Functional Analysis of the HOX11 Target
Genes *ALDH1A1* and *FHL1*

Kim Rice
Bachelor of Science (Hons)
This thesis is presented for the degree of

DOCTOR OF PHILOSOPHY

of
SCHOOL OF VETERINARY AND BIOMEDICAL SCIENCES
MURDOCH UNIVERSITY
WESTERN AUSTRALIA
2004
I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution

Kim L. Rice
ABSTRACT

HOX11 is a developmental regulator that plays a crucial role in the normal development of the spleen and is also aberrantly activated by the t(10;14)(q24;q11) and variant t(7;10)(q35;q24) translocations in a subset of T-cell acute lymphoblastic leukaemias (T-ALLs). The recent finding that HOX11 is deregulated in up to 40% of childhood T-ALLs when abnormalities not detected by cytogenetics are included, suggests that the over-expression of HOX11 and subsequent deregulation of downstream target genes are critical events in the progression of this tumour type. To date, three candidate HOX11 target genes have been reported, two of which are Aldehyde Dehydrogenase 1a1 (ALDH1A1) and Four and a Half LIM domain Protein 1 (FHL1). This investigation focused on two aspects of HOX11 function, namely its roles as a transcriptional regulator and as a nuclear oncoprotein capable of inducing neoplastic transformation. More specifically, we sought to further understand the role of HOX11 in tumorigenesis by 1) Confirming target gene status of ALDH1A1 and FHL1 by assessing whether their proximal promoter regions are transcriptionally regulated by HOX11, 2) Investigating the regulatory elements/transcriptional complexes involved in the response of ALDH1A1 to HOX11 in both a T-cell and an erythroid cell line in order to gain an insight into the mechanism(s) responsible for mediating a HOX11 activity and 3) Assessing the ability of ALDH1A1 and FHL1 to perturb normal patterns of haematopoiesis, on the basis that the transforming capabilities of HOX11 are thought to derive from its ability to affect haematopoietic cell differentiation.

To confirm ALDH1A1 and FHL1 as target genes, they were both characterised in terms of the ability of their proximal promoters to be transcriptionally regulated by HOX11 using luciferase reporter assays. Significant repression of the proximal promoters of ALDH1A1 and FHL1 by HOX11 was observed in PER-117 T-cells which provided further evidence for their status as target genes. In the case of ALDH1A1, a CCAAT box (-74/-70bp) was identified as the primary cis-regulatory element involved in ALDH1A1 transcription and repression by HOX11 appeared to occur, either directly or indirectly, via interactions at the CCAAT box. Electromobility shift assays (EMSAs) revealed the disruption of a specific complex at this site by HOX11, which also altered the formation of complexes at a non-canonical TATA box (a GATA box at -34/-29bp). Significantly,
HOX11 was shown to have the potential to interact with TFIIB, a member of the basal transcriptional complex. This, together with the presence of a TFIIB responsive element immediately 5’ of the GATA box, suggested that HOX11 may repress transcription by interfering with members of a preinitiation complex on the ALDH1A1 promoter. The transcriptional repression by HOX11 demonstrated in T-cells was dependent on DNA binding helix 3 of the homeodomain, suggesting that repression may require DNA binding. Alternatively, this region may be required for stable protein-protein interactions. In support of this, the in vitro association of HOX11 with TFIIB was disrupted upon deletion of helix 3, and the HOX11ΔH3 mutant switched from a transcriptional repressor to a potent activator of transcription. Together, this data supports a model whereby HOX11 represses transcription by interfering with activation complexes at the CCAAT box and at the GATA box possibly via protein-protein interactions involving the homeodomain helix 3, whereas deletion of the region disables repressor-specific interactions, resulting in potent activation by HOX11.

Luciferase reporter gene assays investigating the response of nested deletions of the ALDH1A1 promoter to HOX11 in the HEL900 erythroleukaemic cell line, also identified the CCAAT box (-74/-70bp) as the primary cis-regulatory element involved in ALDH1A1 transcription. However, in stark contrast to its effect in T-cells, HOX11 was shown to activate transcription in the HEL cell line, both from the empty pGL3Basic luciferase reporter vector and from the ALDH1A1 promoter, in a manner independent of the homeodomain DNA binding helix 3. HOX11 thus appears to be a dichotomous regulator, capable of both transcriptional activation and repression depending on the circumstances. The mechanisms underlying these two functions appear to be distinct, with repression but not activation requiring the presence of homeodomain helix 3.

ALDH1A1 encodes an enzyme involved in the irreversible conversion of retinaldehyde to the biologically active metabolite, retinoic acid (RA) and appears to be physiologically regulated by Hox11 in the developing spleen. Since RA is a potent modulator of cellular differentiation, proliferation and apoptosis, the dysregulation of RA synthesis is likely to have severe consequences for the cell and may constitute a mechanism whereby overexpression of HOX11 predisposes T-cells to malignant
transformation. *FHL1* also appears to have potential relevance to tumorigenesis, given that it encodes protein isoforms with suspected roles in transcriptional regulation. As a starting point to investigate a possible link between these HOX11 target genes and leukaemogenesis, the effect of overexpressing *ALDH1A1* and *FHL1* on murine haematopoiesis was assessed following reconstitution of lethally irradiated mice with retrovirally-transduced primary murine bone marrow cells. The enforced expression of *ALDH1A1* in bone marrow was associated with a marked increase in myelopoiesis and a decrease in B and T-lymphopoiesis. By contrast, overexpression of *FHL1* was not associated with perturbations in myelopoiesis or lymphopoiesis, although a slight increase in erythropoiesis was observed in the bone marrow. While further work is required to clarify the possible oncogenic roles of both of these HOX11 target genes, these findings have served to identify *ALDH1A1* in particular, as a gene which could potentially be involved in HOX11-mediated tumorigenesis.
# Table of Contents

**Title page**  
**Declaration**  
**Abstract**  
**Table of Contents**  
**Acknowledgements**  
**Publications and Abstracts**  
**Abbreviations**  
**List of Figures and Tables**

### Chapter 1: General Introduction

1. **T-CELL ACUTE LYMPHOBLASTIC LEUKAEMIA (ALL)**  
   1.1 Biology and Classification of Leukaemias  
   1.2 Molecular Genetics of Childhood T-ALL

2. **HOMEobox GENES**  
   1.2.1 The Role of Homeobox Genes in Development  
   1.2.2 Hox Proteins as Transcription Factors – Specificity of Homeotic Gene Function  
   1.2.3 Homeobox Genes – Roles in Haematopoiesis and Leukaemogenesis

3. **HOX11**  
   1.3.1 The Normal Role of HOX11  
   1.3.2 The Abnormal Role of HOX11  
   1.3.3 HOX11 Target Genes

4. **HOX11 Target Genes**  
   1.4.1 Identification of Downstream Target Genes of HOX11  
   1.4.2 \textit{ALDH1A1} as a Target Gene of HOX11
1.4.2.1 The Role of ALDH1A1 in Normal Development 38
1.4.2.1.1 Retinoid Metabolism 38
1.4.2.1.2 Aldehyde Dehydrogenases (ALDHs) 41
1.4.2.1.3 ALDH1A1 - A Crucial Regulator of Retinoic Acid Biosynthesis During Embryogenesis and Adulthood 44
1.4.2.1.4 The Role of Retinoid Signalling Haematopoiesis 46
1.4.2.2 The Abnormal Role of ALDH1A1 49
1.4.2.2.1 Relevance of ALDH1A1 as a Target Gene of HOX11 49

1.4.3 FHL1 as a Target Gene of HOX11 51
1.4.3.1 LIM Domain Proteins 51
1.4.3.2 FHL1 52
1.4.3.2.1 Family Members and Related Isoforms 52
1.4.3.2.2 Expression and Function 54
1.4.3.2.3 Relevance of FHL1 as a Target Gene of HOX11 55

1.5 THESIS AIMS 56

Chapter 2: Assessment of ALDH1A1 as a Transcriptional Target of HOX11 57
2.1 INTRODUCTION 57
2.2 METHODS 61
2.2.1 pGL3Basic-ALDH1A1 Promoter Vector Construction 61
2.2.1.1 Preparation of pGL3Basic Luciferase Reporter Vector 61
2.2.1.2 Agarose Gel Electrophoresis 61
2.2.1.3 Agarose Gel Purification of Digested DNA Fragments 62
2.2.1.4 Generation of ALDH1A1 Promoter Fragments 62
2.2.1.5 Restriction Digestion and Purification of Purified ALDH1A1 Promoter PCR Products 65
2.2.1.6 Ligation of Promoter Fragments into pGL3Basic Luciferase Reporter Vector 66
2.2.1.7 Preparation of Competent Cells 66
2.2.1.8 Bacterial Transformations 67
2.2.1.9 Small Scale Plasmid Purification 67
2.2.1.10 Screening Recombinant Colonies for the Presence of Promoter Inserts 68
2.2.1.11 DNA Sequencing 69
2.2.1.12 Large Scale Purification of Plasmid DNA 69
2.2.1.13 Standardisation of Plasmid DNA for Transient Transfection Experiments

2.2.1.14 DNA Analysis Using UV Spectrophotometry

2.2.1.15 DNA Analysis by Restriction Digestion

2.2.2 Cell Culture

2.2.2.1 HEL-900 Cell Lines

2.2.2.2 PER-117 Cell Lines

2.2.2.3 The NIH 3T3 Cell Line

2.2.2.4 Cryopreservation of Mammalian Cell Lines

2.2.3 Luciferase Assays

2.2.3.1 Transient Transfection of HEL and PER-117 Cell Lines

2.2.3.2 Transient Transfection of NIH 3T3 Cells

2.2.3.3 Western Blotting

2.2.3.3.1 Preparation of Nuclear Cell Extracts

2.2.3.3.2 Determination of Nuclear Protein Extract Concentration

2.2.3.3.3 Analysis of HOX11 Expression by Western Blotting

2.2.3.3.4 Detection of HOX11 Protein

2.2.3.4 Luciferase and β-Galactosidase Reporter Gene Assays

2.2.3.5 Data Processing

2.2.3.6 Statistical Manipulation of Data

2.2.4 DNase1 Footprinting

2.2.4.1 PCR Labelling of the ALDH1A1 Proximal Promoter

2.2.4.1.1 Radiolabelling the Primer Using 32P-gATP

2.2.4.1.2 Generation of Singly End-Labelled ALDH1A1P Fragments Using PCR

2.2.4.1.3 DNA Quantitation Using Activity

2.2.4.2 Footprinting Reactions

2.2.4.3 Phenol-Chloroform Extractions

2.2.4.4 Preparation of Sequencing Reactions

2.2.4.5 Denaturing Polyacrylamide Gel Electrophoresis

2.2.4.6 Gel Drying and Autoradiography

2.2.5 Electrophoretic Mobility Shift Assays

2.2.5.1 Oligonucleotide Labelling and Annealing

2.2.5.2 EMSA Reactions

2.2.5.2 Native Polyacrylamide Gel Electrophoresis
Chapter 3: *FHL1* as a Target Gene of HOX11: Identification of the *FHL1* Gene Promoter and Assessment of its Transcriptional Regulation by HOX11

3.1 INTRODUCTION

3.2 METHODS

3.2.1 5’ RACE

3.2.1.1 Cell Lines

3.2.1.2 RNA Isolation

3.2.1.3 Quantitative and Qualitative RNA Analysis

3.2.1.4 5’ RACE – System Version 2.0

3.2.1.4.1 SUPERSCRIPT™ II RT

3.2.1.4.2 THERMOSCRIPT™ RT

3.2.1.4.3 Omniscrypt™ RT

3.2.1.5 5’ RACE - GeneRacer™ Protocol

3.2.1.5.1 De-phosphorylation of ALL-SIL and HeLa Total RNA
3.2.1.5.2 Removing the 5' Cap Structure from Full-length mRNA 152
3.2.1.5.3 Ligation of the RNA Oligonucleotide Adapter to Decapped mRNA 152
3.2.1.5.4 Reverse Transcribing mRNA 153
3.2.1.5.5 PCR Amplification 153

3.2.1.6 Cloning of 5' RACE Products 154

3.2.2 Luciferase Reporter Gene Assays 156

3.2.2.1 Preparation of FHL1 Promoter Luciferase Reporter Vectors 156

3.2.2.1.1 PCR Amplification of FHL1 Promoter Insert DNA 156
3.2.2.1.2 Restriction Digestion of FHL1 Promoter Insert DNA 156
3.2.2.1.3 Preparation of pGL3Basic Luciferase Reporter Vector 157
3.2.2.1.4 FHL1 Reporter Construct Preparation 157

3.2.2.2 Luciferase and β-Galactosidase Reporter Gene Assay 158

3.3 RESULTS 158

3.3.1 ALL-SIL RNA Analysis 158

3.3.2 Identification of the Transcriptional Start Sites of the FHL1 Gene 160

3.3.2.1 Analysis of the Human FHL1 Transcripts by 5'RACE (Version 2.0) 162
3.3.2.2 Analysis of the Human FHL1 mRNA Transcripts by 5'RLM-RACE 169

3.3.3 Bioinformatic Analysis of the Human FHL1 Proximal Promoter 169

3.3.4 Comparison of Human-Murine FHL1 Promoter Sequences 173

3.3.5 Identification of a HOX11 Responsive Element in the FHL1 Promoter 175

3.4 DISCUSSION 180

Chapter 4: Protein Binding Partners of HOX11 190

4.1 INTRODUCTION 190

4.2 METHODS 192

4.2.1 GST-Pulldown Assays 192

4.2.1.1 pCIneo CTF1 and TFIIB Vector Construction 192

4.2.1.1.1 Preparation of pCIneo' Mammalian Expression Vector 193
4.2.1.1.2 Generation of ALL-SIL and HEL and cDNA 193
4.2.1.1.3 PCR Amplification of CTF1 and TFIIB Insert cDNA 194
4.2.1.1.4 Restriction Endonuclease Digestion of CTF1 and TFIIB Insert cDNA
4.2.1.1.5 Cloning of CTF1 and TFIIB cDNA
4.2.1.1.6 Purification of Plasmid DNA
4.2.1.2 Production of 35S-Radiolabelled Proteins
4.2.1.2.1 In vitro Transcription/Translation
4.2.1.2.2 SDS-PAGE Analysis of 35S-Methionine Labelled In vitro-Translated Proteins
4.2.1.3 Production of Recombinant GST-HOX11 and GST-HOX11ΔH3 Fusion Protein
4.2.1.3.1 Preparation of Glutathione Sepharose 4B
4.2.1.3.2 Expression and Purification of Recombinant GST, GSTHOX11 and GSTHOX11ΔH3
4.2.1.3.3 Quantitative and Qualitative Analysis of GST Fusion Proteins
4.2.1.4 Verification of the Biological Activity of Recombinant HOX11

4.3 RESULTS
4.3.1 Production of Recombinant GST Fusion Proteins
4.3.2 Verification of the Biological Activity of Recombinant HOX11 Fusion Proteins
4.3.3 In vitro Characterisation of MEIS1, MEIS2A, HOX11 and TFIIB Binding to HOX11

4.4 DISCUSSION

Chapter 5: Effects Of Overexpression Of ALDH1A1 and FHL1 on Murine Haematopoiesis
5.1 INTRODUCTION
5.2 METHODS
5.2.1 Generation of Recombinant Retrovirus
5.2.1.1 MigR1-ALDH1A1/FHL1 Vector Construction
5.2.1.1.1 Preparation of MigR1 Retroviral Expression Vector
5.2.1.1.2 PCR Amplification of ALDH1A1 and FHL1 Insert cDNA
5.2.1.1.3 Restriction Endonuclease Digestion of ALDH1A1 and FHL1 Insert cDNA
5.2.1.1.4 Cloning of ALDH1A1 and FHL1 cDNA into the MigR1 Retroviral Vector
5.2.1.2 Phoenix™ Retrovirus Production
5.2.1.2.1 Growth and Passage of Phoenix™-Eco Cells
5.2.1.2.2 Transfection of Phoenix™-Eco Cells With MigR1 Retroviral Constructs
5.2.2 Bone Marrow Reconstitution

5.2.2.1 Haematopoietic Cell Transduction by Spinofection

5.2.2.2 Reconstitution of Mice With Transduced Bone Marrow Cells

5.2.3 FACS Analysis

5.2.3.1 Tissue Processing

5.2.3.2 Immunophenostaining of Transplant Recipient Cells

5.2.3.3 Flowcytometric and FlowJo Analysis

5.3 RESULTS

5.3.1 Production of Recombinant Retrovirus

5.3.2 Stable Transduction of Bone Marrow Cells with Retrovirus and Effects on Different Haematopoietic Lineages

5.3.3 Effect of ALDH1A1 and FHL1 Expression on Haematopoietic Lineages

5.4 DISCUSSION

Chapter 7: REFERENCES

Appendix 1

Appendix 2:

Appendix 3:
ACKNOWLEDGMENTS

Firstly I would like to thank my Supervisor, Wayne Greene. Without your expertise, guidance and support this thesis would not have been possible. Over the last 5 years under your supervision, I have come to truly appreciate your talents as a molecular biologist, manager, teacher and comedian (!). I have always respected your open-minded and logical approach to analysing every situation - in fields ranging from DNA to the existence of God (!). Thankyou for your encouragement during result droughts - for always seeing a positive side to any experiment - for helping me to make the transition from pessimist to realist!

I would also like to thank Dr Ursula Kees, Dr Paul Watt and Dr Rachael Brake for your support in the form of technical assistance, advice and the use of facilities at the TVW Institute of Child Health Research, Perth, Western Australia.

Thanks to Dr David Izon (I can still hear your voice “Kimmmmiiiiieee – look at meeee” in my ears to this day) for all your help with the bone marrow reconstitution experiments and late night FACs sessions.

Many thanks to Ross Taplin for all your technical assistance with the statistical component of this thesis.

I would also like to thank a number of amazing people at Murdoch University, without whom this work would not have been possible. Special thanks to David Lines, Peter Fallon, Russ Hobbs, Pat Marshall, Luba Sanderson and Ken Chong for putting up with late orders, unlocking our office ten times a day and generally making the Murdoch Vet School a cheery place to be in! Thanks to the going-out crew, Michael, Nat, Caroyln, Susannah, Ryan, Meredith and Nico – thesis writing was always more bearable at the thought of a night out with you guys. It has been a privilege to share 3am Fast Eddie’s sessions and crazy taxi drivers with you all. Special thanks to Meredith and Ryan. To Meredith, for your loyal friendship, for understanding why talking shop at lunch is unhealthy, for the all-night dancing, the pumpkin soup – basically for all the big things and all the small ones that would require a thesis in its own right. To Ryan, for being a
fantastic friend as well as an electrician and mechanic all in one! Many thanks also for the Clustal tutorial - stop glaring - HEE! I shall miss you both terribly. Also thanks to Michael, Ryan and Meredith for bringing the mystery into my life!

To Marion, firstly for being such a caring, wonderful friend and later, a caring, wonderful boss! Thankyou so much for your support and understanding this year.

Thanks to Bianca, Christina, Leanne, Heath and Kate for understanding that weekends do not exist, but always making the time to catch up.

Thankyou to Moira and Will, my surrogate Parents, for your wonderful friendship…and for trusting me with your house, despite my fascination with wood staining, during the writing-up period.

A heartfelt thankyou to my crazy honours students, Brooke, Musfira and Yovina (and I know you would say the same for me!) for your hugs and chocolate.

Lastly and most importantly, thankyou to my family. To my sisters, Sarah and Jade, and Rob (officially christened ‘Wob’), who I love dearly. Your phone calls and the thought of being with you both in Melbourne soon sustained me through many a printer jam and broken photocopier. To Mum and Dad who have lived and breathed this thesis – to the point of speaking of it as a fourth child. I can hardly find words to express what you both mean to me and how much your love and support has contributed to this work. Thankyou.
PUBLICATIONS

The following paper has been published in international peer-reviewed journal during the course of this PhD candidature:


ABSTRACTS

The following abstract has been presented at a conference proceeding during the course of this PhD candidature:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Absorbance</td>
</tr>
<tr>
<td>aa</td>
<td>Amino Acid</td>
</tr>
<tr>
<td>Abd</td>
<td>Abdominal</td>
</tr>
<tr>
<td>ADH</td>
<td>Alcohol Dehydrogenase</td>
</tr>
<tr>
<td>AF-2</td>
<td>Activation Function Domain-2</td>
</tr>
<tr>
<td>ALDH</td>
<td>Aldehyde Dehydrogenase</td>
</tr>
<tr>
<td>Aldh1a1</td>
<td>Aldehyde Dehydrogenase 1a1</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute Lymphoblastic Leukaemia</td>
</tr>
<tr>
<td>AML</td>
<td>Acute Myeloid Leukaemia</td>
</tr>
<tr>
<td>Amp</td>
<td>Ampicillin Resistance Gene</td>
</tr>
<tr>
<td>Antp</td>
<td>Antennapedia</td>
</tr>
<tr>
<td>AP</td>
<td>Anterior-Posterior</td>
</tr>
<tr>
<td>APC</td>
<td>Allophycocyanin</td>
</tr>
<tr>
<td>APL</td>
<td>Acute Promyelocytic Leukaemia</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium Persulfate</td>
</tr>
<tr>
<td>ARC</td>
<td>Animal Resource Centre</td>
</tr>
<tr>
<td>AS</td>
<td>Antisense</td>
</tr>
<tr>
<td>ATRA</td>
<td>All-trans Retinoic Acid</td>
</tr>
<tr>
<td>B-ALL</td>
<td>B-cell Acute Lymphoblastic Leukaemia</td>
</tr>
<tr>
<td>bap</td>
<td>Bagpipe</td>
</tr>
<tr>
<td>BD</td>
<td>Becton Dickinson</td>
</tr>
<tr>
<td>bHLH</td>
<td>Basic Helix-Loop-Helix</td>
</tr>
<tr>
<td>Bio</td>
<td>Biotin</td>
</tr>
<tr>
<td>bp</td>
<td>Base Pairs</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>Bx</td>
<td>Bithorax</td>
</tr>
<tr>
<td>CAT</td>
<td>Chloramphenicol Acetyl Transferase</td>
</tr>
<tr>
<td>CD</td>
<td>Cluster of Differentiation</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CDP</td>
<td>CCAAT-Displacement Protein</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin Dependent Kinase</td>
</tr>
<tr>
<td>C/EBPβ</td>
<td>CCAAT/Enhancer-Binding Protein β</td>
</tr>
<tr>
<td>CFU-GM</td>
<td>Colony Forming Unit-Granulocyte Macrophage</td>
</tr>
<tr>
<td>cGy</td>
<td>Centigrays</td>
</tr>
<tr>
<td>chIP</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>CIP</td>
<td>Calf Intestinal Phosphatase</td>
</tr>
<tr>
<td>CKI</td>
<td>Cyclin Kinase Inhibitor</td>
</tr>
<tr>
<td>CLCR</td>
<td>Children’s Leukaemia and Cancer Research</td>
</tr>
<tr>
<td>CLP</td>
<td>Common Lymphoid Progenitor</td>
</tr>
<tr>
<td>cm</td>
<td>Centimetre</td>
</tr>
<tr>
<td>CMP</td>
<td>Common Myeloid Progenitor</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>Co-IP</td>
<td>Co-Immunoprecipitation</td>
</tr>
<tr>
<td>COOH</td>
<td>Carboxyl Terminus</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>cpm</td>
<td>Counts Per Minute</td>
</tr>
<tr>
<td>cps</td>
<td>Counts Per Second</td>
</tr>
<tr>
<td>CRBP</td>
<td>Cellular Retinol Binding Protein</td>
</tr>
<tr>
<td>CRAB</td>
<td>Cellular Retinoic Acid Binding Protein</td>
</tr>
<tr>
<td>CTF1</td>
<td>CCAAT-box-binding Transcription Factor 1</td>
</tr>
<tr>
<td>CYP450</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA Binding Domain</td>
</tr>
<tr>
<td>DC</td>
<td>Dendritic Cell</td>
</tr>
<tr>
<td>dE</td>
<td>Embryonic Day</td>
</tr>
<tr>
<td>Dfd</td>
<td>Deformed</td>
</tr>
<tr>
<td>dH₂O</td>
<td>Deionised Water</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's Modified Medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl Sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleoside Triphosphate</td>
</tr>
<tr>
<td>ddNTP</td>
<td>Dideoxynucleotide Triphosphate</td>
</tr>
<tr>
<td>dpc</td>
<td>Days Postcoitum</td>
</tr>
<tr>
<td>DPE</td>
<td>Downstream Promoter Element</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EB</td>
<td>Embryoid Bodies</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediamine Tetra Acetic Acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>Ethylene Glycol-bis Tetra Acetic Acid</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electrophoretic Mobility Shift Assay</td>
</tr>
<tr>
<td>en</td>
<td>Engrailed</td>
</tr>
<tr>
<td>EPCD</td>
<td>Epicardially Derived Cell</td>
</tr>
<tr>
<td>ERα</td>
<td>Estrogen Receptor α</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed Sequence Tag</td>
</tr>
<tr>
<td>eve</td>
<td>Even-Skipped</td>
</tr>
<tr>
<td>Exd</td>
<td>Extradenticle</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence Activated Cell Sorting</td>
</tr>
<tr>
<td>FCS</td>
<td>Foetal Calf Serum</td>
</tr>
<tr>
<td>Fhl1</td>
<td>Four and a Half LIM Domain Protein 1</td>
</tr>
<tr>
<td>FRET</td>
<td>Fluorescence Resonance Energy Transfer</td>
</tr>
<tr>
<td>FSC</td>
<td>Forward Scatter</td>
</tr>
<tr>
<td>FTOC</td>
<td>Foetal Thymic Organ Culture</td>
</tr>
<tr>
<td>g</td>
<td>Gram</td>
</tr>
<tr>
<td>gag</td>
<td>Group Antigen</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescence Protein</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>Granulocyte-Macrophage Colony Stimulating Factor</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>GTF</td>
<td>General Transcription Factor</td>
</tr>
<tr>
<td>h</td>
<td>Hour</td>
</tr>
<tr>
<td>HAT</td>
<td>Histone Acetyltransferase</td>
</tr>
<tr>
<td>HBS</td>
<td>Hepes Buffered Saline</td>
</tr>
<tr>
<td>HCA</td>
<td>Human Cardiac Alpha Gene</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone Deacetylase</td>
</tr>
<tr>
<td>Hdg-1</td>
<td>HOX11 Dependent Gene-1</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-2-Hydroxyethylpiperazine-N’-2-Ethanesulfonic Acid</td>
</tr>
<tr>
<td>HOM</td>
<td>Homeotic</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>HOM-C</td>
<td>Homeotic Complex</td>
</tr>
<tr>
<td>HOX</td>
<td>Homeobox</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish Peroxidase</td>
</tr>
<tr>
<td>HSC</td>
<td>Haematopoietic Stem Cell</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>Inr</td>
<td>Initiator</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-d-Thiogalactopyranoside</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal Ribosome Entry Site</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>L</td>
<td>Litre</td>
</tr>
<tr>
<td>lab</td>
<td>Labial</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>LBD</td>
<td>Ligand Binding Domain</td>
</tr>
<tr>
<td>lbe</td>
<td>Ladybird early</td>
</tr>
<tr>
<td>lbl</td>
<td>Ladybird late</td>
</tr>
<tr>
<td>LIM</td>
<td>Lin-11, Isl-1 and Mec-3</td>
</tr>
<tr>
<td>LMO</td>
<td>LIM-Only</td>
</tr>
<tr>
<td>LT-HSC</td>
<td>Long-Term Haematopoietic Stem Cells</td>
</tr>
<tr>
<td>LTR</td>
<td>Long Terminal Repeat</td>
</tr>
<tr>
<td>lym</td>
<td>Lymphoid Progenitor</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>M/E</td>
<td>Myeloerythroid Progenitor</td>
</tr>
<tr>
<td>MED-1</td>
<td>Multiple Start Site Element Downstream-1 Element</td>
</tr>
<tr>
<td>MEIS</td>
<td>Myeloid Ectopic Integration Site</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>min</td>
<td>Minute</td>
</tr>
<tr>
<td>MIS</td>
<td>Mullerian Inhibiting Substance</td>
</tr>
<tr>
<td>mL</td>
<td>millilitre</td>
</tr>
<tr>
<td>MLL</td>
<td>Mixed Lineage Leukaemia</td>
</tr>
<tr>
<td>MLP</td>
<td>Multilineage Progenitor</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-[N-morpholino] Propanesulfonic Acid</td>
</tr>
<tr>
<td>MPD</td>
<td>Myeloproliferative Disease</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MSCV</td>
<td>Murine Stem Cell Virus</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular Weight</td>
</tr>
<tr>
<td>MWM</td>
<td>Molecular Weight Marker</td>
</tr>
<tr>
<td>MZF-1</td>
<td>Myeloid Zinc Finger-1</td>
</tr>
<tr>
<td>NeoR</td>
<td>Nuclear receptor CoRepressor</td>
</tr>
<tr>
<td>NES</td>
<td>Nuclear Export Sequence</td>
</tr>
<tr>
<td>ng</td>
<td>Nanogram</td>
</tr>
<tr>
<td>NGFR</td>
<td>Nerve Growth Factor Receptor</td>
</tr>
<tr>
<td>NH₂</td>
<td>Amino Terminus</td>
</tr>
<tr>
<td>NK</td>
<td>Natural Killer Cell</td>
</tr>
<tr>
<td>NKL</td>
<td>NK-Like</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear Localisation Signal</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometre</td>
</tr>
<tr>
<td>nmol</td>
<td>Nanmoles</td>
</tr>
</tbody>
</table>
nM  Nanomolar
NMR  Nuclear Magnetic Resonance
NP-40  Nonident-40
NS1  Non-Specific Competitor 1
NS2  Non-Specific Competitor 2
NTP  Nucleoside Triphosphate
OD  Optical Density
ONPG  O-N-Galactopyranoside
pb  Proboscipedia
PBS  Phosphate Buffered Saline
Phx1  Pre-B-cell Leukemia Transcription Factor 1
PCR  Polymerase Chain Reaction
PE  Phycoerythrin
PI  Propidium Iodide
PIM  PBX-Interaction Motif
PLZF  Promyelocytic Leukaemia Zinc Finger
PML  Promyelocytic Leukemia
PMSF  Phenylmethylsulfonyl Fluoride
PMT  Photomultiplier Tube
PNK  Polynucleotide Kinase
PolydIdC  Polydeoxycytidylic Acid
PP1C  Protein Phosphatase 1 Catalytic Subunit
PP2AC  Protein Phosphatase 2A Catalytic Subunit
PTC-100  Programmable Thermal Controller-100
PTP  Protein Tyrosine Phosphatase
RA  Retinoic Acid
RAR  Retinoic Acid Receptor
RARE  Retinoic Acid Response Element
RACE  Rapid Amplification of cDNA Ends
RALDH2  Retinaldehyde Dehydrogenase 2
RARE  Retinoic Acid Response Element
RBC  Red Blood Cell
RBP  Retinol Binding Proteins
RDA  Representational Difference Analysis
RLM-RACE  RNA Ligase-Mediated Rapid Amplification of cDNA Ends
RLU  Relative Light Unit
RNA  Ribonucleic Acid
RPH  Royal Perth Hospital
RT  Reverse Transcription
RT-PCR  Reverse Transcriptase-Polymerase Chain Reaction
RXR  Retinoid X Receptor
S  Sense
Sal  Spalt
SCF  Stem Cell Factor
SCL  Stem Cell Leukaemia
SCID  Severe Combined Immunodeficiency
SDH  Short-chain Dehydrogenase Reductase
SDS  Sodium Dodecyl Sulfate
SDS-PAGE  SDS-Polyacrylamide Gel Electrophoresis
sec  Second
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIL</td>
<td>SCL Interupting Locus</td>
</tr>
<tr>
<td>SLIM1</td>
<td>Striated Muscle LIM protein 1</td>
</tr>
<tr>
<td>SLIMMER</td>
<td>SLIM1 with Extra Regions</td>
</tr>
<tr>
<td>slou</td>
<td>Slouch</td>
</tr>
<tr>
<td>SMRT</td>
<td>Silencing Mediator of Retinoid and Thyroid Hormone Receptor</td>
</tr>
<tr>
<td>SRF</td>
<td>Serum Response Factor</td>
</tr>
<tr>
<td>SRY</td>
<td>Sex-Determining Region Y Gene</td>
</tr>
<tr>
<td>SSS</td>
<td>Single Start Site</td>
</tr>
<tr>
<td>SSC</td>
<td>Side Scatter</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris Acetate Ethylenediamine Tetra Acetic Acid</td>
</tr>
<tr>
<td>TALE</td>
<td>Three-Amino-Acid Loop Extension</td>
</tr>
<tr>
<td>T-ALL</td>
<td>T-Cell Acute Lymphoblastic Leukaemia</td>
</tr>
<tr>
<td>TAP</td>
<td>Tobacco Acid Pyrophosphatase</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris Borate Ethylenediamine Tetra Acetic Acid</td>
</tr>
<tr>
<td>TCR</td>
<td>T-Cell Receptor</td>
</tr>
<tr>
<td>TdT</td>
<td>Terminal Deoxynucleotidyl Transferase</td>
</tr>
<tr>
<td>TGIF</td>
<td>5’-TG-3’ Interacting Factor</td>
</tr>
<tr>
<td>tin</td>
<td>Tinman</td>
</tr>
<tr>
<td>TLE</td>
<td>Transducin-like Enhancer of Split</td>
</tr>
<tr>
<td>TSS</td>
<td>Transcriptional Start Site</td>
</tr>
<tr>
<td>U</td>
<td>Units</td>
</tr>
<tr>
<td>Ubx</td>
<td>Ultrabithorax</td>
</tr>
<tr>
<td>µg</td>
<td>Microgram</td>
</tr>
<tr>
<td>µl</td>
<td>Microlitre</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>v/v</td>
<td>Volume/Volume</td>
</tr>
<tr>
<td>WT1</td>
<td>Wilm’s Tumor Gene</td>
</tr>
<tr>
<td>w/v</td>
<td>Weight/Volume</td>
</tr>
</tbody>
</table>
## List of Figures and Tables

<table>
<thead>
<tr>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chapter One</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Figure 1.1</strong> Development of Lymphoid and Myeloid Lineages from Pluripotent Hematopoietic Stem Cells (HSCs)</td>
<td>2</td>
</tr>
<tr>
<td><strong>Figure 1.2</strong> Organisation of the Mammalian HOX Gene Complex</td>
<td>11</td>
</tr>
<tr>
<td><strong>Figure 1.3</strong> HOX11 Belongs to the 93DE/NKL Cluster of Homeobox Genes, Including Highly Related Members, HOX11L1 and HOX11L2</td>
<td>12</td>
</tr>
<tr>
<td><strong>Figure 1.4</strong> Conserved Residues Present in the Third Alpha Helix of the Homeodomain Contact a 5'-T&lt;sub&gt;1&lt;/sub&gt;A&lt;sub&gt;2&lt;/sub&gt;A&lt;sub&gt;3&lt;/sub&gt;T&lt;sub&gt;4&lt;/sub&gt;-3' Core Sequence Present in the Majority of Homeodomain Binding Sites</td>
<td>15</td>
</tr>
<tr>
<td><strong>Figure 1.5</strong> Hox Proteins Bind DNA Cooperatively with Non-Hox Homeodomain Partners</td>
<td>17</td>
</tr>
<tr>
<td><strong>Figure 1.6</strong> Hox Transcription Factors Bind Cooperatively with PBC and Meis TALE Homeodomain Family Members</td>
<td>19</td>
</tr>
<tr>
<td><strong>Figure 1.7</strong> Trimeric Association of Hoxb1, Pbx1a and Meis Factors on the Hoxb2 Gene r4 Enhancer</td>
<td>22</td>
</tr>
<tr>
<td><strong>Figure 1.8</strong> Prediction of the HOX11 Homeodomain Recognition Motif by Analogy to the DNA Base Contacts Posed in the X-Ray Crystal Model of the En Homeodomain-DNA Complex</td>
<td>28</td>
</tr>
<tr>
<td><strong>Figure 1.9</strong> The HOX11 Homeodomain Protein Featuring a Number of Regions Contributing to the Transactivation Potential of HOX11</td>
<td>29</td>
</tr>
<tr>
<td><strong>Figure 1.10</strong> Mechanism of Retinoic Acid Action</td>
<td>40</td>
</tr>
<tr>
<td><strong>Figure 1.11</strong> Aldehyde Dehydrogenases (ALDHs) Catalyse the Oxidation of Aldehyde to Carboxylic Acid</td>
<td>44</td>
</tr>
<tr>
<td><strong>Figure 1.12</strong> Biosynthesis of Retinoic Acid</td>
<td>47</td>
</tr>
<tr>
<td><strong>Figure 1.13</strong> Key Transcription Factors Involved in Lineage Commitment During Haematopoiesis</td>
<td>53</td>
</tr>
<tr>
<td><strong>Figure 1.14</strong> FHL1 Isoforms</td>
<td></td>
</tr>
<tr>
<td><strong>Table 1.1</strong> Oncogenes Known to be Activated by Chromosomal Translocations in T-cell Acute Lymphoblastic Leukaemia (T-ALL)</td>
<td>5</td>
</tr>
<tr>
<td><strong>Table 1.2</strong> Involvement of HOX Genes in Cellular Transformation</td>
<td>26</td>
</tr>
<tr>
<td><strong>Table 1.3</strong> Nomenclature and Chromosomal Locations of Human Aldehyde Dehydrogenases</td>
<td>42</td>
</tr>
<tr>
<td><strong>Chapter Two</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Figure 2.1</strong> Genomic Structure of Human ALDH1A1</td>
<td>58</td>
</tr>
<tr>
<td><strong>Figure 2.2</strong> Nucleotide Sequence of the 5' Flanking Region of the ALDH1A1 Gene Used in Functional Transcriptional Assays</td>
<td>92</td>
</tr>
</tbody>
</table>
Figure 2.3  Schematic Diagram of the Human ALDH1A1 Promoter Luciferase Fusion Constructs Employed in this Study  93
Figure 2.4  Transcriptional Activity of the ALDH1A1 Promoter in the HEL Cell Line  96
Figure 2.5  Transcriptional Activity of the ALDH1A1 Promoter in the PER-117 Cell Line  97
Figure 2.6  HOX11 Represses SV40-Driven β-Galactosidase Expression in HEL and PER-117 Cell Lines  99
Figure 2.7  HOX11 Activates Transcription From the Empty pGL3Basic Vector in the HEL Cell Line  101
Figure 2.8  The Effect of HOX11 on β-Galactosidase Expression From the pSV-b-Gal Internal Control is Independent of the Cotransfected ALDH1A1 Reporter Construct in (A) PER-117 and (B) HEL Cell Lines  103
Figure 2.9  Analysis of the Effect of HOX11 on Nested Deletion Constructs of the ALDH1A1 Promoter in the HEL Cell Line  104
Figure 2.10  Analysis of the Effect of HOX11 on Nested Deletion Constructs of the ALDH1A1 Promoter in the PER-117 Cell Line  105
Figure 2.11  HOX11ΔH3 Activates SV40-Driven β-Galactosidase Activity in HEL and PER-117 Cell Lines  107
Figure 2.12  The Ability of HOX11 to Repress but not Activate the ALDH1A1 Promoter is Disrupted by Deletion of DNA-binding Helix 3 of the Homeodomain  109
Figure 2.13  DNase1 Footprinting Analysis Using HEL Nuclear Extracts on ALDH1A1 Promoter DNA  111
Figure 2.14  Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter CCAAT box (-74/-70bp)  115
Figure 2.15  Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter CCAAT box (-74/-70bp)  116
Figure 2.16  Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 promoter CCAAT box (-74/-70bp)  118
Figure 2.17  Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-29bp)  

Figure 2.18  Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 promoter GATA Box (-34/-29bp)  

Figure 2.19  Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-29bp)  

Figure 2.20  HOX11 Disrupts Complexes at the CCAAT Box Within the ALDH1A1 Promoter in PER-117 T-Cells  

Figure 2.21  Electromobility Shift Analysis of HEL and HELHOX11 Nuclear Protein Binding to the ALDH1A1 Promoter CCAAT Box (-74/-70bp)  

Figure 2.22  Electromobility Shift Analysis of HEL and HELHOX11 Nuclear Protein Binding to the ALDH1A1 Promoter CCAAT Box (-74/-70bp)  

Figure 2.23  Electromobility Shift Analysis of HEL and HELHOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-29bp)  

Figure 2.24  Electromobility Shift Analysis of HEL and HELHOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-29bp)  

Figure 2.25  HOX11 Activates Transcription in the HEL Cell Line  

Figure 2.26  HOX11 Acts as a Bi-functional Transcription Factor in PER-117 Cells  

Table 2.1  ALDH1A1 Promoter Primer and Oligonucleotide Sequences  

Table 2.2  Parental Mammalian Cell Lines Utilised in Transient Luciferase Reporter Analysis of the ALDH1A1 Promoter  

Table 2.3  ALDH1A1 Promoter Footprinting Primer Sequences  

Table 2.4  The Components of a Typical Footprinting Reaction  

Table 2.5  ALDH1A1 Promoter Electromobility Shift Assay (EMSA) Oligonucleotides  

Table 2.6  Expression Status of HOX11 and ALDH1A1 in HEL and PER-117 Cell Lines  

Table 2.7  Nucleotide Sequences of Protected Sites Identified by DNase1 Footprinting for HEL[1], HELHOX11[4], PER-117[1] and PER-117HOX11[11] Nuclear Extracts on the Proximal ALDH1A1 Promoter  

Chapter 3  

Figure 3.1  Assessment of ALL-SIL RNA Integrity for Use in 5’ RACE Analysis
Figure 3.2  Genomic Structure of the Human FHL1 Gene Annotated with Putative Transcription Initiation Sites Identified by 5’RACE

Figure 3.3  Schematic Representation of the 5’ System for the Rapid Amplification of cDNA Ends Version 2.0

Figure 3.4  Schematic Representation of the GeneRacer™ Method for Obtaining 5’cDNA Ends

Figure 3.5  5’ RACE Analysis Using Version 2.0 with SUPERSCRIPT™ II RT to Determine the Transcriptional Initiation Sites of FHL1

Figure 3.6  5’ RACE Analysis Using Version 2.0 with THERMOSCRIPT™ RT to Determine the Transcriptional Initiation Sites of FHL1

Figure 3.7  5’ RACE Analysis Using Omniscript™ RT to Determine the Transcriptional Initiation Sites of FHL1

Figure 3.8  5’ RACE Analysis Using GeneRacer™ to Determine the Transcriptional Initiation Sites of FHL1

Figure 3.9  Nucleotide Sequence of the Putative FHL1 Promoter

Figure 3.10  Sequence Alignment of the Human and Murine FHL1 Exon 1 and Flanking Region

Figure 3.11  Activity of the FHL1 Promoter in PER-117 and HEL Cell Lines

Figure 3.12  Effect of HOX11 on the Proximal FHL1 Promoter in PER-117 and HEL Cell Lines

Figure 3.13  The Ability of HOX11 to Repress the FHL1 Promoter is Affected by Loss of Homeodomain Helix 3 in PER-117 Cells

Table 3.1  FHL1 Primer Sequences

Chapter 4

Figure 4.1  Purification of Recombinant GST Fusion Proteins by Affinity Chromatography Using Glutathione Sepharose

Figure 4.2  Verification of HOX11 Biological Activity by Electrophoretic Mobility Shift Analysis (EMSA)

Figure 4.3  Quantitative Analysis of GST Fusion Proteins Utilised in GST-Pulldown Assays

Figure 4.4  GST-Pulldown Assay to Assess the Interaction of HOX11 with Putative Protein Partners

Figure 4.5  GST-Pulldown Assay to Assess the Interaction of HOX11 and HOX11ΔH3 with TFIIB and HOX11

Figure 4.6  Hypothetical Models for HOX11 Activity

Table 4.1  Primer Sequences for Cloning and EMSA Assays

Chapter 5

Figure 5.1  The MigR1 Retroviral Expression Vector Used to Generate Recombinant Retrovirus Expressing ALDH1A1
or FHL1

Figure 5.2 Production of Recombinant Retrovirus Using the Phoenix™ Eco Packaging Cell Line

Figure 5.3 Schematic Outline of Retroviral Transduction of Murine Bone Marrow and Subsequent Reconstitution

Figure 5.4 Transfection of MigR1 Retroviral Constructs into Phoenix™- Eco Producer Cells

Figure 5.5A Contribution (%) of MigR1-GFP, MigR1-ALDH1A1-GFP and MigR1-FHL1-GFP transduced cells in spleen, bone marrow and thymus tissues 33 days after transplantation (Replicate #1).

Figure 5.5B Contribution (%) of MigR1-GFP, MigR1-ALDH1A1-GFP and MigR1-FHL1-GFP transduced cells in spleen, bone marrow and thymus tissues 36 days after transplantation (Replicate #2).

Figure 5.5C Contribution (%) of MigR1-GFP, MigR1-ALDH1A1-GFP and MigR1-FHL1-GFP transduced cells in spleen, bone marrow and thymus tissues 38 days after transplantation (Replicate #3).

Figure 5.6A Myeloid Immunophenotyping of Cells Obtained From the Spleens of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow

Figure 5.6B Myeloid Immunophenotyping of Cells Obtained From the Bone Marrow of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow

Figure 5.7A B-cell Immunophenotyping of GFP+ve Cells Obtained From the Spleens of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow

Figure 5.7B B-cell Immunophenotyping of GFP+ve Cells Obtained From the Bone Marrow of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow

Figure 5.8 T-Cell Immunophenotyping of GFP+ve Cells Obtained From the Thymus of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow.

Figure 5.9 Trans-effect of ALDH1A1 Expression on Myeloid Differentiation

Figure 5.10A Erythroid Immunophenotyping of GFP+ve Cells Obtained From the Spleens of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow
Figure 5.10B  Erythroid Immunophenotyping of GFP+ve Cells Obtained From the Bone Marrow of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow

Table 5.1  Primers for Recombinant Retrovirus Constructs
Table 5.2  Fluorochromes Used for Immunophenotyping with the BD FACS Calibur

1  Chapter 6
2  Figure 6.1  The Human ALDH1A1 Promoter Contains a Consensus TFIIB Recognition Element (BRE)
Chapter 1

General Introduction

1.1 T-CELL ACUTE LYMPHOBLASTIC LEUKAEMIA (T-ALL)

1.1.1 Biology and Classification of Leukaemias

Haematopoiesis involves the formation of mature, circulating blood cells via the progressive commitment of self-renewing pluripotent haematopoietic stem cells (HSCs) along lineage-specific pathways (reviewed by Yoder, 2004)(Figure 1.1). The requirement for strict control of transcriptional networks directing processes of cell renewal, commitment, maturation and survival within each lineage is underscored by the plethora of haematological disorders that result from the disruption of pathways regulating normal blood cell development (reviewed by Sawyers et al., 1991). The term leukaemia, derived from the Greek ‘white blood’, refers to cancers of the blood-forming or haematopoietic tissues. These arise from the malignant transformation of haematopoietic cells arrested at various stages of development in the bone marrow, which may then infiltrate the blood, lymph nodes, spleen, liver, central nervous system (CNS) and other organs (Rabbitts, 1991, Colby-Graham & Chordas, 2003). The observation that certain non-leukaemic human blood disorders demonstrate only partial leukaemic phenotypes, for example uncontrolled proliferation (myeloproliferative syndromes) or differentiation block (myelodysplastic syndromes), suggests that full leukaemic transformation requires disruptions affecting both cellular growth and maturation (Tefferi, 2001). These cellular defects are the result of genetic alterations in the form of point mutations, gene amplification, gene deletions or chromosomal translocations that affect the expression of growth factors (e.g. GM-CSF, IL-3, IL-7)(Meeker et al., 1990), growth factor receptors (e.g. tyrosine kinase transmembrane receptors)(Heard et al., 1987), cell cycle regulators (e.g. p16, p15)(Kamb et al., 1994), transcriptional regulators involved in cellular differentiation (e.g. E2A, PBX, SCL, RAR, PLZF, PML) and genes involved in mediating programmed cell death (e.g. BCL-2, BCL-XL, MCL-1)(McDonnell et al., 1989). Leukaemias can be divided into two major classes based on the lineage from which each is derived; thus myeloid leukaemias
Figure 1.1. Development of Lymphoid and Myeloid Lineages from Pluripotent Haematopoietic Stem Cells (HSCs). HSCs can be subdivided into long-term self renewing HSCs, short-term self renewing HSCs and multipotent progenitors which give rise to common lymphoid progenitors (CLP), the precursors of lymphoid cells and common myeloid progenitors (CMP), the precursors of myeloid cells.

(Adapted from Reya et al., 2001)
originate from progenitors giving rise to granular leukocytes whereas lymphoid leukaemias are derived from T-cell or B-cell lymphoid progenitors (Colby-Graham & Chordas, 2003). These can be further sub-classified according to the clinical nature of the disease, such that chronic leukaemias, which are characterised by slow onset and the production of more mature, differentiated cells, are distinct from acute leukaemias, which have a rapid onset and a predominance of highly immature (blastic) cells, and consequently represent a more severe form of the disease (Rabbitts, 1991, Colby-Graham & Chordas, 2003). Acute lymphoblastic leukaemia (ALL) represents the predominant type of leukaemia in children, accounting for approximately one-quarter of all juvenile tumours, with a peak incidence occurring between the ages of 2-5 years (Uckun et al., 1998, Pui, 2000). Recent advances in molecular genetic techniques, for example the use of oligonucleotide microarrays to profile patterns of gene expression in the leukaemic blasts of multiple paediatric ALL patients, confirms that ALL is a heterogeneous disease that can be divided into either B-lineage or T-lineage ALL on the basis of cell surface antigen expression differences (B-lineage markers; CD19, CD22 or T-lineage markers; CD2, CD3, CD8). ALL can also be classified on the basis of distinct genetic lesions affecting specific transcriptional pathways (Yeoh et al., 2002, Ross et al., 2003). For example, B-ALL consists of genetically distinct subtypes including t(9;22)[BCR-ABL], t(1;19)[E2A-PBX1], t(12;21)[TEL-AML1], rearrangements affecting the MLL gene at 11q23 and hyperdiploid karyotype (>50 chromosomes), whereas T-ALL involves chromosomal translocations affecting the expression of various transcription factors including TAL1, LMO1, LMO2 and HOX11 (reviewed by Raimondi, 1993). It has been found that the molecular aberrations underlying these individual subtypes can influence the differential response of patients to chemotherapy (Yeoh et al., 2002, reviewed by Faderl et al., 2003). This highlights the usefulness of expression profiling as a powerful diagnostic tool in the treatment of ALL by aiding the accurate stratification of patients into biologically distinct subtypes with significant prognostic value. Such technologies are likely to supercede conventional diagnostic approaches which require a combination of clinical characteristics (age, sex), laboratory analyses (white blood cell count) and characterisation of leukaemic blasts (via immunophenotyping, cytogenetic analysis, cytochemical staining and morphological analyses), since these are labour intensive and costly.
1.1.2 Molecular Genetics of Childhood T-ALL

T-cell acute lymphoblastic leukaemia (T-ALL) is a malignant disease of thymocytes, accounting for approximately 10-15% of paediatric and 25% of adult acute lymphoblastic leukaemia cases (Heerema et al., 1998, Ferrando et al., 2002). T-lineage ALL patients are predominately male, non-white, usually older in age (>10 years old), and are characterised by a number of unfavourable presenting features including high white blood cell count (>50,000/ul), CNS infiltration, a mediastinal mass and enlargement of the spleen, liver and lymph nodes (Uckun et al., 1998, van de Berg et al., 1998, Heerema et al., 1998). Although the majority of ALL cases are immunophenotypically classified as B-lineage (~85%), considerable interest in the aetiology, biology and treatment of T-ALL exists, particularly as children with T-ALL have a worse prognosis compared to children with B-cell ALL (B-ALL)(reviewed by Uckun et al., 1998, van de Berg et al., 1998). Recent therapeutic advances following the use of risk-adapted treatment protocols, which involves the tailoring of doses and combinations of anti-leukaemic drugs according to each patient’s risk of relapse, has significantly improved the clinical outcome of T-ALL in children and adolescents, with 5 year relapse-free survival rates now ranging between 60-75%. However, further advances in diagnosis and treatment will require an improved knowledge of the biology of T-ALL at a molecular level (Schrapppe et al., 2000, Silverman et al., 2001).

T-ALL is commonly associated with recurrent chromosomal translocations and intrachromosomal rearrangements that activate developmentally important transcription factors (Rabbitts, 1991). These translocations typically juxtapose strong promoter or enhancer elements driving high levels of expression from the TCR β locus (7;q34) or the TCR α/δ locus (14;q11), next to transcription factor genes including HOX11/TLX1 (Dube et al., 1991), HOX11L2 (Bernard et al., 2001), TAL1/SCL (Begley et al., 1989), TAL2, LYL1 (Mellentin et al., 1989), BHLHB1 (Wang et al., 2000) and LMO1/LMO2 (Boehm et al., 1991)(Table 1.1). The mechanism underlying these translocations likely involves errors in T-cell DNA recombinase activity during rearrangement of TCR V, D and J segments in developing cortical T-cells, since hallmarks of V(D)J recombinase activity are often found at translocation breakpoints (Visvader & Begley, 1991). These transcription factors are normally expressed in early haematopoietic cells, but are
<table>
<thead>
<tr>
<th>Transcription Factor</th>
<th>Translocation</th>
<th>Motif</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCL (TAL1)</td>
<td>t(1;14)</td>
<td>bHLH</td>
<td>Park et al., 1998, Jonsson et al., 1991</td>
</tr>
<tr>
<td>TAL2</td>
<td>t(7;9)</td>
<td>bHLH</td>
<td>Xia et al., 1991</td>
</tr>
<tr>
<td>LYL1</td>
<td>t(7;19)</td>
<td>bHLH</td>
<td>Xia et al., 1991, Mellentin et al., 1989</td>
</tr>
<tr>
<td>BHLHB1</td>
<td>t(14;21)</td>
<td>bHLH</td>
<td>Wang et al., 2000</td>
</tr>
<tr>
<td>MYC</td>
<td>t(8;14)</td>
<td>bHLH-ZIP</td>
<td>Shima et al., 1986</td>
</tr>
<tr>
<td>LMO1</td>
<td>t(11;14)</td>
<td>LIM</td>
<td>McGuire et al., 1989, Boehm et al., 1991</td>
</tr>
<tr>
<td>LMO2</td>
<td>t(11;14)</td>
<td>LIM</td>
<td>Boehm et al., 1991</td>
</tr>
<tr>
<td></td>
<td>t(7;11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCL1</td>
<td>t(7;14)</td>
<td>unknown</td>
<td>Laine et al., 2000, Pekarsky et al., 2000</td>
</tr>
<tr>
<td>TCL1b</td>
<td>t(7;14)</td>
<td>unknown</td>
<td>Pekarsky et al., 1999</td>
</tr>
<tr>
<td>HOX11 (TLX1)</td>
<td>t(10;14)</td>
<td>homeodomain</td>
<td>Hatano et al., 1991</td>
</tr>
<tr>
<td></td>
<td>t(7;10)</td>
<td></td>
<td>Kennedy et al., 1991</td>
</tr>
<tr>
<td>HOX11L2 (TLX3)</td>
<td>t(5;14)</td>
<td>homeodomain</td>
<td>Bernard et al., 2001</td>
</tr>
</tbody>
</table>

Table 1.1. Oncogenes Known to be Activated by Chromosomal Translocations in T-cell Acute Lymphoblastic Leukaemia (T-ALL).
minimally expressed or absent in immature and mature T-cells, such that aberrant expression of these putative oncoproteins in developing thymocytes may directly interfere with transcriptional programmes regulating normal thymocyte proliferation, differentiation and survival, thus setting the scene for leukaemic transformation. Two distinct pathways are now emerging for the genes translocated in T-ALL. The first involves the deregulation of target genes by transcriptional complexes comprising members of both the basic helix-loop-helix (bHLH)(TAL1, TAL2, LYL1, BHLHB1) and LIM-only (LMO)(LMO1, LMO2) families of transcription factors, while the second involves dysregulation of target genes mediated by the HOX11 family, (HOX11 and the highly related family member, HOX11L2).

HOX11 (10q24) and its paralogs, HOX11L1 (2p13) and HOX11L2 (5q35), belong to the NK-Like (NKL) family of homeobox genes (Dear & Rabbitts, 1994), which are located outside of the four mammalian HOX clusters. Unlike the clustered homeobox genes, which play important roles in segmental patterning or cell type specific differentiation, the NKL genes appear to be required for processes involving organogenesis, growth control and cell fate. HOX11 was originally identified as an overexpressed gene located at the t(10;14)(q24;q11) or t(7;10)(q35;q24) breakpoints found in childhood T-ALL, and it was subsequently shown that chromosomal translocations involving the HOX11 locus, together with that of HOX11L2, constitute the most common cytogenetic abnormalities in T-ALL (Hatano et al., 1991, Dube et al., 1991, Kennedy et al., 1991, Lu et al., 1991, Bernard et al., 2001, Helias et al., 2002, Cave et al., 2004). HOX11 is not detectable in normal T-cells and translocations involving 10q24 cluster tightly at the 5’ end of the gene, such that the coding capacity of the gene remains unaltered (Kennedy et al., 1991). This supports the hypothesis that the overexpression of an intact HOX11 protein is a crucial event in tumour initiation. Indeed, the constitutive expression of HOX11 has been associated with the immortalisation of haematopoietic progenitors in vitro (Hawley et al., 1997, Keller et al., 1998), and recent work by Ferrando et al., (2002, 2003) and Kees et al., (2003) reveal that as many as 30% of T-ALLs exhibit deregulated HOX11 expression when abnormalities other than cis-acting chromosomal rearrangements are included. HOX11L2 was also shown to be aberrantly expressed in a subset of T-ALL cases, following a novel cryptic recurrent translocation t(5;14)(q35;q32)(Bernard et al., 2001). This rearrangement places HOX11L2 within the
vicinity of the CTIP2 gene, which is strongly expressed in the thymus and implicated in T-cell differentiation (Bernard et al., 2001). Recent studies demonstrating that HOX11L2 is expressed in up to 60% of T-ALL cases, further supports the role of this transcription factor as a T-cell oncogene and highlights the role of HOX11 and related family members in T-cell oncogenesis (Bernard et al., 2001, Mauieux et al., 2002, MacLeod et al., 2003).

