGCGTGCCCGCTGGCGCAAGCAGGCAGGAGCTAACCAGCTGGCCGCCTTCAACCACCTTCTGCCGGGAGGTTTCCCACCCACCAGCATGCCCACGCTGCCACCCTACCAGCTGCCGGACTCTACCTACCCCACCACCACCATCTCCCAAGATGGGGCAGCACAGTACACAGGCCCCAGCCCCTCCGCCATCAACCATGCATCAGGGTGGGCTGGCTGCGGCAGCTGCAGCAGCAGACACCAGCTCTGCCTACGGAGCCCGCCACAGCTTCTCCAGCTACTCTGACAGCTTCACTGGGGCTCCCTCCAACCACATGAACCCTGTCAGCAATGGCCTGTCTCCTCAGGTCATGAGCATCCTTAGCAACCCGAGTGCCGTGCCTCCACAGCCCCAGGCCGACTTCTCCATCTCCCCGCTGCATGGAGGCCTGGACTCGGCTTCCTCCATCTCAGCCAGCTGCAGCCAACGGGCCGACTCCATCAAGCCAGGAGACAGCTTGCCCACGTCCCAGTCTTACTGCCCACCCACCTACAGCACCACTGGCTACAGTGTGGACCCTGTGGCTGCCTACCAGTACAGCCAGTATGGCCAAACTGCTGTTGATTACCTGGCCAAAAACGTGAGCCTGTCCACACGCGCCGTATGAAGCTTGGGGAACACTCCGCTGTGCTGGGACTTCTTCCTGTGGAAACGGGACAAGCCTAC

96