A second distinct pathway leading to T-ALL commonly involves SCL, a gene expressed in the erythroid and megakaryocytic haematopoietic lineages (reviewed by Begley & Green, 1999). This suggests a function in haematopoiesis, however SCL is not expressed in normal T-cells (Begley et al., 1989). The aberrant expression of SCL as a result of t(1;14)(p33;q11) or t(1;7)(p33;q35) translocations or interstitial deletion of approximately 90kB of DNA from chromosome 1, which places SCL under the control of the ubiquitously expressed SCL Interrupting Locus (SIL) gene, in developing thymocytes, implicates it in a major pathway leading to the development of T-ALL (Bash et al., 1995, Begley et al., 1989). These effects are mediated via the bHLH domain of SCL, which is capable of DNA binding and dimerisation with other transcription factors. These include the cysteine-rich LMO proteins LMO1 and LMO2, which are also dysregulated in a subset of T-ALLs harbouring the t(11;14)(p15;q11) and t(11;14)(p13;q11) chromosomal translocations, respectively (Boehm et al., 1991, Wadman et al., 1994, Ono et al., 1997). The overlapping expression of LMO2 with SCL in developing erythroid cells (Visvader et al., 1991), coupled with experiments demonstrating that mice lacking both SCL and LMO2 die at embryonic day 9.5 due to the absence of yolk sac erythropoiesis (Warren et al., 1994, Robb et al., 1995), suggests that these factors work in concert to regulate the erythroid cell lineage, and may also function cooperatively to regulate sets of target genes normally quiescent in T-lineage progenitors, resulting in T-cell transformation. Indeed, transgenic mice in which both SCL and LMO1/LMO2 expression is directed to the thymus, develop T-cell tumours with a shortened latency compared with mice expressing SCL alone (Robb et al., 1995, Larson et al., 1996, Aplan et al., 1997). Transcriptional complexes involving SCL, LMO2 and a third binding partner, GATA-1, which is also essential for normal erythropoiesis are therefore likely to regulate not only erythroid-specific, but potentially T-cell oncogenic target genes. However, unravelling these pathways requires a detailed
investigation of downstream targets. The identification of the receptor tyrosine kinase c-kit as a direct target gene of SCL, may provide crucial insights into the mechanisms by which SCL, in conjunction with LMO2, may contribute to the T-ALL phenotype, particularly as this gene plays an instrumental role in normal haematopoiesis (Lecuyer et al., 2002). The gene for retinaldehyde dehydrogenase 2 (RALDH2/ALDH1A2), which is one of a select few enzymes responsible for the biosynthesis of retinoic acid, a potent modulator of cellular proliferation and differentiation, has also been identified as a direct target of SCL and LMO, and coexpression of SCL and LMO in most T-ALL cell lines was also associated with RALDH2 expression (Ono et al., 1998). In this case, however, SCL and LMO act as cofactors for the T-cell specific GATA-3 transcription factor (Ono et al., 1998). Other bHLH family members including BHLHB1, LYL and MYC are also capable of promoting T-cell transformation. Aberrant expression of these proteins by chromosomal translocation may facilitate leukaemic transformation by dominant negative interference of E47 and E12 variants of E2A transcription factors, thereby disrupting the equilibrium of transcription factor heterodimers required for normal thymocyte development (Mellentin et al., 1989, Wang et al., 2000).

Although the clinical features of T-ALL are relatively uniform, translocations in T-ALL involve the activation of distinct oncogenes at various stages of thymocyte development. This molecular heterogeneity likely accounts for the dramatically different responses of T-ALL patients when treated with the same intensive multi-drug regimes (Ferrando et al., 2002). Indeed, recent studies have used DNA microarray analysis to link the activation of specific T-cell oncogenes to defined stages in T-cell development. This has revealed that these gene expression signatures have prognostic relevance, since the stage of thymocyte arrest determined the susceptibility of cells to drug induced programmed cell death (Ferrando et al., 2002). For example, the HOX11+ gene signature was indicative of arrest at the early cortical stage of thymocyte development and was associated with a favourable clinical outcome (Ferrando et al., 2002). This might be due to the lack of expression of antiapoptotic genes at this stage (e.g. BCL2), rendering cells more susceptible to drug-induced apoptosis (Ferrando et al., 2002). Gene signature profiles were also obtained for TAL1+ (late cortical thymocyte) and LYL1+ (pro-T) T-cell oncogenes, which in contrast to HOX11, were associated with less favourable outcomes, possibly due to the upregulation of BCL2 and
other antiapoptotic molecules in these populations (Ferrando et al., 2002). The powerful predictive nature of gene expression profiling is underscored by the discovery of novel T-ALL oncogenic factors, which demonstrate gene signatures closely related to known T-ALL oncogenes. For example, overexpression of \textit{HOX11L2}, which belongs to the \textit{HOX11} family of transcription factors, was associated with \textit{HOX11} negative samples exhibiting a characteristic \textit{HOX11} gene expression signature (Ferrando et al., 2002).

1.2 **HOMEBOX GENES**

1.2.1 The Role of Homeobox Genes in Development

The diversity of body design, which encompasses organisms of varying complexity, is the result of genetic processes directing embryogenesis in different species (McGinnis & Kuziora, 1994). Despite the obvious differences between the distinct members of the animal kingdom, the genes controlling these developmental events retain structural and organisational similarities, reminiscent of a common ancestral cluster (Hayashi & Scott, 1990, Krumlauf, 1994, Gehring \textit{et al.}, 1994, Mark \textit{et al.}, 1997). The discovery of homeobox (\textit{HOX}) genes as ‘master developmental control genes’, regulating spatial and temporal aspects of morphogenesis and cell differentiation along the anterior-posterior (AP) embryonic axis, occurred following an elegant series of homeotic transformation experiments in the genome of the fruit fly \textit{Drosophila} (Lewis, 1978, Kemphues, 1980, Mark \textit{et al.}, 1997). Derived from the Greek word ‘\textit{homeo}’ which means ‘alike’, the homeotic (HOM) complex (or \textit{HOM-C}) genes were named for their ability to transform one insect body segment into the likeness of another, following specific (homeotic) mutations (Castronovo \textit{et al.}, 1994, Mark \textit{et al.}, 1997). For example, a mutation in the \textit{Antennapaedia (Antp)} gene, the first of 25 \textit{Drosophila} developmental control genes to be identified, leads to the transformation of antennae on the head of a fly, to an extra thoracic leg (Scott \textit{et al.}, 1989, Veraksa \textit{et al.}, 2000). Following their discovery in \textit{Drosophila}, similar clusters of homeobox genes have been identified in a wide range of multicellular organisms including sponges, vertebrates, plants and fungi (Mark \textit{et al.}, 1997).

In \textit{Drosophila}, these genes are organised as a bipartite homeotic gene complex, comprising the \textit{Antennapedia (Antp-C; labial (lab), proboscipedia (pb), Deformed}}
(Dfd), Sex combs reduced, Antennapedia (Antp)) and Bithorax (Bx-C; Ultrabithorax (Ubx), abdominal-A (abd-A) and Abdominal-B (Abd-B)) complexes, located on the right arm of chromosome three at positions 84AB and 89E, respectively (reviewed by Jagla et al., 2001). The mammalian homologs of Antp-C and Bx-C genes exist within distinct, unlinked complexes on human chromosomes 7p15 (HOXA), 17q21 (HOXB), 12q13 (HOXC) and 2q31 (HOXD), with paralog groups 1-8 being most closely related to Drosophila Antp, and groups 9-13 most closely related to Drosophila Abd-B genes (Veraksa et al., 2000, Owens & Hawley, 2002). The organisation of the HOX gene complex is such that during embryonic development, the order of expression along the anterior-posterior (AP) embryonic axis (3’ to 5’) is co-linear with the alignment of genes on the chromosome (Paralogous groups 1-13)(Figure 1.2). The preservation of position and the order of expression of genes within these clusters between invertebrates and vertebrates suggests that mammalian HOX genes are likely to have evolved by two successive genome duplications of a hypothetical, ancestral HOX cluster, and it is postulated that their maintenance within these clustered arrangements is linked to the dependence on central, regulatory regions (Jagla et al., 2001, Owens & Hawley, 2002).

Following the identification of these clustered HOX genes, a new family of homeobox genes, referred to as the NK family, were identified in Drosophila (Kim & Nirenberg, 1989). Although genes within this cluster were originally referred to as ‘orphan’ homeobox genes, since they are not associated with the major HOX clusters, molecular mapping and the identification of additional family members has defined a new cluster of homeobox genes, known as 93DE/NK cluster (Kim & Nirenberg, 1989, Pollard & Holland, 2000, Coulier et al., 2000, reviewed by Jagla et al., 2001)(Figure 1.3). The 93DE/NK cluster, which is also located on the right arm of the third chromosome, is the second largest cluster of physically linked genes in Drosophila, and is composed of six homeobox genes that have retained tight linkage over a region of 180kB, including (in proximal to distal order) NK4/tinman (tin), NK3/bagpipe (bap), ladybird late (lbl), ladybird early (lbe), 93bal/Hox11 and NK1/slouch (slou)(reviewed by Jagla et al., 2001, Harvey, 1996). The discovery of vertebrate homologs of these genes (NKL cluster), suggests that these genes are extremely ancient, and the co-localisation of some vertebrate homologs, for example, HOX11/LBX1 genes to chromosome 10q24 and HOX11L1/LBX2 genes to chromosome 2p13 is reminiscent of such clustering. This
Figure 1.2. Organisation of the Mammalian HOX Gene Complex. During embryonic development, the order of expression along the anterior-posterior (AP) embryonic axis (3’ to 5’) is spatially and temporally co-linear with the alignment of genes on the chromosome (Paralogous groups 1-13).

implies that the vertebrate NKL genes arose from duplications of an ancestral NK cluster, although further lateral dispersal has certainly occurred throughout evolution (Holland, 2001). Additional homeobox genes have since been mapped to nearby chromosomal bands in humans and mice, for example EMX2 maps to 10q25-26, adjacent to the HOX11/LBX1 cluster, implying that these genes were also originally part of a large NKL cluster (Holland, 2001).

Similar to the clustered HOX genes, which are required for cell fate specification during mesoderm development, genes of the 93DE/NKL cluster are involved in mesoderm differentiation programs and fulfil distinct roles in specifying the fate of muscle and heart-forming cells (Jagla et al., 2001). For example, the NNX genes play important roles in muscle patterning, lung and cardiac development, respectively (Knirr et al., 1999, Zhu et al., 2004), LBX family members fulfil important functions in spinal cord patterning, hindbrain development and skeletal muscle development (Jagla et al., 1995, Mennerich et al., 1998, Schubert et al., 2001), and HOX11, HOX11L1 and HOX11L2 have overlapping functions in central nervous system development as well as individual roles in spleen organogenesis (Roberts et al., 1994), enteric nervous system maintenance (Hatano et al., 1997) and respiratory development (Shirasawa et al., 2002), respectively. The dispersal of the vertebrate NKL cluster throughout evolution, however, suggests that clustering may not be related to their regulation as for the HOX genes, which demonstrate spatial and temporal co-linearity, although there is evidence to suggest that temporal co-linearity of NKL genes exists in the developing mesoderm (Jagla et al., 2001).

Given that most models of HOX target selectivity, structure and function have come from investigations into the clustered HOX genes, studies of genes within the evolutionarily distinct 93DE/NKL family are likely to yield a more accurate insight into HOX11-mediated mechanisms of gene regulation. For example, members of the 93DE cluster harbour a short sequence motif (*PFSI*DIL***) in the amino terminal region, which bears homology to the Eh1 repression domain required for the interaction of the Engrailed (En) homeodomain protein with transcriptional corepressors, such as Groucho (Tolkunova et al., 1998). Thus, transcriptional repression via this domain may represent a common mechanism of target gene regulation by members of 93DE family,
including HOX11, which also shares this motif. Moreover members of the 93DE family demonstrate variations in their affinity for the canonical TAAT homeobox core recognition motif. For example, Tin preferentially binds a sequence with a CAAG core as a result of the R to Y substitution at position 54 within the homeodomain (Damante et al., 1996), and HOX11 recognises the TAAG core sequence in addition to the TAAT consensus, due to the substitution of isoleucine for threonine at position 47 within the homeodomain (Dear et al., 1993). Thus, target gene regulation by members of the 93DE family may occur via the specific binding to these variant homeodomain sequences or possibly via competition with HOX proteins for TAAT sites (Jagla et al., 2001).

1.2.2 Hox Proteins as Transcription Factors – Specificity of Homeotic Gene Function

1.2.2.1 Binding of HOX Homeodomains

Members of the HOM/HOX gene family share a common sequence motif comprising 183 nucleotides, that encodes a 61 amino acid homeodomain capable of sequence-specific DNA binding and subsequent transactivation of specific target genes (Ford, 1998). To elucidate the three dimensional structure of the homeodomain, a combination of X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy was employed to investigate the complex structure formed between operator DNA sequences and the homeodomains of engrailed (en), even-skipped (eve), MATα2 and Antp proteins (Kissenger et al., 1990, Wolberger et al., 1991, Billeter et al., 1993). These studies revealed the homeodomain to consist of three alpha helices folded into a tight globular structure and a flexible N-terminal arm (Gehring et al., 1994, Mann, 1995). In the context of DNA binding, base-specific contacts are mediated by Helix 3, which is positioned roughly parallel to the major groove of the DNA and by the flexible N-terminal arm which makes additional base-specific contacts in the minor groove; whereas Helices 1 and 2 are aligned in an antiparallel fashion above the DNA, such that the 5 amino acid loop between these helices interacts with the DNA backbone to stabilise binding.
In conjunction with structural analyses, the DNA binding properties of homeodomain proteins have been analysed by *in vitro* electromobility shift assays (EMSA), DNase1 footprinting and transactivation studies. These have revealed that Hox proteins demonstrate relatively low specificity for a 5'-T<sub>1</sub>A<sub>2</sub>A<sub>3</sub>T<sub>4</sub> -3' core that occurs at a high frequency throughout animal genomes (Treisman *et al.*, 1989, Gehring *et al.*, 1994). Conserved residues present in the third alpha helix of many homeodomain proteins, in particular an isoleucine or valine at position 47 and an asparagine at position 51, contact bases 4 and 3 of this recognition sequence respectively, whilst arginines at positions 5 and 3 of the N-terminal arm contact bases 1 and 2 of this sequence in the minor groove (Figure 1.4)(Wilson *et al.*, 1996). Although differences in adjacent N-terminal residues of the homeodomain may provide the basis for some sequence discrimination, thereby accounting in part for the distinct regulatory properties of these proteins *in vivo*, these subtle differences are unlikely to account for the exquisite specificity demonstrated by these crucial regulators of development.

![Figure 1.4](image)

**Figure 1.4.** Conserved Residues Present in the Third Alpha Helix of the Homeodomain Contact a 5'-T<sub>1</sub>A<sub>2</sub>A<sub>3</sub>T<sub>4</sub> -3' Core Sequence Present in the Majority of Homeodomain Binding Sites.

1.2.2.2 *Hox Proteins and Cofactors*

The question arises - how is target gene specificity achieved allowing homeodomain proteins to fulfil their distinct regulatory roles in development, if all homeodomains demonstrate similar DNA binding specificity? Increasing evidence suggests that the most likely mechanism for the selective regulation of Hox target genes is through
cooperative interactions with protein cofactors that increase the specificity and affinity for larger compound binding sites (Wilson & Desplan, 1995, Biggin & McGinnis, 1997). In support of such a co-selective model in which other factors are required, chimaeric and deletion analyses of HOM proteins in vivo suggest that at least some functional specificity of HOM proteins resides in non-DNA contacting regions within or immediately flanking the homeodomain (McGinnis & Kuziora, 1994). Numerous studies have since demonstrated that HOX proteins form DNA binding complexes with members of the TALE (Three-Amino-acid Loop Extension) subclass of homeodomain-containing proteins including the PBC (mammalian PBX, Drosophila melanogaster Extradenticle (Exd) and Caenorhabditis elegans CEH-20 proteins), MEIS (Myeloid Ectopic Integration Site)(mammalian MEIS and PREP1 and Drosophila Hth proteins) and TGIF (5’-TG-3’ Interacting Factor) families, whose members contain an extension of three amino acids between α-helices 1 and 2 within the homeodomain (Saleh et al., 2000, Burglin, 1997, Mann, 1995). Genetic evidence from Drosophila reveals a functional role for Exd and Hth proteins in Hox-regulated developmental pathways, since mutants that lack exd or hth have homeotic transformations resembling the loss of multiple Hox genes (Rieckhof et al., 1997, Kurant et al., 1998). The importance of these interactions in the regulation of gene transcription and maintenance of normal HOX functions involving axial patterning and haematopoietic cell differentiation, is further supported by the presence of mutations affecting the stoichiometry or stability of members comprising HOX transcriptional complexes identified in a range of haematological malignancies (Lawrence et al., 1996).

Exd was first identified in a screen for zygotic lethal mutations that cause patterning defects in Drosophila and encodes a homeodomain protein that interacts cooperatively with HOM gene products such as Ubx to modify target gene selection (Rauskolb et al., 1993). In a similar manner to exd, the mammalian orthologue Pbx1 (pre-B-cell leukaemia transcription factor 1 gene), which was first identified as a proto-oncogene disrupted by the chromosomal translocation t(1;19) in human paediatric pre-B-cell leukaemias, also binds DNA cooperatively with Hox proteins (Paralog groups 1-10)(Figure 1.5). This enhances their DNA binding specificity and affinity, resulting in the selective regulation of target genes harbouring combinatorial responsive elements (Kamps et al., 1990, Rauskolb et al., 1993, Popperl et al., 1995). Other family members
including Pbx2 and Pbx3 were subsequently identified by DNA cross-hybridisation (Monica et al., 1991). The cooperative interaction of HOX and PBX proteins is mediated in part by a conserved stretch of amino acids located 5 to 50 amino acids N-terminal to the homeodomain in HOX proteins from paralogous groups 1-8, referred to as the ‘pentapeptide’, ‘hexapeptide’ or ‘YPWM’ (Tyr/Phe-Pro-Trp-Met) motif (Phelan et al., 1995, Mann & Chan, 1996), in addition to the homeodomain and adjacent C-terminal 16 amino acids of Pbx1 (Mann & Chan, 1996, Lu & Kamps, 1996)(Figure 1.6). Binding is thought to occur through a ‘lock and key’ mechanism, involving hydrophobic interactions between the YPWM motif of HOX proteins and the hydrophobic pocket of TALE proteins, which is formed by a three amino acid insertion in between helices 1 and 2 of the homeodomain (Piper et al., 1999, Sprules et al., 2003). According to studies performed by Phelan et al., (1995), cooperative DNA binding by HOXD4 and PBX1A occurs via direct protein-protein contacts mediated by the HOXD4 pentapeptide, although the possibility that the pentapeptide also plays a role in DNA recognition of HOX/PBX binding sites can not be ruled out. Further support for the role of this motif in mediating heterodimer formation derive from studies in which the binding of the HoxB1 hexapeptide in a pocket of the Pbx1 homeodomain facilitates the binding of HoxB1 and Pbx1 to overlapping binding sites located on opposite sides of the DNA (Piper et al., 1999). In addition to this role of mediating HOX/PBX heterodimerisation, the YPWM motif has also been shown to inhibit DNA binding by the Drosophila HOX protein labial (Lab), which is relieved by interaction with EXD or PBX (Chan et al., 1996). Although HOX proteins contained within paralogous groups 9-13 do not possess a typical ‘YPWM’ motif, a number of these proteins are still capable of interacting with PBX through conserved tryptophan residues in a conserved ANW sequence located N terminal to the homeodomain (Chang et al., 1996).

A model for the PBX-HOX-DNA ternary complex was proposed following binding site selection, protein-DNA and mutagenesis experiments by several independent groups in which PBX and HOX homeodomains are arranged as a head to tail heterodimer at the consensus binding sequence 5’-[T_{1}G_{2}A_{3}T_{4}T_{5}N_{6}A_{7}T_{8}G_{9}G_{10}]-3’(Chan & Mann, 1996, Lu & Kamps, 1996). According to this model, PBX occupies the 5’ half site (TGAT) and the HOX partner binds the 3’ half site (TNATGG), with differential binding specificity by HOX proteins occurring via interactions with the variable nucleotide at position 6.
Figure 1.6. Hox Transcription Factors Bind DNA Cooperatively with PBC and Meis TALE Homeodomain Family Members. Protein-protein interactions are mediated via conserved motifs shared between related members.

(Adapted from Allen et al., 2000)
The formation of HOX-PBX heterodimers alters the conformation of the HOX homeodomain N-terminal arm, which upon restructuring leads to extended DNA contacts, thereby refining DNA sequence specificity (Mann & Chan, 1996). Indeed, specific residues contained within the N-terminal arm of the homeodomain (Arg3), that are important for binding of HOX monomers, are not required for DNA binding in heterodimers with PBX. This demonstrates the effect of homeodomain structural rearrangement due to complex formation on target specificity (Phelan & Featherstone, 1997).

In addition to PBX, the MEIS subfamily of homeodomain proteins, which include MEIS1, MEIS2, MEIS3, TGIF and PREP1 have also been shown to interact and bind cooperatively to DNA with HOX gene products (Shen et al., 1997)(Figures 1.5, 1.6). Meis1 was originally identified as a gene whose expression was activated by proviral integration in myeloid leukaemias occurring in BHX-2 strain of mice, suggesting a link between these genes and the regulation of cell proliferation (Moskow et al., 1995, Nakamura et al., 1996). Murine Meis2 and Meis3 were subsequently isolated by cross-hybridisation with the Meis1 homeodomain (Nakamura et al., 1996). The concomitant proviral activation of Hoxa7 or Hoxa9 with Meis1 in malignant cells derived from these mice provided the first evidence of biological synergy between Hox proteins and Meis1 (Nakamura et al., 1996). Using a series of in vitro binding assays, it has since been shown that Meis1 forms heterodimeric DNA binding complexes with Hoxa9, in addition to other members of the AbdB-like subset of Hox proteins, on a DNA sequence harbouring both Meis1 (TGACAG) and AbdB-like Hox (TTTACGAC) sites (Shen et al., 1997). Unlike the interactions described for Pbx1, which serve to increase the binding affinity and specificity of Hox proteins (Paralog groups 1-8), the affinity and specificity of AbdB-like Hox (9-13) proteins is not enhanced by interactions with either Pbx or Meis1 (Shen et al., 1997). Instead, the relatively weak affinity of Meis1 for DNA is greatly stabilised by interactions with AbdB-like Hox proteins, involving the N-terminal region of the Hox protein (Shen et al., 1997).

This model is not universal for all Hox-Meis interactions however, since there have been documented cases where the Meis component does indeed contribute to the binding specificity of Hox complexes. For example, the association of a ternary
complex involving Hoxb1, Pbx and Meis on the Hoxb2 hindbrain enhancer, r4, requires the specific binding of Meis to a DNA sequence located upstream of a consensus Hox-Pbx binding site and mutations of this distal Meis site recapitulate the effect of mutating the Pbx-Hox site (Jacobs et al., 1999)(Figure 1.7). These studies revealed that DNA binding by Meis (as opposed to being tethered to a functional complex via protein-protein interactions with Pbx) increases the binding site selectivity of ternary complexes involving Hox and Pbx, and is required for the r4 enhancer-directed expression of a reporter transgene in the developing hindbrain (Jacobs et al., 1999). In the absence of DNA binding, Meis proteins may also function to negatively affect the binding of Pbx-Hox heterodimers to dimeric sites, by interacting with regions adjacent to the Pbx homeodomain, causing alterations in the DNA binding capabilities of Hox-Pbx complexes (Jacobs et al., 1999). In addition to the role of TALE homeodomain proteins in directing the binding specificity of ternary complexes involving HOX members, a functional role for Hth (MEIS) proteins in the nuclear localisation of Exd (PBX) has also been documented in Drosophila (Pai et al., 1998). This suggests that MEIS proteins may also control HOX function via the regulation of subcellular distributions of PBX cofactors, in addition to mechanisms involving the formation or disruption of gene-specific transcription complexes - yielding the potential for even greater transcriptional diversity (Chang et al., 1997, Shanmugam et al., 1999).

It should be noted however, that most of our knowledge regarding HOX target selectivity derives from studies involving the clustered HOX genes, and may not be applicable to HOX genes located in evolutionarily distinct clusters, for example HOX11, which belongs to the NKL family of proteins. Thus, HOX11 may act via mechanisms distinct from models of ‘co-selective binding’. An alternative, but not mutually exclusive model proposed by Galant et al., (2002), postulates that Hox proteins may regulate targets in the absence of cofactors, by binding to multiple monomer sites within cis-regulatory elements. For example, in Drosophila, the Hox protein Ultrabithorax suppresses wing development by repressing a cis-regulatory element of the spalt (sal) gene in the halteres, through multiple monomer Ubx-binding sites (Galant et al., 2002). The selective regulation of certain HOX target genes in the absence of PBC/MEIS cofactors appears to be restricted to genes involved in the development of distal appendages, a theme previously described in Drosophila (Mann
Figure 1.7. Trimeric Association of Hoxb1, Pbx1a and Meis Factors on the Hoxb2 Gene r4 Enhancer. Pbx-Meis-Hoxb1 heterotrimeric complexes assemble on the Hoxb2 enhancer containing a Meis site located upstream and in the reverse orientation with respect to the Pbx-Hox consensus core sequence. Pbx and Meis components heterodimerise via their amino terminal motifs, allowing their homeodomains to bind DNA in a flexible configuration. This enables Pbx-Hoxb1 complexes to bind the more stringently spaced Hox-Pbx core sequence. Interaction between Pbx and Hox components occurs via the lock and key mechanism involving the homeodomain of Pbx (pink shaded oval) and the YPWM motif of Hox (blue shaded oval).

(Adapted from Jacobs et al., 1999)
& Morata, 2000), other arthropods (Gonzalez-Crespo & Morata, 1996) and vertebrates (Sellieri et al., 2001). Examples of this widespread-binding model have since been observed in mammals, with the NKL family members NKX2.1 and NKX2.5 both having been shown to regulate target genes directly via single or multiple monomeric sites (Ramirez et al., 1997, Sepulveda et al., 1998, Kim et al., 2002). In addition, the ability of HOX proteins, particularly those from NKL and POU families to form functional homodimers, provides yet another mechanism by which HOX proteins can regulate target gene expression. For example NKX2.5 homodimerises on the atrial natriuretic factor promoter (Kasahara et al., 2001) and PIT1 regulates promoters controlling the expression of the prolactin and growth hormone genes via the formation of functional homodimers (Day et al., 1998).

1.2.3 Homeobox Genes – Roles in Haematopoiesis and Leukaemogenesis

The process of haematopoiesis involves the generation of highly specialised effector cells constituting the lymphoid and myeloid lineages, from a small pool of multipotent, self-renewing HSCs. A complete understanding of the genes that play a definitive role in the initiation and maintenance of this tightly regulated process, is crucial in order to gain further insight into disrupted mechanisms underlying diseases of the blood. In addition to the role of growth factors in signalling pathways directing haematopoiesis, a number of nuclear transcriptional regulators have been shown to participate directly in haematopoietic processes of stem cell-self renewal, lineage commitment, progenitor cell expansion and terminal differentiation. These include genes belonging to Notch, Wnt and Hox families (Austin et al., 1997, Carlesso et al., 1999, van Oostveen et al., 1999). In particular, members of the Hox family of transcription factors have emerged as strong candidates as regulators of haematopoiesis, following studies demonstrating that these genes are expressed in haematopoietic progenitors and are capable of affecting the proliferative, differentiative and phenotypic characteristics of haematopoietic cells (Shen et al., 1989, Kongsuwan et al., 1998).

The expression of Class I HOX genes from paralogous groups A, B and C has been detected in primary human haematopoietic cells, with enhanced expression of genes from the 3’ regions of the A and B complexes (for example HoxB3) found in the most
primitive subsets of haematopoietic stem cells, whereas genes at the 5’ end of the cluster (e.g. HoxA10) were expressed at equal levels in all the CD34+ populations tested (Sauvageau et al., 1994). Given that these genes were no longer detectable in CD34-fractions, it appears that the expression of specific HOX genes is required during different stages of blood cell progenitor development and that commitment to myeloid or erythroid lineages is accompanied by global downregulation of HOX gene expression (Sauvageau et al., 1994, Look, 1997, Pineault et al., 2002). This ordered 3’ to 5’ wave of HOX gene activation has also been observed during T-cell activation (Care et al., 1994) and bears striking resemblance to the 3’ to 5’ sequential activation of HOX genes along the AP axis in embryogenesis and in retinoic acid treated EC cells. It has been suggested that the expression of Hox genes in primitive bone marrow sub-populations is required for the enhanced proliferation and self-renewal capacity of these cells, perhaps by inhibiting differentiation. In support of this hypothesis, overexpression of certain Hox genes in bone marrow is associated with proliferative syndromes demonstrating various differentiation blocks.

In contrast to normal haematopoiesis, the involvement of HOX genes from all clusters has been demonstrated in leukaemic cells (reviewed by Magli et al., 1997). The role of individual HOX genes in regulating the proliferation and differentiation of specific haematopoietic lineages is supported by various studies in which perturbations in normal haematopoiesis were observed following modulation of Hox gene expression. This occurred via the overexpression of specific genes in transgenic mouse models or from retroviral vectors in murine haematopoietic stem cells and targeted disruption of genes in knockout mouse models or by using antisense oligonucleotides (Lawrence et al., 1997, reviewed by Buske & Humphries, 2000). Initial evidence for the role of HOX genes in haematopoiesis stemmed from the discovery that Hoxb8 is dysregulated in myeloid leukaemia cells (Blatt et al., 1988), and it was subsequently revealed that overexpression of Hoxb8 in murine bone marrow causes leukaemia in irradiated hosts (Perkins et al., 1990). It has since been shown that deregulated expression of a number of homeobox genes including HOXA10, HOXB3 and HOXA9 results in the disruption of normal transcriptional pathways directing haematopoiesis, in some cases facilitating leukaemic transformation. For example, over-expression of Hoxa10 leads to disrupted myeloid and B-cell differentiation, and the selective expansion of megakaryocytic
progenitors, ultimately leading to acute myeloid leukaemia (AML) (Thorsteinsdottir et al., 1997). Similarly, mice over-expressing Hoxb3 exhibit disrupted B-cell development and develop myeloproliferative disease (MPD) (Savageau et al., 1997), whilst over-expression of Hoxa9 results in perturbations in T-cell development and AML (Izon et al., 1998).

The deregulation of specific HOX genes via chromosomal translocations has been directly implicated in human AML. Specifically, HOXA9 and HOXD13 are involved through the t(7;11) and t(2;11) translocations respectively, and there are numerous examples of cancers which exhibit alterations in HOX gene expression (Table 1.2). Taken together these data demonstrate that the lineage- and stage-specific expression of various combinations of HOX genes is required for normal haematopoietic development. Future investigations in this area are likely to involve more detailed analysis of expression patterns of HOX genes during haematopoietic differentiation, the level and significance of functional redundancy between related members in controlling haematopoietic processes and the specific genes regulated by HOX proteins and associated cofactors.

1.3 HOX11

1.3.1 The Normal Role of HOX11

1.3.1.1 The HOX11 Family

HOX11 belongs to a subfamily of divergent homeobox genes including HOX11L1/Ncx/Enx and HOX11L2/Rnx, that are characterised by an unusual threonine in place of the more common isoleucine or valine at position 47, within the highly conserved third helix of the homeodomain (Dear et al., 1993, Hatano et al., 1997, Cheng & Mak, 1993). Like other homeodomain proteins, the HOX11 family are involved in the regulation of cellular growth and differentiation. Given the high degree of sequence conservation between these members in the homeodomain and in motifs found in known transcription factors, it is likely that this family functions similarly at the molecular level (Allen et al., 1991, Lawrence & Largman, 1992, Higashijima et al., 1992, Cheng & Mak, 1993). Analysis of the expression pattern of Hox11, Hox11L1 and
<table>
<thead>
<tr>
<th>Gene</th>
<th>Cancer</th>
<th>Alteration</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$PAX2$</td>
<td>Wilms’</td>
<td>OE</td>
<td></td>
</tr>
<tr>
<td>$PAX3$</td>
<td>Rhabdomyosarcoma</td>
<td>trans</td>
<td></td>
</tr>
<tr>
<td>$PAX5$</td>
<td>Glioblastoma</td>
<td>OE</td>
<td></td>
</tr>
<tr>
<td>$PAX7$</td>
<td>Rhabdomyosarcoma</td>
<td>trans</td>
<td></td>
</tr>
<tr>
<td>$PAX8$</td>
<td>Wilms’ thyroid</td>
<td>OE</td>
<td></td>
</tr>
<tr>
<td>$HOXA1$</td>
<td>Breast</td>
<td>OE</td>
<td>Chariot &amp; Castronovo (1996)</td>
</tr>
<tr>
<td>$HOXA9$</td>
<td>Myeloid Leukaemia</td>
<td>trans</td>
<td></td>
</tr>
<tr>
<td>$HOXA10$</td>
<td>Myeloid Leukaemia</td>
<td>OE</td>
<td></td>
</tr>
<tr>
<td>$HOXB5$</td>
<td>Renal</td>
<td>UE</td>
<td></td>
</tr>
<tr>
<td>$HOXB7$</td>
<td>Colorectal</td>
<td>mRNA size</td>
<td></td>
</tr>
<tr>
<td>$HOXB8$</td>
<td>Myeloid Leukaemia</td>
<td>OE</td>
<td>Lawrence et al. (1996)</td>
</tr>
<tr>
<td>$HOXC11$</td>
<td>Wilms’</td>
<td>OE</td>
<td></td>
</tr>
<tr>
<td>$HOXD4$</td>
<td>Renal, colorectal</td>
<td>mRNA size</td>
<td></td>
</tr>
<tr>
<td>$HOX11$</td>
<td>T-cell leukaemia</td>
<td>trans</td>
<td>Hatano et al. (1991)</td>
</tr>
<tr>
<td>$HOX11L2$</td>
<td>T-cell leukaemia</td>
<td>trans</td>
<td>Bernard et al., (2001)</td>
</tr>
<tr>
<td>$HSIX1$</td>
<td>Breast</td>
<td>OE</td>
<td>Ford et al. (1998)</td>
</tr>
<tr>
<td>$Cdx1, 2$</td>
<td>Colorectal</td>
<td>UE</td>
<td>Mallo et al. (1997)</td>
</tr>
<tr>
<td>$GBX2$</td>
<td>Prostate</td>
<td>OE</td>
<td>Gao et al. (1996)</td>
</tr>
<tr>
<td>$PBX1$</td>
<td>Pre-B ALL</td>
<td>trans</td>
<td>Nourse et al. (1990)</td>
</tr>
</tbody>
</table>

**Table 1.2. Involvement of HOX Genes in Cellular Transformation.** (OE; overexpressed, UE; underexpressed, trans; translocated).

(Taken from Ford, 1998)
Hox11L2 in early mouse and Xenopus development reveals significant overlap within the cranial ganglia, and since individual knockouts for each of these genes do not demonstrate CNS abnormalities, it is likely that a level of functional redundancy exists between these family members in CNS development (Roberts et al., 1994, Shirasawa et al., 1997, Patterson & Krieg, 1999, Shirasawa et al., 2000). As development proceeds, however, distinct patterns of expression emerge and these functions are exemplified by nullizygous phenotypes. Thus, mice deficient in Hox11 are asplenic, in accordance with high expression in the primitive spleen (Roberts et al., 1994) and Hox11L1−/− mice develop myenteric neuronal hyperplasia and megacolon, consistent with expression in the neuronal subset of neural crest-derived tissues including the dorsal root ganglia, cranial ganglia, sympathetic ganglia, adrenal medulla and enteric ganglia (Shirasawa et al., 1997). By contrast, Hox11L2 knockout mice develop a syndrome resembling congenital central hypoventilation, consistent with expression in the ventral medullary respiratory centre (Shirasawa et al., 2000).

1.3.1.2 The Hox11 Protein

The HOX11 gene encodes a 34kDa homeoprotein that possesses a number of features required of a transcription factor, in that it is predominantly localised to the nucleus and demonstrates both sequence specific DNA-binding and transactivation/repression potential (Dear et al., 1993, Zhang et al., 1996, Owens et al., 2002). Using a random oligonucleotide selection technique, an optimal monomeric binding sequence (TAAGTG) for the full length Hox11 protein was determined in vitro (Tang & Breitman, 1995). The compatibility of this motif with the predicted HOX binding sequence (TAANTG), which is based on the engrailed homeodomain-DNA model, highlights the consistency of homeodomain-DNA interactions, even when the homeodomains share little sequence homology (En and Hox11 homeodomains share 38% protein sequence identity; Tang & Breitman, 1995). Despite this lack of homology, a number of critical DNA contacting residues present in the homeodomain N-terminal arms and the recognition helices of both proteins are conserved including Arg64, Gln350 and Asn351. The Hox11 protein is unusual, however, in that it contains a threonine at position 47 (Thr47) within helix 3 of the homeodomain, in place of the more common isoleucine or valine residues. This suggests the involvement of this residue in
determining a distinct specificity of homeodomain binding by Hox11 and its other family members, since Thr\textsubscript{47} contacts DNA via a hydrophilic hydroxyl group in its side chain in addition to the hydrophobic methyl group present in other Hox proteins (Dear \textit{et al.}, 1993, Tang & Breitman, 1995)(Figure 1.8).

\begin{figure}
\centering
\includegraphics[width=0.7\textwidth]{hox11_homeodomain_dna.png}
\caption{Prediction of the HOX11 Homeodomain Recognition Motif by Analogy to the DNA Base Contacts Proposed in the X-ray Crystal Model of the En Homeodomain-DNA Complex. According to the En homeodomain-DNA model (left), conserved residues within the homeodomain play an important role in DNA recognition including Arg\textsubscript{5} to thymine at the first position, Arg\textsubscript{3} to the complementary thymine at the second base pair, Asn\textsubscript{51} to adenyl at the third position, Ile\textsubscript{47} to thymine at the fourth position and either Gln\textsubscript{50} or Lys\textsubscript{50} to the TG or CC dinucleotide following the TAAT core. In the HOX11 homeodomain (right), the substitution of Thr\textsubscript{47} for Ile\textsubscript{47}, affects the DNA binding specificity. (Adapted from Tang & Breitman, 1995)}
\end{figure}

\textit{In vitro} reporter assays investigating the transactivation potential of Hox11, revealed that optimal transactivation of reporter genes through promoters in both yeast and mammalian cells requires the involvement of three distinct domains, including a glycine-proline rich region at the NH\textsubscript{2} terminus, the homeodomain and a glutamine rich region at the COOH terminus (Zhang \textit{et al.}, 1996)(Figure 1.9). The importance of these domains in the HOX11-dependent transcriptional activation of an endogenous target gene, \textit{HOX11 dependent gene-1} (Hdg-1/Aldh1a1) in NIH 3T3 cells, was subsequently tested by two independent groups (Masson \textit{et al.}, 1998, Owens \textit{et al.}, 2002). The role of the homeodomain and the glycine-proline rich NH\textsubscript{2} terminus, in particular the first 50 amino acids which harbour a conserved FIL motif (Jagla \textit{et al.}, 2001), in the transcriptional activation of Hdg-1 by HOX11, was confirmed by both groups (Masson \textit{et al.}, 1998, Owens \textit{et al.}, 2002). The contribution of the COOH-terminus domain to
transactivation remains unclear however, since deletion of this region had only a minor effect on *Aldh1a1* induction in studies performed by Masson and co-workers (1998), whereas deletion of this region by Owens and co-workers (2002) abrogated *Aldh1a1* expression.

**Figure 1.9.** The HOX11 Homeodomain Protein Featuring a Number of Regions Contributing to the Transactivation Potential of HOX11.

HOX11 appears not to be a simple activator of transcription since it is also capable of repressing transcription in a DNA binding-independent manner (Owens *et al.*, 2003). Structure-function analyses, in which the ability of HOX11 mutants to repress basal transcription from the SV40 early promoter in HeLa cells was assessed, revealed that the NH₂ and COOH terminal domains of HOX11, that also contribute to the transactivation of *Aldh1a1* were likewise necessary for transcriptional repression (Owens *et al.*, 2003). However, although HOX11-mediated repression required a structurally intact homeodomain, point mutations within helix 3 (T₄₁I, K₅₅Q) affecting the DNA binding specificity of HOX11, had no significant effect on repressor activity, suggesting that repression may be mediated via protein-protein interactions as opposed to sequence-specific DNA binding (Owens *et al.*, 2003). In contrast, these point mutations, which change the *in vitro* HOX11 DNA target sequence from TAAC/TAAT to TAAT alone, significantly affected the ability of HOX11 to induce *Aldh1a1* gene expression (Owens *et al.*, 2003). Taken together, these results suggest that the activation and repression functions of HOX11 may occur via distinct mechanisms, in which the transactivation of *Aldh1a1* in NIH 3T3 cells requires DNA binding involving a TAAC-type DNA sequence, whereas repression may occur in the absence of homeodomain-specific DNA binding, possibly via protein-protein interactions with members of the
basal transcriptional complex, by competing with factors for a common binding site or by quenching (Owens et al., 2003).

In accordance with other Hox family members, HOX11 also possesses a functional FPWME PBX-interaction motif (PIM), related to the Y/F-PWM-K/R pentapeptide motif of Class I HOX proteins, required for cooperative interactions with Pbx/exd proteins. However, this motif did not appear to be required for activation of Aldh1al or for inhibition of promoter activity, implying that if HOX11 does require interaction with a cofactor, other protein interaction motifs must be involved (Masson et al., 1998, Owens et al., 2003). In addition to these studies, which were designed to dissect the role of functional domains of HOX11 required for transcriptional activation and/or repression of downstream targets, the involvement of these regions in HOX11-mediated immortalisation was also assessed using in vitro haematopoietic progenitor immortalisation assays (Owens et al., 2003). These experiments highlighted differences between transcriptional pathways resulting in the activation of Aldh1al and those contributing to the ability of HOX11 to give rise to IL-3 dependent cell lines, since deletion of the NH2 terminal 50 amino acids and the Gln-rich domain near the COOH terminus abrogated Aldh1al induction, but not the immortalising capacity of HOX11 (Owens et al., 2003). In addition, although Aldh1al induction and immortalisation by HOX11 required DNA binding activity, the T47 point mutation only affected the ability of HOX11 to transactivate Aldh1al, suggesting distinct mechanisms for HOX11-mediated transactivation (via a ‘TAAC’ pathway) as opposed to immortalisation (via a ‘TAAT’ pathway)(Owens et al., 2003). Finally, the involvement of the PIM in HOX11-mediated immortalisation, and associated dispensability in Aldh1al transactivation, further supports the distinction between these transcriptional pathways and highlights the multiple activities of HOX11 in the regulation of target genes controlling processes of hematopoietic differentiation, progenitor cell immortalisation and leukaemogenesis (Owens et al., 2002). It should be noted, however, that although Aldh1al was identified as a gene transcriptionally upregulated by HOX11 in NIH 3T3 cells, the effect of HOX11 on ALDH1A1 expression in the context of human T-ALL remains unknown, and as such mechanisms of regulation derived from studies by Owens et al., (2003) may not be applicable to models of leukaemogenesis involving modulation of ALDH1A1 expression. Moreover, the relevance of an immortalisation assay whereby HOX11
induces the development IL-3 dependent cell lines with an immature myeloid phenotype, as the model system to discriminate the regulation of ALDH1A1 by HOX11 from HOX11-mediated leukaemogenesis is questionable. These discrepancies highlight the requirement for a model that accurately reflects the tumorigenic activity of HOX11 in T-cells.

1.3.1.3 HOX11 and Cofactors

In concordance with other members of the HOX protein family, experimental data suggests that HOX11 also interacts with specific members of the TALE homeodomain family in order to direct cell-type specific gene expression. Expression profiling of PBX and MEIS homeobox genes in a panel of murine haematopoietic cell lines and human T-ALL cell lines harbouring the t(10;14) translocation (T-ALL Sil, K3P), revealed that HOX11-expressing murine cells were always associated with expression of at least one PBX member (PBX1, PBX2, PBX3), whereas in the case of the two human T-ALL cell lines, only PBX2 transcripts were detected (Allen et al., 2000). Evidence of a functionally relevant interaction with PBX2 was revealed following the discovery that overexpression of both HOX11 and PBX2 led to reduced contact-dependent growth inhibition in NIH 3T3 cells (Allen et al., 2000). Furthermore, this same group identified the in vitro formation of a trimeric complex involving HOX11, PBX and MEIS1 on a DNA sequence harbouring a HOX/PBX binding site (Allen et al., 2000). This suggested the potential for in vivo regulatory transcription by higher order, multi-component transcriptional complexes involving HOX11, although the role of this complex in leukaemic transformation remains unclear since no expression of MEIS genes was detected in the two T-ALL cell lines tested (Allen et al., 2000). Given that a PBX-containing complex larger than a HOX11/PBX heterodimer was identified in the HOX11 expressing T-ALL Sil and K3P cell lines, Allen et al., (2000) postulated the involvement of a ubiquitously expressed MEIS family member, PREP1 in this complex. Cooperative interactions between HOX11 and members of the PBX homeodomain family may also play important roles in embryonic development as well as in the deregulation of normal growth control mechanisms, since Pbx1 is also required for spleen development. Strikingly, Pbx1\(^{-/-}\) knockout mice recapitulate the asplenic phenotype of nullizygous Hox11 mice, and double heterozygotes of Pbx1\(^{+/-}\) and
Hox11−/+ have reduced spleen size (Roberts et al., 1994, Dear et al., 1995, Selleri & Cleary, 1999).

In addition to members of the TALE family of homeodomain transcription factors, HOX proteins also cooperate with additional cofactors to regulate transcription. For example HOX proteins from 11 paralog groups were shown to bind CBP or p300 through interactions mediated by the homeodomain (Shen et al., 2001). Unlike interactions involving HOX, PBX and MEIS, which regulate transcription via sequence-specific DNA binding, HOX proteins do not bind DNA cooperatively with CBP, but instead inhibit the histone acetyltransferase (HAT) activity, thereby repressing transcription (Shen et al., 2001). The cardiac homeodomain factor Nkx2.5 has also been shown to physically associate with GATA-4 to activate the cardiac alpha-actin (alphaCA) promoter in CV-1 fibroblasts (Sepulveda et al., 1998). With regard to association with non-HOX factors, HOX11 has been shown to cooperatively bind CTF1, a ubiquitously expressed CCAAT-box-binding transcription factor involved in the regulation of a variety of eukaryotic genes (Santoro et al., 1988, Zhang et al., 1999). The identification of consensus CTF1 binding sites in a DNA binding site selection assay using Flag-tagged HOX11, and the subsequent demonstration that CTF1 and HOX11 physically associate in vivo in a HOX11-immortalised haematopoietic precursor cell line HX3, suggests a collaborative role for these proteins in transcriptional regulation (Zhang et al., 1999). Indeed, transduction of recombinant retrovirus expressing an antisense CTF1 cDNA into HX3 cells, resulted in the reduction of colonies when cells were grown in soft agar, suggesting that HOX11 may collaborate with CTF1 to mediate haematopoietic precursor cell immortalisation (Zhang et al., 1999). CTF1 also recruits TFIIB, a member of the preinitiation complex, thus a model in which the inappropriate expression of HOX11 in haematopoietic precursor cells leads to complex formation between HOX11, CTF1 and TFIIB and subsequent deregulation of CTF1 target genes has been postulated (Zhang et al., 1999).

In a completely novel and as yet uncorroborated role, HOX11 has also been identified as a protein partner for the serine-threonine phosphatase 2A (PP2AC) and protein phosphatase 1 (PP1C) catalytic subunits in mammalian cells (Kawabe et al., 1997). PP1 and PP2A are capable of maintaining the G2 meiotic arrest of Xenopus oocytes and
thus, may represent targets for oncogenic transformation (Kawabe et al., 1997). The possibility that HOX11 may disrupt a G2 checkpoint maintained by PP2A and PP1, by allowing cells to proceed inappropriately through M phase in Xenopus oocytes and in the Jurkat T-cell line, suggests a model where HOX11 contributes to genomic instability by altering cell cycle regulation (Kawabe et al., 1997). Support for the possible role of HOX11 in cell cycle progression also comes from studies demonstrating that levels of HOX11 RNA and protein fluctuate during different phases of the cell cycle, with the highest levels coinciding with a G1 restriction point, suggesting that HOX11 may play a role in cell cycle progression into the S phase (Zhang et al., 1993).

### 1.3.1.4 The Role of HOX11 in Development

The role of HOX11 in development has been studied in mice, which express the murine HOX11 orthologue, Tlx1/Hox11 (Raju et al., 1993). Hox11 is expressed in various embryological tissues including the muscle plates of the branchial arches and subsequently in the cranial nerves that innervate them, the developing hindbrain, spinal cord, outflow tracts of the heart, pharynx, genticulate, pancreas, thymus and within a portion of the splenic mesoderm (Cheng & Mak, 1993, Raju et al., 1993, Roberts et al., 1994, Dear et al., 1995). Despite the expression of Hox11 in these tissues, homozygous null mutant mice (Hox11−/−) are only affected in spleen organogenesis and are otherwise phenotypically normal, suggesting that other genes may functionally compensate for the lack of Hox11 in these tissues or that the expression of Hox11 in these areas is merely reminiscent of its homeobox evolutionary origin but has no residual function (Roberts et al., 1994, Hatano et al., 1997, Shirasawa et al., 1997). In support of the former hypothesis, the expression of Hox11/Tlx1, Hox11L1/Tlx2 and Hox11L2/Tlx3 overlaps considerably in the central nervous system, suggesting that a degree of functional redundancy exists between these highly related family members.

Hox11 is initially expressed in the splanchnic mesoderm at embryonic day 11.5 (dE11.5). Continued expression is observed in the developing spleen through dE13.5, but is downregulated thereafter (Roberts et al., 1994). In order to dissect the function of Hox11 in spleen organogenesis, the fate of the cells contributing to the developing spleen was analysed in Hox11−/− mice and Hox11−/+ chimeras (Kanzler & Dear, 2001).
In the absence of *Hox11*, cells of the spleen anlage were unable to develop normally past dE13, and eventually contributed to the pancreatic mesenchyme and the splenic blood vessels, suggesting that although *Hox11* is not required for the formation of the spleen primordium, expression is required for cells to proceed past a specific developmental stage (Kanzler & Dear, 2001). The asplenic phenotype observed in *Hox11*−/− mice is therefore a result of halted spleen morphogenesis and subsequent redirection of *Hox11*−/− splenic precursors to neighbouring organs, as opposed to organ involution (Kanzler & Dear, 2001). The role of *Hox11* in spleen development from dE13 onwards is cell autonomous, since wild-type cells were unable to rescue *Hox11*-null cells from spleen exclusion in *Hox11*−/−+/+ chimeras (Kanzler & Dear, 2001).

*Hox11* null mice displayed normal B-cell and T-cell profiles in thymus, lymph node and peripheral blood tissues and normal red blood cell counts, however, increased numbers of white blood cells, including neutrophils and lymphocytes were observed (Roberts *et al*., 1994). In addition, erythrocytes from *Hox11*−/− mice were found to contain Howell-Jolly bodies, typically found in asplenia.

### 1.3.2 The Abnormal Role of HOX11

#### 1.3.2.1 Expression of HOX11 in Normal T-cells

In addition to the role of HOX11 in spleen organogenesis, it has been postulated that *HOX11* may be expressed in normal lymphocytes (Lu *et al*., 1992, Yamamoto *et al*., 1995). Analysis of *HOX11* mRNA expression in selected tissues and cell lines by quantitative reverse transcriptase polymerase chain reaction (RT-PCR), however, yielded mixed results (Lu *et al*., 1992, Zhang *et al*., 1993, Yamamoto *et al*., 1995). According to studies by Lu *et al*., (1992), *HOX11* mRNA was detected in two leukaemic T-cell lines, Jurkat and HPB-ALL, as well as in normal T-cells - although expression in the latter was significantly lower.

It has been suggested that the low level of *HOX11* expression observed in transformed cell lines such as Jurkat, in which cells expressed <1/100 of the amount of *HOX11* mRNA found in embryonic spleens (dE17.5), may simply be the result of abnormalities
in these leukaemic lines. As such, it may not reflect the physiological conditions of their normal counterparts (Yamamoto et al., 1995). In addition, Yamamoto et al., (1995), were unable to detect HOX11 mRNA in normal T-cells using RT-PCR analysis and HOX11 was not induced in T-cells stimulated with PMA and ionomycin, or following pro-apoptotic stimulation with glucocorticoids, gamma irradiation and anti-FAS, suggesting that HOX11 is not involved in major signal transduction or cell death pathways in normal thymocytes (Yamamoto et al., 1995). Thus, in conjunction with evidence from other independent groups, the consensus view is that HOX11, like several other T-ALL oncogenes, is not normally expressed in T-cells at any stage of development, and that ectopic expression of HOX11 in T-cells is a primary event in the development of T-lineage tumours (Zutter et al., 1990, Hatano et al., 1991, Salvati et al., 1995, Yamamoto et al., 1995).

1.3.2.2 Involvement of HOX11 in T-ALL

The HOX11 proto-oncogene was originally identified from chromosomal translocation breakpoints in T-ALL, which involve a reciprocal exchange between the HOX11 locus at 10q24 and either TCRδ (14q11) or TCRβ (7q35)(Hatano et al., 1991, Kennedy et al., 1991, Dube et al., 1991, Lu et al., 1992). These translocations, which are present in approximately 5-10% of patients with T-ALL, occur within a 15kb breakpoint cluster region at the 5’ end of the HOX11 gene, such that the coding capacity of the gene remains intact, resulting in the inappropriate expression of HOX11 in T-cells (Kennedy et al., 1991). Deregulated expression of HOX11 has since been identified in as many as 30% of T-ALLs, when abnormalities not detectable by cytogenetics are included (Salvati et al., 1995, Kees et al., 2003). One proposed mechanism for the reactivation of HOX11 in T-ALL specimens lacking the translocation, involves CpG demethylation of the proximal HOX11 promoter, possibly in concert with other derepression mechanisms (Watt et al., 2000). Given the precedent for deregulation of transcription factors in T-cell leukaemia by translocation into TCR loci, it is likely that the ectopic expression of HOX11 in T-cells is a crucial event in tumour initiation or progression (McGuire et al., 1989, Mellentin et al., 1989, Begley et al., 1989). The localisation of HOX11 protein in the nucleus of T-cells, in conjunction with the DNA binding and transcriptional transactivating potential of HOX11, suggests that HOX11 may function in vivo as an
oncogenic transcription factor by perturbing normal patterns of gene expression in T-cell tumour precursors, thus leading to the development of leukaemia (Dear et al., 1993, Zhang et al., 1993, Dear et al., 1995).

In order to investigate the mechanism by which a nuclear oncogene like HOX11 can contribute to the genesis of human T-cell tumours, transgenic mice in which the expression of HOX11 was targeted to the thymus using CD2 regulatory elements were created, however no tumours were observed even after an extended time period (20 months)(Greene & Rabbits, unpublished observations). Furthermore, attempts to ectopically express HOX11 in the developing thymocytes of LCK-HOX11 transgenic mice, resulted in an apparent embryonic lethality (Hough et al., 1998). Thus, to date, attempts to accurately recapitulate the tumorigenic activity of HOX11 in T-cells have been unsuccessful. Despite this, several groups have demonstrated that HOX11 harbours the capacity for widespread immortalisation of various haematopoietic precursors, which may be a significant contributory factor in T-cell leukaemogenesis, (Hawley et al., 1994, Hough et al., 1998). For example, constitutive expression of HOX11 in murine haematopoietic precursors blocks the differentiation of IL-3-responsive progenitors, resulting in the generation of myeloid progenitor cell lines (Hawley et al., 1994). DNA binding was shown to be essential for the immortalisation function of HOX11 in this study. This is in keeping with a model whereby the transforming capability of HOX11 involves dysregulation of specific HOX11 target genes involved in the regulation of normal haematopoietic differentiation programmes. Although over-expression of HOX11 resulted in the immortalisation of myeloid progenitors, the inability of MSCV-HOX11 bone marrow transplant recipients to develop leukaemias during the 6-month observation period, highlighted the requirement for secondary mutations in addition to HOX11 over-expression for full acquisition of the malignant phenotype. This might involve, for example, mutations promoting cellular survival or proliferation (Hawley et al., 1994). The tumorigenic activity of HOX11 in lymphoid progenitors was supported by studies in which transgenic mice expressing low levels of HOX11 under the control of the Ig heavy chain regulatory sequence in B-lymphocytes, led to the development of large cell lymphomas (IgM+IgD+ mature B-cells) in the spleen, that were not transplantable into SCID recipients (Hough et al., 1998). The immortalisation capacity of HOX11 in adult bone marrow-derived precursors has also
been extended to embryonic haematopoietic precursors (Yu et al., 2002, Keller et al., 1998). Overexpression of \textit{HOX11} in embryoid bodies (EB), which represent the embryonic or yolk sac stage of haematopoiesis, resulted in the establishment of novel haematopoietic cell lines (EBHX), a number of which were capable of generating primitive and definitive erythroid progeny (Keller et al., 1998). In addition, the ectopic expression of \textit{HOX11} was associated with an immature phenotype in J2E erythroid cells, supporting the concept that the transforming capacity of \textit{HOX11} may stem from its ability to alter haematopoietic cell differentiation (Greene et al., 2002).

Taken together, these results suggest that although the ectopic expression of \textit{HOX11} is clinically associated with the development of T-cell tumours, the immortalisation activity of \textit{HOX11} is non-lineage specific. Indeed, \textit{HOX11} mRNA has also been detected in blast cells from patients with acute myeloid tumours in addition to the erythroleukaemic cell line, K562 (Lu et al., 1992, Hawley et al., 1994). The disruption of haematopoietic differentiation via a \textit{HOX11}-mediated pathway may therefore predispose cells of various lineages to leukaemic transformation, however the specific mechanisms by which \textit{HOX11} induces tumour growth remain enigmatic.

1.4 \textbf{HOX11 TARGET GENES}

1.4.1 Identification of Downstream Target Genes of \textit{HOX11}

The transforming capacity of \textit{HOX11} likely involves the dysregulation of target genes involved in transcriptional networks that normally regulate thymocyte proliferation, differentiation and survival during T-cell development. Three putative \textit{HOX11} target genes have been identified so far, namely \textit{aldehyde dehydrogenase 1a1} (\textit{Aldh1a1}) and the Wilm’s Tumour gene (\textit{Wt1}), which appear to be physiologically regulated by \textit{Hox11} in the developing spleen, and \textit{Four and a half LIM domain protein 1} (\textit{Fhll}), which encodes a LIM domain protein normally highly expressed in skeletal muscle (Greene et al., 1998, Koehler et al., 2000). \textit{Aldh1a1} encodes an enzyme involved in retinoic acid synthesis and was originally identified as a gene activated by \textit{HOX11} in NIH 3T3 cells in a cDNA representational difference analysis (RDA) screen for \textit{HOX11} target genes (Greene et al., 1998). Conversely, \textit{Aldh1a1} appears to be repressed by \textit{Hox11} during
spleen organogenesis since Aldh1a1 mRNA levels are elevated in the spleen anlagen of Hox11-null mouse embryos (Greene et al., 1998). Furthermore, a study of ALDH1A1 expression in cell lines stably transfected with HOX11, revealed that ALDH1A1 is positively regulated by HOX11 in the immature T-cell line, PER-117 and negatively regulated by HOX11 in the human erythroleukaemic cell line, HEL (Greene et al., 1998, Greene unpublished observations). Thus, it appears that HOX11 is a bifunctional transcriptional regulator, capable of both transcriptional activation or repression depending on the cellular context. Subsequently it was shown that Wt1, which is also required for spleen development, represents a physiological target of Hox11, and furthermore, WTI is over-expressed in 75% of ALL cases and is also associated with various other cancer subtypes (Hastie, 1994, Loeb et al., 2001). Although the specific relevance of WTI, ALDH1A1 and FHL1 as targets of HOX11 in neoplastic transformation has yet to be established, several studies suggest functional interplay between highly related gene members during normal development. For example, both WTI and RALDH2/ALDH1A2 are coexpressed during cardiac development in avian epicardially derived cells (EPCDs), and it is postulated that WTI is responsible for maintaining the EPCDs in an undifferentiated, retinoic acid (RA)-synthesising state, required for normal differentiation of the ventricular myocardium (Perez-Pomares et al., 2002). Moreover, the LIM-only coactivator, FHL2, has been shown to cooperate in vivo with WTI to regulate the expression of Mullerian Inhibiting Substance (MIS) and the nuclear receptor DAX1, which control the differentiation of the Mullerian duct during sex determination (Du et al., 2002). A detailed understanding of the network of genes affected by HOX11 is therefore crucial in order to gain insights into the molecular pathways perturbed by HOX11 in the leukaemic transformation process.

1.4.2 ALDH1A1 as a Target Gene of HOX11

1.4.2.1 The Role of ALDH1A1 in Normal Development

1.4.2.1.1 Retinoid Metabolism

Vitamin A (retinol) is the parent compound of a family of signalling molecules (retinoids) which are responsible for regulating diverse aspects of cellular
differentiation, proliferation, apoptosis, reproduction and embryonic development, through its biologically active metabolite, RA (reviewed by Napoli, 1999, Maden, 1999, Haselbeck et al., 1999). The process of retinoid metabolism begins with the acquisition of retinol, in the form of retinyl esters (animal fat) or as a provitamin carotenoid (plants), which are mobilised from the liver to target cells via high affinity retinol binding proteins (RBP)(Gaines & Berliner, 2003). Upon internalisation, retinol is bound to the cellular retinol binding proteins (CRBP, CRBPII) and is subsequently converted to the biologically active derivatives, retinal and RA, which are released by CRBP in order to enter the nucleus. The pleiotropic effects exerted by RA (which include the structural variants, all-trans-RA (ATRA) and 9-cis-RA) in various cell types may then occur via the regulation of gene transcription mediated by a family of ligand-dependent receptors, which include the retinoic acid receptors (RAR) and retinoid X receptors (RXR)(McCaffery & Drager, 1995, Maden, 1999, Haselbeck et al., 1999, Collins, 2002). These nuclear receptors, which comprise a DNA-binding domain (DBD) connected to a high affinity ligand-binding domain (LBD) which also harbours an activation function (AF-2) domain, form transcriptionally active heterodimers, that recognise specific RA Responsive Elements (RAREs) located in the regulatory regions of RA-induced target genes (Bourguet et al., 2000, Gaines & Berliner, 2003). Upon binding of cognate ligand (RAR can be activated by all-trans RA or 9-cis RA, whereas RXRs are activated by 9-cis RA alone), these receptors undergo conformational changes that facilitate the release of co-repressors, and subsequent interaction with transcriptional co-activators and members of the basal transcriptional machinery in order to activate gene transcription (Torchia et al., 1998)(Figure 1.10). Thus, in the absence of ligand binding, the RAR/RXR heterodimer recruits nuclear co-repressors NcoR (Nuclear receptor CoRepressor) and SMRT (Silencing Mediator of Retinoid and Thyroid Hormone Receptor), which in turn bind histone deacetylases (HDACs), leading to the establishment of repressive chromatin structure; whereas in the case of RA binding, the recruitment of co-activators and associated histone acetyl transferases (HATs) is favoured, resulting in transcriptionally permissive chromatin structure and gene activation (Chen & Evans, 1995, Chen et al., 1997, reviewed by Xu et al., 1999). The existence of three receptor subtypes (RAR/RXRα, β and γ), each of which contains multiple isoforms generated by alternative splicing or utilisation of different promoters, increases the potential for various heterodimeric combinations responsible for complex
RA-mediated target gene regulation (Chambon, 1996). Given the complexity of receptor family members and the differential spatio-temporal expression patterns of these receptors in various cell types, it is therefore not surprising that the focus in unravelling the diversity of effects mediated by RA, has concentrated primarily on aspects involving RAR/RXR signalling (Maden, 1999).

![Diagram of Retinoic Acid Action](image)

**Figure 1.10. Mechanism of Retinoic Acid Action.** (A) In the absence of retinoic acid (RA; all-trans-RA, 9-cis-RA, didehydro-RA or 4-oxo-RA), the RAR/RXR heterodimer recruits nuclear co-repressors, which bind histone deacetylases (HDACs), leading to transcriptional repression. (B) Upon ligand binding, gene transcription is initiated by the recruitment of co-activators and associated histone acetyl transferases (HATs). The various RAR and RXR subtypes (α, β and γ) are encoded by distinct genes and encode multiple isoforms via alternative splicing, resulting in 48 possible heterodimeric combinations, accounting for the diversity of RA-mediated effects. In addition, RXR is also a partner for other receptors belonging to the steroid/retinoid/vitamin D/thyroid hormone superfamily.

(Taken from Napoli, 1999)

The discovery by Niederreither et al., (1999) that a single mutation affecting the RA synthesising enzyme, RALDH2 in embryonic mice yielded a phenotype bearing striking
similarities to those described for RAR/RXR knockouts, emphasised the importance of retinoid metabolism in contributing to the local diversity in RA actions. Such aspects include the processes of retinol circulation and cellular uptake mediated by CRBPs, bioavailability of RA as regulated by cellular retinoic acid binding proteins (CRABPs), mechanisms of RA degradation and of particular interest within the scope of this project, the control of RA biosynthesis (Leid et al., 1992, McCaffery & Drager, 1995, Maden, 1999).

1.4.2.1.2 Aldehyde Dehydrogenases (ALDHs)

Aldehyde dehydrogenases (ALDHs) are a group of enzymes present in organisms ranging from bacteria to humans, that are responsible for catalysing the pyridine nucleotide-dependent oxidation of aldehydes to their corresponding carboxylic acids (Yoshida et al., 1998, Sladek, 2002). As a general rule, NAD serves as the hydrogen acceptor in this essentially irreversible reaction, although in the case of some enzymes, NADP may perform this function (Figure 1.11)(Sladek, 2003).

\[
\text{O} \quad \text{ALDH} \quad \text{R-C-H} \quad \text{ALDH} \quad \text{R-C-OH}
\]

\[
\text{NAD(P)}^+ \quad \text{NAD(P)H}
\]

**Figure 1.11. Aldehyde Dehydrogenases (ALDHs) Catalyse the Oxidation of Aldehyde to Carboxylic Acid.** ALDH exists as an active enzyme in dimeric form, with each subunit harbouring an NAD(P)-binding pocket, a catalytic and an arm-like bridging domain.

(Taken from Liu et al., 1997)

In mammals, the general detoxification of aldehydes, in addition to other specific catalytic functions are performed by members of the human aldehyde dehydrogenase superfamily, which to date, comprise seventeen enzymes (Sophos et al., 2001, Sladek, 2003)(Table 1.3). These enzymes, which are critical for normal biological development and physiological homeostasis, are predominantly located in the cytosol and mitochondrial compartments and can be broadly sub-divided into two groups. The first
<table>
<thead>
<tr>
<th>Class</th>
<th>Family</th>
<th>Gene Name</th>
<th>Trivial Name</th>
<th>Chromosomal Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>ALDH1</td>
<td>ALDH1A1</td>
<td>ALDH1/RALDH1</td>
<td>9q21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALDH1A2</td>
<td>ALDH1A7/RALDH2</td>
<td>15q21.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALDH1A3</td>
<td>ALDH6/RALDH3</td>
<td>15q26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALDH1B1</td>
<td>ALDH5/ALDHx</td>
<td>9p13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALDH1L1</td>
<td>FOLATEDH</td>
<td>3q21.3</td>
</tr>
<tr>
<td>Class 2</td>
<td>ALDH2</td>
<td>ALDH2</td>
<td>ALDH1/ALDH2</td>
<td>12q24.2</td>
</tr>
<tr>
<td>Class 3</td>
<td>ALDH3</td>
<td>ALDH3A1</td>
<td>ALDH3/ALDH-III</td>
<td>17p11.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALDH3A2</td>
<td>ALDH5/ALDH10</td>
<td>17p11.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALDH3B1</td>
<td>ALDH-7</td>
<td>11q13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALDH3B2</td>
<td>ALDH-8</td>
<td>11q13</td>
</tr>
<tr>
<td>Class 4</td>
<td>ALDH4</td>
<td>ALDH4A1</td>
<td>ALDH-4</td>
<td>1p36</td>
</tr>
<tr>
<td>Class 5</td>
<td>ALDH5</td>
<td>ALDH5A1</td>
<td>E5/SSDH</td>
<td>6p22</td>
</tr>
<tr>
<td>Class 6</td>
<td>ALDH6</td>
<td>ALDH6A1</td>
<td>MMSDH</td>
<td>14q24.3</td>
</tr>
<tr>
<td>Class 7</td>
<td>ALDH7</td>
<td>ALDH7A1</td>
<td>ATQ1</td>
<td>5q31</td>
</tr>
<tr>
<td>Class 8</td>
<td>ALDH8</td>
<td>ALDH8A1</td>
<td>ALDH-12</td>
<td>6q24.1-25.1</td>
</tr>
<tr>
<td>Class 9</td>
<td>ALDH9</td>
<td>ALDH9A1</td>
<td>ALDH-9/E3</td>
<td>1q22-q23</td>
</tr>
<tr>
<td>Class 18</td>
<td>ALDH18</td>
<td>ALDH18A1</td>
<td>P5cS</td>
<td>10q24.3-q24.6</td>
</tr>
</tbody>
</table>

(Compiled from Sladek, 2003)

Table 1.3. Nomenclature and Chromosomal Locations of Human Aldehyde Dehydrogenases.
category comprise ‘dedicated’ enzymes that are crucial for the relatively specific conversion of a single endobiotic to an essential metabolite in the absence of insult. This includes ALDHs 1A1, 1A2, 1A3, 8A1, 1L1, 3A2, 4A1, 5A1, 6A1, 9A1 and 18A1). By contrast, the second group encapsulates those enzymes which are relatively non-substrate specific and provide protection against the plethora of potentially cytotoxic, mutagenic, genotoxic and carcinogenic aldehydes encountered as either physiologically derived intermediates of metabolism (e.g. during lipid peroxidation) or when the organism encounters a hostile environment (e.g. ultraviolet (UV) light, toxic chemicals). Enzymes in this category include ALDH3A1 and ALDH5A1 (Lui et al., 1997, Sladek, 2003).

It should be noted, however, that exceptions to this general classification scheme do exist, such as in the case of ALDH1A1, which catalyses the oxidation of retinaldehyde to retinoic acid (a bioactivation function necessary for normal development), in addition to performing a detoxification function by catalysing the oxidation of the anti-cancer pro-drug intermediates aldophosphamide and aldoifosfamide (Sladek, 2003). In addition, a number of ALDH enzymes remain to be categorised. This is because their identification as members of the ALDH superfamily in the wake of the Human Genome sequencing project preceded a complete understanding of the physical properties, tissue distribution, sub-cellular localisation and substrate/co-factor preferences demonstrated by these enzymes (Sladek, 2003).
1.4.2.1.3 ALDH1A1 – A Crucial Regulator of Retinoic Acid Biosynthesis During Embryogenesis and Adulthood

The biosynthesis of RA from retinol involves two sequential steps (Figure 1.12).

Figure 1.12. Biosynthesis of Retinoic Acid. The reversible interconversion of retinol to retinal is catalysed by members of the alcohol dehydrogenase (ADH) and short-chain dehydrogenase/reductase (SDR) enzyme family, whereas the members of the ALDH and cytochrome P450 families (CYP450) catalyse the irreversible oxidation of retinal to RA.

The first step encompasses the reversible oxidation of retinol to retinaldehyde, and is catalysed by either cytosolic retinol dehydrogenases (members of the alcohol dehydrogenase (ADH) family) or by microsomal retinol dehydrogenases (members of the short-chain dehydrogenase/reductase (SDH) family)(Haselbeck et al., 1999). The rate-limiting second step however, involves the irreversible oxidation of both all-trans and 9-cis retinal to their corresponding RA isomers, and is catalysed by a select group of enzymes within the family of ALDHs, namely human Class 1 Aldehyde Dehydrogenase 1A1 (ALDH1A1), 1A2 (ALDH1A2), 1A3 (ALDH1A3) and ALDH8A1, in addition to members of the cytochrome P450 family (CYP450) including P4501A1 and 1A2 (Duester, 1996, Raner et al., 1996, Sladek, 2003). In mice these activities are performed by the murine homologs, which include Aldh1a1 (Raldh1), Aldh1a2 (Raldh2) and Aldh1a3 (Raldh3)(Zhao et al., 1996, Haselbeck et al., 1999).

ALDH1A1 is a cytosolic enzyme of approximately 54.7kDa, that exists as a homotetrameric present in a wide range of human foetal and adult tissues including the liver, kidney, erythrocytes, skeletal muscle, lung, breast, lens, stomach mucosa, brain,
pancreas, testis, prostate and ovary (Sladek, 2003). Analysis of the crystal structure of ALDH1A1 isolated from sheep liver, revealed the presence of a unique 'substrate tunnel' capable of specifically binding retinaldehyde (Moore *et al.*, 1998). Further studies revealed that this channel was also present in another RA synthesising enzyme, ALDH1A2 but absent in ALDH2 and ALDH3A1 enzymes, suggesting that the major role for ALDH1A1 is in the biosynthesis of RA - the tightly controlled synthesis of which is crucial to ensure correct gene expression at appropriate developmental stages (Yoshida *et al.*, 1992, Lamb *et al.*, 1999).

During embryogenesis, RA gradients are likely to play a crucial role in AP patterning by regulating members of the *Hox* gene family, which are RA-inducible and selectively expressed in posterior tissues (trunk and posterior hindbrain) along the AP axis (Haselbeck *et al.*, 1999). In addition to this role in AP patterning, RA is also involved in other aspects of embryonic development and co-localisation studies examining the distribution of RA and RA-synthesising enzymes in mice, reveal distinct roles for two different enzymes in RA synthesis. The first enzyme, Aldh1a2, demonstrates a high catalytic efficiency for RA synthesis and is the predominant RA synthesising enzyme in the trunk and frontonasal region of mid-gestation mouse embryos; whereas the second enzyme, Aldh1a1, which is of particular interest in the context of this project, displays a lower catalytic activity and is important for regulating low-level RA synthesis in the cranial tissues including the optic vesicles, ventral mesencephalon, otic vesicles and thymic primordia (Maden *et al.*, 1996, Dickman *et al.*, 1997, Haselbeck *et al.*, 1999, Niederreither *et al.*, 1999). Further studies in *Xenopus* demonstrate that the synthesis of RA by ALDH1A1 is crucial during embryogenesis for the development of certain anterior structures including the olfactory placodes, dorsal retina and lens placode as well as for trunk structures, which include the pronephros and pronephric duct (Ang & Duester, 1999). A number of adult tissues are also dependent on ALDH1A1 for RA-mediated transcriptional regulation, including the skin (keratinocyte differentiation), testes (spermatogenesis), brain, liver and kidney (McCaffery & Drager, 1995) and the specific expression of Aldh1a1 in adult mouse organs suggests that this enzyme may also function as the predominant regulator of RA synthesis in adult organisms (Haselbeck *et al.*, 1999).
The conserved expression patterns of the various RA synthesising enzymes reflects the importance for strict regulation of RA synthesis required for normal embryo morphogenesis and distinct adult functions (Ang & Duester, 1999, Russo et al., 2002). Indeed, the requirement for such fine tuning in the synthesis of this highly potent, lipophilic signalling molecule is exemplified by the pleiotrophic effects on embryonic limb development observed following the application of ATRA at different stages of embryonic development. When pharmacological doses of ATRA were administered 5.5 days postcoitum (dpc), limb duplication was observed, however if ATRA was applied later in development at 10 and 12.5dpc limb development was stunted, further emphasising the requirement for correct spatio-temporal regulation of RA synthesis (Kochhar, 1980, Niederreither et al., 1996).

In addition to its function as a RA synthesising enzyme, ALDH1A1 also plays a role in alcohol detoxification. Additional detoxification roles for ALDH1A1 include the elimination of peroxidic aldehydes produced in the eye lens, as well as protection of human and murine cells against the anti-cancer agents cyclophosphamide and ifosfamide which were commonly used in tumour cell purging regimens during bone marrow transplantation (Yoshida et al., 1998, Sladek, 2003).

1.4.2.1.4 The Role of Retinoid Signalling in Haematopoiesis

Haematopoiesis involves the continuous generation of blood forming cells, which include granulocytes, macrophages, platelets and red blood cells (RBCs), as well as cells comprising the immune system, which include B-cells, T-cells, dendritic cells (DC) and natural killer cells (NK), from a limited pool of pluripotent HSCs (Reya et al., 2001, Russo et al., 2002). Interest in the processes governing normal HSC development, particularly from the perspective of gaining insights into mechanisms governing ‘cancer stem cell’ self renewal, has led to the identification of specific growth factors and transcriptional regulators required for the survival, differentiation and proliferation of HSCs (Metcalf, 1998, Herault, 2002). A number of transcription factors are exclusively expressed in haematopoietic cells, and roles for SCL (Shivdasanl et al., 1995), LMO2 (Warren et al., 1994), PU.1 (McKercher et al., 1996), GATA-1 (Weiss et al., 1994), E2A (Zhuang et al., 1994), C/EBP and Ikaros (Georgopoulos et al., 1994), in lineage
commitment during haematopoiesis have been proposed following gene targeting experiments in mouse embryonic stem cells and by generating knockout mice, despite the inherent limitations of such studies (e.g. lack of phenotype due to functional redundancy, loss of function resulting in early effects thereby obscuring functional roles later in development)(Figure 1.13).

![Figure 1.13. Key Transcription Factors Involved in Lineage Commitment During Haematopoiesis. HSC, pluripotent haematopoietic stem cell, Lym, lymphoid progenitor, M/E, myeloerythroid progenitor.](adapted from Orkin, 1995).

In addition to these lineage-specific transcription factors, the retinoic acid receptors, RAR and RXR, function as ligand inducible transcriptional regulators within retinoid signalling pathways that have established roles in haematopoiesis. Gene targeting experiments designed to dissect the specific involvement of various RAR subtypes in haematopoiesis, reveal crucial roles for RARα and RARγ isoforms in granulopoiesis, since mice lacking both receptors display a block to granulocytic differentiation (Labrecque et al., 1998). The role of retinoids and RARs in myelopoiesis has been extensively studied, following the discovery that pharmacological doses [1μM] of ATRA are able to induce the terminal differentiation of immature, malignant promyelocytes which accumulate in the bone marrow and peripheral blood of patients
with acute promyelocytic leukaemia (APL) (Collins, 2002). APL is a subtype of acute myeloid leukaemia (AML), in which patients typically harbour the t(15;17) chromosomal translocation (~90%), resulting in the production of a hybrid fusion protein, comprising the DNA/ligand binding and transactivation portion of the retinoic acid receptor alpha (RARα)(17q21) and the transactivating N-terminal region of the promyelocytic leukemia (PML) transcription factor (15q22)(Kakizuka et al., 1991). This chimaeric oncprotein acts by antagonising wild type RARα and constitutively repressing RARα target genes via the recruitment of histone deacetylases that are not released at physiological concentrations of ATRA (Grignani et al., 1998). In addition to these clinical observations, studies in which murine bone marrow cells were transduced with a mutant RARα demonstrating dominant negative activity, generated myeloid cell lines stunted at various stages of differentiation, providing further evidence that retinoid signalling may also be involved in normal granulopoiesis (Collins, 2002).

The effects of ATRA on the development of various haematopoietic lineages from pluripotent stem cells appears complex, with some studies reporting that ATRA enhances the growth of erythroid and myeloid progenitors, as well as the long- and short-term repopulating activity of cells derived from lin^-c-kit^-Sca-1^+ precursors; while other reports suggest an inhibitory role for ATRA in HSC proliferation and differentiation (Douer & Koeffler, 1982, Smeland et al., 1994, Purton et al., 1999). Although these varied effects of RA on haematopoiesis may be due to the pleiotrophic effects of RA on cultured target cells present at varying maturational stages, Tocci and co-workers (1996) also demonstrate the dual action of RA on human foetal haematopoietic progenitor cells. In keeping with other related studies, this group demonstrate that ATRA and 9-cis RA inhibits the proliferative capacity of HSCs and modulates the growth of purified HSCs by redirecting cells from the multipotent/erythroid, monocytic to the granulocytic-neutrophilic differentiation lineage programme (Tocci et al., 1996). To add a further layer of complexity, the effects of RA on the survival of adult human marrow CD34^+ stem cells also appear to be dependent on the presence of growth factors, such that in the presence of growth factors, ATRA promotes apoptosis of CD34^+ bone marrow (BM) cells, whereas ATRA exerts a protective effect on CD34^+ cells deprived of growth factors and hence subject to apoptotic stress (Herault et al., 2002). Although the exact mechanisms by which ATRA
is able to exert these opposing effects have yet to be clarified, it is clear that ATRA alone has a direct effect on the survival of adult HSCs, and in keeping with the notion that RA may regulate HSC development in vivo, highly elevated levels of ALDH1A1 have been documented in these multipotent progenitors (Kastan et al., 1990, Jones et al., 1995, Storms et al., 1999).

The role of retinoic acid signalling in lymphopoiesis, is less well documented, however a number of studies report that physiological levels of retinoids inhibit the proliferation of normal peripheral blood B-cells and B-cell precursors from bone marrow, by inhibiting cell cycle machinery responsible for the progression of cells from G1 into the S phase (Blomhoff et al., 1992, Fahlman et al., 1995, Lomo et al., 1998, Naderi & Blomhoff, 1999). Progression through the cell cycle involves the phosphorylation of the retinoblastoma protein (pRB) by cyclin D, E and A- associated kinases, thereby stimulating the release of transcriptional regulators like E2F from complexes with pRB, facilitating the transcription of S-phase genes (Weintraub et al., 1992). ATRA inhibits B-cell proliferation by rapidly inducing dephosphorylation of pRB, through decreased levels of cyclins E and A and increased levels of p21Cip1 kinase inhibitor (Naderi & Blomhoff, 1999). Conversely, ATRA stimulates the proliferation of peripheral blood T-cells by increasing pRB phosphorylation, through increased levels of cyclin D3, E and A and decreased levels of p27Kip1. However, since the effect on cell cycle machinery was delayed (~ 20h post treatment), the effect of RA on T-cell proliferation is thought to be indirect and possibly due to increased production of IL-2 (Ertesvag et al., 2002).

1.4.2.2 The Abnormal Role of ALDH1A1

1.4.2.2.1 Relevance of ALDH1A1 as a Target Gene of HOX11

The importance of retinoid signalling in processes governing haematopoietic cellular proliferation and differentiation is exemplified by the block to granulopoiesis as a result of overexpressing a truncated form of RARα or the PML/RARα fusion gene, typically found in a subtype of APL (de The et al., 1991). However, although RARs and RAR-binding proteins are key players in retinoid signalling, the availability of RA, which is determined by the expression of specific RA-synthesising enzymes, also plays a crucial
role in determining the biological function of RA within the cell. The second step in RA synthesis, catalysed by ALDH1A1, is particularly important since it represents the irreversible step at which the active metabolite is made available to the cell (McCaffery & Drager, 1995). Previous studies reveal that changes in ALDH activity, in particular an increase in the activities of Class 1 and 3 ALDHs, coincide with experimental liver and bladder carcinogenesis and have also been observed in a number of human tumours including mammary, liver and colon cancers (Lindahl, 1992). This increase in ALDH activity is associated with enhanced tumour cell growth and drug resistance, and inhibition of ALDH1/3 activities in hepatoma cells resulted in a 15% reduction in cell number, suggesting that these isozymes play important roles in processes governing cell death (Canuto et al., 2001).

The ectopic expression of HOX11 in T-cell leukaemias bearing t(10;14)(q24;q11) or t(7;10)(q35;q24), is a significant contributory factor in the initiation of tumorigenesis (Kennedy et al., 1991). To date, the regulatory effects of HOX11 on ALDH1A1 expression have been examined in a number of cell types including murine spleen cells, NIH 3T3 murine fibroblasts, the T-ALL cell line PER-117, the human erythroleukaemic cell line, HEL and the murine erythroleukaemic cell line J2E (Greene et al., 1998, Greene et al., 2002). Although HOX11 was found to activate Aldh1a1 expression in NIH 3T3, J2E and PER-117 cells, levels of Aldh1a1 were significantly reduced in cells of the developing spleen and in HEL cells transfected with HOX11 (Greene et al., 1998; Greene, unpublished observations). A possible link between ALDH1A1 downregulation and tumorigenesis has been previously demonstrated in mouse hepatocellular tumours, human prostate tumours and other various tumour cell lines (Ryzlak et al., 1992, Dragani et al., 1996). It is therefore plausible that a similar mechanism of tumour progression may exist in T-cells ectopically expressing HOX11.

It appears that the repression of Aldh1a1 in splenic precursor cells is essential for normal spleen development, however the inappropriate activation/repression of ALDH1A1 in developing T-cells may potentially result in a block in cell differentiation, thereby facilitating leukaemogenesis. The potential significance of altered Class 1 ALDH activity in the development of T-ALL is also exemplified by recent studies in which the TAL1 and LMO1/LMO2 oncoproteins were shown to synergistically activate
the transcription of the related RA-synthesising enzyme, \textit{ALDH1A2 (RALDH2)} in leukaemic T-cells by acting as cofactors for GATA-3 (Ono \textit{et al.}, 1998). In the erythroid lineage, these factors function synergistically in conjunction with erythroid-specific GATA-1 to regulate erythroid cell differentiation. The inappropriate expression of SCL and LMO proteins in the T-lineage provides a plausible mechanism for dysregulated gene expression, by mimicking an erythroid-specific transcriptional complex, where GATA-1 is replaced by T-cell specific GATA-3 (Ono \textit{et al.}, 1998). In light of the key roles played by retinoic acid, it is evident that \textit{ALDH1A1}, the enzyme regulating the delicate balance of RA within the cell, fulfils a crucial role in maintaining normal processes affecting proliferation, quiescence and differentiation. Thus, alterations affecting the cellular availability of \textit{ALDH1A1}, may disrupt a homeostatic control point in RA synthesis, with potentially serious consequences for the cells affected.

\section{1.4.3 \textit{FHL1} as a Target Gene of HOX11}

\subsection{1.4.3.1 LIM Domain Proteins}

LIM domain proteins can be classified as either 1) LIM-homeodomain (LIM-HD) proteins, 2) LIM-functional domain proteins, 3) C-terminal LIM proteins and LIM-only (LMO) proteins, and derive their name from the presence of a consensus amino acid sequence \((C-X_2-C-X_{17-19}H-X_2-C)-X_2(C-X_2-C-X_{16-20}C-X_2-H/D/C)\) originally identified in the homeodomain-containing transcription factors Lin-11, \textit{Isl1} and \textit{Mec-3} (Freyd \textit{et al.}, 1990). The LIM domain comprises 50-60 amino acids and encodes a cysteine-rich, double zinc finger motif reminiscent of the DNA-binding zinc fingers of GATA1-like transcription factors (Brown \textit{et al.}, 1999). Despite the similarity between these motifs, however, LIM domains have not been shown to bind DNA directly and instead mediate protein-protein interactions between transcription factors, signalling and cytoskeleton-associated proteins (Dawid \textit{et al.}, 1998, Curtiss & Heilig, 1998). The presence of multiple LIM domains in most LIM proteins enables them to act as molecular scaffolds to bridge the interactions between transcription factors, thus providing a backbone for the coordinated assembly of specific protein complexes involved in the regulation of processes such as embryonic development, cellular differentiation and cytoskeletal
structure (Schmeichel & Beckerle, 1994). The involvement of LIM domain proteins in such diverse processes may be attributed to the ability of LIM domains to interact with multiple protein families; LIM domains have been shown to form homo- and heterodimers, bind tyrosine-containing motifs, PDZ domains, and helix-loop-helix domains in transcription factors, kinases and cytoskeletal proteins (Kuroda et al., 1996, Tu et al., 1999).

1.4.3.2 FHL1

1.4.3.2.1 Family members and Related Isoforms

FHL1, previously SLIM1, is a member of the four and a half LIM (FHL) domain protein family, which also includes FHL2/SLIM3/DRAL (47% amino acid homology), FHL3/SLIM2 (45% amino acid homology), FHL4 and ACT (Activator of the cAMP Response Element Modulator in Testis), that fulfil essential roles in the development and maintenance of the cardiovascular system and striated muscle (Morgan & Madgwick, 1996, Chu et al., 2000). Each member contains an N-terminal zinc finger followed by four complete LIM domains and no other associated motifs (Morgan & Madgwick, 1996). To date, three alternative splice variants of FHL1 have been identified, including the murine isoform, named KyoT2 (Taniguchi et al., 1998), as well as a human variant, designated SLIMMER/FHL1B (SLIM1 with Extra Regions)(Brown et al., 1999, Lee et al., 1999)(Figure 1.14). KyoT2 contains the NH2-terminal two and a half LIM domains of FHL1 followed by 27 novel C-terminal amino acids, which mediate binding to RBP-J/RBP-Jk/Su(H), a DNA-binding protein involved in the Notch signalling pathway. The displacement of RBP-J from DNA as a result of binding KyoT2, leads to a block in differentiation suppression mediated by Notch (Taniguchi et al., 1998). FHL1C, the transcript corresponding to KyoT2 in humans, was subsequently identified in a human testis cDNA library, and shares 91% homology at the nucleotide sequence level and 96% homology at the amino acid sequence level with its murine counterpart (Ng et al., 2001). Given that FHL1C contains the RBP-J binding region, it is likely that this isoform also interacts with RBP-J to repress the transcriptional activity of RBP-J (Ng et al., 2001). SLIMMER/FHL1B contains the NH2-terminal three and a half LIM domains of FHL1 followed by a novel C-terminal 96 amino acids
**Figure 1.14. FHL1 Isoforms.** (A). Splice variants of FHL1. Exons are represented by open boxes. Splicing of the predominantly expressed FHL1 transcript is indicated by a solid line. FHL1B which contains exon 4b, results in a translation frameshift to generate a novel C-terminus that does not include the fourth LIM domain but encodes a nuclear localisation signal, a nuclear export signal and an RBP-J binding region and is represented by a solid blue line. FHL1C lacks exon 4 and 4b resulting in a frameshifted exon 5 to produce an RBP-J binding region and is represented by a dashed line. (B). The human cDNAs encoding FHL1, FHL1B, FHL1C and their corresponding translated proteins. Amino acid domains are represented as follows; single zinc finger (Z), LIM Domains (LIM1-4), Nuclear Localisation Signals (NLS), Nuclear export signal (NES), RBP-J binding region (R).
encompassing three putative bipartite nuclear localisation signals (NLS), a leucine rich sequence consistent with a nuclear export sequence (NES) and 27 amino acids identical to the RBP-J binding region of KyoT2 (Brown et al., 1999, Lee et al., 1999). The localisation of SLIMMER to the nucleus of myoblasts and the cytoplasm of myotubes suggests a role for this isoform in differentiation associated transcriptional regulation of muscle cells (Brown et al., 1999).

1.4.3.2.2 Expression and Function

FHL1 was first identified as a 2.3kb transcript highly expressed in skeletal muscle, and was later shown to be expressed in cardiac muscle and at relatively lower levels in a range of tissues including prostate, testis, ovary, small and large intestine, placenta and lung (Morgan et al., 1995, Greene et al., 1999). In the developing mouse embryo, Fhl1 is expressed specifically in the cardiac outflow tract of the heart and is subsequently restricted to the aortic arch and atria in adult rabbit heart tissue (Brown et al., 1999). The structural and regulatory roles played by FHL1 in the differentiation of human heart muscle is further supported by microarray studies which suggest that FHL1 may play a role in the pathogenesis of cardiomyopathy (Yang et al., 2000, Lim et al., 2001). Thus, FHL1 expression was shown to be significantly reduced in patients with ischemic dilated cardiomyopathy and conversely, increased in cases of human congenital hypertrophic cardiomyopathy (Yang et al., 2000, Lim et al., 2001). Intriguingly, the localisation of FHL1 to human chromosome Xq27.2 may also provide evidence of the important functional role played by FHL1 in the differentiation of the cardiac outflow tract to form the aortic and pulmonary outflow tracts. Specifically, in the case of Turner Syndrome (45, XO) in which females lack a complete or partial X chromosome, the most common cardiac malformations are those involving aortic function (Lin et al., 1998).

Despite increasing evidence that FHL1 has an important role in regulating skeletal and cardiac muscle cytoskeleton, very few studies have addressed the intracellular location and normal cellular function of FHL1. Studies by Brown et al., (1999) demonstrated that FHL1 is predominantly expressed in the cytosol of skeletal muscle-derived Sol8 myocytes and localised to focal adhesions and actin cytoskeleton in COS-7 cells,
suggesting a regulatory role in cytoskeletal structure. More recently, studies by McGrath et al., (2003) revealed that FHL1 demonstrates integrin-mediated localisation in both the nucleus and within the cytoplasm, specifically at focal adhesions and stress fibers within skeletal myoblasts, where it acts to regulate integrin-mediated cellular functions. Integrin signalling regulates muscle migration, differentiation and proliferation, and it is postulated that FHL1 functions by inhibiting integrin-mediated myoblast adhesion, to promote spreading and migration (McGrath et al., 2003).

Although a function for FHL1 in the regulation of integrin signals to the cytoskeleton is apparent, the nuclear localisation of FHL1 suggests that it may also play a role in transcriptional regulation, however there is no evidence to support this hypothesis to date. The observation that both FHL1 isoforms, KyoT2/FHL1C and SLIMMER, harbour a domain capable of binding the transcription factor RBP-J, however, strengthens the notion that transcriptional regulation may be shared mechanism of action between the three isoforms (Brown et al., 1999).

1.4.3.2.3 Relevance of FHL1 as a Target Gene of HOX11

In addition to ALDH1A1, FHL1 was also identified as a possible target for regulation by HOX11, following cDNA RDA experiments to compare differential gene expression profiles between NIH 3T3 cells and NIH 3T3 cells ectopically expressing HOX11 (Greene et al., 1998). Unlike ALDH1A1 however, FHL1 was not consistently upregulated by HOX11 in subsequent screening experiments of additional NIH 3T3 HOX11-expressing clones (activated in ~50%), suggesting that factors other than HOX11 expression may be required for FHL1 upregulation (Greene et al., 1998). The possibility that FHL1 may represent an oncogenically relevant target, is highlighted by the fact that the LIM-only proteins, LMO1 and LMO2, are also implicated in T-cell leukaemogenesis, and intriguingly, FHL1 was shown to be expressed in 7 of 12 T-ALL cell lines examined by Northern blot (Greene et al., 1998). Moreover, FHL1 maps to the long arm of chromosome Xq26, which is frequently amplified in B-cell neoplasms (Werner et al., 1997). Further characterisation of the expression patterns and regulation of FHL1 by HOX11 is necessary to investigate the potential role of FHL1 in haematopoietic malignancies.
1.5 THESIS AIMS

The identification and characterisation of relevant target genes is crucial if we are to completely understand the mechanism(s) by which nuclear oncogenes like HOX11 cause cellular transformation. Two such genes were identified as being transcriptionally upregulated by HOX11 in NIH 3T3 murine fibroblasts, namely aldehyde dehydrogenase 1a1 (Aldh1a1), which also appears to be physiologically regulated by Hox11 in the developing spleen, and four and a half LIM domain protein 1 (Fhl1) (Greene et al., 1998). Although HOX11 has been shown to modulate endogenous ALDH1A1 expression in numerous cell types, the mechanism by which HOX11 regulates ALDH1A1 and FHL1 gene expression remains unclear. Thus, in order to confirm the status of ALDH1A1 and FHL1 as transcriptional targets of HOX11, and to investigate a possible link between these targets and leukaemogenesis, this study aimed to:

(1) Assess the transcriptional response of the 5’ regulatory regions of ALDH1A1 and FHL1 to HOX11.

(2) Investigate the regulatory elements/transcriptional complexes involved in the response of ALDH1A1 to HOX11.

(3) Determine the effect of overexpressing ALDH1A1 and FHL1 on murine haematopoiesis, as a starting point to investigate whether either of these genes might account for the tumorigenic potential of HOX11.
Chapter 2

Assessment of \textit{ALDH1A1} as a Transcriptional Target of HOX11

2.1 INTRODUCTION

The ectopic expression of HOX11 in T-cell leukaemias bearing the t(10;14)(q24;q11) or variant t(7;10)(q35;q24) translocations, is a significant contributory factor to the progression of tumorigenesis (Kennedy \textit{et al.}, 1991). In its normal role, HOX11 is also crucial for the development of the spleen (Roberts \textit{et al.}, 1994, Dear \textit{et al.}, 1995). The identification and study of relevant target genes is therefore essential in order to elucidate the mechanism(s) by which HOX11 induces tumour growth and controls normal development. Two genes have been identified by Greene \textit{et al.}, (1998) as being HOX11 dependent - \textit{Four and a half LIM domain protein 1} (\textit{Fhl1}, formerly \textit{Slim1}), which encodes a LIM domain protein involved in muscle development and function (Yang \textit{et al.}, 2000, Lim \textit{et al.}, 2001, Robinson \textit{et al.}, 2003) and \textit{Class 1 Aldehyde Dehydrogenase} (\textit{Aldh1a1}), which appears to be physiologically regulated in the developing spleen and codes for an enzyme responsible for the irreversible oxidation of retinaldehyde to retinoic acid (RA)(Duester, 1996, Raner \textit{et al.}, 1996, Sladek, 2003). RA is a key regulator of transcriptional pathways governing processes of cellular proliferation, differentiation and apoptosis, thereby controlling various aspects of vertebrate embryo patterning, organogenesis and haematopoiesis. It is therefore evident that ALDH1A1, an enzyme regulating the balance of RA within the cell, fulfils a crucial role in maintaining normal cellular processes. Thus, alterations affecting the expression of \textit{ALDH1A1}, and therefore the rate-limiting step in RA synthesis, may have adverse effects on cellular development, as exemplified by studies in which the perturbation of RA concentrations leads to homeotic transformations and morphological abnormalities (Mavilio, 1993, Duester, 1996).

The human cytosolic \textit{ALDH1A1} gene is located on chromosome 9q21 and comprises 13 exons, spanning a region of approximately 53 kilobases (kb)(Hsu \textit{et al.}, 1989)(Figure 2.1). Primer extension analysis defined the transcription initiation start site (+1) to a position 53 base pairs (bp) upstream of the ATG initiation codon, yielding a 2,116bp
transcript (Hsu et al., 1989). Northern blot analysis of a panel of human tissues demonstrated that a transcript of this size is highly expressed in liver and pancreas, moderately expressed in skeletal muscle and kidney and expressed at low levels in the brain, heart and lung (Yanagawa et al., 1995). Preliminary studies attempting to delineate cis-regulatory elements responsible for the tissue-specific gene expression of ALDH1A1, were subsequently performed by Yanagawa et al., (1995), by analysing the expression of nested deletions of a 2,536bp region of the ALDH1A1 promoter (pCAT-2536/+42 to pCAT-50/+42) in ALDH1A1-expressing Hep3B cells. These studies identified a minimal regulatory region (-91/+53bp) responsible for the cell type-specific expression of ALDH1A1, containing several putative regulatory elements conserved across human, marmoset, mouse and rat species, including a CCAAT box (-74/70bp), Octamer sequence (-58/-51bp) and an ATAAA box (-32/-27bp). Further mutational analysis and gel retardation assays revealed that the CCAAT box was the primary cis-regulatory element required for basal promoter activity in Hep3B cells, possibly via binding the nuclear factor NF-Y, whereas the non-canonical TATA box (ATAAA) did not appear to be a necessary for promoter activation (Yanagawa et al., 1995). Subsequent analysis of the ALDH1A1 promoter (pCAT-2536/+42) by Elizondo et al., (2000), revealed that RARα and the CCAAT/Enhancer-Binding Protein β (C/EBPβ) stimulate ALDH1A1 reporter gene expression in the presence of endogenous levels of RA in Hepa-1 mouse hepatoma cells through a Retinoic Acid Response-like Element (-91/-75bp) and the previously identified CCAAT box (-74bp). Exogenously applied RA was capable of inhibiting this activation, in a negative feedback mechanism to control RA biosynthesis in Hepa-1 cells, possibly by decreasing cellular levels of C/EBPβ.

Thus far, the regulatory effects of HOX11 on ALDH1A1 expression have been examined in a number of cell types including developing spleen, NIH 3T3 fibroblasts, the T-ALL cell line PER-117, the human erythroleukaemic cell line HEL and the murine erythroleukaemic cell line J2E (Greene et al., 1998, Greene et al., 2002, Greene, unpublished observations). Although HOX11 was consistently shown to upregulate the expression of ALDH1A1 in NIH 3T3 and PER-117 cells, a significant downregulation was observed in HEL cells and cells of the developing spleen, reflecting the ability of HOX11 to act as both a positive and negative regulator of transcription, possibly as a
result of differential interactions with cell-type specific cofactors (Pinsonneault et al., 1997, Greene et al., 1998).

Despite the ability of HOX11 to affect ALDH1A1 expression in a variety of different cell lines, screening of a panel of T-ALL cell lines failed to reveal a correlation between HOX11 and ALDH1A1 expression levels (Greene et al., 1998). Thus, the significance of ALDH1A1 as an oncogenically relevant target gene of HOX11 remains unclear. The notion that HOX11 may alter retinoic acid synthesis in T-cell tumorigenesis by modulating ALDH1A1 expression is given added intrigue by the observation that ALDH1A2/RALDH2, a second enzyme involved in the irreversible conversion of retinaldehyde to RA, is a target gene of a transcriptional complex involving SCL/LMO and GATA-3 in leukaemic T-cells (Ono et al., 1998). Like HOX11, SCL (TAL1) and LMO1/2 are transcription factors specifically implicated in childhood T-ALL. This hints at a possibility that a common mechanism for the development of T-ALL involving disrupted retinoic synthesis may exist. However, regardless of whether the regulation of ALDH1A1 by HOX11 is confined to normal developmental processes or also incorporates pathways leading to neoplastic transformation, delineating the molecular mechanism(s) by which HOX11 affects ALDH1A1 expression would greatly add to our understanding of HOX11 function.

In order to begin to elucidate the mechanism(s) by which HOX11 affects the transcriptional activity of ALDH1A1, the effect of HOX11 was examined on a 2.2kb region of the ALDH1A1 promoter in both an ALDH1A1-expressing (HEL900) and non-expressing (PER-117) cell line. The specific aims of this chapter were 1) to perform luciferase reporter assays to investigate the activity of the ALDH1A1 promoter and delineate any HOX11-responsive elements and 2) to analyse the transcriptional complexes involved in any HOX11-mediated effect by electrophoretic mobility shift assays.
2.2 METHODS

2.2.1 pGL3Basic-ALDH1A1 Promoter Vector Construction

2.2.1.1 Preparation of pGL3Basic Luciferase Reporter Vector

The pGL3Basic luciferase reporter vector (4.8kb) lacks eukaryotic promoter and enhancer sequences, allowing the cloning of putative regulatory sequences upstream of the reporter gene (deWet et al., 1987)(Promega; Genbank®/EMBL Accession Number: U47295)(Appendix 2). The reporter gene is a modified coding region for the firefly (Photinus pyralis) luciferase that has been optimised for monitoring transcriptional activity in transfected eukaryotic cells. The vector contains the F1 origin of replication, the SV40 late poly(A) signal and the ampicillin (Amp) gene as a selectable marker. The pGL3Basic Luciferase Reporter Vector was prepared for sub-cloning by restriction digestion of supercoiled, pGL3Basic plasmid DNA with Sac I/Nhe I and Sac I/Hind III restriction endonucleases (Promega). Reactions were assembled in 1.5ml polypropylene microcentrifuge tubes in a final volume of 100µl, comprising MULTICORE™ Buffer [1x](Promega), 0.1mg/ml nuclease-free acetylated bovine serum albumin (BSA) (Promega), 2µg of supercoiled pGL3Basic plasmid DNA, and either 60U of Sac I/20U of Nhe I restriction endonucleases or 60U of Sac I/20U of Hind III restriction endonucleases for 3.5h at 37°C. Following digestion, restriction endonucleases were heat inactivated by incubation at 65°C for 20min.

2.2.1.2 Agarose Gel Electrophoresis

The digested pGL3Basic plasmid vector was analysed by agarose gel electrophoresis on 0.6-1.5% (w/v) agarose gels containing 0.2µg/ml ethidium bromide (Bio-Rad Laboratories). Agarose gels were prepared in 1x Tris-Acetate-EDTA (TAE) Buffer [50mM Tris-acetate (pH 8.0), 1mM EDTA (pH 8.0)] using RNase-DNase-free agarose (Progene). DNA samples were combined with 5x Nucleic Acid Sample Loading Buffer [1x; 10mM Tris-HCl (pH 8.0), 5% (v/v) glycerol, 1mM EDTA, 0.04% (w/v) bromophenol blue, 0.04% (w/v) Xylene Cyanole FF](Bio-Rad Laboratories) immediately prior to loading and electrophoresis was performed in 1x TAE buffer
containing 0.1µg/ml ethidium bromide using a Mini-sub horizontal gel apparatus (Bio-Rad Laboratories) at constant voltage (70-80V) for 1h. Following electrophoresis, DNA was visualised using an ultraviolet UV transilluminator (Gel Doc 1000) and Molecular Analyst software (Bio-Rad Laboratories). The concentration and molecular size of electrophoresed DNA samples was estimated by direct comparison with known molecular weight standards, PhiX174/Hae III and/or Lambda/Hind III molecular weight markers (Promega) [250ng/lane], which were electrophoresed in parallel with samples being analysed.

### 2.2.1.3 Agarose Gel Purification of Digested DNA Fragments

Digested pGL3Basic vector DNA was gel purified using the Qiaex II Gel Purification Kit (Qiagen) according to the manufacturer’s instructions. DNA fragments were resolved by agarose gel electrophoresis (Section 2.2.1.2) on a 1.3% (w/v) agarose gel. The desired fragments were then excised from the gel using a sterile scalpel blade and transferred to 1.5ml polypropylene microcentrifuge tubes (~300mg of gel/tube). In order to recover the DNA, Buffer QX1 (3 volumes), dH2O (2 volumes) and QIAEX II beads (15ul) were combined with the gel slices and incubated at 50°C for 10min with vortexing at regular intervals to aid solubilisation of agarose and binding of DNA to QIAEX II beads. Following this, the QIAEX II beads were pelleted by centrifugation (2,720 x g, 1min, room temperature)(PMC-060, Capsulefuge) and washed with 500µl Buffer QXI (x1) and 500µl Buffer PE (x2). The pellet was air-dried and the DNA was eluted in 50-100µl dH2O depending on the starting quantity of product by incubating at room temperature for 10min. The QIAEX II beads were pelleted and the supernatant containing DNA removed into a fresh tube and stored at -20°C until required. DNA concentration was estimated by electrophoresis of a sample of the purified DNA fragment on a 1.3% (w/v) agarose gel (Section 2.2.1.2) alongside quantitative PhiX174/Hae III molecular weight markers.

### 2.2.1.4 Generation of ALDH1A1 Promoter Fragments

For the construction of pGL3BasicALDH1A1P luciferase reporter constructs, nested deletions of the ALDH1A1 proximal promoter ranging from -2159/-146bp to +42bp
with respect to the transcription initiation start site (TSS) identified by Yanagawa et al., (1995), were generated by polymerase chain reaction (PCR) using ADPF-2159, -1968, -978, -303, -201 and -146 forward primers, in conjunction with a common reverse primer (ADPR+42)(Table 2.1). ADPF forward and ADPR+42 reverse primers contain Sac I and Nhe I restriction endonuclease sites respectively, to enable cohesive ligation into the Sac I/Nhe I site of the pGL3Basic luciferase reporter vector (Table 2.1; underline). The ALDH1A1P-91ΔCAT deletion construct was generated using the ADPF-91ΔCAT forward primer, which contains a deletion of the CAAT box (-74/ -70bp) and the ADPR+42 reverse primer described above (Table 2.1). Fragments were amplified from human genomic DNA using the high fidelity DNA polymerase PfuTurbo® (Stratagene). Reactions were performed in thin-walled, 0.5ml polypropylene tubes in a final volume of 50µl comprising 10x Cloned Pfu Reaction Buffer [1x; 20mM Tris-HCl (pH 8.8), 10mM KCl, 10mM (NH₄)₂SO₄, 2mM MgSO₄, 0.1% (v/v) Triton® X-100, 0.1mg/ml BSA](Stratagene), 0.2mM dNTPs (Fisher Biotech), 20pmol each of forward and reverse primers, 530ng of human genomic DNA and 2.5U of PfuTurbo® DNA Polymerase (Stratagene). For fragments -2159/+42 and -1968/+42, samples were denatured for 2min at 94°C and then amplified by 10 cycles (94°C for 10s, 60°C for 30s and 72°C for 2min), with additional cycle extension times increasing by 20s up to 8:40s, followed by a final extension of 72°C for 7 min (EXTEND72; Appendix 1). For fragments -978/+42, -303/+42 and -201/+42, samples were denatured for 3min at 95°C and then amplified by 30 cycles (95°C for 40s, 60°C for 40s and 72°C for 3min), with a final extension of 72°C for 8min (KIM60; Appendix 1). All reactions were performed using a PTC-100 (Programmable Thermal Controller-100; MJ Research Inc). PCR products were analysed by agarose gel electrophoresis on 0.6-1.5% (w/v) agarose gels as described in Section 2.2.1.2. ALDH1A1P promoter fragments were subsequently gel purified as described in Section 2.2.1.3.

For the generation of pGL3BasicALDH1A1P-91/+42, the pGL2BasicALDH1A1P-91/+42 vector was digested with Sac I and Hind III restriction endonucleases (Promega), to facilitate sub-cloning of the ALDH1A1P-91/+42 promoter fragment into the Sac I/Hind III site of the pGL3Basic vector. Digestion reactions were performed in 1.5ml polypropylene microcentrifuge tubes in a final volume of 50µl comprising 10x MULTICORE™ Reaction Buffer [1x; 25mM Tris acetate (pH 7.8), 0.1M potassium
TABLE 2.1
acetate, 10mM magnesium acetate, 1mM DTT](Promega), 20µg of pGL2BasicALDH1A1P-91/+42 plasmid DNA, 0.1mg/ml nuclease-free acetylated BSA (Promega), 60U of Sac I restriction endonuclease and 20U of Hind III restriction endonuclease and incubated at 37°C for 3h. Restriction endonucleases were subsequently heat inactivated by incubation at 65°C for 20min. The ALDH1A1P-91/+42 promoter fragment was subsequently gel purified as described in Section 2.2.1.3 and stored at -20°C prior to ligation.

ALDH1A1P-50/+42, -36/+42 and +1/+42 promoter inserts were generated by annealing complementary oligonucleotides encompassing the promoter region of interest (Table 2.1). Antisense (AS) oligonucleotides were synthesised with Sac I/Nhe I overhangs to facilitate cohesive ligation into the pGL3Basic luciferase reporter vector. Equimolar quantities of sense (S) and antisense DNA strands were combined in thin-walled, 0.5ml polypropylene microcentrifuge tubes and annealed in TES Buffer [1x; 10mM Tris-HCl (pH 7.4), 50mM NaCl, 1mM EDTA] by heating to 90°C for 5min and cooling to 22°C at a rate of 1°C/min in a PTC-100 thermocycler and stored at -20°C prior to ligation.

**2.2.1.5 Restriction Digestion and Purification of Purified ALDH1A1 Promoter PCR Products**

Gel purified ALDH1A1P DNA fragments generated in Section 2.2.1.4 were prepared for sub-cloning into the pGL3Basic luciferase reporter vector by digestion with Sac I and Nhe I restriction endonucleases (Promega). Reactions were assembled in 1.5ml polypropylene microcentrifuge tubes in a final volume of 100µl, comprising 10x MULTICORE™ Reaction Buffer [1x](Promega), 0.1mg/ml nuclease-free acetylated BSA (Promega), 50µl of gel purified DNA fragment, 60U of Sac I restriction endonuclease and 20U of Nhe I restriction endonuclease for ~16h at 37°C to ensure complete digestion. Restriction endonucleases were subsequently heat inactivated by incubation at 65°C for 20min. Digested fragments were subsequently purified from unwanted digestion products, polymerases and salts using the QIAquick PCR Purification Kit (Qiagen) according to the manufacturer’s instructions. In order to purify the DNA, Buffer PB (5 volumes) was combined with the digestion reaction (1 volume) in a 1.5ml polypropylene microcentrifuge tube and the sample was applied to the
QIAquick spin column and centrifuged at 10,000 x g for 1 min at room temperature (Biofuge Pico, Heraeus Instruments). The flow-through was discarded and the column was washed by applying 750 µl Buffer PE to the column and centrifuging at 10,000 x g for 1 min at room temperature. The flow-through was discarded and the column was re-centrifuged (10,000 x g, 1 min, room temperature) to remove traces of residual ethanol. The DNA was then eluted in a fresh 1.5ml polypropylene microcentrifuge tube by applying 50 µl sterile dH2O to the column, incubating for 1 min, followed by centrifugation (10,000 x g, 1 min, room temperature) and stored at -20°C. An aliquot of the purified insert was then electrophoresed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) to estimate DNA integrity and concentration prior to ligation.

2.2.1.6. Ligation of Promoter Fragments into pGL3Basic Luciferase Reporter Vector

Ligation reactions were typically performed in a 20 µl reaction volume comprising a linearised plasmid to insert molar ratio of 3:1. Thus, for a ligation reaction using 40 ng pGL3Basic vector plasmid, the following equation was utilised to determine the amount of insert DNA required to achieve the required molar ratio:

<table>
<thead>
<tr>
<th>Insert (ng)</th>
<th>Insert size (bp)/Plasmid size (bp) x 120</th>
</tr>
</thead>
</table>

Ligation reactions were performed in 0.5 ml polypropylene microcentrifuge tubes and comprised 40 ng of linearised pGL3Basic vector DNA, insert DNA [3:1 molar ratio], 10x T4 DNA Ligase Buffer [1x; 30 mM Tris-HCl (pH 7.8), 10 mM MgCl2, 10 mM DTT, 1 mM ATP] (Promega), 6 U of T4 DNA Ligase (Promega) and incubated at 15°C for 16-24 h. Mock ligations replacing insert DNA with dH2O were also performed in parallel to assess background levels due to incomplete vector digestion or re-circularisation.

2.2.1.7 Preparation of Competent Cells

All recombinant cloning procedures were performed using the Escherichia coli (E. coli) strain dDH5α-T1R [FØ80LacZΔM15Δ (lacZYA-argF) U169 deoR recA1 endA1 hsdR17 (rK- mK+) phoA supE44 thi-1 gyrA96 relA1 tonA1](Invitrogen Life Technologies). Chemically competent DH5α-T1R cells were prepared using the calcium
chloride method as described in Sambrook and Russell (2001). In brief, a single bacterial colony from a freshly streaked Luria-Bertani (LB) agar plate [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar] was inoculated into a 50ml polypropylene tube containing 5ml of 2YT media [1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl] and incubated for ~16h at 37°C with vigorous shaking (225rpm). The starter culture was subsequently diluted (1:100) in fresh 2YT media and cultured to mid-logarithmic phase (Optical Density (O.D) ~ 0.45-0.55) at 37°C with vigorous shaking (225rpm; ~2-3h). The cells were transferred to a 50ml polypropylene tube and recovered by centrifugation (820 x g, 10min, 4°C)(IEC Centra GP8R). The cells were then resuspended in 50ml of ice cold 100mM CaCl₂ and incubated on ice for 20min. Following centrifugation (460 x g, 10min, 4°C), the cells were resuspended in 10ml ice-cold 100mM CaCl₂ and 15% (v/v) glycerol. The competent cells were dispensed in 240µl aliquots in 1.5ml polypropylene tubes, snap frozen in liquid N₂ and stored at -80°C.

2.2.1.8 Bacterial Transformations

Ligation reactions were transformed into chemically competent DH5α E. coli prepared in Section 2.2.1.7. To transform recombinant plasmid DNA, pre- aliquoted competent cells (240µl) were thawed on ice and gently mixed with ligation or mock ligation reactions (5-10µl). Transformation reactions were incubated on ice for 1h, followed by heat shock at 42°C for 90s and immediately placed on ice for a further 2min. To facilitate the growth of transformed bacteria, 300µl of pre-warmed 2YT growth media was added to the cells, which were then incubated at 37°C for 30min. A sample of the recombinant culture was spread onto LB agar plates containing 0.1mg/ml ampicillin (Austrapen®)(125µl, 375µl/plate) and the bacterial colonies were incubated at 37°C for 16-20h to allow growth of the transformant bacterial colonies. Plates were stored at 4°C for up to 1 month.

2.2.1.9 Small Scale Plasmid Purification

Plasmid Mini-Preps of selected colonies were prepared for restriction digest analysis and sequencing using the GenElute™ Plasmid Miniprep Kit (Sigma) according to the
manufacturer’s specifications. In brief, randomly selected colonies were used to inoculate 5ml of 2YT media containing 0.1mg/ml ampicillin and grown with vigorous shaking (225rpm) for 16h. Cells were harvested (2ml) by centrifugation (13,000 x g, 1min, room temperature)(Biofuge Pico, Heraeus Instruments) and the supernatant was removed using a pipette. The bacterial pellet was then completely homogenised in 200µl Resuspension Solution and the cells were lysed by adding 200µl Lysis Solution. The tubes were inverted gently between 8-10 times to ensure complete mixing and incubated at room temperature for 5min. The reaction was then neutralised by the addition of 350µl Neutralisation Buffer and the tubes inverted gently between 4-6 times. The lysate was subsequently cleared of bacterial debris by centrifugation (13,000 x g, 10min, room temperature) and loaded onto a pre-assembled GenElute™ Miniprep binding column. Following centrifugation (13,000 x g, 1min, room temperature), the flow through liquid was discarded and the column was washed with 750µl Wash Solution. The flow through liquid was discarded and the column was centrifuged for an additional 2min (13,000 x g, room temperature) to remove residual ethanol. The GenElute™ Miniprep binding column was then transferred to a fresh collection tube and the DNA was eluted by applying 100µl of sterile dH₂O to the column followed by centrifugation (13,000 x g, 1min, room temperature).

2.2.1.10 Screening Recombinant Colonies for the Presence of Promoter Inserts

Cloning verification was performed by screening small-scale DNA preparations generated in Section 2.2.1.9 (between 10-20/construct) by restriction digest analysis; the number of colonies screened depending upon the level of background as indicated by comparing colony numbers on test versus mock ligation plates. Restriction digests were performed in 1.5ml polypropylene microcentrifuge tubes in a final volume of 30µl comprising MULTICORE™ Restriction Buffer [1x](Promega), 1µg of plasmid miniprep DNA, 0.1mg/ml nuclease-free acetylated BSA (Promega), and either 60U of Sac I/20U of Nhe I restriction endonucleases (Promega) or 60U of Sac I/20U of Hind III restriction endonucleases (Promega) for 1.5h at 37°C. The products of digestion were electrophoresed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) and preparations positive for an insert of the correct size were sequenced verified.
2.2.1.11 DNA Sequencing

All DNA sequencing was performed by the DNA Sequence Service located at Royal Perth Hospital (RPH), on a fully automated ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) using the ABI PRISM™ Big Dye™ Version 3.1 Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems). The ABI PRISM™ system is based on a modified dideoxy method of DNA sequencing in which a complementary primer is annealed to the DNA template of interest and temperature cycling is used to generate a fluorescently labelled complementary cDNA strand by random incorporation of deoxynucleotide triphosphates (dNTPs) and fluorescently labelled dideoxynucleotide triphosphates (ddNTPs). These ddNTPs lack the 3’-hydroxyl group required for phosphodiester bond formation during strand elongation by DNA Polymerase I, such that incorporation of a ddNTP into the growing strand results in the termination of complementary strand synthesis. Sequence verification was performed using GL3For/GL3Rev primers flanking the multiple cloning site of the pGL3Basic luciferase reporter vector (Table 2.1).

Analysis of raw sequence data was performed using SeqEd™ (V1.0.3) sequence analysis software from PE Applied Biosystems (Australia). Comparative genome database (NCBI GenBank; http://www.ncbi.nlm.nih.gov) and sequence alignment analyses were performed using SeqEd™ and Clustal Alignment programmes (http://www.ebi.ac.uk/clustalw/).

2.2.1.12 Large Scale Purification of Plasmid DNA

Large scale preparations of highly purified, supercoiled plasmid DNA were isolated using the QIAGEN Plasmid Midi/Mega Kits (Qiagen) which are based on a modified alkaline lysis procedure, followed by the binding of plasmid DNA to QIAGEN Anion-Exchange Resin (via gravity flow) under low salt and pH conditions. Following the removal of RNA, proteins, dyes and other impurities by a medium salt wash, high quality plasmid DNA is recovered by elution with a high salt buffer and concentrated and desalted by isopropanol precipitation (Qiagen, 2000). Recombinant bacterial cultures were propagated for plasmid purification by inoculating a 5ml starter culture in
2YT media containing 0.1mg/ml ampicillin from a freshly streaked selective LB agar plate and incubating at 37°C for ~8h with vigorous shaking (225rpm). Starter cultures were then diluted 1:500 in 2YTA media and grown for a further 12-16h at 37°C with vigorous shaking (225rpm). Bacterial cells were harvested from late log-phase cultures in 35ml/500ml polypropylene microcentrifuge tubes (30ml, 200ml) by centrifugation at 6000 x g for 15min at 4°C (AvantiTM J-25I Centrifuge, Beckman, JA-25-50, JA-14) and all traces of culture medium were removed. Pelleted cells were resuspended in Buffer P1 [50mM Tris.Cl (pH 8.0), 10mM EDTA, 100ug/ml RNase A](4ml, 25ml) and subsequently lysed with the addition of Buffer P2 (4ml, 25ml) [200mM NaOH, 1% (w/v) SDS] and the reaction was allowed to proceed at room temperature for 5min. The lysis reaction was then neutralized with Buffer P3 (4ml, 25ml)[3.0M potassium acetate (pH 5.5)] and incubated on ice for 30min to enhance the precipitation of from the DNA. Following this incubation, the reaction was gently mixed and the lysate was cleared of bacterial debris by centrifugation at 20,000 x g for 30min at 4°C (JA-25-50, JA-14). The supernatant was removed to a fresh tube and re-centrifuged at 20,000 x g for 15min at 4°C to ensure removal of all particulate debris and immediately applied to a QIA-tip column equilibrated with Buffer QBT (4ml, 35ml)[750mM NaCl, 50mM 3-[N-morpholino] propanesulfonic acid (MOPS)(pH 7.0), 15% (v/v) isopropanol, 0.15% (v/v) Triton®X-100]. The column was then washed with Buffer QC (2 x 10ml, 2 x 100ml)[1.0M NaCl, 50mM MOPS (pH 7.0), 15% (v/v) isopropanol] to remove unbound contaminants and the purified plasmid DNA eluted with Buffer QF (5ml, 17.5ml)[1.25M NaCl, 50mM Tris.Cl (pH 8.5), 12% (v/v) isopropanol]. Plasmid DNA was then precipitated by the addition of 0.7 volumes of isopropanol and pelleted by centrifugation in 2ml/35ml polypropylene microcentrifuge tubes at 15,000 x g for 30min at 4°C. The recovered DNA was then washed with 70% (v/v) ethanol, air dried under sterile conditions and resuspended in sterile dH₂O (120µl, 1ml).

2.2.1.13 Standardisation of Plasmid DNA for Transient Transfection Experiments

In order to standardise the quality of DNA prepared in Section 2.2.1.12, which was to be used in transfection experiments (Section 2.2.1.3.1), all DNA preparations were re-precipitated according to Sambrook and Russell (2001). In brief, DNA samples were combined in a 1.5ml polypropylene microcentrifuge tube with 0.1 volumes of 3M NaAc
(pH 5.2) and 2 volumes of ice-cold ethanol. The samples were mixed thoroughly and incubated at -20°C for 2h to enhance DNA precipitation. The DNA was then recovered by centrifugation at 12,000 x g for 15 min at 4°C (Sigma 1-15 centrifuge) and the pellets were washed with 1ml of 70% (v/v) ethanol. The supernatant was removed under sterile conditions and the pellets were air-dried and resuspended in sterile dH$_2$O (250µl/pellet).

2.2.1.14 DNA Analysis Using UV Spectrophotometry

All plasmid DNA preparations used in transient transfection experiments were routinely checked for concentration and purity using a combination of UV spectrophotometry and agarose gel electrophoresis. DNA samples in solution were diluted 1/80 to a final volume of 160µl in dH$_2$O. The spectrophotometer (Shimadzu UV Mini 1240 Spectrophotometer) was blanked with a 160µl aliquot of the sample diluent. Ultraviolet light absorbance of the diluted DNA sample was then determined at 260 nm and 280 nm wavelengths, and DNA concentration and purity was determined using the calculations described below.

\[
\text{DNA Concentration [mg/ml]: } O.D_{260nm} \times \text{dilution factor} \times 50
\]

where (1 O.D unit = 50µg/ml of double stranded DNA)

\[
\text{DNA Purity: } O.D_{260nm} / O.D_{280nm}
\]

2.2.1.15 DNA Analysis by Restriction Digestion

DNA samples were qualitatively analysed by restriction endonuclease digestion prior to use in transfections. The applicable DNA sample (1-2µg) was digested in a final volume of 30µl comprising 10x MULTICORE™ Restriction Buffer [1x](Promega), 0.1mg/ml nuclease-free acetylated BSA (Promega), either 20U of Sac I/20U of Nhe I restriction endonucleases or 20U of Sac I/20U of Hind III restriction endonucleases at 37°C for 2-
4h or overnight (~16h). The products of digestion were subsequently analysed by agarose gel electrophoresis on a 1.3% (w/v) agarose gel (Section 2.2.1.2).

2.2.2 Cell Culture

All cultured cell lines (summarised in Table 2.2) were provided by Dr Ursula Kees, Division of Children’s Leukaemia and Cancer Research (CLCR) of the TVW Telethon Institute for Child Health Research, Western Australia and were shown to be Mycoplasma-free at the level of PCR (performed by Jette Ford at the Division of CLCR). Cells were typically cultured in 75cm² filter-top culture flasks (NUNC™) and maintained in appropriate complete growth media at 37°C with 5% CO₂.

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Phenotype</th>
<th>HOX11 Translocation</th>
<th>HOX11 Expression</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEL900</td>
<td>Foetal Erythroleukaemia</td>
<td>-</td>
<td>-</td>
<td>Martin &amp; Papayannopoulo 1982</td>
</tr>
<tr>
<td>PER-117</td>
<td>T-ALL</td>
<td>-</td>
<td>-</td>
<td>Kees et al., 1987</td>
</tr>
<tr>
<td>NIH 3T3</td>
<td>Murine embryonic fibroblast</td>
<td>-</td>
<td>-</td>
<td>Jainchill et al., 1969</td>
</tr>
</tbody>
</table>

Table 2.2. Parental Mammalian Cell Lines Utilised in Transient Luciferase Reporter Analysis of the ALDH1A1 Promoter.

2.2.2.1 HEL-900 Cell Lines

HEL-900 (HEL) is a human erythroleukaemic cell line (Martin & Papayannopoulo, 1982). Cells were subcultured twice weekly (1:10) in RPMI-1640 Multi-cel Medium (Trace) supplemented with 10% (v/v) foetal calf serum (FCS), 2mM L-glutamine (Trace) and 50µg/ml of penicillin/streptomycin (Trace). Clones HEL[1] and HEL-HOX11[4] were generated by stable transfection of the parent HEL-900 cell line with the empty pEFBOS and pEFBOS-HOX11 plasmid constructs, respectively. Cells were subcultured twice weekly (1:10) as per the parental cell line except that puromycin
[1µg/ml](Trace) was also included in the growth media to maintain selection for stable transfectants.

### 2.2.2.2 PER-117 Cell Lines

PER-117 is a human, childhood T-ALL cell line bearing markers of a T-cell precursor phenotype (Kees et al., 1987). Cells were subcultured twice weekly (1:5) in RPMI-1640 Multi-cel Medium supplemented with 10% (v/v) FCS, 2mM L-glutamine, 50µg/ml of pencilllin/streptomycin, 2mM non-essential amino acids (Trace) and 2mM pyruvate (Trace). Clones PER-117[1] and PER-117HOX11[11] were generated by stable transfection of the parent PER-117 cell line with the empty pEFBOS and pEFBOS-HOX11 plasmid constructs respectively. Cells were subcultured twice weekly (1:5) in RPMI-1640 Multi-cel Medium supplemented with 20% (v/v) FCS, 2mM L-glutamine, 50ug/ml of pencilllin/streptomycin, 2mM non-essential amino acids and 2mM pyruvate. Puromycin [0.2µg/ml] was also included in the growth media to maintain selection for stable transfectants.

### 2.2.2.3 The NIH 3T3 Cell Line

NIH 3T3 is an adherent, murine fibroblast cell line established from NIH Swiss mouse embryo cultures (Jainchill et al., 1969). Cells were subcultured twice weekly (1:10) in Dulbecco’s Modified Medium (DMEM) supplemented with 10% (v/v) FCS, 2mM L-glutamine and 50µg/ml of penicillin/streptomycin. To passage cells, medium was removed and the cells were trypsinised with a 0.05% (v/v) Trypsin/0.53mM EDTA solution (Trace) until cells detached from flask surface. Trypsinisation was then quenched using an equal amount of complete DMEM medium and the cells were repeatedly aspirated to achieve a single cell suspension and sub-cultured into fresh complete DMEM medium.

### 2.2.2.4 Cryopreservation of Mammalian Cell Lines

Mammalian haematopoietic cell lines were maintained in long-term storage by cryopreservation. Cells were cultured to mid-log phase (4.6 x 10^5 cells/ml) in complete
growth medium and recovered by centrifugation in 50ml polypropylene tubes at 453 x g for 5min at room temperature (Eppendorf 5810). Cells were resuspended at approximately 5 x 10^6 cells/ml in freshly prepared cryopreservation medium [90% (v/v) FCS, 10% (v/v) Dimethyl Sulfoxide (DMSO)] and 1ml aliquots were dispensed into 1.5ml polypropylene cryovials (NALGENE Cryoware). Cryovials were cooled for >24h at -80°C in a polystyrene container and subsequently transferred into liquid nitrogen for long-term storage. To re-establish cryopreserved cell cultures, a vial of stock culture was thawed rapidly in a 37°C waterbath and the cells were transferred to pre-warmed complete growth media (10ml) in a 50ml polypropylene tube and washed with complete growth media to remove traces of DMSO. The cells were then recovered by centrifugation at 290 x g for 5min at room temperature, resuspended in complete growth media [suspension cell lines (5ml); adherent cells lines (10ml)] and transferred to a 25cm² filter top culture flask (NUNC™). Cells were incubated at 37°C, 5% CO₂ and expanded upon reaching confluence.

2.2.3 Luciferase Assays

2.2.3.1 Transient Transfection of HEL and PER-117 Cell Lines

The conditions employed for transfection of HEL and PER-117 cell lines were as previously determined (Brake, 1998). Cells were cultured to mid-logarithmic stage (~ 4-5 x 10⁵ cells/ml). Each transfection used 1 x 10⁷ cells, which were washed (x2) in 5ml of RPMI only prior to transfection and resuspended in 365µl of the same solution. Cells were electroporated in a 0.4cm gap electroporation cuvette (Bio-Rad Laboratories) with 15µg of test luciferase reporter plasmid and 5µg of control pSV-β-Galactosidase plasmid using the Gene Pulser System (Bio-Rad Laboratories) set at 300V, 960µF capacitance. Post-electroporation, cells were immediately resuspended in 7ml of complete RPMI-1640 Multi-cel Medium and incubated in a humidified atmosphere at 37°C in 5% CO₂. Cells were then harvested approximately 22h post-transfection for analysis of HOX11 expression (Section 2.2.1.3.3) or for luciferase/β-Galactosidase assays (Section 2.2.1.3.6) as required.
2.2.3.2 Transient Transfection of NIH 3T3 Cells

FuGENE 6 Transfection Reagent (Roche) is a multi-component, lipid-based transfection reagent that complexes with and transports DNA into the cell to efficiently transfect a wide variety of cell types. In order to achieve a cell monolayer of 50-80% confluency on the day of transfection, approximately 2x10^5 cells/well were seeded in a final volume of 2ml of complete DMEM Multi-cel Medium in 6-well plates (35mm diameter)(NUNC™) and incubated for 24h at 37°C, 5% CO₂. Transfections were performed based on a FUGENE 6 reagent (µl):DNA (µg) ratio of 3:1 by combining 93µl serum free DMEM medium only, with 7µl of FuGENE 6 Reagent in a sterile 1.5ml polypropylene tube. DNA solutions containing luciferase plasmid reporter and pSV-β-Galactosidase internal control constructs (2.3µg total) were then added to the pre-diluted FuGENE 6 reagent and the solution was incubated for 15min at room temperature. The complex mixture was then added dropwise to the cells and the plates were swirled to ensure even dispersal. Cells were incubated for 24h at 37°C, 5% CO₂ and harvested for luciferase assays as outlined in Section 2.2.1.3.6.

2.2.3.3 Western Blotting

2.2.3.3.1 Preparation of Nuclear Cell Extracts

Nuclear extracts were prepared as outlined by Schreiber et al., (1989). Cells were harvested at mid-log phase (approximately 5 x 10^5 cells/ml). Approximately 1 x 10^7 cells were collected by centrifugation in 50ml polypropylene tubes at 820 x g for 5min at room temperature (Centra GP8R, IEC). The cells were washed once with 10ml of cold Phosphate Buffered Saline (PBS) and resuspended in 1ml PBS. The cells were then transferred to 1.5ml polypropylene microcentrifuge tubes and pelleted at 17,900 x g for 30s at 4°C (Eppendorf Centrifuge 5417C, Crown Scientific). Following the removal of PBS, the pellet was resuspended in 400µl of Buffer A [10mM Hepes pH 7.9, 10mM KCl, 0.1mM EDTA, 0.1mM EGTA; supplemented with freshly added proteinase inhibitors 0.5M DTT, 1µg/ml aprotinin, 1µg/ml leupeptin, 0.5M PMSF, NaVO₄] by gentle pipetting, and allowed to swell on ice for 30min. Cells were then lysed by brief vortexing following the addition of 0.5% (v/v) Nonident-40 (NP-40) detergent.
The cells were pelleted by centrifugation at 17,900 x g for 1 min at 4°C and supernatants were completely removed. The cell pellet was immediately resuspended in 100 µl (25 µl for Western Blot analysis) of cold Buffer C [20 mM Hepes (pH 7.9), 0.42 M NaCl, 1 mM EDTA, 1 mM EGTA, 20% (v/v) glycerol, containing freshly added proteinase inhibitors 0.5 mM DTT, 1 µg/ml aprotinin, 1 µg/ml leupeptinin, 0.5 mM PMSF, NaVO₄] and nuclear proteins were extracted by vigorous shaking for 30 min at 4°C. Cell debris was pelleted by centrifugation at 17,900 x g for 5 min at 4°C and the supernatants were removed, pooled and aliquoted into single use volumes (20-40 µl) in 0.5 ml polypropylene microcentrifuge tubes on ice. Nuclear extracts were snap frozen in liquid nitrogen and stored at -80°C.

2.2.3.3.2 *Determination of Nuclear Protein Extract Concentration*

Nuclear protein concentration was determined using the Bio-Rad *DC* Protein Assay Kit (Bio-Rad Laboratories) according to the Microplate Protocol. A standard curve within the range (0.2 mg/ml-1.44 mg/ml) was generated using the BSA provided. Nuclear extract concentrations were determined by performing the protein assay over a range of dilutions (neat, 1/3, 1/5 and 1/6). Both standards and tests were prepared using Buffer C (protein extraction solution, see Section 2.2.3.3.1) as the diluent. In brief, a 5 µl aliquot of sample (BSA standards, nuclear extracts, Buffer C blank) was applied in duplicate to the wells of a transparent 96-U shaped well microtitration multiwell plate (Linbro®/Titretek®, ICN Biomedical). 25 µl of Reagent A’ was then added to each well, followed by the addition of 200 µl of Reagent B. Upon addition of reagents, the microplate was gently agitated and incubated at room temperature for 15 min. The protein content of each sample was determined by measurement of absorbance at 655 nm (A₆₅₅nm) using a Bio-Rad Model 3550-UV microplate reader (Bio-Rad Laboratories). The mean absorbance values for the BSA standards were then used to generate a standard curve from which the total protein content of each nuclear extract was calculated.
Nuclear extracts of pEFBOSHOX11-transfected cells [~1x10^7 cells in 20μl] were combined with 2x SDS loading dye [1x; 31.25mM Tris (pH 6.8), 12.5% (v/v) glycerol, 1% (w/v) SDS, 0.005% (w/v) bromophenol blue, 2.5% (v/v) β-mercaptoethanol] in 1.5ml screwcap polypropylene microcentrifuge tubes and proteins were denatured by boiling samples for 5min at 100°C and placed on ice. Biotinylated molecular weight markers (MWM)(Bio-Rad Laboratories) were prepared by combining 1μl MWM with 5x SDS loading dye [1x] in a final volume of 45μl and heating at 100°C for 5min. Prestained low range MWM (Bio-Rad Laboratories) were thawed and heated to 60°C. Samples were then resolved by SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)[loading order; biotinylated MWM (10μl), prestained MWM (5μl), nuclear extract samples (40μl), in 5x TANK buffer (pH 8.3)[1x; 3g/L Tris, 14.4g/L glycine, 1g/L SDS] on a 12% polyacrylamide gel run at 180V for 1h.

SDS-PAGE was performed using the Mini-PROTEAN 3 Cell and 10.1 x 7.3/8.2cm glass plates separated by 0.75mm spacers (Bio-Rad Laboratories). To prepare the 12% (v/v) resolving gel, 3.2ml of 30% acrylamide solution (37.5:1)(Bio-Rad Laboratories) was combined with 2ml of Lower Gel Buffer (pH 8.8)[1.5M Tris, 0.4% SDS (w/v)] and 2.8ml dH_2O. Gel polymerisation was initiated by the addition of 25μl of 10% APS (w/v)(Univar Laboratories) and 10μl of TEMED (Bio-Rad Laboratories) immediately prior to pouring. Resolving gels were poured to approximately 0.5cm below the inserted comb, gently overlaid with 20% (v/v) ethanol in 1 x TANK Buffer and allowed to set for 1h at room temperature. To prepare the 4.5% stacking gel, 440μl of 30% acrylamide solution (37.5:1) was combined with 780μl of Upper Gel Buffer (pH 6.8)[0.5M Tris, 0.4% (w/v) SDS] and 1.78ml dH_2O. The stacking gel was polymerised by adding 25μl of 10% (w/v) APS and 5μl of TEMED, then layered over the resolving gel and allowed to set for 1h at room temperature.

Following SDS-PAGE, proteins were transferred to Hybond ECL membrane (Amersham Biosciences) using a Trans-Blot® Semi-Dry Transfer Cell (Bio-Rad Laboratories). The gel sandwich was disassembled and the stacking gel was removed using a scalpel blade. The resolving gel and nitrocellulose membrane were then
equilibrated in Towbin transfer buffer (pH 8.3) [25mM Tris, 192mM glycine, 20% (v/v) methanol] and the transfer was assembled by applying Whatman® 3MM filter paper (Whatman) (3 sheets pre-wetted in Towbin transfer buffer), nitrocellulose membrane, polyacrylamide gel and Whatman® 3MM filter paper (3 sheets pre-wetted in Towbin transfer buffer). Transfer was performed at 15V for 30min and the membrane was divided through the pre-stained MWM lane to facilitate incubation of nuclear extracts and biotinylated MWM with appropriate antibodies. Membranes were then blocked for 1h at room temperature in 1x TBS-Tween buffer [1mM Tris, 7.5mM NaCl, 0.05% (v/v) Tween-20] containing 3.5% (w/v) gelatin, with gentle rocking.

2.2.3.3.4 Detection of HOX11 Protein

For the detection of transfected HOX11 proteins, the blocked membrane was washed (3 x 5min) in 1x TBS-Tween buffer and incubated with 1° anti-human HOX11 monoclonal antibody (from mouse) (Santa Cruz Biotechnology) diluted 1:1200 in 1xTBS Buffer containing 0.1% (v/v) Tween-20 and 0.5% (w/v) gelatin, for 1.5h at room temperature with gentle agitation. The membrane was washed (3 x 5min) in 1x TBS-Tween buffer and incubated with 2° anti-rabbit-IgG, peroxidase-linked species-specific antibody (from donkey) (Amersham Pharmacia Biotech), diluted 1:5000 in 1x TBS Buffer containing 0.1% (v/v) Tween-20 and 0.5% (w/v) gelatin, for 1h at room temperature with gentle agitation. The membrane containing the biotinylated protein MWM was washed (3 x 5min) in 1x TBS-Tween buffer and incubated with streptavidin-horse radish peroxidase (SA-HRP) (Amersham Pharmacia Biotech), diluted 1:2000 in 1x TBS Buffer containing 0.1% (v/v) Tween-20 and 0.5% (w/v) gelatin, for 1.5h at room temperature with gentle agitation. The membranes were then washed (3 x 5min) in 1x TBS-Tween buffer and ECL detection was performed by incubating with 5ml/10cm² of ECL substrate solution (Amersham Pharmacia Biotech) for 1min at room temperature. The membranes containing biotinylated MWM and nuclear extracts were then reassembled in Saran wrap and exposed to ECL Hyperfilm (Amersham Pharmacia Biotech) for 10s-2min, depending on the intensity of signal required and the film was developed using an SPM 300 developer (Fuji).
2.2.3.4 Luciferase and β-Galactosidase Reporter Gene Assays

The Tropix® Dual-Light® System (Applied Biosystems) employed for these assays, is a chemiluminescent reporter gene assay system, which allows for the sequential detection of luciferase and β-Galactosidase within the same sample volume. Harvesting of transformed cells was performed as recommended by the manufacturer’s instructions. For adherent lines (NIH 3T3), cells were washed (x2) in 2ml of PBS and lysed directly by incubating with Dual-Light® kit Lysis Solution containing freshly added DTT [0.5mM] to cover cells (125µl/well) for 10min at room temperature. The cells were then detached from the plate with a 23cm² blade cell scraper (NUNC™) and lysates were transferred to 1.5ml polypropylene microcentrifuge tubes and pelleted by centrifugation at 17,900 x g for 5min at 4°C (Eppendorf Centrifuge 5417C, Crown Scientific). The supernatants were collected on ice, and stored at -70°C. For suspension lines (HEL, PER-117), transformed cells were transferred to 15ml polypropylene tubes and washed (x3) with 5ml of ice-cold PBS. The cells were then resuspended in 250µl of Dual-Light® kit Lysis Solution containing freshly added DTT [0.5mM]. The cells were lysed passively in the hypotonic Lysis Solution by incubating on ice over a 10min period and pelleted by centrifugation at 17,900 x g for 5min at 4°C. The supernatants were collected on ice, and stored at -70°C. Prior to assaying, cell lysates were thawed and briefly centrifuged to pellet any cellular debris. In the case of HEL and NIH 3T3, lysates were diluted 1:2 with freshly prepared Dual-Light® Lysis solution containing 0.5mM DTT, whereas PER-117 lysates were assayed neat. 20µl of lysate or lysis solution alone (blank) was then assayed in an opaque 96-well microtitre plate (Ascent) using the Luminoskan Ascent® and Ascent™ software (Labsystems). The assay was performed according to the Chemiluminescent Detection Protocol (C). Buffers A and B were equilibrated to room temperature and the Galacton Plus® substrate was diluted (1:100) in the appropriate amount of Buffer B required for the entire assay. In order to maintain consistent timing for the addition of Buffers A and B to each sample on the plate, appropriate quantities for Buffer A (25µl/well) and Buffer B (100µl/well) were premixed and the relative light units (RLUs) of each luciferase reaction were read in counts per second (cps) following automatic injection (125µl/well). The microtitre plate was then incubated at room temperature for 1h and RLU measurements of the β-
Galactosidase light reaction (cps) were taken using the same multi-label counter following automatic injection of Accelerator II (100µl/well).

2.2.3.5 Data Processing

Transfections were normalised using a standard method. For each transfection being assayed, the activity (cps) was first corrected for noise by subtracting background (determined by assaying 20µl of Dual-Light® Lysis Solution). The corrected luciferase values were then normalised for transfection efficiency by dividing by the corresponding corrected β-Galactosidase value to obtain normalised RLU's for each transfection.

2.2.3.6 Statistical Manipulation of Data

Transfections were conducted a minimum of three times for each construct. Due to the inherent variability encountered in luciferase experiments, which include both transfection (cell specific variables) and assay parameters, each experiment was ‘centred’ prior to data processing. This involved taking the mean average RLU value for that experiment which was then used to divide the raw RLU values. Only those constructs that were common to all experiments in a given cell line were used to calculate the average RLU value. Centred data was then used to calculate the mean and standard errors for each construct using Microsoft Excel. Statistical analysis was also performed where indicated, using random effect models in SPLUS, by Dr Ross Taplin, Division of Science and Engineering, Murdoch University.

2.2.4 DNase1 Footprinting

2.2.4.1 PCR Labelling of the ALDH1A1 Proximal Promoter

2.2.4.1.1 Radiolabelling the Primer Using $^{32}$P-$\gamma$ATP

In order to generate radiolabelled ALDH1A1 promoter fragments for use in footprinting reactions, amplification was performed using radiolabelled sense (ALDHPF) or
antisense (ALDHPR) primers (Table 2.3). The labelling reactions were performed in 0.5ml polypropylene microcentrifuge tubes in a final volume of 15µl, containing 50ng of either ALDHPF/ALDHPR primer, 5x Forward Reaction Buffer [1x; 70mM Tris-HCl (pH 7.6), 10mM MgCl₂, 100mM KCl, 1mM 2-mercaptoethanol](GIBCOBRL®), 100 µCi of ³²P-γATP (Geneworks) and 10U of T4 Polynucleotide Kinase (PNK)(GIBCOBRL®), added in the order described. This reaction mixture was incubated at 37°C for 1h, and the enzyme was subsequently heat inactivated by incubation at 85°C for 20min. The reaction volume was adjusted to a final volume of 50µl total with sterile dH₂O, and the radiolabelled primer was purified from non-labelled primer using a Microspin G25 spin column, in accordance with the manufacturer’s specifications (Amersham, Pharmacia Biotech). Briefly, the resin within the G25 spin column was resuspended by inversion several times and a flicking motion. The bottom of the column was removed, the cap loosened 1/4 turn and the column was placed in a 1.5ml polypropylene microcentrifuge tube. The resin was then pelleted by centrifugation at 725 x g for 1min at room temperature (Biofuge Pico, Heraeus Instruments) and the column was transferred to a fresh tube. The radiolabelled sample was then applied to the centre of the angled surface of the compacted resin bed and centrifuged at 725 x g for 2min at room temperature. The exact volume collected was determined for subsequent DNA quantitation calculations and stored at -20°C.

2.2.4.1.2 Generation of Singly End-Labelled ALDH1A1P Fragments Using PCR

To generate double stranded, radiolabelled ALDH1A1 promoter fragments, high fidelity PCR was performed using primer sets (ALDH1PF-344/ALDH1PR2-123; generating an 225bp product)(ALDH1PR+42/ALDH1PF2-193; generating a 237bp product) containing radiolabelled and non-labelled primers (Table 2.3). PCR Reactions were performed in thin-walled 0.5ml polypropylene tubes in a final volume of 50µl, comprising 10x Pfu™ Polymerase Reaction Buffer [1x; 20mM Tris-HCl (pH 8.8), 10mM KCl, 10mM (NH₄)₂SO₄, 2mM MgSO₄, 0.1% Triton® X-100, 0.1mg/ml nuclease-free bovine serum albumin (BSA)](Stratagene), 25ng of either ALDH1PF-344/ALDH1PR+42 radiolabelled primer, 30ng of either ALDH1PR2-123 or ALDH1PF2-193 unlabelled primer, 0.2mM dNTPs (Fisher Biotech), 100ng of pGL2-484 DNA template and 1.25U of Pfu™ Turbo (Stratagene). Amplification was
performed using a PTC-100 (MJ Research Inc.) using a Touchdown programme comprising 30 cycles of 94°C for 1min, 63-55°C for 1min (temperature drops by 1°C every second cycle, with the last 14 cycles at 55°C and 72°C for 1.5min) (TD1; Appendix 1). The radiolabelled PCR product was purified using a Qiaquick PCR Purification Kit (Qiagen) as described in Section 2.2.1.1.6, with the exception that the DNA was eluted with 50µl Buffer EB. The exact volume retrieved from the column was determined for DNA quantitation calculations and stored at -20°C.

2.2.4.1.3 DNA Quantitation Using Activity

To measure the activity (reported in cpm/µl) of ³²P in the primers and ALDH1A1/P PCR product, dilutions of each test sample (1:10 dilution) were prepared using sterile dH₂O in 1.5ml polypropylene microcentrifuge tubes and counted in a scintillation counter (Bioscan/QC 2000). In order to determine the concentration of ALDH1A1/P PCR product the following equation was employed, where;

\[
\begin{align*}
    a &= \text{total activity of primer used per PCR reaction} \\
    b &= \text{total activity of labelled ALDH1A1/P PCR product} \\
    c &= \text{pmol of labelled primer (3.155)} \\
    d &= \text{pmol of double stranded ALDH1A1/P PCR product} \\
    0.66 &= \text{MW}_{\text{Average}} \text{ of a base pair} \\
    n &= \text{total base pairs of PCR generated product}
\end{align*}
\]

\[
\frac{b}{a} \times c = d \\
\frac{d \times 0.66 \times n}{\text{volume of PCR product retrieved}} = \text{ng}
\]

2.2.4.2 Footprinting Reactions

A typical footprinting reaction is outlined in Table 2.4. Control reactions, containing no nuclear extract, were treated in the same manner as tests. Reactions were performed in 1.5ml polypropylene microcentrifuge tubes in a final volume of 50µl comprising 5x Binding Buffer [1x; 20mM Hepes (pH 7.9), 100mM KCl, 0.2mM EDTA, 1mM DTT,
8 mM MgCl₂, 2% (v/v) glycerol, 1 µg of polydeoxycytidylic acid (polydIdC)(ICN Biomedicals) and 5-50 µg of nuclear extract and incubated at 22°C for 25 min. Following this, 15 ng of end-labelled target DNA (³²P-ALDH1A1) was added to all the tubes, which were then incubated at 22°C for a further 10 min. The target DNA was then subjected to DNase1 digestion. Tubes were digested in sets of 3-4 by adding 5x DNase1 Buffer [1x; 25 mM NaCl, 10 mM Hapes (pH 7.9), 5 mM MgCl₂, 1 mM CaCl₂] immediately prior to digestion, then initiating digestion with the addition of 0.01U-0.3U of DNase1 (Sigma), followed by a gentle vortex. The digestions were allowed to proceed for precisely 1 min 20 s and terminated with the addition of 10 µl of Stop Buffer [100 mM MES, 165 mM EDTA, 4.95% (w/v) SDS, 0.025% (w/v) herring sperm DNA], followed by thorough vortexing.

<table>
<thead>
<tr>
<th>Tube</th>
<th>Typical Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>5X Binding Buffer (µl)</td>
<td>10</td>
</tr>
<tr>
<td>PolydIdC (µg)</td>
<td>1</td>
</tr>
<tr>
<td>Nuclear Extract (µg)</td>
<td>10</td>
</tr>
<tr>
<td>dH₂O (µl)</td>
<td>Up to 50µl</td>
</tr>
</tbody>
</table>

**Incubate at Room Temperature for 25 min**

| Labelled ALDH1P (ng)        | 15               |

**Incubate at Room Temperature for 10 min**

| 1x DNase1 Buffer (µl)      | 50               |
| DNase1 (U)                 | (0.01U – 0.3U)   |
| Stop Buffer (µl)           | 10               |

**Table 2.4. The Components of a Typical Footprinting Reaction.** The order of addition and incubation times and temperatures are indicated. Reactions were incubated at 22°C on a heating block to ensure reproducibility of DNase1 digestion conditions.

2.2.4.3 **Phenol-Chloroform Ex extractions**

DNA was extracted with 100 µl of phenol:chloroform:isoamyl alcohol (25:24:1)(ICN Biomedical) and centrifuged to separate the phases at 15,000 x g for 3 min at room temperature (Biofuge Pico, Heraeus Instruments). The upper aqueous layer was
removed and transferred to a new 1.5ml polypropylene microcentrifuge tube containing 0.1 volumes of 3M sodium acetate, 2 volumes of cold absolute ethanol and 20µg of glycogen (ICN Biomedical). The tubes were incubated for 30min on dry ice and centrifuged at 15,000 x g for 15min at room temperature. The supernatants were carefully removed and discarded and the radioactive DNA pellets were overlaid with cold 200µl of cold 70% (v/v) ethanol and centrifuged for an additional 5min at 15,000 x g at room temperature. The ethanol was removed and the pellets were dried at 37°C in a preheated block and dissolved in the appropriate volume of formamide loading buffer [10mM EDTA, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanole and 98% (v/v) deionized formamide]. Reactions were loaded onto gels immediately or stored at -20°C until required.

2.2.4.4 Preparation of Sequencing Reactions

Footprints were localised with respect to the DNA sequence using the Cycle Sequencing Kit according to the manufacturer’s instructions (Pharmacia Biotech). Cycle sequencing reactions for ‘A’, ‘C’, ‘G’ or ‘T’ were performed in 0.5ml polypropylene microcentrifuge tubes in a final volume of 7µl, comprising Termination Mix, either [222.8µM ddATP], [6.86µM ddGTP], [85.7µM ddCTP] or [242.9µM dTTP](Pharmacia Biotech), 5x Sequencing Buffer [1x; 70mM Tris-HCl (pH 9.0), 1.5mM MgCl₂, 10.4% (v/v) DMSO](Pharmacia Biotech), 11/4µM dNTP Mix (Pharmacia Biotech), 100fmol of pGL2-484 DNA template, 0.6pmol of labelled primer and 0.2U of Taq DNA polymerase (Pharmacia Biotech). Reactions were overlaid with one drop (~20µl) of paraffin oil and centrifuged for 20s to remove air bubbles. Cycle sequencing reactions were performed using a PTC-100 thermocycler (MJ Research Inc) and involved 25 cycles of 95°C for 30s, 55°C for 36s and 72°C for 84s (SEQ1; Appendix 1). Reactions were terminated by the addition of 3µl Stop Solution [97.5% (v/v) deionized formamide, 10mM EDTA, 0.3% (w/v) xylene cyanol, 0.3% (w/v) bromophenol blue](Pharmacia Biotech) to each tube and stored at -20°C for use within three days of preparation.
2.2.4.5 Denaturing Polyacrylamide Gel Electrophoresis

DNase1 footprinting reactions were analysed under denaturing conditions by polyacrylamide gel electrophoresis using Sequi-Gen® GT gel apparatus (Bio-Rad Laboratories). Sequencing gel glass plates (38 x 50cm) were washed thoroughly with detergent, rinsed and dried before use. Both plates were cleaned twice with 96% (v/v) ethanol using lint-free towels (Kimwipes), and one plate was treated with a siliconising agent (Rain-X, UNELKO Corp.) in between ethanol washes. A 6% polyacrylamide sequencing gel consisting of 30% Acrylamide/Bis (19:1; w/v), 10x TBE Buffer [1x; 89mM Tris, 89mM Boric Acid, 2mM EDTA (pH 8.0)] and 7.5M urea was freshly prepared and 750µl of a 10% (w/v) APS solution and 60µl of TEMED were added per 150ml gel solution, immediately prior to casting. The apparatus was assembled with 0.4mm spacers and the gel was cast according to the manufacturer’s instructions. A 0.4mm, 32-well comb was inserted and the gel was left to set for a minimum of 3h before use. Electrophoresis was performed in 1x TBE buffer and pre-electrophoresed to a temperature of 50°C. Prior to electrophoresis, footprinting samples were denatured by heating at 70°C for 5min on a pre-heated block. Sequencing reactions were prepared by removing the overlay of paraffin oil and heating at 80°C for 3min. Footprinting samples (4µl/lane) and sequencing reactions (3µl/lane) were loaded in preformed wells and samples were electrophoresed at constant power (120W, 50°C) for 1.5-4.5h depending on the length of the run required.

2.2.4.6 Gel Drying and Autoradiography

Following electrophoresis, the gels were transferred to Whatman® 3MM filter paper (Whatman) and covered with Saran wrap. Gels were dried at 80°C for 2h in a pre-heated gel dryer (Model 583, Bio-Rad Laboratories), and autoradiographed using an intensifying screen (Fuji G-8 EC-AWU Cassette, 27.9x 35.6cm; Fuji Medical X-ray film super HRG30) at -70°C for the appropriate length of time. Prior to developing, cassettes were allowed to thaw for 1h at room temperature and films were developed using an SPM 300 developer (Fuji).
2.2.5 Electrophoretic Mobility Shift Assays (EMSAs)

2.2.5.1 Oligonucleotide Labelling and Annealing

Oligonucleotides used for EMSA experiments are outlined in Table 2.5. Labelling reactions were performed in 0.5ml polypropylene microcentrifuge tubes a final volume of 20µl comprising 5x Forward Buffer [1x; 70mM Tris-HCl (pH 7.6), 10mM MgCl₂, 100mM KCl, 1mM 2-mercaptoethanol](GIBCOBRL®), 20pmol of oligonucleotide, 40 µCi of γ³²P-ATP (Geneworks) and 15U of T4 Polynucleotide Kinase (PNK)(GIBCOBRL®) at 37°C for 30min. The PNK was subsequently heat inactivated by incubation at 65°C for 10min. The reaction volume was adjusted to a final volume of 50µl with sterile dH₂O and the radiolabelled oligonucleotide was purified using Microspin G25 spin columns (Amersham Pharmacia Biotech) as described in Section 2.2.4.1.1. The reaction volume was adjusted to a final volume of 100µl with sterile dH₂O and stored at -20°C until required. Annealing reactions were performed in 0.5ml polypropylene microcentrifuge tubes in a final volume of 30µl, with equimolar amounts of both radiolabelled and unlabelled sense and antisense oligonucleotides respectively [4pmol] in 5 x Annealing Buffer [1x; 10mM Tris-HCl (pH 7.4), 50mM NaCl, 1mM EDTA] by heating the reaction to 90°C for 5min and cooling to 22°C at a rate of 1°C/min using a PTC-100 thermocycler (MJ Research Inc). Following this, the reaction was adjusted to a final concentration of 0.04pmol/µl with 1x Annealing Buffer.
TABLE 2.5
2.2.5.2  **EMSA Reactions**

All EMSA reactions were performed in 1.5ml polypropylene microcentrifuge tubes using freshly thawed nuclear extracts. Binding reactions were performed in a final volume of 15µl comprising 5 x EMSA Binding Buffer [1x; 20mM Hepes (pH 7.6), 50mM KCl, 1mM EDTA (pH 8.0), 5% (v/v) glycerol, 1mM DTT (added fresh)], 0.5µg of polydIdC, 0.12pmol of ³²P-end labelled duplex oligonucleotide, 10µg of nuclear extract and 1µl of antibody where appropriate. Competition experiments involved the addition of either unlabelled duplex related competitor or unrelated competitor in molar excess (5-35 fold) relative to the radiolabelled probe prior to the addition of nuclear proteins. The reactions were incubated at 22-24°C for 30min prior to electrophoresis.

2.2.5.3  **Native Polyacrylamide Gel Electrophoresis**

Electrophoresis of EMSA reactions was performed at 4°C, using the PROTEAN II Gel System (20cm cell)(Bio-Rad Laboratories). Glass plates (inner plate; 20cm x 20cm, outer plate; 22.3cm x 20cm) were washed thoroughly with detergent, rinsed and dried before use. Both plates were cleaned twice with 96% (v/v) ethanol using lint- free towels, and the short plate was treated with a siliconising agent (Rain-X, UNELKO Corp.) in between ethanol washes. The gel apparatus was assembled using 1.5mm spacers and a 1.5mm, 15 well comb according to the manufacturer’s instructions. A 4% polyacrylamide gel consisting of 30% Acrylamide/Bis (37.5:1; 13.3%v/v) and 10x TBE [0.5x; 44.5mM Tris, 44.5mM Boric acid, 1mM EDTA (pH 8.0)] was prepared and 525ul of a 10% (w/v) APS solution and 52.5ul of TEMED were added per 70ml of gel solution, immediately prior to casting. The gel was left to set for a minimum of 2h and stored at 4°C overnight. Freshly prepared 0.5x TBE Buffer and the central cooling core were also pre-incubated at 4°C until required. Prior to loading, the gel was prepared by pre-running with chilled 0.5x TBE Buffer at 17mA (constant current) for 1h. The entire sample volume (~15µl/lane) was electrophoresed at 17mA constant current for 1.5-2h, until the dye front measured ~2cm from the end of the gel and transferred to Whatman® 3MM filter paper. Gels were dried for 2h at 80°C under vacuum using a Bio-Rad Model 583 Gel Dryer (Bio-Rad Laboratories) and subject to autoradiography using an intensifying screen (Fuji G-8 EC-AWU Cassette, 27.9x 35.6cm; Fuji Medical X-ray
film super HRG30) at -70°C for the appropriate length of time. Prior to developing, cassettes were allowed to thaw for 1h at room temperature and films were developed using an SPM 300 developer (Fuji).

2.3 RESULTS

2.3.1 Identification of a HOX11 Responsive Element Within the ALDH1A1 Promoter

Previous studies have demonstrated that HOX11 consistently affects the expression of a ~2.1kb ALDH1A1 transcript in both the NIH 3T3 and HEL cell lines, as assessed by Northern blot (Greene et al., 1998, Greene, unpublished observations). In the case of NIH 3T3 cells there was a consistent upregulation, while in HEL cells there was a consistent downregulation of ALDH1A1 expression. Moreover, HOX11 was shown to upregulate ALDH1A1 expression at the level of RT-PCR in the T-ALL cell line, PER-117 (Greene, unpublished observations). Given that the TSS of human cytosolic ALDH1A1 had been previously determined by Hsu et al., (1989) and corresponds to the 2.1kb transcript observed in NIH 3T3 and HEL cell lines, preliminary investigations utilised Chloramphenicol Acetyl Transferase (CAT) assays to determine the effect of HOX11 on a 2536bp region 5’ to the transcriptional start site of ALDH1A1 (designated +1)(Greene, unpublished observations). Overexpression of HOX11 in NIH 3T3 cells stimulated ALDH1A1 promoter activity, suggesting the presence of a HOX11 responsive element within this region. Thus, in order to identify regions of the ALDH1A1 promoter involved in the cell type-specific expression of ALDH1A1 in HEL (human erythroleukaemia) and PER-117 (immature T-cell) cell lines, and to delineate the specific element(s) responsible for mediating the effects of HOX11, with the intention of gaining further insight into the molecular mechanism(s) by which HOX11 alters the transcriptional activity of ALDH1A1 (and possibly other target genes), a series of luciferase reporter constructs containing progressive deletions of a 2,536bp region of the ALDH1A1 promoter were transfected into the HEL and PER-117 cell lines. The reporter gene activity of the various constructs was then assessed in both the presence and absence of HOX11.
The nested deletions of the *ALDH1A1* promoter within the region -2159bp to +42bp (relative to the TSS +1; Figure 2.2) were amplified from human genomic DNA and cloned upstream of the firefly (*Photinus pyralis*) luciferase cDNA of the pGL3Basic luciferase reporter vector (Figure 2.3). The basis for this assay lies in the ability of the luciferase enzyme (a 61kDa protein) to oxidise the substrate D-luciferin in the presence of ATP, oxygen and Mg²⁺, yielding a fluorescent product that may be quantified by detecting released light. Since the production of the luciferase enzyme is under the control of the regulatory sequence of interest, the emission of light is therefore a measure of the activity of the promoter under a given set of conditions. The sensitivity of this assay, coupled with its relative simplicity and broad linear range, which spans over eight orders of magnitude, makes the luciferase assay an ideal system for preliminary investigations of gene regulatory sequences. The pGL3Basic vector utilised in these studies, lacks eukaryotic promoter or enhancer sequences, allowing the cloning and functional assessment of putative regulatory sequences, and also contains a polyadenylation site upstream of the multiple cloning site to prevent read-through of aberrant transcripts potentially initiating within the vector backbone, thereby reducing background luciferase noise.

In order to assess the effect of HOX11 on *ALDH1A1* promoter expression, individual promoter reporter gene constructs were cotransfected with either the empty pEFBOS mammalian expression vector, which served as a negative control for HOX11 expression, or with the pEFBOSHOX11 expression vector (kindly supplied by Dr T.H. Rabbitts, MRC Laboratory of Molecular Biology, Cambridge, U.K). The pEFBOS mammalian expression vector (5.8kb) is derived from the pUC119 plasmid and employs the promoter of polypeptide chain elongation factor 1α (*EF-1α*) to drive the expression of genes cloned into the multiple cloning site (Mizushima & Nagata, 1990)(Appendix 2). The vector also contains the polyadenylation signal from the human G-CSF gene as well as the ampicillin (*Amp*) gene as a selectable marker. The pEFBOSHOX11 construct is derived from the pEFBOS parental vector and contains a 1044bp human HOX11 cDNA for the expression of full-length HOX11 (Masson *et al.*, 1998). All transfections also included the pSV-β-Galactosidase (pSV-β-Gal) vector as the internal control to facilitate the normalisation of raw luciferase output for individual samples.
FIGURE 2.2
Figure 2.3. Schematic Diagram of the Human *ALDH1A1* Promoter Luciferase Fusion Constructs Employed in this Study. Nested deletions of the human *ALDH1A1* promoter were amplified by high fidelity PCR using *PfuTurbo*® (Stratagene) and inserted into the multiple cloning site of the pGL3Basic luciferase vector (Promega), upstream of the cDNA encoding the modified firefly luciferase (LUC).
The use of pSV-β-Gal allowed the efficiency of transfection of reporter constructs into HEL and PER-117 cells and other variations to be taken into account. pSV-β-Gal (6.8kb)(Promega; Genbank®/EMBL Accession Number: X65335; Appendix 2) uses the SV40 early promoter and enhancer sequences to drive the high-level expression of bacterial lacZ. This encodes an enzyme (β-Galactosidase) responsible for catalysing the hydrolysis of β-galactoside sugars such as the colorimetric substrate, o-nitrophenyl-β-D-galactopyranoside (ONPG), which can be spectrophotometrically quantitated at a wavelength of 410nm. Using the Dual Light® reporter gene assay system (Applied Biosystems), which incorporates luciferin and Tropix® Galacton Plus® substrates in provided assay buffers, it was possible to concomitantly assay both experimental (luciferase) and internal control (β-Galactosidase) gene products within the same cell extract.

2.3.1.1 Optimisation of Luciferase Reporter Gene Assays

ALDH1A1 was originally identified as a gene upregulated following the overexpression of HOX11 in NIH 3T3 murine fibroblasts (Greene et al., 1998), and subsequent CAT reporter assays revealed the presence of a HOX11-responsive element within a 2159bp fragment (-2159/+42bp) relative to the TSS of the ALDH1A1 promoter in NIH 3T3 cells (Greene, unpublished observations). Therefore, prior to commencing ALDH1A1 promoter analyses in the HEL and PER-117 cell lines, experiments were performed to demonstrate the reproducibility of results obtained using the luciferase reporter assay as compared to those obtained in previous studies utilising the CAT reporter assay. In accordance with previous studies, HOX11 was shown to transactivate the ALDH1A1 promoter following cotransfection of the full-length pGL3BasicALDH1A1P-2159/+42 luciferase construct and the HOX11-expressing construct, pEFBOSHOX11 in NIH 3T3 cells (data not shown). Having confirmed the responsiveness of the ALDH1A1 promoter to HOX11 in NIH 3T3 cells, all subsequent analyses were performed in the human erythroleukaemic cell line HEL, which strongly expresses ALDH1A1, together with the non-ALDH1A1 expressing cell line, PER-117, established from the leukaemic cells of a patient with T-ALL. Although both of these cell lines do not normally express HOX11, their leukaemic origins were likely to provide a relevant background with respect to specific cofactors, growth factors and signal transduction pathways in which to study the differential effects of HOX11 on ALDH1A1 expression. In stable transfection
experiments, HOX11 was shown to repress the expression of the endogenous ALDH1A1 gene in HEL cells and upregulate it in PER-117 cells (Table 2.6).

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>HEL</th>
<th>HELHOX11</th>
<th>PER-117</th>
<th>PER-117HOX11</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOX11 Expression</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+++</td>
</tr>
<tr>
<td>ALDH1A1 Expression</td>
<td>++++</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

**Table 2.6. Expression Status of HOX11 and ALDH1A1 in HEL and PER-117 Cell Lines.** HELHOX11 and PER-117HOX11 refer to clones produced by stable transfection with the HOX11-expressing construct pEFBOSHOX11. The expression status of HOX11 and ALDH1A1 was compared to negative clones of HEL and PER-117, which were produced by stable transfection with the empty pEFBOS vector alone (Greene, unpublished observations).

### 2.3.1.2 ALDH1A1 Promoter Activity in PER-117 and HEL Cell Lines

To identify the region(s) important for ALDH1A1 promoter transactivation in HEL and PER-117 and cells, the transcriptional activities of various deletions of the ALDH1A1 promoter were analysed using luciferase assays. Initial investigations tested constructs ranging from -2159/+42 to -146/+42 (Figure 2.4, 2.5). Data is presented in the conventional format which involves the normalisation of raw luciferase (cps) to raw β-Galactosidase (cps)(raw luciferase - lysis solution blank)/raw β-Gal - lysis solution blank), to yield normalised RLUs. Consistent with previous work (Yanagawa et al., 1995), these analyses revealed that the main transcriptional activity resided within a proximal region of the ALDH1A1 promoter, in this case within 146bp relative to the TSS (+1) in both HEL and PER-117 (Figure 2.4A, 2.5A). Constructs containing 5’ extensions of the promoter sequence (-1968 and -2159) were associated with a modest reduction in luciferase activity in both HEL and PER-117 cell lines, suggesting that distal negative regulatory elements may exist between -146/-2159bp (Figure 2.4A, 2.5A). An additional negative element between -201 and -303 was also apparent in both cell lines. In order to delineate the specific element(s) responsible for promoter activity, a series of constructs ranging from -91 to +1 were subsequently tested (Figure 2.4B, 2.5B). This included the -91ΔCAT construct, which harbours a deletion of the CCAAT box (located at -74/-70bp), since this element has previously been demonstrated to be
FIGURE 2.5
crucial for ALDH1A1 promoter transactivation in Hep3B human hepatocellular carcinoma cells (Yanagawa et al., 1995). Luciferase activity dramatically declined following deletion of the CCAAT box in both HEL and PER-117, confirming the importance of this cis-regulatory element for ALDH1A1 promoter activation in the HEL and PER-117 cell lines (Figures 2.4B, 2.5B).

2.3.1.3 The Effect of HOX11 on the pGL3Basic Luciferase Vector and on the pSV-β-Galactosidase Internal Control

Initial investigations into the effect of HOX11 on the activity of the ALDH1A1 promoter in the HEL and PER-117 cell lines, involving the transient introduction of promoter constructs, pEFBOS, pEFBOSHOX11 and the pSV-β-Gal internal control, revealed an unexpected phenomenon. It appeared that raw β-Galactosidase (β-Gal) values (cps) were typically lower in HOX11-transfected cells as compared to their matched pEFBOS-transfected counterparts, although this effect was more pronounced in the HEL cell line (Figure 2.6A) as compared to PER-117 cells (Figure 2.6B). This observation raised concerns that HOX11 was affecting the expression of the pSV-β-Gal internal control, either by actively repressing the SV40 promoter driving β-Gal expression, by sequestering positive trans-acting factors to other regulatory elements or by quenching mechanisms. The ability of certain transcription factors to affect the transcriptional activity of β-Gal co-reporter plasmids driven by various constitutively expressed promoters (RSV, CMV, pTK) has been previously documented (Huszar et al., 2001). It was suggested that the use of internal control plasmids to normalise luciferase activity in transient transfections is questionable and may lead to errors in data analysis. Moreover, a recent publication by Owens et al., (2003), demonstrated the ability of HOX11 to repress the SV40 viral promoter, thereby substantiating concerns regarding the use of pSV-β-Gal as an unbiased, internal normalisation control. Based on our evidence, which suggests that HOX11 represses the SV40 promoter driving β-Gal expression, this method is likely to affect the data in two ways depending on whether HOX11 activates or represses ALDH1A1 gene transcription. Thus, in a situation in which HOX11 activates ALDH1A1 promoter activity, the fold activation of HOX11-transfected samples compared to matched control samples is likely to be overestimated, since normalised RLUs will be artificially large due to the effect of dividing HOX11-transfected raw luciferase values by smaller β-Gal values. Conversely, in a situation
FIGURE 2.6
in which HOX11 represses \textit{ALDH1A1} promoter activity, the fold repression of HOX11-transfected samples compared to matched pEFBOS-transfected samples will be underestimated, this time due to the fact that normalised RLUs for these samples will be artificially large.

As an additional complication, it appeared that in HEL cells, HOX11 was capable of affecting luciferase activity, even in the absence of functional promoter sequences (Figure 2.7). HOX11 was demonstrated to activate transcription approximately 10-fold over background (pEFBOS) from the pGL3Basic reporter vector alone in the HEL cell line, raising concerns that the effect of HOX11 on luciferase activity may be independent of inserted promoter sequences (Figure 2.7). Indeed, although the pGL3Basic vector has been engineered with polyadenylation sites upstream and downstream of the luciferase coding region in order to minimise the effect of background read-through via cryptic promoter and enhancer elements, the ability of the osteoblast-specific transcription factors Cbfa1 and estrogen receptor \( \alpha \) to activate the pGL2- and pGL3Basic luciferase reporter vectors, possibly via randomly occurring \textit{cis}-acting elements in the vector backbone, has been previously demonstrated (Thirunavukkarasu \textit{et al.}, 2000).

Thus, in order to determine whether the effect of HOX11 on the \textit{ALDH1A1} promoter constructs was significantly different from the effect of HOX11 on the pGL3Basic vector itself, the data was statistically analysed in order to take into account the effect of HOX11 on the pSV-\( \beta \)-Gal internal control as well as on the pGL3Basic vector. Before undertaking this analysis, it was first necessary to assess whether \( \beta \)-Galactosidase expression was equally affected by HOX11 regardless of the pGL3Basic or \textit{ALDH1A1} promoter construct used in both the HEL and PER-117 cell lines. Statistical analysis was performed by Dr Ross Taplin (Division of Science and Engineering, Murdoch University). Raw \( \beta \)-Galactosidase values were logarithmically transformed (\( \log_2 \)) and expressed as a ratio of the promoter with HOX11 versus the promoter without HOX11. Each ratio was then expressed by subtracting the corresponding ratio for the pGL3Basic empty vector control. The mean was taken for triplicate experiments and two sided p-value testing for the null hypothesis that the values were equal to zero demonstrated at the \( P < 0.05 \) level of significance that the values were not significantly different from
zero. This analysis demonstrated that HOX11 affected the β-Galactosidase internal control similarly for pGL3Basic as for all ALDH1A1 promoter constructs in both the PER-117 (Figure 2.8A) and HEL (Figure 2.8B) cell lines, such that the pGL3Basic empty vector control may be included in subsequent normalisation calculations.

2.3.1.4 Modulation of ALDH1A1 Promoter Transactivation by HOX11 in HEL and PER-117 Cells

HOX11 has been shown to activate the expression of endogenous ALDH1A1 in PER-117 cells, and repress the expression of endogenous ALDH1A1 in HEL cells (Greene, unpublished observations). We therefore examined whether HOX11 could modulate transactivation of the ALDH1A1 promoter constructs in these cell lines. As described in the previous section (Section 2.3.1.3), the net effect of HOX11 on ALDH1A1 promoter activity was calculated by taking a ratio of raw luciferase activity (cps) over raw β-Galactosidase activity (cps) for each well to obtain normalised RLU values. Following log₂ transformation, values were then expressed as a ratio of the ALDH1A1 promoter construct with HOX11 vs no HOX11, and all values were expressed by subtracting the corresponding ratio for the pGL3Basic vector control. This value was then expressed graphically as a log fold change of individual promoter constructs relative to the pGL3Basic empty vector control. Two sided p-value testing was then undertaken for each promoter construct relative to pGL3Basic empty vector and between individual promoter constructs, using the null hypothesis of no difference. Western blot analyses were performed on HEL and PER-117 cell lysates, confirming the expression of the 36kDa HOX11 protein in pEFBOSHOX11-transfected cells (Figure 2.9B, 2.10B).

Expression of HOX11 in HEL cells resulted in a modest upregulation of the ALDH1A1 promoter within the smallest construct +1 by a factor of 1.6 fold (P < 0.05). This activation by HOX11 was sustained in constructs ranging between +1 and -91 and did not require the CCAAT box at -70, since deletion of this element had no effect (Figure 2.9A). The introduction of additional promoter sequence (-146) abrogated the effect of HOX11 (P < 0.005), and this was maintained to a greater or lesser extent in all the progressively longer constructs (Figure 2.9A). These results suggested that HOX11 may activate transcription through an element located between +1/+42bp of the ALDH1A1 promoter, in a manner subject to negative regulation by factors acting on a region
between -91/-146bp. In contrast, expression of HOX11 in PER-117 cells resulted in the significant repression of the -91 construct and of progressively larger constructs encompassing this region by a factor of 2-3.5 fold (P < 0.001). However, no repression by HOX11 was demonstrated using constructs -50 or those progressively smaller (Figure 2.10A). Significantly, HOX11-mediated repression was relieved following deletion of the CCAAT box (P < 0.001), suggesting that the CCAAT box is not only crucial for ALDH1A1 promoter activity, but may also represent a HOX11 responsive element in PER-117 cells.

2.3.1.5 Effect of a Homeodomain Mutation on HOX11 Transactivation/Repression Potential

In order to assess whether DNA binding might be required for HOX11 to mediate transactivation in HEL cells and repression at the CCAAT box in PER-117 cells, a DNA binding mutant of HOX11 (pEFBOSHOX11ΔH3) was cotransfected with the empty pGL3Basic and pGL3Basic-91 constructs for both HEL and PER-117 cell lines. HOX11ΔH3 contains a 10 amino acid (aa) deletion (DAQVKTWFQN) within helix 3 of the HOX11 homeodomain, which has been replaced by the dipeptide GT (Masson et al., 1998). It has been shown by EMSAs to be incapable of binding a HOX11 consensus site (Heidari et al., 2002). In the first instance, the effect of HOX11ΔH3 on β-Galactosidase activity in HEL and PER-117 was assessed (Figure 2.11). This analysis revealed that unlike HOX11, which represses the expression of β-Galactosidase in HEL and to a lesser extent in PER-117 (Figure 2.6A, 2.6B), HOX11ΔH3 stimulated expression of β-Galactosidase in both HEL and PER-117 compared to the pEFBOS control (Figure 2.11A, 2.11B). These results confirmed the ability of HOX11 to modulate the expression of β-Galactosidase, and suggested that DNA binding is required for HOX11-mediated repression. By contrast, the ability of HOX11 to transactivate transcription is apparently independent of DNA binding, possibly occurring via protein-protein interactions. Regardless of the mechanism of action, normalisation of luciferase values through the use of β-Galactosidase would clearly distort the final interpretation of data when using the HOX11ΔH3 mutant. For this reason, the effect of HOX11ΔH3 on the ALDH1A1 promoter is presented as raw luciferase (cps), without normalisation to β-Galactosidase.
In the case of HEL, these analyses revealed that like HOX11, HOX11ΔH3 was able to stimulate luciferase activity from both the pGL3Basic and pGL3Basic-91 construct (2.12A). Conversely, in the case of PER-117, the removal the DNA binding helix (ΔH3) disabled the ability of HOX11 to repress transcription from the -91 ALDH1A1 promoter construct, and intriguingly, HOX11ΔH3 actually stimulated transcription approximately 5-fold compared to the pEFBOS control (Figure 2.12B). Together, these results support the hypothesis that HOX11 modulates gene transcription via two distinct mechanisms – the first involves the ability of HOX11 to repress gene transcription which apparently requires DNA binding, whereas the second mechanism involves transcriptional activation which appears to be independent of DNA binding.

2.3.2 Identification of Transcription Factor Binding Sites within the Proximal Promoter of ALDH1A1 by DNase1 Footprinting

DNase1 footprinting is a widely used method for studying the binding of regulatory proteins to their cognate DNA recognition sequences (Galas & Schmitz, 1978). DNase1 interacts with DNA in the minor groove and in the presence of Ca\(^{2+}\) and Mg\(^{2+}\), cleaves a single strand of double stranded DNA to yield 5'-PO\(_4\) and 3'-OH products. This technique involves the partial digestion by DNase1 of a \(^{32}\)P-end-labelled fragment containing putative regulatory sequences of interest, to generate a unique pattern of fragments, which are subsequently separated on a denaturing acrylamide gel such that their positions represent the distance from the end label to the points of cleavage. Sequence specific binding of proteins to DNA sequences on the radiolabelled probe, protects the region from DNase1 cleavage, thus generating a ‘footprint’ in the pattern of cleaved fragments. The exact position of the protected region may then be identified by running a DNA sequencing ladder adjacent to the footprint. Having demonstrated clear differences in the transcriptional activity of the ALDH1A1 promoter in HOX11-transfected cells by \textit{in vitro} luciferase reporter assays, we attempted to characterise the sequence-specific binding of nuclear proteins to the proximal ALDH1A1 promoter by DNase1 footprinting and to determine the effect of HOX11 on protein binding within this region. On the basis of the functional promoter assays performed in Section 2.3.1, in addition to studies conducted by Yanagawa et al., (1995), two overlapping fragments (-344/-123bp and -193/+42bp) encompassing the proximal ALDH1A1 promoter were
FIGURE 2.12
selected to further investigate the regulatory elements involved in \textit{ALDH1A1} expression. Stably transfected cell clones were used for this part of the investigation. HEL clone 1 and PER-117 clone 1 are HOX11-negative and act as negative controls, while HEL clone 4 and PER-117 clone 11 are positive for HOX11 expression. Crude nuclear fractions from HEL[1], HELHOX11[4] and PER-117[1], PER-117HOX11[11] cell lines were incubated with genomic fragments $^{32}$P-radiolabelled at a single terminus, such that upon separation of digested fragments during electrophoresis only the DNA strand extending from the radiolabelled primer was visualised by autoradiography.

\textbf{2.3.2.1 \textit{DNaseI} Footprinting Using \textit{HEL} and \textit{PER-117} Nuclear Extracts}

The results of footprinting experiments for HEL[1], revealed consistent protection at two distinct positions (Figure 2.13) and the nucleotide sequence of these sites and the putative transcription factor binding sites pertaining to these regions are summarised in Table 2.7.
Table 2.7. Nucleotide Sequences of Protected Sites Identified by DNase1 Footprinting for HEL[1], HELHOX11[4], PER-117[1] and PER-117HOX11[11] Nuclear Extracts on the Proximal ALDH1A1 Promoter. Potential transcription factor binding sites contained within the footprint regions were identified using the TRANSFAC (Version 1.3) database (http://transfac.gbf.de/index.html), and are indicated by bold type.

The most distal site of protection identified, ALDHDIST, spanned nucleotides ranging from -257/-246bp (Figure 2.13). Analysis of ALDHDIST on the sense strand (extending from radiolabelled primer ALDH1PF-344) revealed strong protection characterised by adjacent hypersensitive sites that was also visualised on the antisense strand (data not shown). The sequence of ALDHDIST was notable for three reasons. Firstly the sequence has an axis of symmetry, in which the five flanking bases either side of a central TA dinucleotide are a direct mirror image (ACACT TA TCACA). Secondly the sequence harbours one of four putative HOX11 binding sites within the 2.2kb region of the ALDH1A1 promoter. Finally the sequence contains a putative GATA site.

The second region of protection observed for HEL[1], ALDHPROX, extended protection from -36 to -27bp (Figure 2.13). Analysis of ALDHPROX on the antisense
strand (extending from radiolabelled primer ALDH1PR+42) revealed relatively strong protection characterised by significant hypersensitivity extending from -47/-37bp and between -26/-22bp and was also visualised on the sense strand (data not shown). Sequence analysis of ALDHPROX identified a putative GATA binding site (-34/-29bp) overlapping a non-canonical TATA box (ATAAAA)(-32/-27bp), potentially capable of binding TFIID. This appeared to be a GATA box (-34/-28bp), a composite TATA/GATA element, capable of binding both GATA and TFIID proteins (Barton et al., 1993). The GATA box is therefore distinct from a GATA site, which is only capable of binding GATA family members. Footprinting experiments using HELHOX11[4] nuclear extracts revealed identical regions of protection in terms of nucleotide sequence and associated regions of hypersensitivity, thus HOX11 appeared to have no discernable effect on protein binding at the two identified sites. No protection was observed within the region +1/+42bp for either HEL[1] or HELHOX11[4], which was somewhat surprising, given that this was the minimal region required for the stimulation of transcription by HOX11 observed in Section 2.3.1.2.

Footprinting experiments with PER-117[1] nuclear extracts revealed identical regions of protection at -257/-246bp (ALDHDIST) and -36/-27bp (ALDHPROX). For this reason the results are not shown. Analysis of ALDHDIST on the sense strand revealed identical protection and associated hypersensitivity to that seen for HEL[1] and similar to HEL[1], protection was also observed on the antisense strand (data not shown). A comparison of footprints obtained for PER-117[1] vs PER-117HOX11[11] nuclear extracts also revealed no observable differences in terms of nucleotide sequences or associated hypersensitive sites on either sense or antisense strands for either ALDHDIST/PROX (data not shown). Surprisingly, DNase1 protection was not observed at the CCAAT box for either HEL[1] or PER-117[1], despite this region being identified as an important cis-regulatory element involved in the expression of ALDH1A1 in HEL and PER-117 cell lines, as well as representing a possible HOX11-responsive element in PER-117.
2.3.3 Characterisation of the Protein Binding at the CCAAT Box by EMSA

It was demonstrated by *in vitro* luciferase reporter assays that the CCAAT box located at -74/-70bp is the primary *cis*-acting element responsible for *ALDH1A1* promoter activity in both HEL and PER-117 cell lines, since the specific deletion of this element dramatically reduced its transactivation potential. Moreover, the CCAAT box was also shown to be involved in HOX11-mediated repression in the T-ALL cell line, PER-117. DNase1 footprinting of the proximal *ALDH1A1* promoter (spanning -344 to +42bp) encompassing the CCAAT box, however, did not reveal protein binding at this site. Thus, in an attempt to correlate the functional activity of the CCAAT box with protein-DNA binding, and to assess the effect of HOX11 on complex formation at this site, electromobility shift assays (EMSAs) were performed using an oligonucleotide probe spanning the CCAAT box (*ALDH*CAT). Nuclear extracts from HOX11 negative cell lines HEL[1] and PER-117[1] were compared to their HOX11-expressing counterparts HELHOX11[4] and PER-117HOX11[11]. In addition to the CCAAT box, two other regions of interest identified by DNase1 footprinting experiments were also included in EMSA analyses, in order to develop a basic model of transcription factor binding at these sites. These include a region of protection at -257/-246bp (*ALDH*DIST), in addition to a region of protection surrounding a putative GATA box at -34/-28bp (*ALDH*PROX).

2.3.3.1 EMSA Assays Using PER-117 Nuclear Extracts

Having demonstrated that HOX11 is capable of repressing the *ALDH1A1* promoter, possibly via interactions at the CCAAT box, we sought to investigate the effect of HOX11 on protein-DNA binding complexes within this region, which may constitute a transcriptional mechanism for the differential regulation observed in the previous Section 2.3.1.4. Incubation of the *ALDH*CAT with PER-117 and PER-117HOX11 nuclear extracts identified eight protein-DNA complex species (A-H), three of which were deemed specific (A, C and D) following elimination of complexes B, E, G and H with a 35-fold molar excess of either non-specific competitor 1 (NS1) or 2 (NS2) and inability of a 35-fold molar excess of unlabelled *ALDH*CAT probe to efficiently compete complexes F and G (Figures 2.14, 2.15). Although binding of complexes C and
Figure 2.14. Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter CCAAT Box (-74/-70bp). Radiolabelled ALDHCAT oligonucleotide probes (0.12pmol) were incubated with 10µg of nuclear extract to reveal complexes A-H (Lanes 2, 3). Specific complexes (A, C and D) were identified following competition experiments using either a 5 or 35-fold molar excess of unlabelled non-specific competitor (ALDHDIST; Lanes 4-7; NS1; Lanes 12, 13) or specific competitor (ALDHCAT; Lanes 8-11).
Figure 2.15. Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDHI1A1 Promoter CCAAT Box (-74/-70bp). Radiolabelled ALDHCAT oligonucleotide probes (0.12pmol) were incubated with 10μg of nuclear extract to reveal complexes A-H (Lanes 2, 3). Specific complexes (A, C and D) were identified following competition experiments using either a 5 or 35-fold molar excess of unlabelled specific competitor (ALDHCAT; Lanes 4, 5) or non-specific competitor (ALDHDIST; Lanes 6, 7, NS2; Lanes 8, 9). Supershift analysis was performed using monoclonal anti-HOX11 antibody (1-3ul)(Lanes 10-13).
D appeared similar for both PER-117 and PER-117HOX11, the much weaker complex A was consistently shown to be specific for PER-117 only, suggesting that the expression of HOX11 is associated with loss of a transcriptional complex at the CCAAT box (Figures 2.14, 2.15). In order to investigate the potential of ALDHDIST, which contains an in vitro HOX11 binding site, to bind factors at the CCAAT box, competition experiments were also performed using the ALDHDIST oligonucleotide, however complexes A, C and D remained intact, suggesting the binding of unique factors at each site (Figures 2.14, 2.15).

Whilst the expression of HOX11 in PER-117 extracts was not associated with the acquisition of a HOX11-specific complex, supershift experiments were performed with monoclonal anti-HOX11 antibody (SantaCruz) in the event that a masked HOX11-specific complex may be perturbed by antibody binding (or that disruption of HOX11 function by antibody binding may result in the formation of a unique complex). However, no evidence of a complex involving HOX11 was observed (Figure 2.15). This may have been due to the HOX11 antibody itself, since although demonstrated to be specific, it has not been shown to be capable of causing a supershift. In order to confirm the specific involvement of the CCAAT box in the formation of complexes A, C and D, competition experiments were performed using the mutant oligonucleotide (ALDHCATMut), in which the CCAAT box sequence was replaced with random nucleotides (GACTG), whilst flanking sequences were left intact (Figure 2.16). Unlike the specific competitor, the mutant ALDHCATMut oligonucleotide was unable to efficiently compete complexes A, C and D, confirming that these complexes are indeed specific to nucleotides comprising the CCAAT box.

Despite previous studies which indicated that the ATAAA sequence located at -32/-29bp is not a primary regulatory element in Hep3B cells (Yanagawa et al., 1995), our DNase1 footprinting assays indicated transcription factor binding within this region, suggesting that it may be involved in the transcriptional regulation of ALDH1A1 in both the PER-117 and HEL cell lines. In order to investigate protein binding at this site, PER-117 and PER-117HOX11 nuclear extracts were incubated with the ALDHPROX oligonucleotide probe encompassing the GATA box (Figure 2.17). Gel retardation assays identified five complexes (I-M) and competition with unlabelled ALDHPROX
Figure 2.16. Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter CCAAT Box (-74/-70bp). Radiolabelled ALDH CAT oligonucleotide probes (0.12pmol) were incubated with 10ug of nuclear extract to reveal complexes A-H (Lanes 2, 3). Specific complexes were identified following competition experiments using a 35-fold molar excess of unlabelled specific competitor (ALDH CAT; Lanes 4, 5). The specific involvement of the CCAAT box in complex formation was assessed by competing binding to the CCAAT box with a 35-fold excess of unlabelled ALDH CAT Mut oligonucleotide (Lanes 6, 7).
**Figure 2.17. Electromobility Shift Analysis of PER-117 and PER-117 HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-29bp).** Radiolabelled *ALDH*PROX oligonucleotide probes (0.12pmol) were incubated with 10ug of nuclear extract to reveal complexes I-M (Lanes 2, 3). Specific complexes were identified following competition experiments using either a 5 or 35-fold molar excess of unlabelled specific competitor (*ALDH*PROX; Lanes 8-11) and non-specific competitor (NS1; Lanes 12, 13). The involvement of *ALDHDIST* in complex formation was assessed by competing binding to the GATA box with a 35-fold excess of unlabelled *ALDHDIST* oligonucleotide (Lanes 4-7).
and unrelated competitor oligonucleotides confirmed the specificity of complexes I, J, K and M. Of particular interest was the substantial reduction in the intensity of complexes I and J in PER-117HOX11 compared to PER-117, whereas the formation of complexes K and M appeared enhanced in PER-117HOX11 compared to PER-117. Thus, the expression of HOX11 appears to be associated with quantitative (and possibly qualitative) differences in transcription factor binding profiles at the GATA box, which were not detectable by DNase1 footprinting. Although the DNase1-protected region identified at -257/-246bp (ALDHDIST) was not linked to significant changes in transcriptional activity in vitro luciferase experiments for either HEL or PER-117, a possible link between ALDHDIST and transcription factor binding at the GATA box was identified following competition experiments using unlabelled ALDHDIST. ALDHDIST was shown to compete efficiently with specific complexes I, J and M and to a lesser extent K (Figure 2.17). This implicates the involvement of GATA factors in these complexes, since the only sequence that ALDHDIST and ALDHPROX have in common is a canonical GATA site (WGATAR). The specific involvement of nucleotides -32/-28bp (ATAAAA) in the formation of these complexes was subsequently confirmed by the inability of an unlabelled mutant competitor oligonucleotide (ALDHPROXMut) to effectively remove complexes from the wild type ATAAA sequence (Figure 2.18). In this mutant oligonucleotide, the sequence ATAAA was replaced with CTCAGC, whilst flanking sequences were left intact. Although, as expected, competition experiments involving NS2 oligonucleotide did not affect formation of complexes I-M, addition of anti-HOX11 antibody resulted in the appearance of the otherwise absent complex I for PER-117HOX11 (Figure 2.19). These results imply that HOX11 may be directly involved in the disruption of complexes at the GATA box, and that inhibition of HOX11 function by antibody binding may enable the specific reformation of complex I. A summary of protein binding for PER-117 vs PER-117 HOX11 cell lines at the CCAAT and GATA boxes is outlined in Figure 2.20.

2.3.3.2 EMSA Assays Using HEL Nuclear Extracts

Since the behaviour of the ALDH1A1 promoter in response to HOX11 was radically different to that observed in PER-117 T-cells, we next sought to examine protein-DNA complexes in the HEL cell line. EMSA of the ALDHCAT oligonucleotide using HEL and HELHOX11 nuclear extracts yielded a slightly different DNA binding profile
Figure 2.18. Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-28bp). Radiolabelled ALDHPROX oligonucleotide probes (0.12pmol) were incubated with 10ug of nuclear extract to reveal complexes I-M (Lanes 2, 3). Specific complexes were identified following competition experiments using a 35-fold molar excess of unlabelled specific competitor (ALDHPROX; Lanes 4, 5). The specific involvement of the GATA box in complex formation was assessed by competing binding to the GATA box with a 35-fold excess of unlabelled ALDHPROXMut oligonucleotide (Lanes 6, 7).
Non-specific

<table>
<thead>
<tr>
<th>Competitor/Ab</th>
<th>ALDH PROX</th>
<th>NS2</th>
<th>ALDH DIST</th>
<th>α-HOX11 Ab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extract</td>
<td>1ul</td>
<td>3ul</td>
<td>1ul</td>
<td>3ul</td>
</tr>
</tbody>
</table>

Figure 2.19. Electromobility Shift Analysis of PER-117 and PER-117HOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-28bp). Radiolabelled ALDHPROX oligonucleotide probes (0.12pmol) were incubated with 10ug of nuclear extract to reveal complexes I-M (Lanes 2, 3). Specific complexes were identified following competition experiments using a 35-fold molar excess of unlabelled specific competitor (ALDHPROX; Lanes 4, 5) and non-specific competitor (NS2; Lanes 6, 7). Binding activity of proteins to ALDHDIST was assessed by competing binding to the GATA box with a 35-fold excess of unlabelled ALDHDIST oligonucleotide (Lanes 8, 9). Supershift analysis was performed using monoclonal anti-HOX11 antibody (1-3ul)(Lanes 9-12).
FIGURE 2.20
compared to PER-117 (Figure 2.21). HEL nuclear extracts did not exhibit complexes A and C, although complex D remained in common between both cell lines. In addition, HEL nuclear extracts resulted in the gain of two additional complexes, complex O and a high molecular weight complex N (Figure 2.21, 2.22). Notably, no differences were observed between the HEL and HELHOX11 cell lines. In a similar manner to PER-117, supershift experiments with anti-HOX11 antibody did not indicate involvement of HOX11 binding at the CCAAT box (Figure 2.22). Competition experiments involving the mutant ALDHCATMut oligonucleotide confirmed the specific involvement of the CCAAT element in the formation of complexes D, N and O (Figure 2.21).

Gel retardation of the ALDHPROX oligonucleotide with HEL and HELHOX11 extracts also yielded a DNA binding profile similar to PER-117, including specific complexes I, K and M, although complex J was absent in HEL and HELHOX11, suggesting that proteins binding this region may be involved in T-cell specific regulation (Figure 2.23). Quantitative differences between the degree of complex M binding were also observed between the erythroid and T-cell lines, with stronger binding activity observed for HEL regardless of HOX11 expression status, compared to PER-117/HOX11. The specificity of complex I, K and M binding to the ATAAA residues was subsequently confirmed following competition experiments with unlabelled ALDHPROXMut mutant competitor, which was unable to contest binding of proteins to the wild type sequence, even at a 35-fold molar excess (Figure 2.23). Supershift experiments with HOX11 antibody did not affect complex formation, suggesting that HOX11 is not involved in DNA-binding complexes at the GATA box, in accordance with the observation of no difference in I, K or M complexes between HEL and HELHOX11 (Figure 2.24). Intriguingly, however, unlabelled ALDHDIST oligonucleotide was capable of competing complexes I, K and M as effectively as unlabelled specific competitor. This again highlights the possibility of interactions of GATA factors at ALDHDIST and at the putative GATA box (Figure 2.24).

In summary, the results presented in this Chapter indicate that the expression of ALDH1A1 in PER-117 and HEL cells is regulated, at least in part, by the binding of trans-acting factors at the CCAAT box (-74/-70bp). Although no DNA footprint was discerned at this site, EMSA analysis of protein-DNA complexes identified binding of
Figure 2.21. Electromobility Shift Analysis of HEL and HELHOX11 Nuclear Protein Binding to the *ALDH1A1* Promoter CCAAT Box (-74/-70bp). Radiolabelled *ALDH*CAT oligonucleotide probes (0.12pmol) were incubated with 10ug of nuclear extract to reveal complexes B-O (Lanes 1, 2). Specific complexes were identified following competition experiments using a 35-fold molar excess of unlabelled specific competitor (*ALDH*CAT; Lanes 3, 4). The specific involvement of the CCAAT box in complex formation was assessed by competing binding to the CCAAT box with a 35-fold excess of unlabelled *ALDH*CATMut oligonucleotide (Lanes 5, 6).
<table>
<thead>
<tr>
<th>Competitor/Ab</th>
<th>ALDHCAT</th>
<th>ALDH DIST</th>
<th>NS2</th>
<th>α–HOX11 Ab</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
</tr>
</tbody>
</table>

Figure 2.22. Electromobility Shift Analysis of HEL and HELHOX11 Nuclear Protein Binding to the ALDH1A1 Promoter CCAAT Box (-74/-70bp). Radiolabelled ALDH1PCAT oligonucleotide probes (0.12pmol) were incubated with 10μg of nuclear extract to reveal complexes B-O (Lanes 2, 3). Specific complexes were identified following competition experiments using a 35-fold molar excess of unlabelled specific competitor (ALDH1PCAT; Lanes 4, 5) or non-specific competitor (NS2; Lanes 8, 9). Binding activity of proteins to ALDHDIST DNA sequences was assessed by competing binding to the CCAAT box with a 35-fold excess of unlabelled ALDHDIST oligonucleotide (Lanes 6, 7). Supershift analysis was performed using monoclonal anti-HOX11 antibody (1-3ul)(Lanes 10-13).
Figure 2.23. Electromobility Shift Analysis of HEL and HELHOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-28bp). Radiolabelled ALDHPROX oligonucleotide probes (0.12pmol) were incubated with 10ug of nuclear extract to reveal complexes I-M (Lanes 1, 2). Specific complexes were identified following competition experiments using a 35-fold molar excess of unlabelled specific competitor (ALDHPROX; Lanes 3, 4). The specific involvement of the GATA box in complex formation was assessed by competing binding to the GATA box with a 35-fold excess of unlabelled ALDHPROXMut oligonucleotide (Lanes 5, 6).
Figure 2.24. Electromobility Shift Analysis of HEL and HELHOX11 Nuclear Protein Binding to the ALDH1A1 Promoter GATA Box (-34/-28bp). Radiolabelled ALDHPROX oligonucleotide probes (0.12pmol) were incubated with 10μg of nuclear extract to reveal complexes I-M (Lanes 2, 3). Specific complexes were identified following competition experiments using a 35-fold molar excess of unlabelled of specific competitor (ALDHPROX; Lanes 4, 5) and non-specific competitor (NS2; Lanes 6, 7). Binding activity of proteins to ALDHDIST sequences was assessed by competing binding to the GATA box with a 35-fold excess of unlabelled ALDHPHOX11 oligonucleotide (Lanes 8, 9). Supershift analysis was performed using monoclonal anti-HOX11 antibody (1-3ul)(Lanes 10-13).
common factors (complex D). Cell-type specific complexes (PER-117; A, C)(HEL; N, O) were also capable of binding the CCAAT box. In keeping with the luciferase results which demonstrated that HOX11 affects \textit{ALDH1A1} expression via interactions at the CCAAT box in the PER-117 but not HEL cell line, expression of HOX11 was associated with the loss of a PER-117 specific complex (complex A) in EMSA analyses. In the case of HEL, where only a small specific effect on \textit{ALDH1A1} promoter activity was observed in luciferase assays which did not depend on the presence of the CCAAT box or the GATA box, there were no observable changes in the protein-DNA binding complexes due to HOX11 at either of these sites. Analysis of the GATA box also revealed binding of common factors between PER-117 and HEL (complexes I, K and M), however, binding of additional factors was also observed in the case of PER-117 (complex J). Significantly, the presence of HOX11 was associated with quantitative changes in the formation of complexes I and J in PER-117, whereas protein binding at this site was again unaffected by HOX11 in the HEL cell line. The ability of \textit{ALDHDIST} to effectively compete with \textit{ALDHPROX} for protein binding, suggests that these two probes share a common binding site. Given that the only sequence that these probes have in common is WGATAA, the nuclear protein(s) capable of binding to both may belong to the GATA family.

2.4 DISCUSSION

Following the identification of \textit{ALDH1A1} as a gene transcriptionally activated by HOX11 in the murine fibroblast NIH 3T3 cell line and physiologically repressed by Hox11 in the developing spleen, the purpose of this study was firstly to identify important \textit{cis}-acting DNA elements within the promoter of \textit{ALDH1A1} and to examine associated \textit{trans}-acting regulatory proteins and secondly, to investigate the effect of HOX11 on the transcriptional regulation of \textit{ALDH1A1} in both an \textit{ALDH1A1}-expressing (HEL) and non-expressing (PER-117) cell line.

A transient reporter assay involving the cotransfection of both promoter and HOX11 expression constructs into both the HEL and PER-117 cell lines, was chosen as the system to initially characterise \textit{cis}-acting DNA sequences regulating \textit{ALDH1A1} gene expression for several reasons. Firstly, the transient system is a simple and rapid method that allows maximal flexibility with regards to the cell line and promoter constructs.
being used to study the regulation of a given gene. Secondly, the introduction of HOX11 into the cells of interest by transient transfection, as opposed to analysing the effects of HOX11 on ALDH1A1 promoter activity in stably transfected HOX11-expressing versus non-expressing clones, reduces the likelihood of analysing promoter effects that may be caused by clonal differences as opposed to differential HOX11 expression. Analysis of nested deletions of the ALDH1A1 promoter in both HEL and PER-117 cell lines revealed that, in accordance with previous studies in the ALDH1A1-expressing Hep3B cell line and in Hepa-1 mouse hepatoma cells (Yanagawa et al., 1995, Elizondo et al., 2000), the main activity of the ALDH1A1 promoter pertains to a functional CCAAT box located at -74/-70bp, since deletion of this element led to a reduction in luciferase activity to baseline levels. The CCAAT box is an element essential for the regulation of a wide variety of eukaryotic promoters and may bind several nuclear factors including CP1 (NF-Y), CP2, CCAAT box transcription factor (CTF/NF1), which is thought to play a role in the constitutive expression of certain genes and the CCAAT/enhancer binding protein (C/EBP), which is thought to activate genes specifically expressed in the liver.

Although DNase1 footprinting of the proximal ALDH1A1 promoter encompassing this element did not identify transcription factor binding at this site, EMSA of the ALDH1CAT oligonucleotide using PER-117 and HEL nuclear extracts identified three specific protein-DNA binding complexes for PER-117 (A, C and D) and three specific complexes for HEL (D, N and O). Since this probe also harbours an RARα binding site, which has previously been shown to bind RARα/RXR heterodimers required in part for ALDH1A1 expression (Elizondo et al., 2000)(Figure 2.2), competition experiments were performed using a mutant CCAAT oligonucleotide probe. This confirmed that these complexes were indeed specific to nucleotides comprising the CCAAT box (-74/-70bp). Aside from complex D which was common between the two cell lines, the differential binding activities demonstrated by PER-117 (complexes A and C) and HEL (complexes N and O), suggested that ALDH1A1 is regulated at the CCAAT box by additional cell-type specific transcriptional complexes. This may account for the differential expression of ALDH1A1 in these cell lines. Previous analyses of protein-DNA interactions at the CCAAT box demonstrated binding by NF-Y in the Hep3B liver cell line (Yanagawa et al., 1995) and by C/EBPβ binding in liver
Hepa-1 cells (Elizondo et al., 2000). Since there is precedence for combinatorial protein-protein interactions between C/EBP and GATA family members (Du et al., 2002, Tremblay et al., 2002), it is possible that regulation at the CCAAT box involves a common factor (e.g. C/EBP) in addition to T-lineage (in the case of PER-117) and erythroid (in the case of HEL)-specific GATA family members (GATA-3 and GATA-1, respectively). Future experiments involving supershift analyses, using antibodies specific for C/EBP, GATA-1 or GATA-3 members, may provide an insight into the lineage-specific factors important for regulating promoter activity at this site in both the HEL and PER-117 cell lines.

The identification of specific protein-DNA binding complexes at the CCAAT box by EMSA but not DNase1 footprinting, highlights the advantage of utilising a combination of both assays when attempting to characterise trans-acting protein-DNA binding in vitro, since some DNA binding proteins may only be detected using one method (Carey & Smale, 2000). For example, in certain instances, physiological levels of a crucial transcription factor may be at too low a concentration to protect sufficient probe in a DNase1 footprinting assay to yield a discernable footprint, in which case EMSA is useful, since binding of protein factors to a small percentage of probe will result in visible changes in the migration of the probe. Protein binding may also remain undetected by DNase1 footprinting analysis if the binding site resides within a DNA region that is relatively devoid of DNase cleavage sites, whereas EMSAs may successfully identify binding activities in these areas. In certain cases where protein-DNA complexes are unable to withstand gel electrophoresis, however, DNase1 footprinting may be advantageous, since this method relies on interactions in solution without requiring maintenance of complex formation.

Previous reports demonstrated that the non-canonical TATA box (ATAAA) is not a primary regulatory element in the ALDH1A1-expressing Hep3B liver cell line (Yanagawa et al., 1995). Indeed our functional promoter analyses indicated that this element alone was unable to support ALDH1A1 promoter activity in either HEL or PER-117 cells. Nevertheless, DNase1 footprinting assays revealed protein binding at this site (ALDHPROX) for both the HEL and PER-117 cell lines. Moreover, EMSAs using the ALDHPROX oligonucleotide probe encompassing this site, identified binding activities specific for ATAAA residues in both HEL and PER-117. In the case of PER-
117, four specific DNA binding complexes were identified (complexes I, J, K and M), compared with three specific complexes for HEL (complexes I, K and M). Complex J may therefore represent a T-cell specific binding activity in a transcriptional complex involving other common factors. Intriguingly, the ALDHPROX probe encompasses a putative GATA binding site and an oligonucleotide probe spanning the DNase1 protected site ALDHDIST, which also harbours a canonical GATA-1/3 binding site, effectively competed binding of specific complexes I, J and M to the ALDHPROX oligonucleotide. This indicated the presence of a common binding site in both probes. Taken together, these results suggest that GATA factors may be involved in the specific regulation of ALDH1A1 in HEL and PER-117 cell lines at both the ALDHDIST and ALDHPROX sites. Given the roles of GATA-1 in the differentiation of the erythroid lineage (Pevny et al., 1991) and of GATA-3 in the regulation of T-cell specific genes (Ko et al., 1991), it is possible that the differential binding observed between HEL (erythroid lineage) and PER-117 (T-cell lineage) cell lines may reflect binding of specific members of the GATA family to these sites. Supershift analyses are now required to confirm the identity of proteins binding to these elements. It is possible that the lack of an effect on ALDH1A1 promoter activity upon mutation of the ATAAA box (Yanagawa et al., 1995), or upon removal of the ALDHDIST site in transient transfection assays (e.g. between constructs -303 and -201/42), may be due to a degree of functional redundancy between these sites. Indeed, the initial failure to identify important control elements in the Ig µ enhancer was partly due to the functional redundancy of several regulatory elements within this enhancer (Dang et al., 1998). The effect of simultaneous mutation of both ALDHDIST and ALDHPROX sites may therefore provide an insight into the importance of these elements in promoter function.

Given that the GATA site is located at the canonical TATA location in the ALDH1A1 promoter (at -30bp), it is likely that it represents a specialised TATA box, capable of binding both GATA and members of the basal transcriptional machinery. The existence of such a dual function element overrepresented among erythroid-specific promoters has been previously documented, where it is referred to as a GATA box (Fong & Emerson, 1992, Barton et al., 1993, Aird et al., 1994, Surinya et al., 1997, Tsuchiya et al., 1997). One model derived from studies of the rat platelet factor 4 (PF4) gene, suggests that GATA1 or 2 binds at the GATA box to prevent transcription by sterically
inhibiting the formation of preinitiation complexes, and suggests that the relative levels of GATA and basal factors may account for tissue specificity of gene expression (Aird et al., 1994). In a separate study, GATA-1 was shown to regulate the chick β-globin gene, through GATA sites present in the distal 3’enhancer and at GATA box motif (-30bp) in the promoter (Fong & Emerson 1992). The binding of both elements mediated interactions between the enhancer and promoter elements (Fong & Emerson 1992). It is postulated that TFIID and associated adaptor proteins subsequently displace GATA from the -30bp box and interact with proteins at the erythroid-specific enhancer to initiate transcription (Fong & Emerson 1992). In the case of ALDH1A1, which is not normally highly expressed in haematopoietic tissues, it is possible that GATA-1, -2 or -3 factors may bind the GATA box of the ALDH1A1 promoter in their respective tissues to inhibit expression of ALDH1A1, whereas in other tissues expressing GATA-4, -5 or -6 (e.g. heart, gut, urogenital, brain), this GATA site is unoccupied and TFIID is able to bind and initiate transcription, thereby accounting for the expression of ALDH1A1 in these tissues. This theory does not explain the high level of endogenous expression of ALDH1A1 in HEL cells, however this may not be reflective of a normal physiological situation given that HEL is a transformed cell line.

Having identified regions of the ALDH1A1 promoter important for the activation in both HEL and PER-117 cell lines, it was then pertinent to assess the effect of HOX11 on ALDH1A1 promoter activity and to identify the specific element(s) responsible for mediating a HOX11-induced response. Initial experiments were performed in HEL cells and a comparison of endogenous ALDH1A1 expression in HEL cells stably transfected with either pEFBOS or pEFBOSHOX11 revealed a significant downregulation of ALDH1A1 in response to HOX11. This suggested that HOX11 represses ALDH1A1 promoter activity in this cellular background. In contrast, transient transfection of HEL cells with HOX11 and ALDH1A1 promoter constructs ranging from +1 to -2159/+42bp, resulted in an apparent activation of the ALDH1A1 promoter in all constructs tested. Since HOX11 was capable of activating transcription from the pGL3Basic luciferase vector in the HEL cell background, even in the absence of ALDH1A1 promoter sequences, statistical analyses were performed. This was done in order to assess whether the effect of HOX11 on the ALDH1A1 promoter was above the background effect of HOX11 on the pGL3Basic vector alone. Although the mechanism by which
HOX11 is able to spuriously activate luciferase expression remains unclear, independent studies have reported high background luciferase expression in the pGL2 and pGL3Basic series of reporter vectors resulting from the interaction of general transcription factors (de Wet et al., 1987), or tissue specific factors such as the osteoblast specific transcription factor Cbfal and the estrogen receptor α (ERα), with cryptic promoter elements in the vector backbone (Grimm & Nordeen, 1999, Thirunavukkarasu et al., 2000).

Statistical analyses revealed that the effect of HOX11 on the ALDH1A1 promoter was higher in constructs ranging from +1 to -91/+42bp, than the effect of HOX11 on the pGL3Basic vector alone. This suggested that HOX11 was capable of transactivating the ALDH1A1 promoter either directly or indirectly via an element located between +1/+42bp (Figure 2.25A). Although no consensus binding sites for HOX11 or known cofactors were identified in this region, future studies involving EMSAs may shed insight into the mechanism of HOX11-mediated transactivation within this region. The ability of HOX11 to activate transcription above background was sustained for ALDH1A1 promoter constructs between +1 and -91 inclusive. However, following the addition of 55bp to create the -146 construct, the activation potential of HOX11 significantly declined, with a 1.8-fold decrease in activity observed between constructs -91 and -146 (P=0.001). This suggested that a factor binding between -91/-146bp is capable of exerting a negative influence on the weak HOX11 transactivation potential within the +1 construct (Figure 2.25B). Analysis of this region in more detail identified several elements including an Ets-1 site which is conserved between human, mouse, marmoset and rat species, however, the specific mechanism by which this occurs will require further analysis of trans-acting factors within this region and due to time constraints was not investigated in this study. Notably, the effect of HOX11 on ALDH1A1 promoter regulation in the HEL cell line did not appear to involve the CCAAT box (-74/-70bp) or GATA box (-34/-28bp), and in keeping with these observations, we did not observe any differences in DNA binding activities between HEL and HELHOX11 nuclear extracts at either the CCAAT box or the GATA box by EMSA.
Figure 2.25. HOX11 Activates Transcription in the HEL Cell Line. (A) HOX11 stimulates transcription from the pGL3Basic vector alone, as well as from the ALDH1A1 promoter (+1/+42bp) (B) Factor(s) binding the region -146 to -91bp inhibit the transactivation potential of HOX11 (C) HOX11 stimulates transcription following deletion of the DNA binding helix (∆H3), suggesting that HOX11 may function as a transcriptional activator by via protein-protein interactions with members of the basal transcriptional machinery.

The ability of a HOX11 DNA-binding mutant (HOX11∆H3) to transactivate transcription from the ALDH1A1 promoter in the HEL cell line was subsequently investigated using the -91 reporter construct, since this construct demonstrated the greatest transactivation by HOX11 (1.84-fold over the effect of HOX11 on pGL3Basic; P < 0.001). These experiments revealed that the DNA-binding helix was not required for the activation function of HOX11, suggesting that HOX11 may stimulate transcription in a DNA-binding independent manner, possibly via protein-protein interactions. Intriguingly, HOX11 has been implicated in an in vivo association with a member of the basal transcriptional complex, TFIIB, following immunoprecipitation of HOX11 from the T-ALL cell line ALL-SIL and analysis of HOX11-bound factors by mass spectrophotometry (MALDI-TOF)(Heidari & Greene, unpublished observations). This suggests that HOX11 may transactivate transcription via interactions with components of the basal transcriptional machinery (Figure 2.25C). The interaction of
general transcription factors (GTFs) with sequence-specific DNA binding proteins is thought to constitute a mechanism for establishment of high-level gene expression, and interactions may involve various regulatory domains of the transcription factor as well as the DNA binding domain (Sauer et al., 1995). In further support of the notion that HOX11 may modulate gene transcription by interacting with components of the basal transcriptional machinery, in vitro transcription assays performed by Owens et al., (2003) demonstrated that HOX11 is capable of repressing transcription at the adenovirus major late promoter, by interacting with one or more components of the RNA polymerase holoenzyme.

Following these studies, the effect of HOX11 on ALDH1A1 expression was studied in the T-ALL cell line, PER-117. Although expression of HOX11 induces upregulation of endogenous levels of ALDH1A1 in stably transfected PER-117 cells (Greene, unpublished observations), HOX11 repressed ALDH1A1 promoter activity by approximately 2.3-fold (P < 0.001), via an element contained within the -91 construct in transient reporter assays. Thus, the transient transfection studies with the ALDH1A1 promoter constructs consistently yielded an effect by HOX11 opposite to that of the endogenous ALDH1A1 gene. Further refinement of this responsive element by analysis of progressively smaller deletion constructs of the ALDH1A1 promoter (-50 to +1), in conjunction with deletion analysis (-91ΔCAT), identified the CCAAT box (-74/-70bp) as the specific cis-regulatory element through which HOX11 mediates repression. Analysis of protein binding at this site by EMSA, revealed the loss of a specific protein-DNA binding complex (complex A) in HOX11-expressing cells. This suggested that HOX11 repressed transcription by interfering with the binding of a specific transcription factor at the CCAAT box. In contrast to the activation function of HOX11, which occurred independently of DNA binding, deletion of the DNA binding helix of HOX11 (ΔH3) abolished repressor function, as evidenced for both the -91 and the pSV-β-Gal internal control reporter constructs. Moreover, this relief in transcriptional repression was coupled with a concomitant increase (~5-fold) in promoter activity. Taken together, these results suggest that HOX11 is capable of acting as both a transcriptional repressor and activator via two potentially independent mechanisms. In the case of transcriptional repression HOX11 may act either by:
1) Passive Repression – HOX11 may interfere with positive trans-acting factors at the CCAAT box (for example those in DNA binding complex A; Figure 2.26B), for example by displacement (occlusion), as is the case for the CCAAT-Displacement Protein/Cut homeobox (CDP/Cux) transcription factor, which is reported to repress a large number of genes by competing for binding site occupancy (Barberis et al., 1987). The hypothesis whereby HOX11 disrupts binding of a factor(s) to the CCAAT box was supported by EMSA of an oligonucleotide encompassing the CCAAT box using PER-117 and PER-117HOX11 nuclear extracts, which revealed the loss of a specific complex (A) in the presence of HOX11. One potential candidate for an interacting factor at this site is CTF1, a ubiquitously expressed CCAAT box binding transcription factor which is capable of recruiting TBP-DNA complexes and has previously been shown to interact with HOX11 (Zhang et al., 1999).

2) Active Repression – This may involve inhibition of the formation of a functional preinitiation complex, inhibition of adjacent bound transcriptional activators (quenching), or recruitment of repressor complexes including histone deacetylases (HDACs), which modify local chromatin structure causing a reduction in gene expression (Figure 2.26C). Notably, the transcriptional activation domain at the NH₂ terminus of HOX11, harbours a binding domain for members of the Groucho/transducin-like Enhancer of split (Gro/TLE) co-repressor family, which are capable of repressing transcription via the recruitment of HDACs (Yao et al., 2001) and/or by interacting with members of the basal transcriptional machinery, specifically, TFIIE (Gromoller & Lehming, 2000). It is of interest to note that the expression of HOX11 in PER-117 was not only associated with the loss of a complex at the CCAAT box (complex A), but was also associated with a loss of PER-117-specific complexes I and J at the GATA box, and the enhancement of complexes K and M. Moreover, inclusion of anti-HOX11 antibody into EMSA binding reactions reduced the inhibition of one of these complexes (I), implying that HOX11 is directly involved in the disruption of complex I at the GATA box. Thus, HOX11 is associated with perturbation of specific protein complexes at the GATA box. Given that the expression of HOX11 in our transient assays resulted in transcriptional repression, this data may fit a model in which HOX11 acts by inhibiting the formation of a functional preinitiation complex.
**Figure 2.26. HOX11 Acts as a Bi-functional Transcription Factor in PER-117 Cells.** (A) *Trans*-acting factors stimulate transcription of *ALDH1A1* via binding to the CCAAT box. HOX11 is capable of repressing transcription at the CCAAT box (-74/70bp) in a DNA-binding dependent manner associated with a loss in complex A binding, either by passive mechanisms (B) by precluding binding of positive-acting factors at the CCAAT box or active mechanisms (C) possibly by recruiting repressor complexes including histone deacetylases (HDACs). This change in activity is also associated with changes in complex formation at the GATA box. Upon loss of DNA binding (ΔH3), HOX11 stimulates transcription above that effected by complexes A, C and D, possibly via protein-protein interactions with members of the basal transcriptional machinery (D).
The models of transcriptional activation and repression described in Figures 2.25 and 2.26 portray the effects of HOX11 acting directly in concert with the basal transcriptional machinery or via direct interactions at the CCAAT box. It is also possible that HOX11 functions to alter ALDH1A1 transcription indirectly by affecting the expression of genes higher up in a hierarchical transcriptional cascade. Given that HOX11 is capable of affecting the expression of ALDH1A1 in a plethora of cellular backgrounds of human and murine origin including murine spleen, NIH 3T3, HEL, PER-117 and J2E cell lines, however, it seems more likely that ALDH1A1 represents a direct target of HOX11.

Despite the fact that the ALDH1A1 promoter is repressed by HOX11 in our in vitro reporter assays in PER-117, which contrasts with the induction of ALDH1A1 by HOX11 in stably transduced PER-117 cells, these studies nevertheless provide insight into cis-regulatory elements and potential mechanisms by which HOX11 is able to affect gene transcription. The differences in these in vitro reporter assays as opposed to in vivo expression levels may reflect the limitations of transient transfection assays and may relate to the fact that the regulatory sequences under investigation are maintained in an artificial configuration. That is, they are constrained within a plasmid construct that does not conform to the appropriate chromatin configuration, which may affect normal mechanisms of gene regulation. In addition, artificially high copy numbers of the regulatory sequence are present following transfection into the cell, resulting in disequilibrium between the ratio of transcription factors to regulatory elements, which may also account for this discrepancy. Alternatively, more distal promoter elements not included in these promoter constructs may be required for appropriate activity, in terms of repression or activation.

The results of this Chapter indicated that HOX11 represses transcription in a DNA binding-dependent manner, since deletion of helix 3 of the homeodomain (HOX11ΔH3) abolished the ability of HOX11 to repress transcription from the ALDH1A1 promoter. However, Owens et al., (2003) demonstrated that homeodomain point mutations shown to alter or abrogate DNA binding specificity (by mutating residues T₄₇ and K₅₅ of the homeodomain) had no significant effect on ability of HOX11 to repress transcription from the SV40 promoter. They concluded that HOX11-mediated repression occurs in
the absence of DNA binding, possibly via protein-protein interactions. A structurally intact homeodomain was nevertheless still required for repression, since deletion of helix 3 abolished repressor function (Owens et al., 2002). Support for this model derives from studies of the murine homeodomain protein Msx-1 (Catron et al., 1995, Zhang et al., 1996). Msx-1 represses transcription in a DNA binding-independent manner, via direct interactions with the TATA binding protein (TBP)(Zhang et al., 1996). This interaction is mediated through residues in the N-terminal arm of the homeodomain that are also required for transcriptional repression, suggesting that the processes of TBP binding and transcriptional repression are linked (Catron et al., 1995, Zhang et al., 1996). Thus, although our results suggest that repression of ALDH1A1 by HOX11 requires DNA binding, it is possible that deletion of helix 3 of the HOX11 homeodomain perturbs the structure of the homeodomain required for protein-protein interactions with members of the basal transcriptional machinery. Alternatively, although the ability of HOX11 to repress transcription was not affected by T47 and K55 DNA binding-mutant proteins in the SV40 viral promoter model utilised by Owens et al., (2003), these mutations may abrogate HOX11 repressor function in the context of physiological target genes, such as ALDH1A1. Indeed, these mutants affected the ability of HOX11 to induce Aldh1a1 expression in NIH 3T3 cells (Owens et al., 2003). Since our studies did not incorporate the T47 (specificity) and K55 (DNA contact, repressor region) HOX11-DNA binding mutants, however, we were unable to assess the effect of these mutations on the ability of HOX11 to repress the ALDH1A1 promoter and incorporation of such mutants in future studies may clarify the mechanism by which HOX11 represses the ALDH1A1 promoter in PER-117 T-cells.

In summary, the work described in this Chapter has shown that HOX11 is capable of modulating the activity of the ALDH1A1 promoter in a transient transfection assay involving the cotransfection of ALDH1A1 promoter constructs into HEL and PER-117 cell lines. This strengthens the status of ALDH1A1 as a bona fide target gene of HOX11. In the case of HEL, expression of HOX11 was associated with a weak transactivation of the ALDH1A1 promoter within +1/+42bp, however this activation declined with the addition of promoter sequence between -91 and -146bp, suggesting that factors binding in this region may exert a negative regulatory influence on HOX11. A significant activation function for HOX11 was also observed on the empty pGL3Basic vector
construct and was not affected following deletion of the DNA binding helix (HOX11ΔH3). This suggested that activation may occur in a DNA binding-independent manner, possibly via protein-protein interactions. Conversely, HOX11 repressed ALDH1A1 promoter activity in the PER-117 cell line, possibly by interfering with positive-acting factors at the CCAAT box (-74/-70bp) and at the GATA box (-34/28bp). In contrast to the activation function of HOX11, which was not affected by removal of the DNA binding helix, the HOX11ΔH3 mutant was unable to repress transcription, and instead switched to a potent transcriptional activator. These results imply that HOX11 mediates repression by participating in a repressor complex that requires helix 3 either for (1) DNA binding to a specific promoter element or (2) for maintaining stable interactions with components of the repressor complex. Thus, the inability of the ΔH3 mutant to repress transcription may result from an abrogation of DNA binding potential or alternatively, the ΔH3 mutant may act in a dominant negative manner to sequester transcriptional repressors whilst simultaneously participating in a transcriptional activation complex. Although the activation and repression activities of HOX11 appear to differ with respect to the requirement for DNA binding, both mechanisms are likely to involve interactions with basal transcriptional machinery, with the differential activity of HOX11 directed by cell-specific factors in the HEL and PER-117 cell lines. Indeed, there is significant precedent for HOX proteins to display this dual functionality, for example the pancreatic homeodomain-containing factor PDX1, may act as either a transcriptional repressor or activator, depending on the specific PBX isoform with which it associates (Asahara et al., 1999).
Chapter 3

*FHL1* as a Target Gene of HOX11: Identification of the *FHL1* Gene Promoter and Assessment of its Transcriptional Regulation by HOX11

3.1 INTRODUCTION

FHL1 is a 32kDa protein that belongs to a specific subclass of LIM-only (LMO) proteins containing four complete and one amino-terminal half LIM (FHL) domains, including FHL2, FHL3 and FHL4. FHL family members have proposed roles in mediating the protein-protein interactions of transcription factors, signalling and cytoskeleton-associated proteins (Brown *et al*., 1999, Fimia *et al*., 2000, Kong *et al*., 2001). *FHL1*, previously known as *SLIM1* (Striated Muscle LIM protein 1), is expressed at high levels in adult skeletal muscle and at intermediate levels in the cardiac outflow tracts of the developing embryonic heart (Brown *et al*., 1999). It is thought to play a role in the differentiation of this structure, which gives rise to the aortic and pulmonary outflow tracts and is later limited to the aorta and atria in the adult heart (Brown *et al*., 1999). In addition, *FHL1* is also detected as a low abundance transcript in a panel of other tissues including brain, placenta, lung, liver, pancreas, small intestine, colon and the ovaries, suggesting a wider expression than previously reported (Brown *et al*., 1999, Lee *et al*., 1999, Greene *et al*., 1999, Ng *et al*., 2001).

To date, two additional alternatively spliced human isoforms of *FHL1* have been identified. *FHL1B/SLIMMER* shares the amino-terminal three and a half LIM domains with *FHL1*, but differs at the C-terminus with the addition of three putative bipartite nuclear localisation signals (NLS), a putative nuclear export sequence (NES) and an RBP-J binding region (Brown *et al*., 1999, Lee *et al*., 1999). Like *FHL1*, *FHL1B/SLIMMER* is highly expressed in skeletal muscle, however, analysis of the intracellular location of these splice variants in the murine skeletal muscle cell line, Sol8, revealed that whilst *FHL1* expression is predominantly cytosolic, *FHL1B/SLIMMER* shuttles between the nucleus and cytoplasm in myoblasts and myotubes respectively (Brown *et al*., 1999). The association of FHL1 with focal adhesions and actin cytoskeleton in COS-7 cells and the differentiation-dependent intracellular localisation of FHL1B/SLIMMER, suggests that these isoforms fulfil
distinct roles in regulating muscle cell function, where FHL1 is required for muscle cell cytoskeletal structure and FHL1B/SLIMMER acts as a transcriptional regulator in undifferentiated muscle cells (Brown et al., 1999). A second alternatively spliced isoform of FHL1, FHL1C (the human homolog of murine KyoT2) (Tang et al., 1998), comprises the N-terminal two and a half LIM domains of FHL1, followed by a C-terminal RBP-J binding region (Ng et al., 1999). FHL1C is expressed in human testis and skeletal muscle and at lower levels in the heart, and is predicted to interact with RBP-J DNA binding protein, to negatively regulate RBP-J mediated transcription, as the murine KyoT2 does (Tang et al., 1998, Ng et al., 1999).

The identification of functionally relevant target genes and elucidation of the transcriptional complexes controlling the expression of such genes in both physiological and tumorigenic contexts, is crucial to our understanding of the mechanisms by which nuclear oncogenes such as HOX11 can induce tumorigenesis. Fhl1 was identified as a gene that can be transcriptionally upregulated by HOX11 in NIH 3T3 cells (Greene et al., 1998), however the relevance of this relationship with respect to the normal regulatory roles of FHL1, or in leukaemogenesis, remains unclear. The requirement for Hox11 in spleen development during mouse embryogenesis, suggested that Fhl1 and Aldh1a1, which was also identified as a potential HOX11 target in cDNA RDA experiments, may play functional roles in spleen development. Intriguingly, the related family member, FHL2, has recently been found to modulate the activity of another transcription factor required for spleen organogenesis, WT1 (Herzer et al. 1999). Furthermore, Hox11 was shown to regulate the expression of Wt1 in the developing spleen (Koehler et al., 2000), highlighting the possibility that LIM-only transcription factors may be involved in transcriptional networks regulating spleen development. However, although an inverse relationship between Hox11 and Aldh1a1 mRNA levels was identified in the developing spleen, suggesting that Aldh1a1 may represent a physiological target gene of Hox11 involved in spleen organogenesis, the expression of Fhl1 did not appear to be affected by Hox11 in wild type or Hox11 null embryos (Greene et al., 1998). Thus, Fhl1 appears not to be a physiological target gene of HOX11, at least in the spleen. By contrast, it remains formally possible that FHL1 may represent an oncogenically relevant target gene, particularly in view of the observed expression of FHL1 mRNA in eight of twelve T-ALL cell lines as determined by
Northern blot analysis (Greene et al., 1998). This included two cell lines, ALL-SIL and PER-255 that overexpress HOX11 as a result of chromosomal translocations involving the \textit{HOX11} locus. Furthermore, a role for FHL members in tumorigenesis has been suggested following the discovery that the human \textit{ACT} gene is expressed in numerous human tumour cells lines derived from melanomas, squamous cell carcinomas and leukaemias and that \textit{FHL2} and \textit{FHL3} were also overexpressed in several of these cell lines (Morgan & Whawell, 2000). The possibility that \textit{FHL1} may be an oncogenically relevant target of HOX11 is given added weight by the fact that there is precedent for the dysregulation of other LIM-only transcription factors in T-ALL, namely LMO1 and LMO2 (Sanchez & Rabbitts, 1993).

The significance of the observations with respect to the role of \textit{FHL1} in tumorigenesis to date, are unclear. While correlations in the expression patterns of HOX11 and \textit{FHL1} in NIH 3T3 cells and human T-ALL cell lines provide some evidence to support the case that \textit{FHL1} may be a HOX11 target, functional experiments are required to confirm whether HOX11 has the ability to transcriptionally regulate the \textit{FHL1} promoter. To address the role of \textit{FHL1} as a transcriptional target of HOX11, therefore, the ability of HOX11 to regulate the putative promoter of this gene was assessed in the PER-117 T-cell and HEL erythroid cell lines using luciferase assays involving transient transfection. Since neither the transcriptional start site nor promoter of \textit{FHL1} had been previously characterised, a PCR strategy was first employed to map the startpoint of the gene and thus reveal its promoter region. Specifically the aims of this Chapter were 1) to identify the transcriptional start site of the human \textit{FHL1} gene by 5’ rapid amplification of cDNA ends (5’ RACE), 2) to characterise the proximal promoter of \textit{FHL1} by identifying predicted \textit{cis}-acting regulatory elements and other relevant core promoter features and 3) to perform luciferase reporter assays to investigate the activity of the \textit{FHL1} proximal promoter and its possible mechanism of regulation by HOX11 in order to formally determine its status as a HOX11 target gene.
3.2 METHODS

3.2.1 5’ RACE

3.2.1.1 Cell Lines

ALL-SIL is an immature, human leukaemic cell line of T-cell phenotype (T-ALL) that harbours a t(10;14) chromosomal translocation and thus overexpresses HOX11. The cell line was provided by the Division of CLCR of the TVW Telethon Institute for Child Health Research (Perth, Western Australia) and was shown to be Mycoplasma-free at the level of PCR (performed by Jette Ford at the Division of CLCR). Cells were subcultured twice weekly (1:5) in RPMI-1640 Multi-cel Medium (Trace) supplemented with 10% (v/v) FCS (Trace), 2mM L-Glutamine (Trace) and 50µg/ml Penicilllin/Streptomycin (Trace). Cells were typically cultured in 75cm² filter-top culture flasks (NUNC™) and were maintained at 37°C with 5% CO₂.

3.2.1.2 RNA Isolation

Total RNA was extracted from the ALL-SIL cell line using TRIZOL® reagent (Invitrogen), which is based on the guanidinium isothiocyanate/acid-phenol method originally described by Chomzynski and Sacchi (1987), according to the manufacturer’s instructions. Briefly, cells were cultured to mid-log phase (~5 x 10⁵ cells/ml) and approximately 2 x 10⁷ cells were collected by centrifugation in 50ml polypropylene tubes (200 x g, 5min, room temperature)(Eppendorf Centrifuge 5810). The cells were combined in one tube and subsequently lysed in 4ml of TRIZOL® reagent by repetitive pipetting and incubated at room temperature for 5min to allow the complete dissociation of nucleoprotein complexes. The cells were transferred to 1.5ml polypropylene microcentrifuge tubes (0.5 x 10⁶ cells/ml/tube) and 200µl of chloroform/tube was added. The tubes were then shaken vigorously by hand for 15s and incubated at room temperature for 3min. Phase separation was achieved by centrifugation (12,000 x g, 15min, 4°C)(Sigma 1-15) and the upper aqueous layer was removed and transferred to a fresh 1.5ml polypropylene tube. The RNA was precipitated by mixing with 500µl of isopropyl alcohol/tube and incubating at room temperature for 10min. The RNA was
subsequently recovered by centrifugation (12,000 x g, 5min, 4°C) and following removal of the supernatant, the pellet was washed with 75% (v/v) ethanol and pelleted by centrifugation (7,500 x g, 5min, 4°C). Pellets were allowed to air dry for 10min and dissolved in sterile, RNase-free dH₂O (30µl/pellet). The dissolved RNA pellets were combined and stored in 30µl aliquots at -80°C.

3.2.1.3 Quantitative and Qualitative RNA Analysis

RNA integrity and concentration was assessed using the Agilent 2100 Bioanalyzer instrument and Agilent 2100 biosizing software (Agilent, Palo Alto, CA). Samples were processed at the Lotteries State MicroArray Facility (Sir Charles Gairdner Hospital, Perth, WA) using the RNA 6000 LabChip® Kit (Agilent, Palo Alto, CA). RNA integrity was also examined by agarose gel electrophoresis as described in Section 2.2.1.2. RNA samples (1µl) were combined with 5x TBE-Urea Denaturing Sample Buffer [1x](Bio-Rad Laboratories) immediately prior to loading and electrophoresis was performed in 1x TAE buffer containing 0.1µg/ml ethidum bromide using a Bio-Rad Mini-sub horizontal gel apparatus (Bio-Rad Laboratories) at constant voltage (70-80V) for 1h. Following electrophoresis, RNA was visualised using an ultraviolet UV transilluminator (Gel Doc 1000) and Molecular Analyst software (Bio-Rad Laboratories). Qualitative analysis of separated RNA samples was assessed by comparison of the intensity of the 28S rRNA band (size ~ 5kB) with the 18S rRNA band (size ~2kB), where a ratio of 2.0 is indicative of high quality RNA.

The quantitation of RNA samples in aqueous solution was achieved by UV spectrophotometry. RNA samples in solution were diluted 1/80 to a total volume of 160µl in dH₂O. The spectrophotometer (Shimadzu UV Mini 1240 Spectrophotometer) was blanked with a 160µl aliquot of the sample diluent. Ultraviolet light absorbance of the diluted RNA sample was then determined at 260nm and 280nm wavelengths, and RNA concentration and purity was determined using the calculations below;
where (1 O.D unit = 40μg/ml of ds RNA)

RNA Concentration [mg/ml]: O.D_{260nm} x dilution factor x 40

where a ratio of 2.0 indicated a pure RNA solution

RNA Purity: O.D_{260nm}/O.D_{280nm}

3.2.1.4 5’ RACE – System Version 2.0

First strand cDNA was synthesised from 2.5μg total ALL-SIL RNA using one of three methods.

3.2.1.4.1 SUPERSCRIPT™ II RT

Briefly, 2.5μg of total ALL-SIL RNA was combined with 2.5pmol of SLIM1(1) primer (Table 3.1), in a thin-walled, 0.5ml RNase-free polypropylene tube in a final volume of 15.5μl. RNA was denatured by incubation at 70°C for 10min and the reaction was placed on ice. Reverse transcription was performed in a final volume of 25μl comprising 10x PCR Buffer [1x; 20mM Tris-HCl (pH 8.4), 50mM KCl](Invitrogen), 2.5mM MgCl₂ (Invitrogen), 0.4mM dNTPs (Invitrogen), 0.01M DTT (Invitrogen) and 200U of SUPERSCRIPT™ II Reverse Transcriptase (Invitrogen), at 42°C for 50min. The enzyme was then heat inactivated at 70°C for 15min, and the RNA template was subsequently removed by incubation with 1μl of the supplied RNase Mix (a combination of RNase H and RNaseT1)(Invitrogen) at 37°C for 30min and placed on ice.

Unincorporated dNTPs, SLIM1(1) primer and proteins were then separated from the cDNA using a GLASSMAX DNA Isolation Spin Cartridge (Invitrogen) supplied with the kit. The first strand reaction was mixed with 120μl of room temperature equilibrated Binding Solution (6M NaI)(Invitrogen) and the solution was transferred to a
TABLE 3.1
GLASSMAX spin cartridge and centrifuged (13,000 x g, 20s, room temperature)(Biofuge Pico, Heraeus Instruments). The cartridge was then washed by applying cold (4°C) 1x Wash Buffer (Invitrogen) to the column and centrifugation (13,000 x g, 20s, room temperature)(x4 washes). These wash steps were subsequently repeated with cold (4°C) 70% (v/v) ethanol (x2 washes). Following the final ethanol wash, the column was re- centrifuged (13,000 x g, 1min, room temperature) to remove traces of residual ethanol. The purified cDNA was then eluted into a fresh 1.5ml polypropylene microcentrifuge tube, by applying 50µl of sterile, preheated (65°C) dH2O to the spin cartridge and centrifuging (13,000 x g, 20s, room temperature).

A homopolymeric oligo-dC tail was then added to the 3’ end of the cDNA using terminal deoxynucleotidyl transferase (TdT)(Invitrogen). GLASSMAX purified cDNA was initially denatured in a 0.5ml polypropylene tube in a final volume of 24µl, comprising 10µl of cDNA, 200µM dCTP (Invitrogen), 5x Tailing Buffer [1x; 10mM Tris-HCl (pH 8.4), 25mM KCl, 1.5mM MgCl2](Invitrogen), by incubating at 94°C for 3min and subsequently chilled on ice. 1µl of TdT (Invitrogen) was then added to the reaction, and tailing was allowed to proceed for 10min at 37°C. The enzyme was then heat inactivated at 65°C for 10min and the tailed cDNA was subsequently amplified in a primary PCR using a nested gene specific primer SLIM1(2) and the supplied deoxynosine 5’ Abridged Anchor Primer (dI5’AAP)(Invitrogen)(Table 3.1). PCR reactions were performed with *Tth* Plus® DNA Polymerase (Fisher Biotech). Reactions were assembled in a 0.5ml, thin walled polypropylene tube in a final volume of 50µl, containing 10x PCR Reaction Buffer [1x; 67mM Tris-HCl (pH 8.8), 16.6mM [NH4]2SO4, 0.45% (v/v) Triton® X-100, 0.2mg/ml gelatin](Fisher Biotech), 1.5mM MgCl2 (Fisher Biotech), 0.2mM dNTP Mix (Fisher Biotech), 20pmol of SLIM1(2) primer, 20pmol of dI5’AAP (Invitrogen), 5µl of dC tailed cDNA and 2.75U of *Tth* Plus® DNA Polymerase (Fisher Biotech). Samples were denatured for 2min at 94°C and then amplified using 30 cycles (94°C for 30s, 52°C for 30s and 72°C for 1min) followed by a final extension of 72°C for 7 min (5’RACE)(Appendix 1) in a PTC-100 (MJ Research Inc.). The specificity of products generated was confirmed by subsequent nested PCR using SLIM1(3) and the supplied Abridged Universal Amplification Primer (AUAP)(Invitrogen)(Table 3.1), in a 50µl reaction volume containing 10x PCR Reaction Buffer [1x], 1.5mM MgCl2, 0.2mM dNTP Mix, 10pmol of SLIM1(3), 0.2µM
AUAP (Invitrogen), 0.1% (v/v) of the original PCR reaction and 2.75U of Tth Plus® DNA Polymerase. Samples were denatured for 1:30min at 94°C and then amplified (94°C for 30s, 63°C for 1min and 72°C for 1:30min; 30 cycles) with annealing temperatures decreasing by 1°C every 2 cycles, and an annealing temperature of 55°C for the final 14 cycles, followed by a final extension of 72°C for 8min (TD1)(Appendix 1).

PCR products were analysed by agarose gel electrophoresis on 0.6-1.5% (w/v) agarose gels, as described in Section 2.2.1.2. 5’ RACE PCR products were gel purified using the Qiaex II Gel Purification Kit (Qiagen)(Section 2.2.1.3) and directly sequenced using SLIM1(3) and AUAP primers as outlined in Section 2.2.1.11.

3.2.1.4.2 THERMOSCRIPT™ RT

Briefly, 2.5µg total ALL-SIL RNA was combined with 15pmol of SLIM1(1) primer in a thin-walled 0.5ml RNase-free polypropylene tube in a final volume of 20µl. RNA was denatured by incubation at 70°C for 10min and immediately transferred to 50°C. Reverse transcription was performed in a final volume of 40µl comprising 5x cDNA Synthesis Buffer [1x; 50mM Tris acetate (pH 8.4), 75mM potassium acetate, 8mM magnesium acetate](Invitrogen), 5mM DTT (Invitrogen), 40U of RNaseOUT™ (Invitrogen), 1mM dNTP Mix (Invitrogen) and 15U of THERMOSCRIPT™ RT (Invitrogen) with a Master Mix pre-incubated to 50°C, at 65°C for 1h. The reaction was terminated by incubation at 85°C for 5min and the RNA template subsequently removed by incubating with 1µl of RNase Mix at 37°C for 30min and placed on ice. The cDNA was then purified using a GLASSMAX Spin cartridge and tailed as outlined in Section 3.2.1.4.1. Primary and nested PCR was performed using the primer sets described in Section 3.2.1.4.1. Modifications to the existing PCR Programmes utilised in these procedures involved increasing denaturation temperatures to 97°C (5’RACE*, TD2) and extending the final extension of 72°C to 10min (TD2)(Appendix 1). PCR products were analysed by agarose gel electrophoresis on 0.6-1.5%(w/v) agarose gels, as described in Section 2.2.1.2.
3.2.1.4.3 Omniscrypt™ RT

Briefly, 2µg of total ALL-SIL RNA was denatured in a thin-walled 0.5ml RNase-free polypropylene tube in a final volume of 12µl, by incubation at 65°C for 5min and then immediately transferred to 42°C. Reverse transcription was performed in a final volume of 20µl comprising 10x Buffer RT [1x](Qiagen), 0.5mM dNTP Mix (Qiagen), 1µM SLIM1(1) primer, 10U of RNASin (Promega) and 4U of Omniscrypt™ RT (Qiagen) with a Master Mix pre-incubated to 42°C, at 42°C for 2h. The RNA was subsequently removed by incubating with 1µl of RNase Mix at 37°C for 30min and placed on ice. The cDNA was then purified using a GLASSMAX Spin cartridge and tailed as outlined in Section 3.2.1.4.1. Primary and nested PCR was performed using the primer sets described in Section 3.2.1.4.1 and 5’RACE* and TD2 PCR programmes (Appendix 1). PCR products were analysed by agarose gel electrophoresis on 0.6-1.5% (w/v) agarose gels, as described in Section 2.2.1.2.

3.2.1.5 5’ RACE - GeneRacer™ Protocol

3.2.1.5.1 De-phosphorylation of ALL-SIL and HeLa Total RNA

In the first step of this protocol, total ALL-SIL RNA and control HeLa RNA was de-phosphorylated with calf intestinal phosphatase (CIP) to remove the 5’ phosphate group from truncated mRNA and non-mRNA species, leaving capped, full-length mRNA. Removal of the 5’ phosphate group renders these truncated mRNAs incapable of ligating with the GeneRacer™ RNA Oligo adapter, thereby eliminating the amplification of truncated messages and facilitating the amplification of only full-length transcripts.

Approximately 5µg of ALL-SIL RNA and 1µg of HeLa RNA was combined in a thin-walled 0.5ml RNase-free polypropylene tube in a final volume of 10µl, comprising 10x CIP Buffer [1x; 50mM Tris-HCl (pH 8.5), 0.1mM EDTA (pH 8.0)](Invitrogen), 40U of RNaseOUT™ and 10U of CIP (Invitrogen) at 50°C for 1h. The reaction was subsequently precipitated by dilution to 100µl with DEPC treated dH2O and extraction with 1 volume [1:1; 100ul] of Tris-buffered phenol:chloroform (pH 8.0). The reaction
was centrifuged (15,000 x g, 5min, room temperature) and the upper aqueous phase containing RNA was transferred to a fresh 1.5ml polypropylene microcentrifuge tube and precipitated with 2µl of 10mg/ml mussel glycogen (Invitrogen), 10µl of 3M NaAc (pH 5.2)(Invitrogen) and 220µl of 95% (v/v) ethanol. Precipitation was enhanced by incubating on dry ice for 10min and RNA was pelleted by centrifugation (15,000 x g, 20min, 4°C) and washed once with 500µl 70% (v/v) ethanol. Following centrifugation (15,000 x g, 2min, 4°C) the ethanol was removed and the sample was re-centrifuged to collect residual ethanol. The RNA pellet was briefly air-dried for 1-2min and resuspended in 7µl of sterile DEPC-treated water.

3.2.1.5.2 Removing the 5’ Cap Structure from Full-length mRNA

The second part of the GeneRacer™ protocol involves the removal of the 5’ cap structure from the full-length RNA generated in Section 3.2.1.5.1 with tobacco acid pyrophosphatase (TAP), in order to expose the 5’ phosphate group for required for subsequent ligation of the GeneRacer™ RNA Oligo. De-capping was performed in a final volume of 10µl comprising 10x TAP Buffer [1x; 50mM NaAc (pH6.0), 1mM EDTA, 0.1% (v/v) β-mercaptoethanol, 0.01% (v/v) TritonX®-100] (Invitrogen), 40U of RNaseOUT™ and 0.5U of TAP (Invitrogen) by incubation at 37°C for 1h. Following incubation, the RNA was precipitated as outlined in Section 3.2.1.5.1 and resuspended in 7µl of DEPC-treated water.

3.2.1.5.3 Ligation of the RNA Oligonucleotide Adapter to De-capped mRNA

The third step of this protocol involves the ligation of GeneRacer™ RNA Oligo to the 5’ end of the mRNA using T4 RNA ligase in order to provide a known priming site for the GeneRacer™ PCR primers after the mRNA is transcribed into cDNA. The mRNA generated in Section 3.2.1.5.2 was incubated with 0.25µg of pre-aliquoted, lyophilised GeneRacer™ RNA Oligo (Invitrogen)(Table 3.1) at 65°C for 5min to relax the RNA secondary structure and briefly chilled on ice. The ligation was performed in a final volume of 10µl incorporating 10x T4 RNA Ligase Buffer [1x; 33mM Tris-acetate (pH 7.8), 66mM KAc, 10mM MgAc, 0.5mM DTT](Invitrogen), 40U of RNaseOUT™ and
5U of T4 RNA Ligase (Invitrogen) at 37°C for 1h. RNA was precipitated as outlined in Section 3.2.1.5.1 and resuspended in 10µl of sterile DEPC-treated water.

3.2.1.5.4 Reverse Transcribing mRNA

First strand cDNA was synthesised from RNA adapter-ligated ALL-SIL and HeLa mRNA using THERMOSCRIPT™ RT (Invitrogen) and the primers SLIM1(1) and GeneRacer™ Oligo dT respectively (Table 3.1). RNA was combined with 15pmol of SLIM1(1) primer or 1µl of 50µM GeneRacer™ Oligo dT (Invitrogen)(Table 3.1) and 1µl of 10mM dNTPs (Invitrogen) in a 0.5ml thin-walled RNase-free polypropylene tube in a final volume of 13µl, and incubated at 70°C for 10min to denature RNA. The reaction was then immediately transferred to 50°C. Reverse transcription was performed in a final volume of 20µl incorporating 5x cDNA Synthesis Buffer [1x; 50mM Tris acetate (pH 8.4), 75mM KAc, 8mM MgAc], 40U of RNaseOUT™, 5mM DTT and 15U of THERMOSCRIPT™ RT with a Master Mix pre-incubated to 50°C, at 65°C for 1h. The reaction was terminated by incubation at 85°C for 5min and chilled on ice. The RNA template was subsequently removed by incubating with 2U of RNase H at 37°C for 20min.

3.2.1.5.5 PCR Amplification

FHL1-specific transcripts were amplified from ALL-SIL derived cDNA by primary PCR using the GeneRacer™ 5’ and SLIM1(2) primer set (Table 3.1). PCR reactions were performed with Tth Plus® DNA Polymerase (Fisher Biotech) in 0.5ml thin-walled polypropylene tubes in a final volume of 50µl comprising 10x PCR Reaction Buffer [1x], 1.5mM MgCl₂, 0.2mM dNTP Mix, 20pmol of SLIM1(2) primer, 0.6µM GeneRacer™ 5’ primer (Invitrogen), 2µl of ALL-SIL cDNA and 2.75U of Tth Plus® DNA Polymerase using the modified 5’RACE® PCR programme (Appendix 1). Nested PCR was then performed using the GeneRacer™ 5’ Nested and SLIM1(3) primers (Table 3.1) in a 50µl reaction volume containing 10x PCR Reaction Buffer [1x], 1.5mM MgCl₂, 0.2mM dNTP Mix, 10pmol of SLIM1(3) primer, 0.2µM GeneRacer™ 5’ Nested primer (Invitrogen), 1µl of original PCR and 2.75U of Tth Plus® DNA Polymerase using the TD2 PCR programme (Appendix 1).
In addition, the 5' RACE positive control β-Actin was amplified from HeLa derived cDNA. Primary PCR was performed as described for FHL1 using the GeneRacer™ 5’ primer [0.6µM] and Control Primer B.1 [0.6µM](Invitrogen)(Table 3.1) and 2µl of HeLa RT template. Primary products were amplified using the modified 5'RACE* PCR programme (Appendix 1). PCR products were analysed by agarose gel electrophoresis on 0.6-1.5% (w/v) agarose gels, as described in Section 2.2.1.2.

3.2.1.6 Cloning of 5' RACE Products

5’ RACE PCR products generated in Sections 3.2.1.4/3.2.1.5 possess single 3’-adenosine (A) overhangs due to the non-template dependent terminal transferase activity of Taq Polymerase. This can be utilised to directly clone PCR products into T/A cloning vectors, which possess 3’ deoxythymidine (T) overhangs capable of participating in cohesive-end ligation. Freshly prepared PCR products were T/A cloned into the pCR®2.1 TA cloning vector (Invitrogen)(Appendix 2) in accordance with the manufacturer’s instructions. Ligation reactions were performed in 0.5ml polypropylene tubes in a final volume of 10µl, comprising 1-4µl of fresh PCR product, 50ng of pCR®2.1 vector (Invitrogen), Ligation Buffer [1x; 6mM Tris-HCl (pH 7.5), 6mM MgCl2, 5mM NaCl, 0.1mg/ml BSA, 7mM β-mercaptoethanol, 0.1mM ATP, 2mM DTT, 1mM spermidine](Invitrogen) and 4U of T4 DNA Ligase (Invitrogen) and incubated at 15°C for 24-48h. Ligation reactions were transformed immediately or stored at -20°C prior to transformation.

3.2.1.7 Bacterial Transformation

Ligation reactions were transformed into the TOP10F’ strain [F’ {lacI°Tn10 (TetR)} mcrA Δ(mrr-hsdRMS-mcrBC) Ø80lacZΔ M15 Δ lacX74 recA1 deoR araD139 Δ (ara-leu)7697 galU galK rpsL (Strr) endA1 nupG (Invitrogen), which over-expresses the Lac repressor (lacI° gene). Previously prepared, chemically competent cells [50µl] were thawed on ice and combined with 2µl of ligation reaction containing recombinant DNA, and incubated on ice for 30min. The cells were then subjected to heat shock at 42°C for 90s and chilled on ice for 2min. To facilitate growth of the transformed bacteria, 250µl of room temperature equilibrated SOC medium [2% (w/v) Tryptone, 0.5% (w/v) Yeast
Extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl$_2$, 10mM MgSO$_4$, 20mM glucose] was added to the cells, which were then incubated for 1h at 37°C. LB agar plates containing 0.1mg/ml ampicillin (Austrapen®) were prepared for transformation by equilibrating at 37°C for 30min and spreading 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal)(Promega)[50mg/ml; 40µl/plate) and IPTG (Progen)[100mM; 40µl/plate] onto plates and allowing to soak for at least 30min. Aliquots (50-200µl) of the transformed bacterial culture were then spread onto selective LB agar plates and incubated at 37°C for 18-20h to allow growth of the recombinant colonies. To facilitate the blue colour development of non-insert containing colonies, plates were subsequently incubated at 4°C for 1-2h.

3.2.1.8 PCR Screening of Recombinant DNA Clones

Selected colonies were screened for the presence of inserts by colony lysis and PCR as described by Sambrook and Russell (2001). Briefly, individual colonies (typically 10-15 colonies per screen) were selected using a sterile pipette tip and used to inoculate a gridded LB agar plate supplemented with 0.1mg/ml ampicillin and a 1.5ml screw-cap polypropylene tube containing 30µl of TE-Triton® X-100 [10mM Tris-HCl (pH 7.5), 1mM EDTA, 1% (v/v) Triton® X-100] respectively. The bacteria were lysed by boiling the tubes for 5min, and cellular debris was pelleted by centrifugation (15,000 x g, 2min, room temperature). The supernatants were removed to fresh tubes and 5µl was used to screen colonies for the presence of inserts in a standard PCR reaction with $T_{th}$ Plus® DNA Polymerase (Fisher Biotech), utilising M13 Forward (M13F-20) and M13 Reverse (M13R) primers (Table 3.1), which flank the cloning site of the pCR®2.1 TA cloning vector. Reactions were performed in a 50µl volume incorporating 10x PCR Reaction Buffer [1x], 1.5mM MgCl$_2$, 0.2mM dNTP Mix, 20pmol of M13F-20 and M13R primers, 5µl of colony supernatant template and 2.75U of $T_{th}$ Plus® DNA Polymerase. Samples were denatured for 1:40min at 97°C and then amplified (97°C for 40s, 55°C for 1min and 72°C for 1min; 30 cycles), followed by a final extension of 72°C for 7min (SEQKIM)(Appendix 1).
3.2.1.10 DNA Sequencing

Small scale plasmid DNA samples were prepared using the GenElute™ Plasmid Miniprep Kit (Sigma) as described in Section 2.2.1.9 and the inserts were sequenced using the M13 Forward (-20) and M13 Reverse primers flanking the cloning site of the pCR®2.1 TA cloning vector (Section 2.2.1.11), in order to confirm the identity of the cloned insert as a FHL1 gene fragment and to determine the most 5’ sequence corresponding to the transcriptional start site/s.

3.2.2 Luciferase Reporter Gene Assays

3.2.2.1 Preparation of FHL1 Promoter Luciferase Reporter Vectors

3.2.2.1.1 PCR Amplification of FHL1 Promoter Insert DNA

FHL1 promoter fragments were created using the sense primers SLIM1F-300, SLIM1F-928 and a common anti-sense primer SLIM1+73SmaI (Table 3.1), which harbour Sac I and Sma I restriction endonuclease sites respectively to enable cohesive ligation into the pGL3Basic Luciferase Reporter Vector. Fragments were amplified from genomic DNA using DyNAzyme™ EXT DNA Polymerase (Finnzymes) in a 50µl reaction incorporating Optimized DyNAzyme™ EXT 10x Buffer [1x; 50mM Tris-HCl (pH 9.0), 1.5mM MgCl₂, 15mM (NH₄)₂SO₄, 0.1% (v/v) Triton® X-100](Finnzymes), 0.2mM dNTP Mix (Fisher Biotech), 20pmol of either SLIM1F-928/SLIM1F-300 primers, 20pmol of SLIM1+73R primer, 10% (v/v) DMSO (Finnzymes), 500ng of genomic DNA and 0.5U of DyNAzyme™ (Finnzymes) using a modified touchdown PCR programme for high GC content (TD2)(Appendix 1).

3.2.2.1.2 Restriction Digestion of FHL1 Promoter Insert DNA

PCR products generated in Section 3.2.2.1.1 were gel purified using the Qiaex II Gel Purification Kit (Qiagen) as described in Section 2.2.1.3 and digested with Sac I and Sma I restriction endonucleases (Promega) in a 50µl reaction volume comprising 10x MULTICORE™ Buffer [1x; 25mM Tris acetate (pH 7.8), 0.1M potassium acetate,
10 mM magnesium acetate, 1 mM DTT (Promega), 0.1 mg/ml nuclease-free acetylated BSA (Promega), 25 µl of gel purified fragments, 60U of Sac I restriction endonuclease (Promega) and 60U of Sma I restriction endonuclease (Promega) at 25°C for 18h and subsequently at 37°C for 2h to ensure complete digestion. Following digestion, restriction endonucleases were heat inactivated by incubation at 65°C for 20min. Digested PCR fragments were subsequently purified using the QIAquick PCR Purification Kit (Qiagen) according to the manufacturer’s instructions (Section 2.2.1.4). An aliquot of the purified insert was electrophoresed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) to estimate DNA integrity and concentration prior to ligation.

3.2.2.1.3 Preparation of pGL3Basic Luciferase Reporter Vector

The pGL3Basic Luciferase Reporter Vector was prepared for sub-cloning by restriction digestion of supercoiled, pGL3Basic plasmid with Sac I and Sma I restriction endonucleases (Promega) in a 100 µl reaction volume comprising 10x MULTICORE™ Buffer [1x], 0.1 mg/ml nuclease-free acetylated BSA, 2 µg of pGL3Basic plasmid DNA, 60U of Sac I restriction endonuclease and 60U of Sma I restriction endonuclease at 25°C for 18h and subsequently at 37°C for 2h. Following digestion, restriction endonucleases were heat inactivated by incubation at 65°C for 20min and the vector (100 µl) was subsequently purified by agarose gel extraction using the QIAquick PCR Purification Kit (Qiagen) according to the manufacturer’s instructions (Section 2.2.1.3). An aliquot of the purified pGL3Basic vector was electrophoresed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) to estimate DNA integrity and concentration prior to ligation.

3.2.2.1.4 FHL1 Reporter Construct Preparation

Ligations, screening, sequencing and large scale DNA preparation of FHL1 Promoter luciferase reporter constructs was performed essentially as described in Sections 2.2.1.6-2.2.1.15.
3.2.2.2 Luciferase and β-Galactosidase Reporter Gene Assays

HEL-900 and PER-117 cells lines and culture conditions are outlined in Section 2.2.2. FHL1 promoter activity was assessed in HEL-900 and PER-117 cell lines by transient transfection with FHL1-promoter luciferase constructs generated in Section 3.2.2.1. Luciferase and β-galactosidase reporter assays were performed as described in Section 2.2.3. Data analysis was performed as outlined in Sections 2.2.3.5-2.2.3.6.

3.3 RESULTS

3.3.1 ALL-SIL RNA Analysis

The T-ALL cell line, ALL-SIL, which co-expresses HOX11 and FHL1 (Greene et al., 1998), was chosen for the characterisation of FHL1 transcripts. Total RNA from this cell line was extracted using TRIZOL® reagent. TRIZOL® is based on the propensity of RNA molecules to remain dissolved in an aqueous solution comprising 4M guanidinium thiocyanate (pH 4.0) in the presence of a phenol/chloroform organic phase. The latter retains DNA, whilst proteins and other cellular macromolecules are retained at the interphase (Gerstein, 2001). This method ensured the efficient isolation of high quality RNA with minimal protein and genomic DNA contamination, thus maximising the synthesis of full-length cDNA in subsequent 5’RACE analysis. The integrity of ALL-SIL RNA was confirmed using conventional agarose gel electrophoresis, by visual comparison of the intensity of the 28S rRNA band (size ~ 5kB) with the 18S rRNA band (size ~2kB)(ratio of ~2)(Figure 3.1A). In addition, ALL-SIL RNA quantity, purity and integrity was assessed using the RNA 6000 LabChip Kit in conjunction with the Agilent 2100 Bioanalyzer and data is presented as a gel-like image (Figure 3.1B) or as an electropherogram (Figure 3.1C). ALL-SIL RNA concentration [~2.2ug/ul] was calculated using associated Agilent software, which compares unknown samples to the ladder fragments of known concentration. Concentration and purity (A_{260}/A_{280}: ~1.67) was subsequently confirmed by UV spectrophotometric analysis. ALL-SIL RNA integrity was confirmed by the presence of distinct 28S and 18S ribosomal peaks and a calculated rRNA ratio of 1.94, as indicated on the electropherogram (Figure 3.1C). The absence of smaller, defined peaks between the two major ribosomal RNAs in addition
Figure 3.1. Assessment of ALL-SIL RNA Integrity for Use in 5’ RACE Analysis. ALL-SIL RNA integrity and concentration was assessed by agarose gel electrophoresis and Agilent 2100 Bioanalyzer using the RNA 6000 LabChip Kit. (A) Total ALL-SIL RNA (1µg) was analysed by 1.2% agarose gel electrophoresis in 1x TAE Buffer stained with 0.2µg/ml ethidium bromide. (B) Lane 1. Agilent - RNA 6000 Ladder Standard (150ng) Lane 2. Agilent - ALL-SIL RNA sample (3µl). 28S and 18S rRNA bands are indicated (C) Electropherogram of ALL-SIL RNA sample. Major peaks indicate 28S and 18S rRNA bands (rRNA ratio: 1.94).
to a lack of peaks between 28s and the 18S ribosomal RNA, which correspond to smaller RNA molecules, are indicative of high quality RNA. 5S and tRNA molecules are represented peaks between 24-28s (Figure 3.1C) and are commonly present in total RNA samples prepared by TRIzol® extraction.

3.3.2 Identification of the Transcriptional Start Sites of the FHL1 Gene

Identification of the FHL1 transcriptional start site was first necessary in order to identify the proximal promoter of FHL1 and to characterise the cis-acting regulatory elements and trans-acting proteins involved in regulating promoter activity. Previous studies by Greene et al., (1999), in which the genomic structure of FHL1 was determined by PCR amplification of human genomic DNA using primers designed at either end of the FHL1 coding sequence, revealed the gene to consist of five coding exons spanning a region of 4.9kb (E3-E7)(Figure 3.2). Interestingly, introns 3-6 (I3-I6) disrupt the coding region of FHL1 near the start of each complete LIM domain, suggesting that exon duplication may have given rise to the tandem LIM domain repeats (Greene et al., 1999). Following these studies, a cDNA library derived from lung large cell carcinoma (Library: NIH_MGC_18) from which FHL1 EST clones were identified, revealed the existence of least two additional non-coding exons (E1, E2) located 58.9kb and 36.5kb upstream of the ATG respectively (Clone: MGC: 15297; Genbank Accession: NM_001449; Strausberg unpublished results, 2001)(Figure 3.2). Using gene specific primers located in FHL1 exon 3 (SLIM1(1), SLIM1(2); Table 3.1, Figure 3.2), the 5’ extent of the human FHL1 transcript was determined utilising a PCR-based strategy which results in the Rapid Amplification of cDNA Ends (RACE). The aim was to map the start site(s) of FHL1 in order to subsequently investigate whether the proximal promoter of FHL1 could be regulated by HOX11 using functional transcription assays.

Rapid Amplification of cDNA Ends (RACE) or ‘anchored’ PCR, is a method for obtaining full-length sequences from a messenger RNA (mRNA) template between a defined internal site and unknown sequences at either the 3’ or 5’ end of the mRNA, and can be used to amplify rare mRNAs (30 copies/cell) that may challenge conventional cDNA cloning methods (Invitrogen, 2001, Carey & Smale, 2000, Loh et
FIGURE 3.2
A number of commercial kits suitable for the rapid amplification of 5’ cDNA ends were employed in this study. Initial investigations utilised the 5’ RACE System Version 2.0 (Invitrogen) which involves the generation of first strand cDNA from total RNA using a gene specific primer (GSP1), and the subsequent addition of a homopolymeric tail to the 3’ cDNA end catalysed by terminal TdT (Figure 3.3). The tail then serves as an ‘anchor’ for the PCR amplification of the 5’ cDNA ends using a nested GSP2 and a complementary homopolymer-containing anchor primer. Subsequent studies employed a variation of this procedure provided by the GeneRacer™ Kit (Invitrogen), which is based on RNA ligase-mediated and oligo-capping rapid amplification of 5’ and 3’ cDNA ends (RLM-RACE)(Maruyama & Sugano, 1994, Volloch et al., 1994, Schaefer, 1995)(Figure 3.4). This technique captures the full-length 5’ ends of the targeted cDNA via the selective ligation of an RNA oligonucleotide adapter to the 5’ ends of de-capped RNA using T4 RNA ligase.

3.3.2.1 Analysis of the Human FHL1 Transcripts by 5’RACE (Version 2.0)

To map the transcriptional start site(s) of the FHL1 gene, 5’ RACE was initially performed using the 5’ RACE System for the Rapid Amplification of cDNA Ends Version 2.0 (Invitrogen), with total RNA isolated from the ALL-SIL cell line. Reverse transcription (RT) was primed in three independent experiments (RT1, RT2, RT3) from a sequence located in exon 3 (Figure 3.5) using the supplied SUPERSCRIPT™ II RT and the SLIM1(1) gene specific primer. The resultant cDNAs were then amplified by two rounds of PCR using the primer sets AAP/SLIM1(2) and nested AUAP/SLIM1(3) to generate major products; RT1(118bp), RT2 (90bp) and RT3 (240bp)(Figure 3.5A). To rule out the possibility of genomic contamination, no RT controls were performed in parallel with all 5’RACE reactions (data not shown). Sequence data was obtained for PCR products generated from RT1 and RT3 which positioned the 5’ FHL1 transcript start point at positions -160 and -35 relative to the ATG translation initiation codon, within exons 1 and 2 respectively (Figure 3.5B). Since both of these putative startpoints were located downstream of the 5’ end of the longest known FHL1 cDNA, it was likely that cDNA synthesis may have been prematurely terminated using SUPERSCRIPT™ II RT This prompted the use of alternative reverse transcriptases, specifically chosen to overcome RNA secondary structure.
Step 1. Anneal the first strand Gene-Specific Primer 1 (GSP1) to mRNA

\[
\begin{align*}
5' & \quad \text{mRNA} \quad (A)_n \\
\text{GSP1} & \quad 3'
\end{align*}
\]

Step 2. Reverse transcribe the mRNA into cDNA using SuperScript™ II RT

\[
\begin{align*}
5' & \quad (A)_n \\
3' & \quad \text{cDNA}
\end{align*}
\]

Step 3. Remove the RNA template in the RNA:cDNA hybrid and single stranded RNA molecules using an RNase Mix (RNase H/RNase T1)

\[
\begin{align*}
3' & \quad \text{cDNA} \\
5' & \quad \text{RNA}
\end{align*}
\]

Step 4. Purify the cDNA from excess nucleotides and primers with GLOSSMAX Spin Cartridge. Tail the purified cDNA with dCTP and Terminal Deoxynucleotidyl Transferase (TdT)

\[
\begin{align*}
3' & \quad \text{CC} \cdots \text{CC} \\
5' & \quad \text{cDNA}
\end{align*}
\]

Step 5. PCR amplify dC-tailed cDNA using the Abridged Anchor Primer and nested GSP2

\[
\begin{align*}
\text{Abridged Anchor Primer} & \quad 5' \\
\text{GSP2} & \quad 3'
\end{align*}
\]

Step 6. Re-amplify primary PCR product using Abridged Universal Anchor Primer (AUAP) and nested GSP3

\[
\begin{align*}
\text{AUAP} & \quad 5' \\
\text{nested GSP} & \quad 3'
\end{align*}
\]

Figure 3.3. Schematic Representation of the 5' System for the Rapid Amplification of cDNA Ends (RACE) Version 2.0 (Invitrogen).
Step 1. RNA is dephosphorylated with calf intestinal phosphatase (CIP) to remove the 5' phosphate group from truncated and non-mRNAs. CIP has no effect on full-length, capped mRNAs.

Step 2. Dephosphorylated mRNA treated with tobacco acid pyrophosphatase (TAP) to remove the 5' cap structure from intact, full-length mRNAs.

Step 3. GeneRacer™ Oligo ligated to the full-length mRNAs using T4 RNA ligase

Step 4. Ligated mRNA is reverse transcribed with Superscript™ II RT and a GSP /GeneRace™ oligo dT primer is used to generate RACE-ready first strand cDNA with known priming sites at the 5' and 3' ends.

Step 5. 5' RACE GSP1 and GeneRacer™ 5' primer used to amplify first-strand cDNA in 5' RLM-RACE to obtain 5' ends. Secondary RLM-RACE is then performed using 5' RACE GSP2 and GeneRacer™ 5' nested primer

Figure 3.4. Schematic Representation of the GeneRacer™ (Invitrogen) Method for Obtaining 5' cDNA Ends.
Early termination of cDNA synthesis in exon 1 using SUPERSCRIPT™ II RT was probably caused by the formation of secondary structures associated with the GC-rich content of FHL1 transcripts. This might explain the different sized products obtained with RT1, RT2 and RT3 replicates (Figure 3.5). To overcome this, two independent reverse transcriptases were employed. THERMOSCRIP™ RT (GibcoBRL®) is an avian RNase H-minus RT, which demonstrates a higher thermal stability (up to 65°C) than the SUPERSCRIPT™ II RT supplied with the 5’RACE Kit (maximum recommended temperature of 50°C). Omniscript™ RT (Qiagen) is a recombinant heterodimeric enzyme expressed in E. coli, which displays a higher affinity for RNA than Moloney Murine Leukemia Virus (MMLV) enzymes – previously considered to be a standard efficient reverse transcriptase (Qiagen News, 1999). Indeed, experiments performed by Korfhage et al., (2000), demonstrated that this high affinity for RNA enabled Omniscript™ Reverse Transcriptase to read through complex RNA secondary structures with a high GC content (maize gl2 transcript, 75% GC) that challenged conventional MMLV RNaseH- RTs.

5’ RACE products obtained utilising THERMOSCRIP™ and Omniscript™ reverse transcriptases yielded multiple products (Figures 3.6A, 3.7A, respectively), however, the largest product obtained independently with both enzymes terminated at a common position. This mapped the FHL1 transcriptional start site to a position -221bp upstream of the initiation codon (ATG), extending the known 5’ end of the gene by 11bp (Figure 3.6B, 3.7B). Although other putative transcriptional start sites corresponding to smaller 5’ RACE products were generated using the THERMOSCRIP™ RT, these were not independently verified using Omniscript™ RT, an enzyme reported to display a higher affinity for RNA than other reverse transcriptases, enabling it to read through complex RNA secondary structures (Korfhage et al., 2000). Given this, and the fact that the largest products generated with both THERMOSCRIP™ RT and Omniscript™ RT enzymes used terminated at -221bp, it is likely that this RACE product represents the major transcriptional initiation site, at least in ALL-SIL T-cells.
FIGURE 3.7
3.3.2.2 Analysis of the Human FHL1 mRNA Transcripts by 5’ RLM-RACE

In order to confirm the position of the major transcriptional start site identified in Section 3.3.2.1 (-221bp upstream of the ATG initiation codon), and to further investigate the possibility that FHL1 gene transcription initiates from multiple start sites, the GeneRacer™ Kit for the full-length RNA Ligase-mediated rapid amplification of 5’ cDNA ends (RLM-RACE) was next employed. This method relies on the principle that a protective methylated cap is added to the initiating 5’ nucleotide of the primary transcript, such that upon dephosphorylation of 5’ RNA ends with CIP, truncated mRNA and non-mRNA species are eliminated from subsequent ligation with the GeneRacer™ RNA Oligo. RLM-RACE therefore ensures the amplification of only full-length mRNA transcripts. Using this method, major products were amplified of 300 to 400bp and 700bp in size (Figure 3.8A), however, only the smaller products were successfully cloned, despite repeated attempts to enrich and gel purify the larger 700bp fragment. Sequencing of independent clones of the smaller products identified putative transcriptional start sites at approximately -225bp, -253bp and -330bp upstream of the of the ATG initiation codon (Figure 3.8B).

In summary, 5’ RACE analyses using a combination of Thermoscript™ RT, Omniscript™ RT and RLM-RACE to ensure the amplification of only full length mRNA transcripts, resulted in the identification of at least four transcriptional initiation start sites at positions -330bp, -253bp, -225bp and -221bp relative to the FHL1 ATG start codon (summarised in Figure 3.2). For technical reasons, the presence of start sites even further upstream cannot be ruled out, however given the positions of these start sites, which are located relative to a CpG island and predicted promoter (Section 3.3.3), it is likely that these represent authentic transcriptional start sites.

3.3.3 Bioinformatic Analysis of the Human FHL1 Proximal Promoter

Computational prediction of a functional promoter for FHL1 was performed using the Promoter Inspector algorithm (http://www.genomatix.de/softwareservices/software/PromoterInspector/PromoterInspector.html)(Scherf et al., 2000). A 12kb region of the human FHL1 gene encompassing the putative transcriptional start sites identified in
Sections 3.3.2.1/3.3.2.2 (6kb in both 5’ and 3’ directions, centred on exon 1) was examined using PromoterInspector. This method bases selection on the genetic context of promoters (transcription factor binding site occurrence and orientation) as opposed to the mere identification of core promoter elements. Using this method, PromoterInspector identified a promoter (spanning 696nt) within the same vicinity as that identified by 5’ RACE analysis conducted within this study (Figure 3.1). Having identified a putative promoter both by transcription initiation start site mapping and in silico promoter prediction methods, the DNA sequence surrounding the transcriptional start sites identified in Sections 3.3.2.1/3.3.2.2 was analysed in further detail, and revealed that the core promoter region of FHL1 lacks conspicuous promoter elements, such as a consensus TATA (TATAAA) or CCAAT box. The lack of a TATA box, in particular, is in keeping with the finding of multiple transcriptional initiation sites. Given the high G/C content and presence of multiple binding sites for the ubiquitous mammalian transcription factor Sp1 in the 5’ DNA sequence flanking the start sites identified in this study, we postulate that FHL1 belongs to class of ‘TATA-less’ promoters that initiate transcription from multiple sites, over a broad initiation window (Figure 3.9). In further support of this hypothesis, a sequence closely resembling a MED-1 element (multiple start site element downstream)(GCTCCS; where S= strong base pair corresponding to a G or C), found in many multiple start site promoters, was identified at +21 relative to the most 3’ start site identified (GCTCAG)(Figure 3.9). MED-1 elements typically reside at +20 to +45 from the 3’ end of the transcriptional initiation window (Ince & Scotto, 1995).

The putative FHL1 promoter identified in this study also exhibited some features of a single start site promoter (SSS), with a consensus Initiator (Inr)(YYANWYY; where Y= pyrimidine, N= any base pair, W= weak base pair corresponding to an A or T) sequence overlapping the most 3’ start site, being identified (Figure 3.9). The initiator element appears to be functionally analogous to the TATA box and can act in conjunction with a downstream promoter element (DPE)(GNNNRGWYGT; where R= purine), which is typically located ~30bp downstream of the initiation site to direct the specific initiation of transcription (Carey & Smale, 2000). In keeping with the presence of a functional Inr, a sequence resembling a consensus DPE (GNNNRGWYGT) was also identified at +24 relative to the most 3’ start site identified (GCTCAGTCCG).
Figure 3.9. Nucleotide Sequence of the Putative FHL1 Promoter. Human FHL1 transcriptional initiation sites identified by GeneRacer™ 5' RLM-RACE Analysis and confirmed by sequencing are indicated by arrowheads (▼). Nucleotides are numbered relative to the major transcriptional start site (+1). Putative transcription factor binding sites are underlined. Partial sequence of exon 1 with highlighted region indicating the extent of 5' cDNA obtained from Genbank under Accession NM_001449; or as an EST Accession BI559753; dbESTId: 9407273 identified in a cDNA library (Strausberg, 2001). Red brackets indicate a broad window where transcription is initiated and is typical of a MSS promoter. Bold letters indicate a consensus initiator (Inr) sequence. Blue brackets represent a putative downstream promoter element (DPE). Forward and reverse primers for the amplification of the FHL1 proximal promoter are indicated by red and green lettering, respectively.
Consensus binding sites present in the 1.7kb sequence located upstream of the putative FHL1 start sites, were identified using MatInspector V2.2, which is based on TRANSFAC™ 4.0 matrices ([http://transfac.gbf.ge/TRANSFAC/index.html](http://transfac.gbf.ge/TRANSFAC/index.html))(Wigender et al., 2000). Putative consensus motifs for several transcription factors including the Myeloid Zinc Finger-1 (MZF-1), TCF11, Sex-Determining region Y gene product (SRY), CCAAT Displacement Protein (CDP) and FOXL1 were identified (Figure 3.9). Not unexpectedly, the binding sites of a number of transcription factors specifically associated with muscle development were also identified, including MyoD, the murine tinman homolog, cardiac specific homeobox transcription factor, Nkx2.5 and HAND2 (Figure 3.9).

### 3.3.4 Comparison of Human-Murine FHL1 Promoter Sequences

Cross-species sequence comparisons represent a useful way to determine the presence of important regulatory elements and/or coding regions, since these are likely to be evolutionarily conserved. Alignment of the human and murine FHL1 promoters was therefore performed over a 1,147bp region, spanning the putative transcriptional start sites identified by 5’ RACE analysis (Section 3.3.2.1/3.3.2.2)(Figure 3.10). This included sequence located immediately 5’ of human FHL1 exon 1 (exon 1 inclusive) and the corresponding murine sequence which was obtained by BLAST search of the murine database with published human FHL1 Exon 1 (81% homology). Comparative pairwise alignment with the murine FHL1 promoter sequence was performed using the sequence alignment programme, ClustalX (1.81) and revealed 63.79% overall sequence identity between the two species. Extensive blocks of homology occur throughout the promoter region, particularly within the first 300bp upstream of the putative transcriptional start site(s). Initial sequence analysis revealed that, like the human FHL1 promoter, the murine Fhl1 promoter lacks TATA and CCAAT box elements, however the MED-1 and Inr elements identified within the human FHL1 promoter (+21, -1, respectively) were conserved in mouse. Moreover, the murine DPE at +24 (GCTCAGTCCT) shared greater homology to the consensus DPE sequence in comparison to its human counterpart. In addition, a number of putative transcription factor binding sites are highly conserved between the two species including consensus motifs for Sp1, MyoD, CDP, TCF11, MOK2 and AP2.
Figure 3.10. Sequence Alignment of the Human and Murine FHL1 Exon 1 and 5' Flanking Region. Nucleotides with sequence identity are indicated with an asterisk (*). Conserved putative transcription factor binding sites are underlined. The four transcriptional initiation sites identified by 5' GeneRacer™ Analysis are indicated by arrowheads (▼). The previously known extent of exon 1 is highlighted.
3.3.5 Identification of a HOX11 Responsive Element in the FHL1 Promoter

In Chapter 2, HOX11 was demonstrated to be capable of both activating and repressing the transcription of the ALDH1A1 promoter in the HEL and PER-117 cell line, respectively. These results supported the hypothesis that ALDH1A1 is a bona fide target gene of HOX11, and provided a starting point for addressing the way in which HOX11 acts as a dual-function regulator. Likewise, in order to confirm the status of FHL1 as a gene transcriptionally regulated by HOX11, following identification of the transcriptional start sites of FHL1, luciferase reporter assays were performed on two pGL3Basic luciferase reporter constructs containing 297bp and 926bp of 5’ genomic DNA flanking FHL1 exon1. The aim was to investigate the response of the FHL1 proximal promoter to HOX11 in two different cell backgrounds, namely T-cell (PER-117) and erythroid (HEL).

In order to firstly assess the activity of the FHL1 promoter in the absence of HOX11, pGL3BasicFHL1-297 and -926 luciferase constructs, together with a cotransfection control constitutively expressing β-Galactosidase, were transiently transfected into the PER-117 and HEL cell lines (neither of which express the endogenous FHL1 gene at the level of Northern blot). As described in Chapter 2, the results are expressed as relative transcriptional activity, which represents the ratio of luciferase activity to β-galactosidase activity. Transfection of the pGL3BasicFHL1-297 construct resulted in a significant stimulation of promoter activity over the pGL3Basic construct alone, indicating that positive cis-regulatory elements are present in the first 297bp of the FHL1 promoter (Figure 3.11). The presence of an additional 629bp of 5’ flanking sequence (pGL3BasicFHL1-926) resulted in a ~2-fold decrease in promoter activity for PER-117, suggesting that the region between -297/-926bp may contain a negative regulatory element(s) that contribute to the cell-type specific silencing of FHL1 in PER-117. This decrease in expression was not observed in the HEL cell line, however, implying that distal elements not included in these analyses may repress FHL1 in HEL cells.
FIGURE 3.11
In order to investigate the ability of HOX11 to regulate the FHL1 promoter, the -297 and -926 constructs were subsequently cotransfected with either the empty expression vector, pEFBOS, or a HOX11-expressing version, pEFBOSHOX11. As described in Chapter 2, the results of luciferase experiments are presented as a log-fold change of normalised individual promoter constructs relative to the pGL3Basic empty vector control (performed by Dr Ross Taplin, Division of Science and Engineering, Murdoch University). Cotransfection of pEFBOSHOX11 with the pGL3BasicFHL1-297 promoter construct resulted in a modest repression of the FHL1 promoter in PER-117 cells which was augmented following the addition of 5’ flanking sequence (pGL3BasicFHL1-926), suggesting that HOX11 is capable of repressing the FHL1 promoter in PER-117 (Figure 3.12). Slight repression was also demonstrated in the HEL cell line, although this was only statistically significant in the case of the -926 construct (Figure 3.12).

Given that HOX11 was capable of repressing transcription from the FHL1 promoter in the PER-117 cell line, the ability of a DNA binding mutant of HOX11, HOX11H3, to repress transcription was subsequently investigated. Similar to the effect observed for the ALDH1A1 promoter in Chapter 2, HOX11H3 was incapable of repressing transcription from the -297 and -926 FHL1 promoter constructs, and instead demonstrated strong transactivation potential (Figure 3.13). These results corroborate the dual function of HOX11 observed in Chapter 2, and confirm the requirement for the DNA binding helix in HOX11-mediated transcriptional repression.
FIGURE 3.13
3.4 DISCUSSION

Although the expression pattern of FHL1 in skeletal and cardiac muscle indicates a functional role in development, the identification of FHL1 as a possible target gene of the T-cell oncprotein HOX11, expressed at high frequency in T-ALL cell lines (Greene et al., 1998), highlighted the need for an increased understanding of the mechanisms of FHL1 gene regulation in both normal and tumorigenic contexts. The identification of cis-acting sequence elements and trans-acting proteins responsible for the regulation of promoter activity in a given cellular environment, is a crucial starting point in understanding gene expression. These regulatory elements, which constitute the proximal promoter of the gene, are located within the vicinity of the initiation site otherwise known as the transcriptional start site (Carey & Smale, 2000). Given the capacity for many DNA sequences to display above base-line levels of activation in artificial promoter assays, particularly luciferase assays which are extremely sensitive, the accurate identification of the FHL1 transcriptional start site was first necessary in order to undertake subsequent promoter analysis.

5’ RACE was utilised in this study to identify the start site of FHL1, since it provides a rapid and extremely sensitive means for determining cDNA ends. Although the enhanced sensitivity of this PCR-based method facilitates the identification of start sites for genes that are inefficiently transcribed, the limitations of RACE are related to artefacts produced as a result of this sensitivity, highlighting the requirement for careful data analyses to avoid misinterpretation of results. For example, the preferential amplification of shorter cDNAs generated in the initial reverse transcriptase step, which may be the result of premature termination or simply represent a shorter authentic transcript, can be misleading since these smaller products will appear to represent the major initiation start site (Carey & Smale, 2000). For this reason, our analyses of FHL1 incorporated the use of four independent reverse transcriptases in order to maximise the probability of obtaining the longest FHL1-specific cDNA present in the ALL-SIL mRNA population.

A standard 5’ RACE procedure (5’ RACE System Version 2.0; Invitrogen) was initially employed to identify the 5’ end of the FHL1 encoded mRNA in order to determine its
transcriptional start site. However, the possibility that the cDNA amplified was prematurely terminated as a result of the high GC content associated with the 5’ exon template of FHL1 or was degraded or cleaved at the 5’ end during handling, resulting in a false start site position, prompted the use of the GeneRacer™ technique (Invitrogen). This method selects for mRNA containing a cap structure, which is added to the 5’ end of transcripts synthesised by RNA polymerase II, thereby increasing the likelihood of amplifying authentic start sites. Indeed, the sequences obtained using this refined method resulted in the identification of transcriptional start sites extending an additional 108bp from the furthest 5’ end identified using the standard 5’ RACE strategy. Overall, the sequencing of 10 clones, which were obtained using a combination of independent reverse transcriptases and two different techniques, resulted in the identification of four transcriptional initiation start sites -329bp, -252bp, -224bp and -221bp upstream of the FHL1 ATG start codon, within a region spanning 109bp. Although these results may reflect the inability of the RT to transverse a CG-rich region, resulting in early and sporadic termination, it was considered more likely that they represent multiple transcription initiation sites - a characteristic feature of some ‘TATA-less’ promoters (Weis & Reinberg, 1992). Known as multiple start site promoters, these typically possess a ‘window’ of start sites, in addition to a high CpG dinucleotide content and numerous Sp1 (CCCGCC) binding sites - characteristics shared by the FHL1 promoter identified in this study. It is postulated that Sp1 may direct the formation of a preinitiation complex containing i) general transcription factors (GTFs) RNA Pol II, TFIIA, TFIIB, TFIID, TFIE, TFIIF and TFIIF and ii) coactivators and corepressors, to a region 40 to 100bp downstream of its binding sites by interacting with a component of TFIID (Slansky & Farnham, 1996, Smale et al., 1990). This class of TATA-less promoter is also associated with a MED-1 element - a protein binding sequence located +20 to +45 downstream of the initiation window that was shown to be necessary, but not sufficient for multiple start site utilisation (Ince & Scotto, 1995). Indeed, a sequence closely resembling a consensus MED-1 element was identified +21 relative to the most 3’ FHL1 start site and this was also shown to be conserved in the corresponding murine Fhl1 sequence. Whilst many TATA-less promoters are associated with ‘housekeeping’ genes, which generally exhibit little cell-type specificity, oncogenes and genes encoding transcription factors and growth factors are also often TATA-less. An example of the regulation of a tissue specific gene by a TATA-less promoter is the human Zipper
protein kinase, which is expressed specifically in the developing brain (Itoh et al., 2004). Thus, it is entirely plausible that the muscle specific FHL1 is regulated by a TATA-less promoter, as identified by the 5' RACE studies.

Alternatively, the identification of highly conserved, consensus Inr (-1) and DPE (+24) elements relative to the major FHL1 start site identified in this study, suggests that FHL1 may be regulated by a promoter with an Inr-DPE and no TATA box. The DPE, which is located between +28 to +34 relative to the transcription initiation site, is thought to represent a downstream counterpart to the TATA box and is conserved from Drosophila to humans (Burke & Kadonaga, 1997). In this case, the initiator element (Inr), which encompasses the start site, acts in conjunction with the DPE to direct the formation of a preinitiation complex in the absence of a TATA box (Burke & Kadonaga, 1996). It has been postulated that TFIID (a multi-subunit protein comprising TATA box binding protein (TBP) and TBP-associated factors), cooperatively binds Inr and DPE motifs, thereby nucleating the binding of additional GTFs required for the initiation of transcription (Burke & Kadonaga, 1996, 1997).

To complement the molecular techniques used to identify the transcriptional start site and corresponding promoter of FHL1, bioinformatic tools were also employed to search for promoter-specific characteristics in a 12kb region surrounding the putative FHL1 start site. The development of computer programmes to predict eukaryotic promoters and identify functional response elements has led to the emergence of a powerful tool that can be used as an adjunct to experimental gene analysis. The ability to perform rapid, large-scale sequence analysis renders promoter prediction algorithms such as PromoterInspector (Scherf et al., 2000), PromFind (Hutchinson, 1996), PromoterScan (Prestridge, 1995), and Autogene (Kondrakhin et al., 1994), especially useful as ‘pre-processing’ steps in promoter identification and annotation studies. The predictive capacity of any given promoter analysis programme is a function of its ‘approach’, which can be heuristic, core promoter or context driven (Scherf et al., 2000). Heuristic approaches are based on the recognition of several transcription factor binding sites in the appropriate orientation and context, however, although this approach detects promoters with high specificity, it is limited to the identification of a well-characterised subset of promoter types and is therefore unsuitable for promoter prediction on a
genome-wide scale (Werner, 1999). In contrast, the core promoter approach demonstrates relatively low specificity, relying on the identification of core promoter elements including the TATA box, CCAAT box and transcriptional initiation sites (Fickett & Hatzigeorgiou, 1997). Not only does this approach exclude TATA-less promoters, but it also yields large proportion of false positive identifications, thereby limiting its use in large-scale analyses (Fickett & Hatzigeorgiou, 1997). The use of an algorithm that takes into account the heterogeneity of promoter structure whilst recognising contextual similarities is therefore crucial in order to accurately delineate true polymerase II promoters.

PromoterInspector is a promoter prediction tool developed by Scherf et al., (2000), that accurately locates eukaryotic polymerase II promoter regions from large input sequences on the basis of genomic ‘context’. The programme is trained using an unsupervised technique, to recognise promoter-specific features that are compatible with nucleosome remodelling and the formation of a preinitiation complex containing RNA polymerase II, as opposed to identifying mere transcription factor binding sequences. These shared features form the definition of ‘context’, and since the specific features of promoter context are unknown, this approach facilitates the prediction of new, unknown regulatory regions, such that the analysis extends beyond our limited understanding of what constitutes a functional eukaryotic promoter (Scherf et al., 2000, Carey & Smale, 2000). A comparison of PromoterInspector with other promoter prediction tools using a dataset reviewed by Fickett & Hatzigeorgiou (1997) demonstrated the accuracy of this heuristic-free approach, which generated a record true positive to false positive ratio of 2.3 using input sequences up to 300kbp in length (Scherf et al., 2000). Thus, the identification of a putative promoter within a region of the FHL1 5’ UTR (exon 1) overlapping the position of several transcriptional start sites mapped during our 5’ RACE analysis, strongly suggests that the sequence 5’ of transcriptional start site 1 (-329bp) represents a bona fide regulatory region for subsequent bioinformatic and functional investigations.

Using the MatInspector V2.2 transcription factor binding site analysis programme, consensus binding sites, conserved in both human and murine FHL1 promoters were identified for a number of transcriptional regulators. Sp1 is a sequence specific DNA-
binding protein that was originally identified as a transcriptional activator of the SV40 early promoter (Dynan & Tjian, 1983), and was later shown to be a widespread activator of both housekeeping and tissue specific polymerase II genes (Emami et al., 1998, Pugh & Tjian, 1990, Philipsen & Suske, 1999). Studies reveal that Sp1 is involved in directly stabilising the binding of TFIID to core promoter elements via interactions with TBP, hTAFII130 and hTAFII55 and can activate transcription through both TATA-containing and TATA-less promoters within the human genome (Emami et al., 1998). In the context of TATA-less promoters, Sp1 specifies the formation of preinitiation complexes involving the general transcription factor TFIID, within a window of 40-100bp downstream of its binding sites at DNA sequences that closely resemble TATA or Inr elements; thereby accounting for the identification of multiple initiation start sites (Carey & Smale, 2000). Roles for Sp1 in processes such as chromatin remodelling and cell cycle regulation have also been demonstrated, highlighting the importance of GC-boxes and the ubiquitously expressed GC box-binding proteins such as Sp1 in maintaining normal cellular function.

In addition to the multiple Sp1 sites identified within the FHL1 promoter, consensus motifs for the muscle specific transcription factor MyoD were also identified. MyoD belongs to the myogenic bHLH family that play an important role in directing the transcription of muscle-specific genes including the muscle creatine kinase and cardiac alpha-actin genes (Weintraub et al., 1990, Biesiada et al., 1999). The formation of a muscle-specific, multi-protein transcriptional complex including MyoD, Sp1, the serum response factor (SRF) and E12 proteins on muscle-specific promoters, highlights the emergence of a common theme in muscle gene regulation, that involves the combinatorial binding of tissue specific factors including members of the myogenic bHLH family, in addition to ubiquitously expressed factors, such as Sp1 (Biesiada et al., 1999). For example, the simultaneous presence of myogenic bHLH, Sp1 and SRF has been shown to be essential for the complex formation and transcriptional activation of human cardiac alpha (HCA) gene (Biesiada et al., 1999). The co-existence of Sp1 and MyoD binding sites on the FHL1 promoter therefore provides a starting point for investigating the specific regulation of FHL1 in skeletal and cardiac muscle cells. Indeed, an in vivo physical interaction between MyoD and Sp1 has already been demonstrated in muscle cells and studies in Drosophila-derived Schneider cells reveal
that Sp1 and MyoD act synergistically to transactivate HCA promoter activity (Sartorelli et al., 1990).

Putative binding sites were also identified for the cardiac-specific NK-Like (NKL) homeobox protein NKX2.5, which contributes to an array of cardiac developmental pathways in mammals (Benson et al., 1999). Given that FHL1 is also expressed in the developing embryonic heart and in adult heart tissue, it is possible that FHL1 may represent a physiological, cardiac-specific target of NKX2.5 (Brown et al., 1999). Intriguingly, NKX2.5 was recently shown to be activated in three paediatric T-ALL cell lines, HPB-ALL, CCRF-CEM and PEER, harbouring the novel cryptic translocation t(5,14)(q35,q32)(Nagel et al., 2003). These studies also demonstrated that within the NKL homeobox gene family, which include 14 genes in total, only NKX2.5, HOX11L2 and HOX11 were activated in a panel of T-ALL cell lines. It is tempting to speculate that these proteins may dysregulate a similar set of oncogenically relevant target genes to generate the T-ALL phenotype, possibly through a domain shared by these family members. In addition it is also possible that these related proteins may cross-regulate tissue-specific target genes, when expressed out of their normal cellular context. This may account for the identification of a NKX2.5-specific target gene (e.g. FHL1) in a screen designed to identify HOX11 targets, following the over-expression of HOX11 in the murine NIH 3T3 fibroblast cell line. Thus, it is plausible that FHL1 may be regulated by NKX2.5 in normal cardiac muscle development in addition to a tumorigenic scenario whereby NKX2.5 is aberrantly expressed in T-cells; or alternatively, FHL1 may represent a true oncogenically relevant target of HOX11, which is also over-expressed in T-ALL malignancies. It would be of interest to test the hypothesis that NKX2.5, HOX11L2 and HOX11 are functionally redundant in terms of regulating the expression of their individual, tissue specific target genes since this might provide a clue regarding the role that FHL1 may play as a target of NK-like homeobox proteins.

A consensus binding site for HAND2, a bHLH transcription factor required for heart development in the chick and mouse (Srivastava et al., 1995) was also identified in the proximal FHL1 promoter. In the mouse, Hand2 is specifically expressed in the right ventricle, and mouse embryos deficient in Hand2 display hypoplasia of the right
ventricle soon after cardiac looping (Srivastava et al., 1997). In the context of FHL1 gene regulation, it is of interest to note that both Hand2 and Nkx2.5, two transcription factors involved in cardiac development, have been shown to cooperatively regulate the ventricular-specific homeodomain protein Irx4 (Yamagishi et al., 2001). It is therefore possible that Nkx2.5/CSX, HAND2 and FHL1 may constitute a transcriptional network involved in differentiation and function of the developing heart. Functional promoter analyses in cardiac cell lines is required in order to determine the importance of these elements in the transcriptional regulation of FHL1, however, and to positively identify trans-acting proteins at these sites.

Ideally, the FHL1 transcriptional start site positions identified in this study should be confirmed by independent mapping methods such as primer extension, S1 nuclease or RNase protection assays. Primer extension involves the elongation of a gene specific 5’ $^{32}$P-labelled primer to the 5’ end of a specific mRNA (Ghosh et al., 1978). By determining the length of the radiolabelled cDNA products generated on a sequencing gel, the distance from the 5’ end of the primer to the transcriptional start site can be estimated with an accuracy of +/- 1 nucleotide (Carey & Smale, 2000). Although this method is relatively simple and efficient, the problem of premature reverse transcriptase termination due to the formation of stable secondary structures at GC rich 5’ ends may still yield incomplete cDNA species. In contrast to 5’ RACE and primer extension analyses, RNase protection is not reliant on reverse transcription to map 5’ cDNA ends and thus provides a truly independent method for identifying transcription initiation sites. This approach involves the generation of an $\alpha$-$^{32}$P[UTP] labelled RNA probe using a plasmid template containing a genomic fragment spanning the predicted transcriptional start site for the gene of interest. The radiolabelled probe is then annealed to the mRNA population to produce thermodynamically stable RNA:RNA hybrid molecules with single stranded RNA overhangs. These hybrids, which contain the putative transcriptional start site, are resistant to subsequent digestion with RNase T1 and RNase A, which cleave the single stranded RNA overhangs such that the length of the protected region can be used to determine the distance of the 5’ end of the probe to the transcriptional start site (Melton et al., 1984). This method is useful for the quantitative analysis of transcripts, however it is still prone to artefacts since the probe can form secondary structures resulting in unwanted RNase-resistant products (Carey &
Smale, 2000). Finally, S1 nuclease analysis can be used to determine the start site of a gene. This approach is similar to RNase protection except that the probe is a single stranded DNA molecule that spans the anticipated start site, as opposed to an RNA molecule and digestion is performed with nuclease that cleaves single stranded RNA and DNA (Berk & Sharp, 1977). This method is less frequently used compared to RNase protection assays, however, since DNA:RNA duplexes are less stable than RNA:RNA hybrids particularly over long AT rich stretches, and aberrant cleavage within these regions may contribute to artefacts (Carey & Smale, 2000). In addition the probe is typically radiolabelled at the 5’ end and not throughout the entire molecule, such that the procedure is less sensitive, although uniformly labelled probes can be generated.

In order to correlate the intracellular expression of FHL1 with functional regulatory elements, and to investigate the ability of HOX11 to regulate the FHL1 promoter in vitro, an initial investigation was undertaken of the FHL1 promoter using the T-ALL cell line PER-117, and the human erytholeukaemic cell line, HEL. Luciferase reporter constructs spanning -297/+73 and -926/+73 of the FHL1 promoter were inserted upstream of the firefly (Photinus pyralis) luciferase cDNA of the pGL3Basic luciferase reporter vector and were transiently transfected into the PER-117 and HEL cell lines. Analysis of the FHL1 promoter revealed that positive cis-regulatory elements may be required for controlling the transcriptional activity of the FHL1 gene in PER-117 and HEL, since the 297bp of sequence flanking the 5’ FHL1 transcriptional start site was active in both the PER-117 and HEL cell lines (Figure 3.11). This region encompasses numerous Sp1 binding sites, which may account for the normal activation of FHL1 in the absence of repressor elements. Extension of this sequence to incorporate 926bp of DNA flanking the 5’ end of Exon 1, however, resulted in a 2-fold decrease in luciferase activity in PER-117, in contrast with the HEL cell line, which showed no decrease in promoter activity. These results suggest the presence of a cell type-specific negative regulatory element located between -297/-926bp, which may function to repress the expression of FHL1 in T-cells. A number of putative regulatory elements conserved between human and mouse species were identified within this region including additional Sp1 sites, as well as binding sites for TCF11, SRY, MOK2 and NKO2.5 and MyoD, however further work is required in order to delineate the specific negative
element involved in silencing \textit{FHL1} expression in PER-117 cells. It is possible that a more distal negative element responsible for the repression of \textit{FHL1} in HEL exists, however, larger promoter constructs were not included in this study.

A primary goal of this Chapter was to determine whether HOX11 can have an effect on \textit{FHL1} promoter activity. This was assessed by cotransfection of the pEFBOSHOX11 expression vector and \textit{FHL1} promoter constructs into the PER-117 and HEL cell lines. Similar to the \textit{ALDH1A1} promoter, overexpression of HOX11 resulted in modest repression of \textit{FHL1} promoter activity in PER-117 cells, whilst the effect in HEL was negligible (Figure 3.12). Although repression by HOX11 was observed in the -297 construct, this effect appeared to be augmented in the -926bp construct, suggesting that HOX11 may act via elements present in both -297 and -926bp constructs. Given the presence of multiple Sp1 binding sites spanning -926/+73bp, one possible mechanism of repression may involve the ability of HOX11 to interfere Sp1-mediated transactivation. The ability of homeodomain proteins to suppress Sp1-dependent transcription has been previously demonstrated for Msx-1, which was shown to interact with TBP, Sp1 and CBP and is thought to mediate autorepression by ‘squelching’ components of the transcriptional machinery and by forming an active repressor complex (Shetty et al., 1999). Experiments performed by Suzuki et al., (2003) also demonstrate that HoxA11 is capable of repressing Sp1-dependent transcription as well as inhibiting HoxA13 enhanced transcription, although the mechanism for this repression is not clearly understood. The effect of the HOX11\textDelta{}H3 mutant on transcriptional repression from the \textit{FHL1} promoter was also assessed in the PER-117 cell line. In keeping with the results in Chapter 2, HOX11\textDelta{}H3 was unable to mediate repression and instead, activated transcription (Figure 3.13). These results support the hypothesis that transcriptional repression by HOX11 occurs via a mechanism requiring the DNA binding helix, and therefore possibly involves DNA binding; whereas transactivation occurs independently of the DNA binding helix, possibly via a mechanism involving protein-protein interactions. The transactivation demonstrated by HOX11\textDelta{}H3 may conceivably result from 1) the loss of an active DNA-binding repressor complex, 2) a dominant negative mechanism whereby HOX11\textDelta{}H3 sequesters repressor cofactors from other regions of the promoter and/or 3) by forming an activation complex. Indeed, the ability of helix 3 mutants of HoxA13 to suppress
transactivation by HoxA13 in a dominant negative manner has been previously documented (Suzuki et al., 2003). Alternatively, repression by HOX11 may be independent of DNA binding, and helix 3 of the homeodomain may represent an interface for the recruitment of corepressors. Thus, upon deletion of this region, HOX11∆H3 is unable to repress transcription. Notably, the repression mediated by HOX11 on the FHL1 promoter, contrasts with upregulation of FHL1 by HOX11 in NIH 3T3 cells and the expression of FHL1 in T-ALL cell lines (including HOX11-expressing cell lines, ALL-SIL and PER-255). Once again, this may reflect the requirement for additional distal elements not included in the FHL1 promoter constructs tested for correct regulation by HOX11 or the need for chromatin-specific conditions.

In summary, the work described in this Chapter suggests that the human FHL1 gene is regulated by a TATA-less promoter that contains several transcriptional initiation start sites within a 108bp window. Functional promoter analyses suggested that the main activity of the FHL1 promoter occurs within the first 300bp from the transcription initiation start site window, which contain multiple sites for the ubiquitously expressed positive trans-acting Sp1. The existence of negative regulatory elements 5’ of this region, is likely to account for the repression of FHL1 in both PER-117 and HEL, and in the case of PER-117, one or more of these elements may exist between the -926 and -297bp region. Finally, HOX11 was capable of repressing the activity of luciferase reporter constructs through elements located in the FHL1 promoter, thus supporting the status of FHL1 as a gene able to be transcriptionally regulated by HOX11.
Chapter 4

Protein Binding Partners of HOX11

4.1 INTRODUCTION

Our current understanding of the molecular events that lead to T-ALL has greatly benefited from the identification of transcription factors that are deregulated via chromosomal translocations involving T-cell receptor loci (reviewed by Hwang & Baer, 1995, Look, 1997). HOX11 is an enigmatic example of a DNA binding protein that contributes to the development of T-ALL (Lichty et al., 1995). The work presented in previous chapters, suggests that HOX11 may not be a representative of a simple activator or repressor of transcription. Through the use of luciferase reporter assays, HOX11 was shown to repress transcription via elements located in the proximal promoters of ALDH1A1 and FHL1 in the T-ALL cell line, PER-117. In contrast, HOX11 activated transcription through the ALDH1A1 promoter in the erythroleukaemic cell line, HEL. These studies also demonstrated that although the DNA binding helix of HOX11 is not required for activation, this region is essential for transcriptional repression, since deletion of the third α-helix switched HOX11 from a transcriptional repressor to an activator.

The means by which HOX11 can modulate gene transcription remains unknown. There is thus a need for more extensive examination of the issue, however our understanding of HOX11 is very much limited by the absence of interacting partner proteins. Transcription factors tend not to operate independently, but rather, are known to mediate their regulatory effects by forming non-covalently associated multi-subunit complexes that bind DNA (reviewed by Wilson, 1996). Therefore, a fruitful way forward to understanding the precise intracellular functions of HOX11, is to identify protein partners with which it interacts. It is known that HOX proteins can interact with various classes of transcription factor to achieve functional specificity in vivo, including other HOX proteins, TALE homeoproteins (e.g. PBX and MEIS), zinc finger proteins (Durocher et al., 1997), and leucine zipper proteins (Kataoka et al., 2001). At the same time, HOX proteins can, directly or indirectly, aid the recruitment of components of the
basal transcriptional machinery and other components that modulate transcriptional events such as histone acetyl transferases and deacetylases (Sur et al., 2000, Choi et al., 1999, Towler et al., 1994).

In previous work, HOX11 was shown to associate with the CCAAT-box binding transcription factor 1 (CTF1) and it was suggested that this leads to an increased proliferative capacity of haematopoietic precursor cells in vitro (Zhang et al., 1999). Previous unpublished work from this laboratory and that of Dr Ursula Kees (Division of CLCR, TVW Telethon Institute for Child Health Research, Western Australia), has suggested that HOX11 may physically interact with MEIS1, MEIS2 and/or TFIIB. MEIS1 and MEIS2 were implicated as HOX11 interactors in a yeast-2-hybrid assay (Milech et al., in preparation), while TFIIB was identified by mass spectrometry in an ALL-SIL cell immunoprecipitate produced with an anti-HOX11 antibody (Heidari et al., in preparation). Human general transcription factor TFIIB is a ubiquitous factor required by RNA polymerase II for transcription initiation through eukaryotic promoters (Sawadogo & Sentenac, 1990). As the second most common target for transcriptional regulators (after TFIID), numerous studies have shown that TFIIB is involved in both gene activation and repression (Blanco et al., 1995, Wang et al. 1996, Breiling et al., 2001). TFIIB is a target in the repression of transcription by the oncoproteins PLZF-RAR, LAZ3/BCL6 and Mad/Mxi-1 (Nagy et al., 1997, Dhordain et al., 1997, Hong et al., 1997). Transcription factor interactions with TFIIB may lock central components the transcription initiation apparatus into a non-functional complex or conformation that is not conducive to transcription (Muscat et al., 1998). Several studies have shown that transcriptional activator proteins can also interact with TFIIB. These include myogenic determination factor, MyoD (Heller & Bengal, 1998), c-myc (McEwan et al., 1996) and c-Jun (Franklin et al., 1995). In support of TFIIB as a potential target for HOX11-mediated transcriptional regulation, Owens et al., (2003) recently showed that HOX11 functionally interacts with the basal transcriptional machinery using artificial reporter assays. This study did not, however, determine which general factor was targeted by HOX11. Alternatively, since CTF1 has been reported to interact with HOX11 (Zhang et al., 1999), and CTF1 itself is known to interact with TFIIB (Kim & Roeder, 1994), the purported interaction of HOX11 with TFIIB may be indirect via CTF1. The preliminary results from our laboratory, suggesting that HOX11
may interact with MEIS proteins and/or TFIIB, are potentially significant, given that these interactions may have functional and clinical significance. They may also shed light on the mechanism(s) by which HOX11 regulates gene expression, including ALDH1A1 and FHL1. An interaction with TFIIB, in particular might help to explain both the work in this thesis and that of Owens et al., (2003), in which HOX11 was implicated in a direct association with the basal transcriptional apparatus. The main aim of the work in this chapter, therefore, was to examine whether MEIS1, MEIS2, TFIIB and HOX11 itself are capable of a direct physical interaction with HOX11. A variety of methods for studying protein interactions exist, including co-immunoprecipitation (Co-IP)(Voncken et al., 1999, O’Neill et al., 2000) and fluorescence resonance energy transfer (FRET)(Cheung et al., 1991, Souerjik & Berg, 2002). In this study, it was decided to focus on confirming direct interactions with HOX11 using the Glutathione-S-Transferase (GST)-pulldown approach. This was due both to the ready availability of recombinant GST-HOX11 (Heidari et al., 2002), as well as the advantage that it eliminates the requirement for a highly specific antibody, which is necessary for Co-IP approaches. GST-pulldown purification has been successfully used to demonstrate interacting partners of homeodomain proteins (Mitchelmore et al., 2000, Yamada et al., 1999, Cvekl et al., 1999) and of HOX11 itself (Zhang et al., 1999), and offers significant advantages in terms of speed, simplicity and flexibility. As the name suggests, it involves the fusion of proteins of interest (‘bait’) to a GST purification tag followed by immobilization on glutathione agarose affinity matrix. The bound fusion protein can then be used to assay the binding of specific ‘test’ proteins radio-labelled with $^{35}$S-methionine by in vitro translation.

4.2 METHODS

4.2.1 GST-Pulldown Assays

4.2.1.1 pCIneo CTF1 and TFIIB Vector Construction

The pCIneo mammalian expression vector (Promega, Genbank®/EMBL Accession number: U47120)(Appendix 2) carries the T7 RNA polymerase promoter, which enables the synthesis of mRNA transcripts in vitro. The pCIneo vector utilised in these
experiments, contained a deletion of sequence within the multiple cloning site (between EcoR I/Acc I).

4.2.1.1.1 Preparation of pCIneo’ Mammalian Expression Vector

pCIneo’ was prepared for subcloning by restriction digestion with Nhe I restriction endonuclease (Promega) in a 1.5ml polypropylene microcentrifuge tube in a final volume of 50µl, containing 10x Nhe I Restriction Digest Buffer [1x; 6mM Tris-HCl (pH 7.5), 50mM NaCl, 6mM MgCl₂, 1mM DTT](Promega), 0.1mg/ml nuclease-free acetylated BSA (Promega), 1µg of pCIneo’ plasmid DNA and 10U of Nhe I restriction enzyme (Promega) at 37°C for 1.5h. Following digestion, restriction endonuclease was heat inactivated by incubation at 65°C for 20min. Linearised 5’ ends were dephosphorylated to prevent vector re-circularisation by incubation with 1U of CIP (Roche) at room temperature for 30min, followed by heat inactivation at 75°C for 10min. An aliquot of the digested pCIneo’ vector (40ng) was electrophoresed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) to confirm complete digestion prior to ligation with insert DNA.

4.2.1.1.2 Generation of ALL-SIL and HEL and cDNA

cDNA for PCR amplification of the CTF1 and TFIIB coding regions was generated using THERMOSCRIPT™ RT-PCR System (Invitrogen) according to the manufacturer’s instructions. 2µg of either ALL-SIL or HEL total RNA was combined with 1µl of Oligo (dT)₂₀ random primer [50µM](Invitrogen) in an RNase-free 0.5ml polypropylene tube in a final volume of 10µl, and incubated at 65°C for 5min. The reaction was then chilled on ice for 5min and reverse transcription was performed at 55°C for 1h in a final volume of 20µl comprising 5x cDNA Synthesis Buffer [1x; 50mM Tris acetate (pH 8.4), 75mM potassium acetate, 8mM magnesium acetate, stabilizer](Invitrogen), 5mM DTT (Invitrogen), 40U of RNASEOUT™ (Invitrogen), 1mM dNTPs (Invitrogen) and 15U of THERMOSCRIPT™ RT (Invitrogen). The reaction was terminated by incubation at 85°C for 5min and RNA was removed by incubation with 2U of RNase H (Invitrogen) at 37°C for 20min. The resultant cDNA was diluted to a final volume of 100µl using sterile dH₂O and stored at -20°C.
For the construction of pCIneo'-CTF1, the 1523bp human CTF1 cDNA was amplified from ALL-SIL cDNA generated in Section 4.2.1.1.2 using Pfu Turbo® DNA Polymerase (Stratagene) according to the manufacturer’s instructions. Reactions were assembled in thin-walled, 0.5ml polypropylene tubes in a final volume of 50µl, comprising 10x Cloned Pfu DNA Polymerase Reaction Buffer [1x; 20mM Tris-HCl (pH 8.8), 10mM KCl, 10mM (NH₄)₂SO₄, 2mM MgSO₄, 0.1% (v/v) Triton®X-100, 0.1mg/ml nuclease-free BSA] (Stratagene), 0.2mM dNTPs (Fisher Biotech), 20pmol each of CTF1NheF/CTF1NheR primers (Table 4.1), 2.5µl of ALL-SIL cDNA and 2.5U of Pfu Turbo® DNA Polymerase (Stratagene). Samples were denatured for 2min at 94°C and then amplified (94°C for 10s, 60°C for 30s and 72°C for 2min) for 10 cycles, with subsequent cycle elongation times increasing by 20s per cycle up to 8:40s, followed by a final extension of 72°C for 7min (EXTEND72, Appendix 1). For the construction of pCIneo'-TFIIB, the 979bp human TFIIB cDNA was amplified from HEL cDNA generated in Section 4.2.1.1.2 in a 50µl reaction comprising 10x Cloned Pfu™ DNA Polymerase Reaction Buffer [1x], 0.2mM dNTPs, 20pmol each of TFIIBForNhe/TFIIBRevNhe primers (Table 4.1), 2.5µl of HEL cDNA and 2.5U of Pfu™ Turbo Polymerase using the TD1 PCR programme (Appendix 1).
4.2.1.1.4 Restriction Endonuclease Digestion of CTF1 and TFIIB Insert cDNA

Full-length CTF1 and TFIIB PCR products generated in Section 4.2.1.1.3 were gel purified using the Qiaex II Gel Purification Kit (Qiagen)(Section 2.2.1.3) and prepared for subcloning by restriction digestion with Nhe I restriction endonuclease (Promega) in a 100µl reaction volume comprising 10x Nhe I Restriction Digest Buffer [1x], 0.1mg/ml nuclease-free acetylated BSA, 50µl of gel purified CTF1/TFIIB DNA fragments and 30U of Nhe I restriction enzyme at 37°C for 18h. Following digestion, restriction endonucleases were heat inactivated by incubation at 65°C for 20min. Digested PCR fragments were subsequently purified from unwanted digestion products using a QIAquick PCR Purification Kit (Qiagen)(Section 2.2.1.4). An aliquot of the purified insert was electrophoresed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) to estimate DNA integrity and concentration prior to ligation.

4.2.1.1.5 Cloning of CTF1 and TFIIB cDNA

Subcloning of PCR amplified CTF1 and TFIIB cDNA into the pCIneo' expression vector was performed essentially as outlined in Section 2.2.1.6. Ligations were performed in 0.5ml polypropylene tubes in a final volume of 20µl, comprising 40ng of linearised pCIneo' vector DNA, 3:1 molar ratio of insert CTF1/TFIIB DNA, 10x T4 DNA Ligase Buffer [1x; 30mM Tris-HCl pH 7.8, 10mM MgCl₂, 10mM DTT, 1mM ATP](Promega) and 6U of T4 DNA Ligase (Promega) and incubated at 15°C for 16-24h. Ligation reactions were used to transform chemically competent DH5α E.coli (Section 2.2.1.8) and plasmid Mini-Preps of selected colonies were prepared for screening using the Sigma GenElute™ Plasmid Miniprep Kit (Sigma)(Section 2.2.1.9). The presence and orientation of CTF1 and TFIIB inserts were determined by restriction digest analysis of a selection of randomly isolated recombinant clones with Nhe I and Bgl II restriction endonucleases (Promega) respectively. Digests were performed in 1.5ml polypropylene microcentrifuge tubes in a 30µl volume comprising 10x Nhe I/Bgl II Restriction Digest Buffer [1x](Promega), 0.1mg/ml nuclease-free acetylated BSA, 0.5µg of plasmid miniprep DNA, 10U of Nhe I/Bgl II restriction enzyme (Promega) at 37°C for 1.5h. Restriction profiles were analysed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) and selected positive clones sequence verified (Section 2.2.1.11). pCIneo'
HOX11, MEIS1 and MEIS2A plasmids were prepared by Darryl D’Souza (Division of CLCR of the TVW Telethon Institute for Child Health Research, Western Australia).

4.2.1.1.6 Purification of Plasmid DNA

Plasmid DNA for pCIneoHOX11, MEIS1, MEIS2A, CTF1 and TFIIB constructs was prepared for downstream in vitro transcription/translation procedures using the QIAGEN Plasmid Midi Kit (Qiagen) as outlined in Section 2.2.1.12 (grey highlight). The recovered DNA was resuspended to a final concentration of 0.5 µg/µl in nuclease-free dH2O and stored at -20°C.

4.2.1.2 Production of ³⁵S-Radiolabelled Proteins

4.2.1.2.1 In vitro Transcription/Translation

HOX11, MEIS1, MEIS2a, CTF1 and TFIIB were [³⁵S]-methionine radiolabelled by in vitro transcription/translation using the TNT® T7 Coupled Reticulocyte Lysate System (Promega) in accordance with the manufacturer’s instructions. Coupled transcription/translation reactions were assembled on ice in RNase-free 0.5ml polypropylene tubes, in a 50µl reaction volume comprising 25µl of TNT® Rabbit Reticulocyte Lysate (Promega), 2µl of TNT® Reaction Buffer (Promega), 1µg of DNA Template, 0.02mM Amino Acid Mixture minus Methionine (Promega), 20µCi of Redivue [³⁵S]-methionine (Amersham), 40U of RNasin® Ribonuclease Inhibitor (Promega) and 1µl of T7 TNT® RNA Polymerase (Promega). Following addition of all the components, the lysate was gently mixed and in vitro transcription/translation of proteins was allowed to proceed for 90min at 30°C. Translated proteins were stored at -20°C until required.

4.2.1.2.2 SDS-PAGE Analysis of ³⁵S-Methionine Labelled In vitro-Translated Proteins

Radiolabelled protein samples (5µl) were prepared for SDS-PAGE analysis by combining with 2x SDS Sample Buffer [1x; 31.25mM Tris (pH 6.8), 12.5% (v/v) glycerol, 1% (w/v) SDS, 0.005% (w/v) bromophenol blue, 2.5% (v/v) β-
mercaptopoethanol] in a 1.5ml screw-cap polypropylene tube and heating at 90-100°C for 5min to denature the proteins. Protein samples (10µl) were loaded adjacent to Prestained Range MWMs (5µl)(Bio-Rad Laboratories) and resolved by SDS-PAGE as described in Section 2.2.3.3.3. The gel was removed and incubated in Fixing Solution [50% (v/v) methanol, 10% (v/v) glacial acetic acid] with gentle agitation for 30min at room temperature, followed by soaking in Gel Drying Solution [7% (v/v) methanol, 7% (v/v) glacial acetic acid] for 5min at room temperature, to prevent the gel from cracking during the drying process. The gel was then placed on a sheet of Whatman® 3MM filter paper, covered with Saran wrap and dried under vacuum at 80°C for 1h using a Bio-Rad Model 583 Gel Dryer (Bio-Rad Laboratories). Autoradiography was performed using an intensifying screen (Fuji G-8 EC-AWU Cassette, 27.9x35.6cm; Fuji Medical X-ray film super HRG30) at -70°C for 18-36h. Prior to developing, cassettes were allowed to thaw for 1h at room temperature and films were developed using an SPM 300 developer (Fuji).

4.2.1.3 Production of Recombinant GST-HOX11 and GST-HOX11ΔH3 Fusion Proteins

4.2.1.3.1 Preparation of Glutathione Sepharose 4B

Glutathione Sepharose 4B (Pharmacia Biotech) is provided as bulk matrix (75% slurry) for batch purification purposes. Prior to affinity purification, a 50% slurry was prepared according to the manufacturer’s instructions. Briefly, the Glutathione Sepharose 4B matrix was resuspended by gentle shaking and a 665µl aliquot (1.33ml 75% slurry/ml of bed volume required) transferred to a 10ml polypropylene tube. The matrix was sedimented by centrifugation (500 x g, 5min, room temperature)(CentraGP8R, IEC) and washed with 5ml of ice-cold PBS (10ml PBS/1.33ml original slurry) by thorough inversion. The beads were collected by centrifugation (500 x g, 5min, room temperature) and resuspended in 500µl of ice-cold PBS to yield a 50% slurry. Equilibrated beads were stored at 4°C and remained functional for up to 1 month.
4.2.1.3.2 Expression and Purification of Recombinant GST, GSTHOX11 and GSTHOX11∆H3

Parent and recombinant expression vectors (pGEX, pGEX-6P-1/HOX11 and pGEX-6P-1/HOX11∆H3) were obtained from Mansour Heidari (Murdoch University, Perth Western Australia). Escherichia coli strain ER2566 was transformed with the expression constructs pGEX, pGEX-6P-1/HOX11 or pGEX-6P-1/HOX11∆H3 and cultured for 15-16h at 37°C in 2YT media containing 100µg/ml ampicillin with vigorous shaking (~225rpm). Starter cultures were subsequently diluted 1:100 in selective 2YT media (2ml starter culture in 198ml of selective 2YT media) and grown at 37°C to mid to late log phase (OD₆₀₀ ~1.35). Fusion gene expression was induced by adding 0.1mM IPTG to the cultures, and incubation was continued for an additional 1.5h at 37°C. The cells were then collected by centrifugation (7 700 x g, 10min, 4°C)(AvantiTM J-25I Centrifuge, Beckman, JA-25-50) and resuspended in 10ml of ice-cold PBS. Prior to cell lysis the cell suspension was incubated in with 0.1mg/ml of lysozyme, 4µg/ml of DNase1, 1mM PMSF and 5mM DTT on ice for 5min. The cells were then lysed by freeze-thaw (3 cycles of liquid N₂, 37°C H₂O bath) followed by ultrasonication on ice (4 cycles of 30s, 1min rest period between each cycle). 1% (v/v) Triton®X-100 was added to aid the solubilisation of protein and the suspension was mixed gently and incubated on ice for 15min. Bacterial debris was then cleared by centrifugation (12,000 x g, 10min, 4°C) and the supernatants collected. For purification of the GST, GST-HOX11 and GST-HOX11∆H3 proteins, clarified bacterial lysates (~10ml) were applied to 200µl of prepared 50% glutathione sepharose 4B slurry (Section 4.2.1.3.1) and the matrix was incubated with gentle rotation at 4°C for 2h. This initial binding step aimed to remove background E.coli host proteins from the bacterial supernatants. The beads were then collected by centrifugation (500 x g, 5min, 4°C)(CentraGP8R, IEC) and the supernatants re-applied to 200µl of fresh 50% glutathione sepharose 4B slurry and incubated with gentle rotation at 4°C for an additional 18-20h. The matrix was then washed (x3) with 1ml of Wash Buffer [50mM Tris-HCl (pH 8.0), 100mM NaCl, 10% (v/v) glycerol, 1mM PMSF, 0.05mM β-Mercaptoethanol]. For GST-pulldown assays which required fusion protein bound to glutathione sepharose 4B beads, the matrix was resuspended in 200µl of Wash Buffer and stored at 4°C for up to 2 weeks. In cases where purified GST fusion proteins were
required, the matrix was washed (x3) with cold PBS and the recombinant protein eluted by the addition of 500µl (1 volume) of Glutathione Elution Buffer [50mM Tris-HCl (pH 8.0), 10mM reduced glutathione](Pharmacia Biotech) to the matrix and incubation at 4°C for 10min. The beads were then collected by centrifugation (500 x g, 5min, 4°C) and the supernatant removed to a fresh tube on ice. The elution steps were repeated twice more and the supernatants were pooled and stored as aliquots in 0.5ml polypropylene tubes at -80°C.

4.2.1.3.3 Quantitative and Qualitative Analysis of GST Fusion Proteins

The concentration of purified GST, GST-HOX11 and GST-HOX11∆H3 was determined by measurement of UV light absorbance of diluted samples. Protein samples in solution were diluted 1/80 to a final volume of 160µl in dH2O. The spectrophotometer (Shimadzu UV Mini 1240 Spectrophotometer) was blanked with a 160µl aliquot of the sample diluent. Ultraviolet light absorbance of the diluted protein sample was then determined at 280nm wavelengths, and protein concentration was determined using the calculations described below;

<table>
<thead>
<tr>
<th>GST Fusion Protein Concentration [mg/ml]: O.D_{280nm} x dilution factor x 0.5</th>
</tr>
</thead>
</table>

where (1 O.D unit = 0.5mg/ml of GST)

Fusion protein integrity was confirmed by SDS-PAGE electrophoresis as described in Section 4.2.1.2.2. Samples (~1µg purified GST fusion protein) were combined with 2x SDS Sample Buffer in a 1.5ml screwcap polypropylene tube and heated at 90-100°C for 5min prior to electrophoresis on a 12.5% SDS polyacrylamide gel at 160V for 1h. Proteins were visualised by staining with Coomassie Blue Dye [0.25% (w/v) Coomassie Brilliant Blue R-250, 45% (v/v) methanol, 10% (v/v) glacial acetic acid] at room temperature for 2h followed by destaining [40% (v/v) methanol, 10% (v/v) glacial acetic acid] for 4h at room temperature. Gel images were taken using the GS-710 Calibrated Imaging Densitometer and associated software (Bio-Rad Laboratories).
In order to perform GST-pulldown experiments with a known quantity of ‘bait’ protein, the amount of GST fusion protein retained by the glutathione sepharose 4B beads was visually estimated by SDS-PAGE. An aliquot of the bound beads was diluted with fresh glutathione sepharose (GST [1:40], GST- HOX11 [1:10], GST- HOX11ΔH3 [1:7]) and 20µl of this bead mix was transferred to a 1.5ml screwcap polypropylene tube and prepared for SDS-PAGE by combining with 10µl of 2x SDS-PAGE loading buffer and heating for 5min at 90-100°C. Samples were centrifuged (13,000 x g, 1min, room temperature) to collect the beads and the supernatants analysed by SDS-PAGE as described in Section 4.2.1.2.2. The amount of protein in a 20µl aliquot of mixed beads was determined by comparison of the intensity of staining between samples and the Bio-Rad MWM of known concentration. In addition, samples were also compared to known quantities of purified GST, GST-HOX11 and GST-HOX11ΔH3 proteins produced and quantitated in sections 4.2.1.3.2/4.2.1.3.3.

4.2.1.4 Verification of the Biological Activity of Recombinant HOX11

4.2.1.4.1 Electromobility Shift Assay (EMSA)

In order test the DNA binding capability of rHOX11 produced in Section 4.2.1.3.2, EMSA was performed using a probe containing HOX11 consensus binding sites (HOXCon)(Table 4.1). Single stranded oligonucleotide probes (HOXCon) were [γ-32P]ATP labelled and annealed as described in Section 2.2.5.1. For EMSA, 1.5µg of purified GST, GST-HOX11 or GST-HOX11ΔH3 was combined in a 1.5ml polypropylene microcentrifuge tube in a final volume of 15µl with radio-labelled, double stranded HOXCon probe (~3.2ng) in 1x Binding Buffer [20mM HEPES (pH 7.6), 50mM KCl, 1mM EDTA, 1mM DTT, 5% (v/v) glycerol] with 0.5µg of polydI-dC (ICN Biomedicals). The specificity of protein-DNA complexes identified by EMSA was assessed using competition assays, which involved the addition of a 33-fold molar excess of unlabelled specific competitor (HOXCon probe) or non-specific competitor (randomised HOXCon probe)(Table 4.1) to the binding reactions. Samples were incubated at 4°C for 15min and complexes were resolved by non-denaturing polyacrylamide gel electrophoresis in 0.5x TBE Buffer (17mA/gel for 2h) on a 4% non-
denaturing polyacrylamide (TBE) gel (Section 2.2.5.3). Gels were dried and developed as outlined in Section 2.2.5.3.

4.2.1.5 GST-Pulldown Assay

4.2.1.5.1 Binding Reactions

For the in vitro binding assay, 15µg of GST, GST-HOX11 or GST-HOX11ΔH3 bound to glutathione sepharose beads (in a final bead volume of 20µl) was combined in a 1.5ml screwcap polypropylene tube with 5µl of radiolabelled protein in 500µl Binding Buffer [150mM NaCl, 50mM Tris-HCl (pH 7.5), 0.5mM EDTA, 0.5% (v/v) NP-40, 1mM DTT, 0.5mM PMSF and 0.05% (w/v) BSA] for 18-20h at 4°C with continuous rotation. Following incubation, the beads were pelleted (12,000 x g, 2min, 4°C) (Sigma 1-15 centrifuge) and washed extensively (x3) with 500µl of ice-cold Binding Buffer. To detect the interaction of radiolabelled proteins with GST, GST-HOX11 and GST-HOX11ΔH3, the bead pellet was resuspended in 10µl of 2x SDS-PAGE loading buffer and heated at 90-100°C for 5min to elute the retained protein. The beads were collected by centrifugation (12,000 x g, 2min, 4°C) and the resulting supernatant was analysed by SDS-PAGE (Section 4.2.1.2.2). Gels were prepared for drying and developed as described in Section 4.2.1.2.2.

4.3 RESULTS

GST-pulldown purification is based on the principle of affinity chromatography and has proven to be a relatively simple, yet powerful tool for identifying interacting protein partners and delineating specific interaction motifs. This technique is often used to complement other methods for detecting protein-protein interactions, including yeast two-hybrid, Co-IP and electrophoretic mobility shift assays (EMSA). Proteins of interest (‘bait’) are fused to a GST purification tag and immobilised on glutathione agarose affinity matrix, which can then be used to assay the binding of specific ‘test’ proteins radio-labelled with [35S]-methionine by in vitro translation. This strategy was employed to further investigate interactions between HOX11 and selected putative partners, namely MEIS1, MEIS2 and TFIIB, in addition to HOX11 itself. Interaction of
these proteins with the HOX11 mutant, HOX11ΔH3, were also examined in order to determine whether helix 3 of the homeodomain was important to any interaction observed.

### 4.3.1 Production of Recombinant GST Fusion Proteins

Cloning of HOX11 and HOX11ΔH3 cDNAs into the pGEX6P-1 vector (Amersham Pharmacia Biotech) facilitated the expression of *Schistosoma japonicum* GST (26kDa) fused to the cDNA in a *lacIq*-controlled Tac promoter. The optimal conditions for the expression and purification of soluble, full-length GST-HOX11 fusion proteins were previously determined by Heidari *et al.*, (2002). Specifically, optimal growth conditions involved growing bacterial cultures to mid-late log phase (OD600 of 1.35), prior to inducing protein expression with the galactose analog, IPTG for 1h only. The GST-HOX11/HOX11ΔH3 fusion proteins were then purified from clarified bacterial supernatants by glutathione sepharose affinity chromatography at 4°C to minimise protein degradation (Heidari *et al.*, 2002). Under these conditions, relatively pure, full-length GST-HOX11 and GST-HOX11ΔH3 fusion proteins migrating at 60kDa and 59kDa respectively were produced at a final yield of approximately 2mg/L of bacterial culture (Figure 4.1).

### 4.3.2 Verification of the Biological Activity of Recombinant HOX11 Fusion Proteins

The functionality of the recombinant HOX11 fusion proteins (rHOX11) was assessed by demonstrating the DNA binding capabilities of GST-HOX11 compared to the DNA binding mutant GST-HOX11ΔH3, by EMSA (Figure 4.2). EMSA was performed by incubating rHOX11 with a radiolabelled, double-stranded DNA probe harbouring a core consensus HOX11 binding site ‘TAAG’ (HOXCon) previously identified by Tang and Breitman (1995). As predicted, rHOX11 but not rHOX11ΔH3 bound strongly to the DNA probe, subsequently confirmed by competition assays using an excess of unlabelled HOXCon probe, which resulted in the significant diminution of the shifted complex, compared to an excess of unlabelled non-specific probe (HOXNsc, a scrambled version of the TAAG core sequence) which did not compete, and in fact
Figure 4.1. Purification of Recombinant GST Fusion Proteins by Affinity Chromatography Using Glutathione Sepharose. *E. coli* strain ER2566 was transformed with either pGEX-6P-1, pGEX-6P-1/HOX11 or pGEX-6P-1/HOX11ΔH3 and following IPTG induction, resulted in the expression of GST (26kDa), GST-HOX11 (60kDa) or GST-HOX11ΔH3 (59kDa) fusion proteins. Crude cell lysates (15µl) and purified proteins (~200ng) were analysed on an 12% SDS-PAGE gel using Broad Range MWMs (~500ng/band)(Bio-Rad Laboratories) and stained with Coomassie brilliant blue R-250.
Figure 4.2. Verification of HOX11 Biological Activity by Electrophoretic Mobility Shift Analysis (EMSA). GST, GST-HOX11 or GST-HOX11ΔH3 fusion proteins (1.5µg) were incubated with a $^{32}$P-labelled double stranded TAAG core DNA probe (HOXCon, 3.2ng) that corresponds to a HOX11 consensus sequence. GST-HOX11 but not GST-HOX11ΔH3 or GST alone specifically bound to the probe resulting in the formation of a single retarded band (arrow). Binding of GST-HOX11 was competed for by addition of a 33-fold molar excess of unlabelled probe (S, HOXCon) but not non-specific competitor (NS, HOXNsc).
Figure 4.3. Quantitative Analysis of GST Fusion Proteins Utilised in GST-Pulldown Assays. The amount of GST fusion protein bound to Glutathione Sepharose beads was estimated by analysis of a 20µl (containing ~15µg protein) aliquot of protein-bound beads (diluted with fresh Glutathione sepharose 4B; GST 1:20, GST-HOX11 1:5, GST-HOX11ΔH3 1:2.5) on a 12% SDS-PAGE gel stained with Coomassie Blue R-250. Protein Broad Range SDS-PAGE MWMs (Bio-Rad Laboratories) containing 9 bands of approximately 500ng protein each, were used to visually verify the quantity of protein utilised in GST-Pulldown assays.
enhanced, protein binding (Figure 4.2).

4.3.3 *In vitro* Characterisation of MEIS1, MEIS2A, HOX11 and TFIIB Binding to HOX11

The aim of this Chapter was to determine whether MEIS1, MEIS2A, TFIIB and HOX11 itself, could physically interact with HOX11 *in vitro*. MEIS1 and MEIS2A were identified in yeast-2-hybrid screens as a potential HOX11 binding partner (Milech *et al.*, in preparation), while TFIIB had been implicated as a HOX11 interactor by a co-immunoprecipitation/MALDI-TOF mass spectrometry approach (Heidari *et al.*, in preparation). To assess these interactions, glutathione sepharose coupled with approximately 15 µg of either GST, GST-HOX11 or GST-HOX11ΔH3 (Figure 4.3) was incubated under stringent conditions with [35S]-labelled *in vitro* synthesised MEIS1, MEIS2A, TFIIB and HOX11 proteins. The TnT® Coupled Reticulocyte lysate System (Promega) utilised for the generation of [35S]-methionine radiolabelled test proteins, employs the use of a circular DNA template in which the gene-specific cDNA is cloned downstream of the T7 promoter within the pCIneo mammalian expression vector (Promega). This template may then be used to generate [35S]-methionine radiolabelled proteins in a one-step process, involving the synthesis of mRNAs in the presence of nucleoside triphosphates (NTPs) and *in vitro* rabbit reticulocyte lysate translation. In order to rule out the possibility of non-specific protein binding to the GST moiety of the ‘bait’ rHOX11 fusion, glutathione sepharose coupled with GST alone (Figure 4.3) served as a negative control. CTF1 was employed as a positive control since HOX11 and CTF1 had previously been shown to physically associate in both *in vitro* GST pull-down assays and Co-IP studies (Zhang *et al.*, 1999).

Specific retention of CTF1, MEIS1, MEIS2A, TFIIB and HOX11 was observed with GST-HOX11 but not with control GST beads, which suggested that all four proteins are capable of physically interacting with HOX11 *in vitro* (Figure 4.4, 4.5). Notably, there was a considerable difference in the binding efficiencies of these proteins, with CTF1 and HOX11 demonstrating relatively strong binding to GST-HOX11 compared to the significantly weaker MEIS1, MEIS2A and TFIIB interactions (Figure 4.4, 4.5). The finding that HOX11 is capable of homodimerisation was completely novel. Based on
evidence that residues within the homeodomain may be important for homodimeric HOX protein interactions (Kasahara et al., 2001), we subsequently investigated whether the helix 3 region of the HOX11 homeodomain crucial for DNA binding, was also required for the physical association of HOX11 with itself. A comparison of the interaction between GST-HOX11, GST-HOX11ΔH3 (which lacks the DNA recognition helix of the homeodomain) and in vitro translated HOX11, revealed no detectable change in binding (Figure 4.5). This suggested that the main region of HOX11 involved in DNA recognition is distinct from those region(s) involved in homodimerisation. Given evidence by Owens et al., (2003), as well as the work presented in Chapter 2 and 3 of this thesis, that HOX11 may regulate gene transcription via interactions with components of the basal transcriptional machinery, in a DNA independent manner, the effect of this deletion on the interaction observed between HOX11 and TFIIB was also assessed (Figure 4.5). This revealed that the interaction of HOX11 with TFIIB was significantly diminished upon deletion of the homeodomain helix 3. These results suggested that this region may be specifically involved in mediating protein interactions with TFIIB or alternatively, that this specific interaction requires a structurally intact homeodomain.
Figure 4.4. GST-Pulldown Assay to Assess the Interaction of HOX11 with Putative Protein Partners. Radiolabelled $^{35}$S-CTF1, $^{35}$S-MEIS2A, $^{35}$S-MEIS1 and $^{35}$S-HOX11 proteins were produced using the TnT® T7 Coupled Reticulocyte Lysate System (Promega). 5µl of translation mix was incubated with 20µl of glutathione sepharose resin (50:50 slurry) containing approximately 15µg of immobilised GST fusion protein at 4°C overnight in Binding Buffer. The resin was washed and analysed by SDS-PAGE followed by autoradiography. (A) 50% of the input ‘test’ proteins used in in vitro binding assays. Arrowheads indicate sizes of test proteins as determined using Prestained Broad Range SDS PAGE MWMs (Bio-Rad Laboratories). (B) Interaction of GST-HOX11 with $^{35}$S-radiolabelled CTF1, MEIS2A, MEIS1 and HOX11 (Lanes 5-8). GST alone (Lanes 1-4) was used as a negative control.
Figure 4.5. GST-Pulldown Assay to Assess the Interaction of HOX11 and HOX11ΔH3 with TFIIB and HOX11. Radiolabelled $^{35}\text{S}$-TFIIB and $^{35}\text{S}$-HOX11 proteins were produced using the TnT® T7 Coupled Reticulocyte Lysate System (Promega). 5µl of translation mix was incubated with 20µl of glutathione sepharose resin (50:50 slurry) containing immobilised GST fusion protein (≈15µg) at 4°C overnight in Binding Buffer. The resin was washed and analysed by SDS-PAGE followed by autoradiography. (A) 50% of the input test proteins used in in vitro binding assays. Arrowheads indicate sizes of test proteins as determined using Prestained Broad Range SDS-PAGE MWMs (Bio-Rad Laboratories). (B) Interaction of GST-HOX11/HOX11ΔH3 with $^{35}\text{S}$-methionine radiolabelled TFIIB and HOX11. GST alone was used as a negative control.
4.4 DISCUSSION

The transforming function of HOX11 is likely to involve collaboration with protein partners and cross-talk with members of the general transcriptional machinery in order to control the expression of target genes regulating processes of cellular proliferation and/or differentiation. The purpose of this Chapter was to assess the potential of HOX11 to physically associate with selected proteins identified previously as potential cofactors, in order to shed light on the transcriptional role of HOX11. In the first instance, both MEIS1 and MEIS2A, members of the TALE superfamily, were confirmed to directly interact with recombinant HOX11 by GST-pulldown, although the strength of these interactions was relatively weak in comparison to those observed for the positive control interaction between CTF1 and HOX11 (Figure 4.4). In regard to MEIS2, the MEIS2A isoform was selected for use in GST-pulldown assays, since this isoform includes all exons (1-4) and is therefore most likely to contain regions involved in mediating protein-protein interactions. Although both MEIS1 and MEIS2A were shown to interact with HOX11, MEIS2A was considered to be of greater interest, since not only did it appear to bind to HOX11 with greater efficiency than MEIS1, but it was also shown to be coexpressed with HOX11 in several T-ALL cell lines and primary T-ALL specimens (D’Souza, Kees and Greene, unpublished observations), and as such may represent a potentially leukaemic isoform.

The possible relevance of this interaction in terms of HOX11-mediated leukaemic transformation is underscored by a number of studies that have revealed intriguing links between Meis1 and Meis2 overexpression and leukaemogenesis. For example, Meis 1 expression is activated by proviral integration in ~10-15% of BHX-2 myeloid leukaemias (Moskow et al., 1995, Nakamura et al., 1996). Moreover, expression studies of Meis family members in BHX-2 myeloid leukaemia cell lines revealed exclusive expression of either Meis1 (due to retroviral integration) or Meis2 and Meis3, suggesting that overexpression of at least one family member may be required for leukaemic transformation (Nakamura et al., 1996). Although Meis2 and Meis3 expression was not correlated with retroviral integration, mouse mapping studies suggest that the human homologs of Meis2 and Meis3 map to human chromosomes 19q13 and 15q11 respectively, regions harbouring chromosomal abnormalities.
associated with human myeloid leukaemia, highlighting the possibility that deregulation of MEIS family members may be a crucial event in leukaemic transformation (Nakamura et al., 1996, Mitelman, 1994). Meis2 is also coactivated with Hoxb8 in the WEHI-3B murine myelomonocytic cell line, and enforced expression of Hox or Meis genes in 32Dcl3 cells inhibits myeloid cell differentiation, providing further evidence that aberrant Meis2 expression may be a contributing factor in leukaemic transformation (Fujino et al., 2001). More recently, the involvement of Meis1 in leukaemogenesis was demonstrated following the identification of Meis1 and Hoxa9 as molecular targets of the oncogenic mixed lineage leukaemia-eleven nineteen leukaemia (MLL-ENL) fusion protein, which is associated with lymphoid and myeloid leukaemias (Zeisig et al., 2004). In addition, MEIS1 and MEIS2 genes were also found to be highly expressed in a panel of neuroblastoma cell lines, suggesting a that the oncogenic potential of MEIS family members may not be restricted to cells of the haematopoietic lineage (Geerts et al., 2003).

The relatively weak binding observed between HOX11 and MEIS1/2A might be explained by the requirement for additional cofactors or specific binding conditions that may stabilise the interaction between HOX11 and MEIS2. Alternatively, DNA binding may be required for transcriptional complex formation. Indeed, it has been demonstrated in vitro that protein-protein interactions between Meis and Pbx can facilitate the formation of Hox-Pbx-Meis trimeric complexes bound to DNA (Jacobs et al., 1999). PBX2 represents a potential candidate for a third binding partner since PBX2 transcripts have been demonstrated to be exclusively expressed in two human T-ALL cell lines harbouring the HOX11 10q24 translocation (Allen et al., 2000). In contrast, other HOX11 overexpressing murine cell lines contained transcripts of two or even three PBX genes, suggesting that PBX2 may be specifically involved in HOX11-induced leukaemic transformation. This is supported by the observation that overexpression of both HOX11 and PBX2 led to a reduction in contact-dependent growth inhibition in NIH 3T3 cells (Allen et al., 2000). Under the ‘co-selective binding’ model, a number of scenarios involving functional interplay between HOX11, MEIS2 and PBX2 transcription factors might therefore be envisaged. HOX11 may form novel, dimeric or trimeric complexes with MEIS2 and PBX2, resulting in the activation or repression of specific target genes harbouring PBX2-HOX11 or PBX2-HOX11-MEIS2
compound binding sites via the recruitment of coactivators (e.g. CBP) or corepressors (e.g. NcoR, SMRT) (Asahara et al., 1999) (Figure 4.6A, B).

Based on murine models, which have been unable to convincingly recapitulate the tumorigenic activity of HOX11 in T-cells, it is evident that additional genetic lesions are required for the appearance of overt disease. Such secondary events may result in the activation of cofactors required for HOX11-mediated transcription. It cannot be discounted, however, that the oncogenic potential of HOX11 stems from it acting alone or as a homodimer/multimer, and that these additional mutations may involve the dysregulation of tumour suppressor genes or other genes not directly related to the transforming activity of HOX11 (Figure 4.6C). Indeed, although the majority of studies focus on HOX/PBX/MEIS interactions in determining target gene selectivity of HOX proteins, there is evidence to suggest that Hox proteins may act through binding to multiple monomer sites via a ‘widespread binding’ model (Biggin & McGinnis, 1997). Moreover, a number of Hox-regulated structures in Drosophila, in particular the distal appendages, do not require Exd or Hth activity for normal development, suggesting that target gene selectivity and activity by HOX proteins may occur via distinct pathways involving novel, unidentified cofactors or by HOX proteins alone (Azpiazu & Morata, 2000). The finding in this study, that HOX11 is capable of strong self-association, raises the possibility that HOX11 may utilise alternative regulatory mechanisms, in addition to those involving co-selective binding through interactions with PBX and MEIS cofactors, in order to regulate gene transcription. It has been reported that the homodimerisation of homeodomain transcription factors can increase the protein-DNA binding affinity for palindromic DNA sites, and add to the complexity of combinatorial gene regulation in a given cellular environment (Wilson et al., 1993). The role of homodimers in transcriptional repression is particularly well characterised for the Drosophila zinc finger transcription factor, Kruppel, which acts as transcriptional activator in monomeric form and at higher concentrations switches to a homodimeric repressor, and is now becoming an emerging theme in the field of transcriptional regulation. For example the mammalian basic helix loop helix factor, Stra 13 actively represses transcription by binding as homodimers to E-box elements (St-Pierre et al., 2002), while the homeobox transcription factor Hesx1 acts as a short range monomeric repressor, in addition to passively repressing transcription by blocking specific
FIGURE 4.6
transactivation sites as a homodimer (Quirk & Brown, 2002). Indeed, there is evidence to suggest that other HOX proteins, particularly those belonging to the NKL/POU cluster, of which HOX11 is a member, including Nkx2.5 (Kasahara et al., 2001), Cdx2 (Suh et al., 1994), Pit1 (Jacobson et al., 1997) and Oct1 (Poellinger et al., 1989), form functional homodimers. Cooperative dimerisation of homeodomain proteins involves residues located within the homeodomain as well as regions located outside this motif (Jacobson et al., 1997, Kasahara et al., 2001). Studies have shown that the presence of Arg^{28} or Arg^{43} conserved in 50% of homeodomain proteins, prevents dimerisation (Wilson et al., 1993). Intriguingly, HOX11 contains Ser^{28} and Ala^{43}, revealing that it belongs to a class of homeodomain proteins capable of cooperative dimerisation.

The finding by Heidari et al. (in preparation) that TFIIB was involved in a protein complex immunoprecipitated by HOX11 antibody, suggested that HOX11 may interact with TFIIB, a member of the basal transcriptional machinery (Figure 4.6D). TFIIB functions in the correct positioning of RNA polymerase II at the promoter DNA, and is a particularly intriguing interactor given the recent finding that transcriptional repression mediated by HOX11 may involve interactions with basal transcription factors (Owens et al., 2003). This was deduced following experiments demonstrating that HOX11 was able to repress transcription from the SV40 early promoter, despite deletion of various cis-regulatory elements (Owens et al., 2003). Further, S-tagged recombinant HOX11 repressed basal transcription from the adenovirus major late promoter, driven by TFIIIS affinity purified RNA polymerase II holoenzyme but not by Gal4-VP16. Since this promoter does not contain any DNA binding sites for HOX11, these results suggested that HOX11 mediates repression via interaction with one or more components of the RNA polymerase II holoenzyme (Owens et al., 2003). The GST-pulldown analyses described in this chapter, suggested a weak interaction between HOX11 and TFIIB, which may indicate sub-optimal binding conditions required for efficient binding. Alternatively, the stable interaction of HOX11 with TFIIB may require additional components of the basal transcriptional machinery or even other transcription factors. Interestingly, HOX11 has also been shown to interact with the CCAAT box binding transcription factor (CTF1), a ubiquitously expressed transcription factor which is capable of recruiting TFIIB during preinitiation complex assembly (Zhang et al., 1999). Since HOX11 and TFIIB were shown to bind distinct interaction
surfaces on CTF1, it is possible that the immunoprecipitation of TFIIB by HOX11 may have required CTF1 as a bridging factor. This hypothesis could be tested in the future by performing GST-pulldown experiments using GST-HOX11, in vitro translated $^{35}$S-radiolabelled TFIIB and non-labelled CTF1. This interaction may have important functional consequences for the cell since inhibition of CTF1 affects the proliferative capacity of HOX11-immortalised hematopoietic precursor cells, suggesting that HOX11 mediated transformation may require cooperative binding with CTF1 and possibly TFIIB (Zhang et al., 1999).

The radically different behaviour of the HOX11ΔH3 mutant in the transcriptional assays outlined in Chapters 2 and 3 prompted its inclusion in the GST-pulldown experiments to determine whether the third helix of the HOX11 homeodomain is required for homodimerisation. The interaction of in vitro translated HOX11 did not appear to be affected by a deletion of the DNA recognition helix of the homeodomain (GST-HOX11ΔH3), suggesting that the ability of HOX11 to form homodimers may involve interaction domains located outside of this region. This is consistent with the formation of other NKL/POU homodimers, which require residues outside of helix 3 at the C-terminal end of the homeodomain. Nkx2.5, for example, requires the basic amino acids Lys$^{57}$-Arg$^{58}$ for dimerisation (Kasahara et al., 2001), which are present as Arg$^{57}$-Arg$^{58}$ in HOX11. A similarly located motif, Arg$^{58}$-Val$^{59}$-Lys$^{60}$, is required for Pit1 dimerisation on DNA (Jacobson et al., 1997).

Intriguingly, the strength of the interaction of TFIIB with HOX11 was greatly diminished upon deletion of the third helix of the homeodomain, suggesting that residues within this region are important for interactions with this member of the basal transcriptional machinery. Alternatively, deletion of helix 3 may alter the overall structure of the homeodomain, which may account for the reduced binding of TFIIB. It is likely that additional regions outside the helix 3 of the homeodomain are also involved in TFIIB binding, however, since the interaction was not completely abolished in the HOX11ΔH3 mutant. The results in Chapters 2 and 3 suggested that the transcriptional repression mediated by HOX11 is dependent on helix 3, since the HOX11ΔH3 mutant was unable to effectively repress transcription from the ALDH1A1 and FHL1 promoters in the PER-117, and instead resulted in a strong activation. Given
that the third α-helix is known to be required for DNA binding, these results initially suggested that repression was dependent on DNA binding by HOX11. However, in light of these results it is also formally possible that repression could occur via a DNA-binding independent mechanism and instead, is dependent upon interactions with TFIIB and amino acids in the third helix of the HOX11 homeodomain. Indeed, there is precedent for a mechanism whereby homeodomain proteins interact with members of the basal transcriptional machinery to repress transcription. For example, the Drosophila homeodomain protein Even-skipped (Eve) represses transcription by directly interacting with TBP to block TFIID binding to promoter sequences, and this interaction is dependent on residues within the Eve homeodomain (Li & Manley et al., 1998). Moreover, Msx-1-mediated transcriptional repression requires interactions between residues within the N-terminal arm of the homeodomain and TBP (Zhang et al., 1996).

Although the GST-pulldown approach is appealing as it is relatively simple to perform and is cost-effective for assessing multiple protein interactions, since it does not require specific antibodies, the limitations of this technique render it suitable for preliminary analysis of protein interactions, and verification using alternative methods for assessing protein-protein interactions is necessary. In particular, the use of informative negative controls is crucial in order to avoid misinterpretation of results. Glutathione agarose beads bound to GST alone are often used as a negative control for non-specific retention of test proteins, however this control does not account for mass action effects, in which non-specific interactions may occur if the ‘bait’ fusion protein is at a high enough concentration. Optimally, GST-pulldown assays should be performed using the wild type bait protein and a variant of the bait protein, which is known not to bind the test protein in vivo. Alternatively, a panel of mutant bait proteins harbouring deletions/substitutions of putative interaction motifs, such as the HOX11ΔH3 mutant employed in these studies, can be used to confirm that positive interactions obtained using the wild type bait protein are due to the presence of specific protein domains and not due to artefacts. Future studies may involve testing the effect of deleting specific regions or interaction motifs on the interactive capacity of HOX11 with selected proteins, including (1) deletion of the NH₂-terminal 50 amino acids, which have been shown to be crucial for activation function of HOX11 (Masson et al., 1998) (2) the
glutamine-rich region at the COOH-terminus (3) the PBX Interaction Motif (PIM)(FPWME), a conserved motif located upstream of the homeodomain that facilitates interactions with other homeodomain proteins such as PBX (Piper et al., 1999, Sprules et al., 2003), or (4) the FIL motif which is similar to a previously described binding domain for members of the Groucho/transducin-like Enhancer of split (Gro/TLE) co-repressor family, and which are capable of repressing transcription via the recruitment of HDACs (Yao et al., 2001).

In summary, the physical association of HOX11 with selected putative interactors, namely CTF1, MEIS1, MEIS2A and TFIIB was confirmed using GST-pulldown assays. These results suggest that, like other homeodomain proteins, HOX11 interacts with members of the TALE homeodomain family, specifically MEIS1 and MEIS2A. Such interactions may function to enhance binding site specificity, thus providing a mechanism for HOX11 target gene selectivity. The observed interaction of HOX11 with TFIIB suggests that HOX11 may also function to regulate gene transcription by interaction with members of the basal transcriptional machinery and this hypothesis is supported by studies by Owens et al., (2003), which demonstrate that HOX11 represses transcription via interactions with components of the RNA polymerase II holoenzyme. Intriguingly, the human ALDH1A1 promoter, which was previously demonstrated to be regulated by HOX11 in Chapter 2 of this thesis, harbours a TFIIB recognition element (BRE)(-40/-34bp) directly upstream of a compound TATA/GATA site (GATA box). In conjunction with the results presented in this Chapter, a model in which HOX11 interacts with TFIIB to contribute to ALDH1A1 gene regulation may be envisaged. Finally, the ability of HOX11 to strongly homodimerise, provides an alternative mechanism by which HOX11 may activate or repress gene expression, via the formation of monomer/dimers on target gene promoters.
Chapter 5

The Effect Of Enforced Expression of ALDH1A1 and FHL1 on Murine Haematopoiesis

5.1 INTRODUCTION

Although the deregulation of HOX11 is believed to be a fundamental event in the development of leukaemia, attempts to accurately model the transforming capabilities of HOX11 in T-cells have, to date, been largely unsuccessful. Indeed, most of our knowledge regarding the mechanism of HOX11 action, derives from retroviral studies performed by Hawley et al., (1994). These demonstrated that enforced expression of HOX11 in primary murine bone marrow cells leads to a block in haematopoietic differentiation, giving rise to IL-3 dependent immature myeloid cell lines. Thus, the transforming potential of HOX11 appears to derive from its ability to activate or repress distinct target genes affecting haematopoietic differentiation programmes. Progression to a fully malignant state, however, is thought to require additional mutations affecting cell survival or proliferation, since lethally irradiated recipient mice reconstituted with MSCV-HOX11 infected bone marrow did not develop tumours after 10 months (Hawley et al., 1994). Despite the lack of an in vivo animal model that accurately parallels the T-cell oncogenic activity of HOX11 in humans, the in vitro retroviral studies performed by Hawley et al., (1994), provide a useful platform for functionally dissecting the immortalisation capacity of HOX11. With regard to the present study, the ability of known HOX11-regulated genes to mirror HOX11-mediated perturbations in haematopoietic differentiation may yield insights into the transcriptional pathways perturbed by this nuclear oncoprotein.

This study focused on understanding the transcriptional regulation of two candidate target genes of HOX11, namely ALDH1A1 and FHL1, in order to gain an insight into the mechanism of HOX11 action. In Chapter 2, it was established that HOX11 is capable of modulating transcription from the proximal promoter of ALDH1A1, thereby strengthening the hypothesis that ALDH1A1 is a specific target gene of HOX11. The confirmation of ALDH1A1 as a target gene of HOX11 is particularly intriguing given
the precedent for disrupted retinoid signalling in haematopoietic malignancies. The most well documented case involves the *promyelocytic leukaemia zinc finger (PLZF)* gene which was originally identified by the variant translocation t(11;17)(q23;q21) in a subset of patients with APL (Chen *et al.*, 1993). PLZF encodes a zinc finger transcription factor expressed in undifferentiated myeloid cell lines and CD34+ haematopoietic progenitor cells and is subsequently down-regulated during lineage commitment, suggesting that PLZF may have a role in the maintenance or survival of early haematopoietic stem cells and that subsequent downregulation is required for normal bone marrow maturation (Reid *et al.*, 1995). Following the translocation, PLZF is fused to RARα, the RA receptor isoform preferentially expressed in haematopoietic cells, to generate the chimaeric PLZF-RARα fusion protein, which is predicted to encode an aberrant retinoid receptor capable of providing a differentiation block due to the effects of permanent gene silencing. The product of the reciprocal translocation, RARα-PLZF, has been shown to activate cell cycle regulators and enhance cell growth, and the combined expression of these proteins, which results in the dysregulation of RARα and PLZF responsive genes respectively, results in the clonal proliferation of cells blocked at the promyelocyte stage of development. More recently, the concept of dysregulated retinoid signalling being involved in T-ALL has been strengthened by the discovery of *ALDH1A2* as a target gene of a transcriptional complex in leukaemic T-cells including SCL (TAL1) and LMO1 or LMO2, which are transcription factors specifically implicated in childhood T-ALL (Ono *et al.*, 1998).

The strong and specific modulation of the *ALDH1A1* promoter by HOX11, observed in the haematological cell lines PER-117 and HEL, prompted us to determine whether overexpression of *ALDH1A1* alone, would result in a perturbation of normal haematopoiesis, consistent with a possible role in HOX11-mediated leukaemogenesis. Our expectation was that if *ALDH1A1* is an oncogenically relevant target gene, then cells enforcibly expressing *ALDH1A1* would adopt at least some features of those cells transduced by HOX11, most notably a block in differentiation (Hawley *et al.*, 1994, Greene *et al.*, 2002, Izon & Greene, unpublished observations). The effect of overexpressing *FHL1* on murine haematopoiesis was also assessed, given that *FHL1* was also demonstrated to be transcriptionally regulated by HOX11, following transcriptional start site mapping and preliminary functional promoter analyses (Chapter
Thus, as a first step in determining the oncogenic relevance of two candidate HOX11 target genes, the effect of constitutively overexpressing the retinoic acid-synthesising enzyme ALDH1A1 and the LIM-domain protein FHL1, on the development of the T-cell, B-cell, myeloid and erythroid haematopoietic lineages was investigated. This was achieved by using a murine retrovirus harbouring either ALDH1A1 or FHL1 cDNA, to stably transduce primary murine bone marrow cells, which were subsequently transplanted into lethally irradiated recipient mice. The specific aims of this chapter were 1) to produce functional recombinant retrovirus capable of efficiently expressing ALDH1A1 and FHL1 genes in primary murine bone marrow cells and 2) to assess the effects of constitutive ALDH1A1 or FHL1 gene expression on haematopoiesis, in order to establish whether either of these genes may account for the ability of HOX11 to predispose T-cells to leukaemic transformation.

5.2 METHODS

5.2.1 Generation of Recombinant Retrovirus

5.2.1.1 MigR1-ALDH1A1/FHL1 Vector Construction

The MigR1 retroviral expression vector utilised in these studies (a gift from Dr Warren Pear), is derived from the murine stem cell virus based vector (MSCV2.2, CLONTECH; Hawley et al., 1994), which was modified to contain an internal ribosome entry site (IRES) and the cDNA of enhanced green fluorescent protein (eGFP) containing solubility mutations (163; Val → Ala) and red shift mutations (64; Phe → Leu, 65; Ser → Thr; Pear et al., 1998, J. Jacob, unpublished data; Figure 5.1). The resultant vector facilitates the production of a bicistronic mRNA allowing secondary translation from a single mRNA transcript, such that green fluorescence protein (GFP) expression may be used a surrogate marker for transgene expression. The validity of this approach is based on previous studies, which demonstrated that GFP expression correlates with intracellular Notch 1 protein expression, following transduction into murine bone marrow cells (Pui et al., 1999, Izon et al., 2001). The use of a single promoter and the IRES also increases the likelihood that both proteins will be expressed at equivalent levels, since this reduces the potential for promoter competition, which may cause
Figure 5.1. The MigR1 Retroviral Expression Vector Used to Generate Recombinant Retrovirus Expressing ALDH1A1 or FHL1. (A) The MigR1 vector contains 5’ and 3’ long terminal repeats (LTRs), which contain all the requisite cis elements for gene expression including the enhancer, promoter, transcription initiation (capping), transcription terminator and polyadenylation signals. The vector also contains the gene for ampicillin resistance (Amp') for bacterial propagation, internal ribosomal entry site (IRES) for the production of bicistronic mRNA and cDNA for enhanced green fluorescent protein (eGFP), which acts as a surrogate marker for transgene expression. (B) ALDH1A1 and FHL1 cDNAs were cloned into the multiple cloning site of the MigR1 retroviral vector, which was used to generate recombinant retrovirus using the Phoenix™ Retroviral System.
downregulation in internal promoter vectors (Schott et al., 1996).

5.2.1.1.1 Preparation of MigR1 Retroviral Expression Vector

The MigR1 retroviral expression vector was prepared for sub-cloning by digestion with Xho I and EcoR I restriction endonucleases (Promega). Reactions were performed in 1.5ml polypropylene microcentrifuge tubes in final volume of 50µl, containing 10x Restriction Digest Buffer D [1x; 6mM Tris-HCl (pH 7.9), 0.15M NaCl, 6mM MgCl₂, 1mM DTT](Promega), 0.1mg/ml nuclease-free acetylated BSA (Promega), 2µg of MigR1 vector plasmid DNA, together with 40U of Xho I and 24U of EcoR I restriction endonucleases (Promega) at 37°C for 2h. Following digestion, the restriction endonucleases were heat inactivated by incubation at 65°C for 20min. An aliquot of the digested MigR1 vector (40ng) was electrophoresed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) to confirm complete digestion prior to ligation with insert DNA.

5.2.1.1.2 PCR Amplification of ALDH1A1 and FHL1 Insert cDNA

For the construction of MigR1-ALDH1A1, the 1506bp ALDH1A1 coding region was amplified from HEL cDNA generated in Section 4.2.1.1.2 using PfuTurbo® DNA Polymerase (Stratagene) according to the manufacturer’s instructions. Reactions were performed in thin-walled 0.5ml microcentrifuge tubes in a final volume of 50µl, comprising 10x Cloned Pfu DNA Polymerase Reaction Buffer [1x; 20mM Tris-HCl (pH 8.8), 10mM KCl, 10mM (NH₄)₂SO₄, 2mM MgSO₄, 0.1% (v/v) Triton® X-100, 0.1mg/ml nuclease-free BSA](Stratagene), 0.2mM dNTPs (Fisher Biotech), 20pmol each of ALDH1XhoF2/ ALDH1EcoR primers (Table 5.1), 2.5µl of HEL cDNA and 2.5U of PfuTurbo® DNA Polymerase (Stratagene). Samples were denatured for 2min at 94°C and then amplified (94°C for 10s, 60°C for 30s and 72°C for 2min) for 10 cycles, with subsequent cycle elongation times increasing by 20s per cycle up to 8:40s, followed by a final extension of 72°C for 7min (EXTEND72, Appendix 1). For the construction of MigR1-FHL1, the 848bp human FHL1 coding region was amplified from ALL-SIL cDNA generated in Section 4.2.1.1.2 in a 50µl reaction comprising 10x Cloned Pfu DNA Polymerase Reaction Buffer [1x], 0.2mM dNTPs, 20pmol each of SLIMXhoF/SLIMEcoR primers (Table 5.1), 2.5µl of ALL-SIL cDNA and 2.5U of
TABLE 5.1
PfuTurbo® DNA Polymerase using the EXTEND72 PCR programme described above (Appendix 1).

5.2.1.1.3 Restriction Endonuclease Digestion of ALDH1A1 and FHL1 Insert cDNA

Full-length ALDH1A1 and FHL1 PCR products generated in Section 5.2.1.1.2 were gel purified using the Qiaex II Gel Purification Kit (Qiagen; Section 2.2.1.3) and prepared for sub-cloning by restriction digestion with Xho I and EcoR I restriction endonucleases. Digestions were performed in 1.5ml polypropylene microcentrifuge tubes in a final volume of 100µl, comprising 10x Restriction Digest Buffer D [1x], 0.1mg/ml nuclease-free acetylated BSA, 50µl of gel purified \textit{ALDH1A1/FHL1} cDNA, 40U of Xho I and 24U of EcoR I restriction endonucleases at 37°C for 18h. Following digestion, restriction endonucleases were heat inactivated by incubation at 65°C for 20min. Digested PCR fragments were subsequently purified from unwanted digestion products using the QIAquick PCR Purification Kit (Qiagen; Section 2.2.1.4). An aliquot of the purified insert was electrophoresed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) to assess DNA integrity and to estimate concentration prior to ligation.

5.2.1.1.4 Cloning of ALDH1A1 and FHL1 cDNA into the MigR1 Retroviral Vector

Sub-cloning of PCR amplified \textit{ALDH1A1} and \textit{FHL1} cDNA into the MigR1 retroviral expression vector was performed essentially as outlined in Section 2.2.1.6. Ligations were performed in 0.5ml polypropylene microcentrifuge tubes in a final volume of 20µl, comprising 40ng of linearised MigR1 vector DNA, a 3:1 molar ratio of insert \textit{ALDH1A1/FHL1} cDNA [31ng, 17.5ng respectively], 10x T4 DNA Ligase Buffer [1x; 30mM Tris-HCl (pH 7.8), 10mM MgCl$_2$, 10mM DTT, 1mM ATP](Promega) and 6U of T4 DNA Ligase (Promega) and incubated at 15°C for 16-24h. Ligation reactions were used to transform chemically competent DH5α \textit{E.coli} (Section 2.2.1.8) and small scale plasmid Mini-Preps of selected colonies were prepared for screening using the Sigma GenElute™ Plasmid Miniprep Kit (Sigma) according to the manufacturer’s instructions (Section 2.2.1.9). The presence of \textit{ALDH1A1} and \textit{FHL1} inserts was determined by restriction digest analysis of a selection of randomly chosen clones with Xho I and EcoR I restriction endonucleases. Digests were performed in 1.5ml polypropylene
microcentrifuge tubes in a final volume of 30µl, comprising 10x Restriction Digest Buffer D [1x], 0.1mg/ml nuclease-free acetylated BSA, 0.5µg plasmid DNA and 10U each of Xho I and EcoR I restriction endonucleases at 37°C for 1.5h. Restriction profiles were analysed on a 1.3% (w/v) agarose gel (Section 2.2.1.2) and selected recombinant clones were sequence-verified using MigR1SeqFor and MigR1SeqRev primers flanking the cloning site of MigR1 (Table 5.1; Section 2.2.1.11). MigR1, MigR1-ALDH1A1 and MigR1-FHL1 vector construct DNA was prepared for retrovirus production using the QIAGEN Plasmid Midi Kit according to the manufacturer’s instructions (Section 2.2.1.12; grey highlight). The recovered DNA was resuspended at a final concentration of 0.5µg/ul in sterile dH₂O and stored at -20°C.

5.2.1.2 Phoenix™ Retrovirus Production

MigR1, MigR1-ALDH1A1 and MigR1-FHL1 recombinant retrovirus was produced by transient transfection of the Phoenix™ retrovirus producer cell line, Phoenix™-Eco (Orbigen), by calcium phosphate co-precipitation as previously described by Pear et al., (1993)(Figure 5.2).

5.2.1.2.1 Growth and Passage of Phoenix™-Eco Cells

The Phoenix™-Eco cell line is based on the 293T human embryonic kidney cell line, and contain episomes for the production of gag-pol (introduced with hygromycin as a co-selectable marker) and envelope protein (with diptheria resistance as a co-selectable marker) for ecotropic viruses. Cells were subcultured twice weekly (1:10) to 70-80% confluence, in DMEM Multi-cel Medium (Trace) supplemented with 10% (v/v) FCS (Trace), 2mM L-glutamine (Trace) and 50µg/ml pencilllin/streptomycin (Trace) in 75cm² filter-top culture flasks (NUNC™) and maintained at 37°C with 5% CO₂. To passage cells, medium was removed and the cells were trypsinised with a 0.05% (v/v) Trypsin/0.53mM EDTA solution (Trace) until cells detached from flask surface (approximately 30s). Trypsinisation was then quenched using an equal amount of complete DMEM medium and cells were sub-cultured into fresh complete DMEM medium. Prior to transfection, Phoenix™-Eco cells were expanded in complete DMEM medium supplemented with 10µg/ml ciprofloxacins (ICN Biomedicals), 500µg/ml
Figure 5.2. Production of Recombinant Retrovirus Using the Phoenix™ Eco Packaging Cell Line. (1) MigR1 and MigR1-ALDH1A1/FHL1 retroviral expression vectors were transfected into the Phoenix™ Eco packaging cell line by CaPO₄ precipitation. (2) Transcription is initiated within the 5’ LTR and the mRNA is polyadenylated and processed using signals in the transcribed regions from the 3’LTR. (3, 4, 5) Viral mRNA, which contains a packaging signal (ψ+) is bound by group antigen (gag)-derived proteins produced by the packaging cell line, and incorporated into a retroviral particle at the cell surface. The Phoenix™ Eco cell line expresses viral proteins gag, which form the viral core structure, pol, a reverse transcriptase which converts the RNA genome to double stranded DNA pre-integrate form and performs integrase functions and env, which resides in the lipid layer of the viral envelope and determines viral tropism.
hygromycin (Sigma) and 1µg/ml diptheria toxin (Sigma) to treat potential mycoplasma infection and to upregulate gag, pol and env viral genes respectively, for a period of two weeks. Following selective treatment, Phoenix™-Eco cells were maintained in long-term storage by cryopreservation as outlined in Section 2.2.2.4, or used to generate retrovirus for up to one month.

5.2.1.2.2 Transfection of Phoenix™-Eco Cells With MigR1 Retroviral Constructs

The initial plating of Phoenix™-Eco cells is a crucial step in order to obtain high retroviral titres. Prior to plating, cells should not be allowed to become over-confluent, since clumping results in an uneven cell distribution following replating, reducing transfection efficiency. Additionally, all procedures involving plated cells (removing and replacing growth medium, addition of CaPO₄ precipitate) were performed gently, since the Phoenix™-Eco cells are easily detached from the flask surface, thereby affecting the production of retrovirus. Cells were evenly plated at a density of 2.5-3 x 10⁶ into 25cm² culture flasks (NUNC™) in 10ml complete DMEM media, 18-24h prior to transfection. MigR1, MigR1-ALDH1A1 and MigR1-FHL1 DNA [10µg] was prepared for application to cells by pre-mixing with 61µl of 2M CaCl₂ in a final volume of 500µl. This solution was added dropwise to 500µl of 2x Hepes Buffered Saline (HBS; pH 7.0) whilst continuously bubbling, and immediately applied to the sub-confluent (60-70%) Phoenix™-Eco cells, by gently pipetting the solution into the growth media, without directly contacting the cell monolayer. The flask was rocked gently to evenly distribute DNA/CaPO₄ particles and incubated for 18h at 37°C/5%CO₂. The media was then removed and the cells were gently overlaid with 3.5ml of fresh complete DMEM media to collect the retrovirus. The retroviral supernatant was harvested 48h post-transfection. The supernatants were pooled and single-use 1ml volumes (viral titre drops by half per freeze-thaw cycle) were aliquoted into 2ml polypropylene tubes on ice. Aliquots were snap-frozen in liquid N₂ and the virus stock was stored at -80°C.
5.2.1.2.3 Determination of Retrovirus Titre

Murine fibroblast NIH 3T3 cells (2 x 10⁵) were seeded in 60mm petri dishes (NUNC™) 12h prior to retroviral infection with the supernatant. The following day, media was removed and replaced with 2ml of complete DMEM media containing 200µl of retroviral supernatant and 8µg/ml Polybrene (Sigma), and the cells were incubated for 24h at 37°C. Following infection, fresh complete DMEM media (4ml) was added and the cells were incubated for a further 24h at 37°C. The cells were then assayed for expression of GFP by fluorescence activated cell sorting (FACS) analysis. The cells were trypsinised, washed with PBS and resuspended in 500µl of the same solution. Cell suspensions were temporarily incubated on ice prior to FACS analysis, or alternatively, fixed in PBS containing 1% (v/v) paraformaldehyde and stored at 4°C for postponed FACS analysis. The percentage of pooled clones that had stable integration of the MigR1 or MigR1-ALDH1A1/FHL1 constructs was determined by applying single cell suspensions (2 x 10⁵ cells) onto a FACS Calibur Flow cytometer equipped with MAC CELLQUEST software (Becton Dickinson, CA, USA; http://www.bd.com). Retroviral supernatants yielding >20% GFP expressing cells were utilised in downstream applications.

5.2.2 Bone Marrow Reconstitution

5.2.2.1 Haematopoietic Cell Transduction by Spinofection

C57BL/6 12-24 week-old female mice were purchased from the Animal Resource Centre (ARC). Bone marrow cells were extracted by flushing tibias and femurs with PBS (8ml/mouse) using a 26” gauge needle (Becton Dickinson) and resuspended to achieve a single cell suspension. The cell suspension was transferred to a 10ml polypropylene tube, pelleted by centrifugation (394 x g, 3min, 25°C; Eppendorf 5810) and resuspended in 2ml PBS. Bone marrow cells were then separated from contaminating red blood cells by equilibrium density gradient centrifugation with Ficoll-Diatrizoate sodium (FicollPaque™) or by lysis method. The first method involved layering 1ml aliquots of the cell suspension on 4ml of human grade FicollPaque™ PLUS (Amersham) in a 10ml polypropylene tube and centrifuging at
1000 x g for 10min at 25°C (no brake). Cells were then collected at the interface and washed once with 8ml of PBS (394 x g, 5min, 25°C) and resuspended in 2ml of complete DMEM medium. The second method involved purification of lymphoid cells by lysis of contaminating red blood cells. The cells were resuspended in 0.165M NH₄Cl (~20ml/mouse) in a 50ml polypropylene tube and incubated on ice for 30min with intermittent mixing. The purified cells were subsequently pelleted (394 x g, 5min, 25°C) to remove lysis solution, washed once with 10ml of complete DMEM media, and filtered using 70uM nylon cell strainers (BD Falcon) to remove debris. Cells were then pelleted (394 x g, 5min, 25°C) and resuspended in 2ml of complete DMEM media. Approximately 4 x 10⁶ cells were seeded/well into a 12 well culture plate (NUNC™) in a final volume of 2.2ml complete DMEM media and pre-incubated for 52h prior to spinofection with 50ng/ml SCF (Peprotech), 6ng/ml IL-3 (Peprotech), 4ng/ml IL-1β (Peprotech) and 1ng/ml IFN-γ (Peprotech). Following this incubation, cells were gently resuspended by pipetting and divided into two wells (~2 x 10⁶ cells/well). Retroviral infection of bone marrow cells was performed by pre-mixing 1ml of retroviral supernatant with 50ng/ml SCF, 6ng/ml IL-3, 1ng/ml IFN-γ, 4ng/ml IL1-β and 6μg/ml polybrene and transferring to the appropriate wells. The cells were then spin infected (spinofection) by centrifugation at 1010 x g for 50min at 25°C (Rotina 46R, Hettich) and incubated for a further 24h at 37°C/5% CO₂.

5.2.2.2 Reconstitution of Mice With Transduced Bone Marrow Cells

Recipient female mice (C57BL/6) were primed for reconstitution with transduced bone marrow cells by incorporating 1.1μg/L of neomycin sulfate (Sigma) and 126mg/L of polymyxin B sulfate (Sigma) antibiotics into the drinking water at least four days prior to irradiation. On the day of reconstitution, mice were subject to lethal irradiation (546.1cGy/1:10s; two doses separated by 3h)(Gammacell® 3000 Elan, MDS Nordion). Retrovirally transduced bone marrow cells were transferred to a 10ml polypropylene tube and prepared for injection by centrifugation (394 x g, 3min, 25°C) to remove growth media, and resuspended at a final concentration of 5 x 10⁵ -1 x 10⁶ cells/300μl PBS. Cells were maintained on ice until injection. Typically, 5 x 10⁵ -1 x 10⁶ transduced bone marrow cells were injected via the tail vein into each irradiated recipient. Mice were then housed under ‘quarantine conditions’ in micro-isolator cages and provided
with autoclaved food and antibiotic supplemented drinking water for a period of 1 month. Experimental retroviral constructs (MigR1-ALDH1A1, MigR1-FHL1) were tested in triplicate and MigR1 control transplant recipients were generated a total of 6 times in 2 independent experiments.

5.2.3 FACS Analysis

5.2.3.1 Tissue Processing

Mice were sacrificed in groups of three (one each of control MigR1 and test MigR1-ALDH1A1 and MigR1-FHL1 transplant recipients), 33, 36 and 38 days post-transplantation, and tissues were immediately harvested. Spleen, thymus and bone marrow were extracted and stored on ice in 1.5ml polypropylene tubes containing PBS. Cell suspensions of spleen and thymus were then prepared by emulsifying tissues in 5ml of PBS, between pre-wetted microscope slides (25.4 x 76.2mm; Sail Brand), using the ground edges of the slides. Bone marrow cells were extracted by flushing tibia and femur sections with PBS (10ml /mouse) using a 26” gauge needle and homogenous cell suspensions were obtained by repetitive pipetting. Spleen, thymus and bone marrow cell suspensions were then transferred to 10ml polypropylene tubes and pelleted by centrifugation (394 x g, 5min, 25°C) and resuspended in 2ml FACS buffer [PBS containing 5% (v/v) BSA, 0.02% (w/v) NaN₃] prior to filtration to remove cellular debris. Cell counts were performed by the trypan blue exclusion method. An aliquot of the cell suspension was diluted in Trypan Blue Stain (GIBCO BRL Life Technologies)(typically, spleen and bone marrow samples were diluted 1:10 and thymus samples were diluted 1:2 with Trypan Blue Stain) and the cells were counted using a haemocytometer counting chamber, excluding contaminating red blood cells (RBCs). Cell suspensions were maintained at 4°C prior to immunophenostaining.

5.2.3.2 Immunophenostaining of Transplant Recipient Cells

Cells obtained from murine spleen, thymus and bone marrow tissues were analysed for expression of TCRβ, CD4, CD8, CD19, B220, Gr1, Mac1, CD71, Ter119, Sca1 and c-kit cell surface markers by multi-parameter flow cytometry, using fluorochrome-
conjugated antibodies according to previously described protocols (Coligan et al., 1992)(Appendix 3). All antibodies were purchased from Pharmingen and diluted to optimised working concentrations in FACS buffer. Approximately 1 x 10^6 cells/well were stained in a 96 well, round-bottomed microtitre plate (NUNCLONTM) and all centrifugation steps were performed at 320 x g for 5min at 6°C (Rotina 46R, Hettich). Cells were pelleted by centrifugation and the supernatant was discarded by flicking the plate and inverting briefly on absorbent paper to remove excess liquid. Non-specific antibody interactions were blocked by incubating the cells with unconjugated anti-Fcγ antibody (30µl/well) for 10min at 4°C. Cells were pelleted by centrifugation and at this point negative controls for each tissue were resuspended in 500µl FACS buffer and transferred to 5ml polystyrene FACS tubes (12 x 75mm; BD Falcon) at 4°C. Test samples were then incubated with specific monoclonal antibody cocktails to antigens present on the surface of mature blood cell lineages (20µl/well; Appendix 3) for 20min at 4°C. Following incubation, primary antibodies were removed by washing the cells with FACS buffer (200µl/well) and centrifugation. The cells were then incubated with the appropriate secondary antibodies (20µl/well) for 10min at 4°C (Appendix 3). At this point, cells not requiring secondary reagents were resuspended in 500µl FACS buffer, transferred to FACS tubes and maintained at 4°C. Secondary antibodies were removed by washing the cells with FACS buffer (200µl/well) and the cells were pelleted by centrifugation and resuspended in 500µl FACS buffer. Propidium Iodide (PI)(Sigma) was added at a final concentration of 1µg/ml to all cell samples in order to exclude dead cells from subsequent analysis. In the case of T-cell analysis, however, PI was not included in the immunostaining procedure since the PI emission spectrum overlaps the FL3 channel, thus interfering with CD8-PerCP.

5.2.3.3 Flowcytometric and FlowJo Analysis

For flow cytometry, single cell suspensions of ~ 1 x 10^6 cells prepared as in Section 5.2.3.2 were run on a FACS Calibur flow cytometer equipped with CELLQuest software (Becton Dickinson, CA, USA; http://www.bd.com). This system permits rapid, highly sensitive flow cytometric analysis for up to six separate parameters including forward scatter (FSC), side scatter (SSC), FL1 (530nm), FL2 (585nm) and FL3 (>650nm),

232
which use an argon laser beam (488nm), in addition to the fluorescence parameter, FL4 which uses the optional red diode laser (635nm)(Table 5.2).

<table>
<thead>
<tr>
<th>Fluorochromes</th>
<th>Emission Peak (nm)</th>
<th>Detector</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFP (green)</td>
<td></td>
<td>FL1</td>
<td>530/30</td>
</tr>
<tr>
<td>R-Phycoerythrin (PE)</td>
<td>575 (orange-red)</td>
<td>FL2</td>
<td>585/42</td>
</tr>
<tr>
<td>Propidium Iodide (PI)</td>
<td>620 (red)</td>
<td>FL2,FL3</td>
<td>585/42</td>
</tr>
<tr>
<td>PerCP</td>
<td>674 (red)</td>
<td>FL3</td>
<td>670LP</td>
</tr>
<tr>
<td>PE-Cy5</td>
<td>670 (red)</td>
<td>FL3</td>
<td>670LP</td>
</tr>
<tr>
<td>APC</td>
<td>660</td>
<td>FL4</td>
<td>661/16</td>
</tr>
</tbody>
</table>

Table 5.2. Fluorochromes Used for Immunophenotyping with the BD FACS Calibur. The emission wavelength maxima of the dye (nm), the detector configured optimally for the dye and the optical filters installed in front of each detector are outlined.

As individual cells pass through the 488-nm light beam of the argon-ion laser, forward light is scattered by the cell in proportion to cell size and 90°C light is scattered according to cellular complexity (cellular granularity, surface texture, refractive index). Cells stained with specific antibodies, which are covalently linked to fluorochromes (fluorescent dyes), emit light over a broad range of wavelengths, and these fluorescence emission signals are detected by a photomultiplier tube (PMT), by virtue of a set of spectrally selective optical filters corresponding to individual fluorescence detectors (FL1, FL2, FL3, FL4). The signal is subsequently amplified (linearly or logarithmically) and acquired by associated computer software.

Prior to sample analysis, instrument settings for non-fluorescent parameters (FSC and SSC detector levels) were adjusted using an unstained bone marrow sample, such that only relevant events displaying characteristics of cells are detected. The forward scatter threshold was adjusted to a low level (52) to enable visualisation of all populations, including dead cells, RBC, lymphocytes and granulocytes, and FSC/SSC amplifier gains were adjusted to place the majority of cells at the midpoint of each axis. A gate was then defined on the live cell population of interest, using the polygon region tool and the FL1 (GFP) PMT voltage was adjusted so that cells were located within the first decade of the dot plot. The fluorescence channels were then adjusted (compensation)
using single colour controls for individual fluorochromes, to eliminate the overlap of fluorescence signals that may occur with brightly positive samples. Data files were acquired with 100,000 events collected for each sample and flow cytometry data was analysed with FlowJo software (Tree Star).

5.3 RESULTS

The hypothesis that the transforming capacity of HOX11 stems from its ability to perturb haematopoietic differentiation is supported by several studies. These have demonstrated that overexpression of HOX11 in murine bone marrow and embryonic stem cells leads to the generation of growth factor-dependent myeloid progenitor and erythroid cell lines, respectively (Hawley et al., 1994, Keller et al., 1998). Moreover, enforced expression of HOX11 in murine J2E erythroid cells resulted in the imposition of an immature phenotype, suggesting that despite being a T-cell oncoprotein, HOX11 is capable of affecting the differentiation of several cell lineages (Greene et al., 2002). Although details of the mechanism(s) by which HOX11 induces leukaemic transformation following dysregulated expression in T-cells are fragmentary, it is likely that the inappropriate expression of downstream target genes contributes to neoplastic transformation. Thus, in order to investigate the ability of HOX11-regulated genes to alter the differentiation potential of haematopoietic precursors, the effects of constitutive ALDH1A1 and FHL1 expression were analysed in lymphoid, myeloid and erythroid cell lineages, following the reconstitution of lethally irradiated mice with retrovirally transduced (MigR1-ALDH1A1, MigR1-FHL1) primary murine bone marrow cells (Figure 5.3). This approach has been used to study the effects of several ectopically expressing genes that may predispose cells of the haematopoietic system to leukaemic transformation. This includes the fli3 ligand (Hawley et al., 1998), the stem cell leukemia protein (SCL)(Kunisato et al., 2004), the Notch target gene, Hes (Kawamata et al., 2002), Hoxa9 (Thorsteinsdottir et al., 2002) and HOXB4 (Schiedlmeier et al., 2003).
5.3.1 Production of Recombinant Retrovirus

**Figure 5.3. Schematic Outline of Retroviral Transduction of Murine Bone Marrow and Subsequent Reconstitution.** Following the production of recombinant retrovirus using the Phoenix™ Retroviral System, primary bone marrow cells were harvested and pre-incubated with cytokines for 52h. The cells were subsequently transduced with retrovirus by spin infection, and incubated for 18h, prior to reconstitution of lethally irradiated recipient by tail vein injection. Mice were sacrificed after 1 month and the development of haematopoietic lineages was analysed by multiparameter flow cytometry.
The production of high-titre retroviral supernatants is crucial in order to maximise the transduction rates of appropriate haematopoietic progenitors, specifically multipotent HSCs. These comprise a rare population and are required in sufficient numbers for the long-term reconstitution of lethally irradiated recipients (Spangrude, 1994). Since the production of high-titre recombinant retrovirus involves a high level of plasmid expression, the efficiency of transfection of control and test MigR1 retroviral constructs into the Phoenix™ Eco packaging cell line was first assessed by estimating the percentage of cells expressing GFP by immunofluorescence light microscopy (MRC-1024 UV Confocal Microscope, Bio-Rad). As shown in Figure 5.4, the transfection of the viral constructs into the Phoenix™ Eco cells was successful. Given this, viral supernatants were collected and titres were subsequently determined by infection of NIH 3T3 murine fibroblasts with a 1:10 dilution of retroviral supernatant, where GFP expression of at least 20% of transduced cells was taken as a lower limit for an acceptable viral titre. The percentage of GFP-expressing cells for individual retroviral constructs was determined by FACS analysis, and viral titres within the desired range were obtained for MigR1 (23%), MigR1-ALDH1A1 (55%) and MigR1-FHL1 (53%) constructs.

5.3.2 Stable Transduction of Bone Marrow Cells with Retrovirus and Effects on Different Haematopoietic Lineages

In order to transduce murine haematopoietic progenitors, the MSCVMigR1 vector was used to express the ALDH1A1 and FHL1 cDNAs (Figure 5.1). This vector employs a specially designed long terminal repeat (LTR) from the murine stem cell PCMV virus, which enables high level, constitutive gene expression in transduced immature haematopoietic cells including myeloid and lymphoid progenitors, in addition to haematopoietic stem cells capable of long term multi-lineage reconstitution of lethally irradiated recipients (Hawley et al., 1994, Hawley et al., 1995, Hawley et al., 1996) (http://www.vectors.clontech.com). Moreover, sustained transgene expression has been observed in human HSCs transduced with an MSCV-based retroviral vector expressing the nerve growth factor receptor (NGFR) reporter gene and in differentiated erythroid
Figure 5.4. Transfection of MigR1 Retroviral Constructs into Phoenix™-Eco Producer Cells. Transfection efficiency of (A) MigR1-GFP (B) MigR1ALDH1A1-GFP and (C) MigR1FHL1-GFP into Phoenix™ Eco producer cells was assessed by visualising GFP expression 48h post-transfection by confocal microscopy, and myeloid cells derived from these progenitors (Cheng et al., 1998).
The bone marrow reconstitution experiments largely employed conditions previously optimised by Pear et al., (1998) to enhance the likelihood that transduced progenitors would stably express the provirus. This included the pre-incubation of donor cells with cytokines 24h prior to retroviral transduction to induce stem cell cycling and transduction of bone marrow cells by spin infection (spinofection). Nevertheless, some preliminary experiments were first required in order to determine the optimal dosage of radiation and number of cells needed for the repopulation and subsequent detection of transgene-expressing cells. The reconstitution of sub-lethally irradiated recipients (mice subjected to a single dose of radiation of 550cGy) with 0.5 x 10⁶ cells/mouse, yielded healthy mice, however, no GFP was detected by FACS analysis of peripheral blood cells 1 month post-transplantation (data not shown). A second method involved reconstitution of lethally irradiated recipients (mice subjected to two doses of radiation of 546.1cGy/1min,10s; separated by 3h) with 0.5 x 10⁶ cells/mouse. However, these mice showed signs of illness after 3 weeks, and analysis of various lineages by immunostaining and FACS analysis revealed perturbations in the development of all major haematopoietic lineages (data not shown). Taken together, these results suggested that depletion of recipient haematopoietic cells by sub-lethal irradiation and reconstitution with 0.5 x 10⁶ cells/mouse, was not sufficient for donor-derived cells to competitively reconstitute recipient haematopoietic lineages. More cells were required in order re-establish even short term reconstitution in lethally irradiated recipients, therefore subsequent experiments utilised a combination of lethal irradiation and injection of 1.5 x 10⁶ transduced bone marrow cells/mouse.

In order to confirm the expression of ALDH1A1 and FHL1 in retrovirally transduced haematopoietic progenitors derived from transplant recipients, the percentage of cells expressing the surrogate marker, GFP, were determined by FACS analysis. Since GFP is coexpressed with ALDH1A1/FHL1 as a bicistronic mRNA, it therefore provides a direct measure of transgene expression. GFP expression as directed by the MSCV 5’ LTR, was analysed in four different cell lineages (B-cells, T-cells, myeloid and erythroid cells) extracted from the bone marrow and spleen tissues of control (MigR1) and test (MigR1-ALDH1A1, MigR1-FHL1) mice (Figure 5.5A, B and C). The percentage of GFP-expressing cells was typically higher in mice reconstituted with
MigR1-ALDH1A1 and MigR1-FHL1 constructs, compared to the MigR1 negative control alone, in accordance with the higher viral titres obtained for MigR1-ALDH1A1 and MigR1-FHL1 retroviral supernatants (Section 5.3.1). The intensity of GFP fluorescence, which is an indication of the level of expression of the integrated virus was also assessed. This analysis revealed that in MigR1-ALDH1A1 mice analysed by flow cytometry, the major peak of GFP-expressing cells (for bone marrow and spleen samples) was contained within the intensity range of 25-500 on the GFP histogram. In contrast, the major peak of GFP-expressing cells from MigR1-FHL1 mice was contained within the intensity range of 100-10,000, suggesting that FHL1 was expressed at a higher level than ALDH1A1. In the case of the control GFP-expressing retroviral vector, MigR1, the expression pattern of GFP encompassed a broader range (40-2000).

Expression of both ALDH1A1-GFP and FHL1-GFP constructs was lower in bone marrow and spleens of mice sacrificed at 38 days compared to mice examined at 33 and 36 days, however no change in the expression of MigR1-GFP was observed (Figure 5.5A, B and C). These results may reflect the inability of transduced cells to sustain long-term transgene expression, possibly due to due to silencing mechanisms. Alternatively, this phenomenon may be attributed to the infection of short-term progenitors as opposed to long-term repopulating stem cells.

5.3.3 Effect of ALDH1A1 and FHL1 Expression on Different Haematopoietic Lineages

The first step in haematopoiesis involves commitment of a pluripotent, haematopoietic stem cell (Lin-negative, c-kit positive, Sca-1 positive, and CD34 negative) to either a common myeloid progenitor (CMP) or a common lymphoid progenitor (CLP). Following this commitment, CMPs and CLPs are guided by specific sets of transcription factors to form all the elements of the blood, with CMPs giving rise to myeloid, erythroid, megakaryocyte and mast cell lineages, and CLPs differentiating into T-cells and B-cells. In order to assess the effects of overexpressing either of two target genes of the T-cell oncoprotein HOX11, namely ALDH1A1 and FHL1 on haematopoietic differentiation, the maturational status of cells from the spleen, bone
Figure 5.5A. Contribution (%) of MigR1-GFP, MigR1ALDH1A1-GFP and MigR1FHL1-GFP transduced cells in spleen, bone marrow and thymus tissues 33 days after transplantation (Replicate #1). GFP Fluoresces in the FL1 channel.

Replicate 2 (36 days post-transplant)
Figure 5.5B. Contribution (%) of MigR1-GFP, MigR1ALDH1A1-GFP and MigR1FHL1-GFP transduced cells in spleen, bone marrow and thymus tissues 36 days after transplantation (Replicate #2). GFP Fluoresces in the FL1 channel.
Figure 5.5C. Contribution (%) of MigR1-GFP, MigR1ALDH1A1-GFP and MigR1FHL1-GFP transduced cells in spleen, bone marrow and thymus tissues 38 days after transplantation (Replicate #3). GFP fluoresces in the FL1 channel.
marrow and thymus tissues of control (MigR1; x 3 mice) and test (MigR1-ALDH1A1 and MigR1-FHL1; x 3 mice each) mice was analysed by immunophenotyping with lineage-specific antibodies.

Firstly, the effect of ALDH1A1 and FHL1 expression on the myeloid lineage was assessed by staining cells from the spleen and bone marrow of recipient mice for the expression of myeloid-specific cell surface markers Gr1 and Mac1. The overexpression of ALDH1A1 in haematopoietic progenitors appeared to be correlated with an increase in mature granulocytes, coexpressing high levels of Gr1 and Mac1, when compared to control MigR1 mice (Figure 5.6). This was evident in both the spleen (mean of 5.8±1.3% vs 23.7±1.8%, n = 3; Figure 5.6A) and bone marrow (mean of 33.3±8.6% vs 56.6±2.5%, n = 3; Figure 5.6B), although the effect was more pronounced in the spleen (~4-fold) compared to the bone marrow (~1.7-fold). Unlike ALDH1A1, overexpression of FHL1 in haematopoietic progenitors was not associated with a significant change in myeloid differentiation compared to MigR1 control mice in either the spleen (mean of 5.8±1.3% vs 10.5±5.6%, n = 3; Figure 5.6A) or bone marrow (mean of 33.3±8.6% vs 26.6±3.3%, n = 3; Figure 5.6B).

B-cell differentiation was next assessed by analysing the cell surface markers CD19 and B220. B-cell development occurs in the bone marrow in the following order; pro-B cells, which represent the first committed B-cell precursor expressing B-cell-specific genes CD19+ and B220+, pre-B-cells and immature B-cells, which enter the splenic follicles where they differentiate into mature plasma cells or germinal centre B-cells (reviewed in Wang & Clarke, 2003). In contrast to the stimulatory effect of ALDH1A1 on the myeloid compartment, B-cell development appeared to be inhibited by the overexpression of ALDH1A1 in haematopoietic progenitors, as evidenced by the reduced numbers of CD19'B220' cells in MigR1-ALDH1A1 mice compared to MigR1 control mice in the spleen (mean of 53±13.6% vs 21.4±3.8%, n = 3; Figure 5.7A) and bone marrow (mean of 27.1±15.1% vs 5.1±0.9%, n = 3; Figure 5.7B). These results suggest that ALDH1A1 may affect the extent of contribution to mature cells from CLP or CMP lineages, since this increase in the myeloid compartment was associated with a concomitant decrease in one of the lymphoid compartments. Indeed, although ALDH1A1 did not appear to affect the differentiation profile of T-cells compared to
Figure 5.6A. Myeloid Immunophenotyping of Cells Obtained From the Spleens of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow. Cells were stained with phycoerythin (PE)-labelled Gr1 and biotin (Bio)-labelled Mac1 mAbs, which fluoresce in the FL2 and FL1 channels respectively. Mice reconstituted with MigR1-ALDH1A1 transduced bone marrow displayed elevated numbers of mature (Mac1+Gr1+) myeloid cells, compared with control MigR1 mice in spleen (mean of 5.83±1.3% vs 23.7±1.8%, n=3). No significant difference in myeloid differentiation was detected in mice reconstituted with MigR1-FHL1 transduced bone marrow compared to MigR1 control mice (mean of 5.8±1.3% vs 10.5±5.6%, n=3).
Figure 5.6B. Myeloid Immunophenotyping of Cells Obtained From the Bone Marrow of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow. Cells were stained with phycoerythrin (PE)-labelled Gr1 and biotin (Bio)-labelled Mac1 mAbs, which fluoresce in the FL2 and FL1 channels respectively. Mice reconstituted with MigR1-ALDH1A1 transduced bone marrow displayed elevated numbers of mature (Mac1+Gr1+) myeloid cells, compared with control MigR1 mice bone marrow (mean of 33.3±8.6% vs 56.6±2.5%, n=3). No significant difference in myeloid differentiation was detected in mice reconstituted with MigR1-FHL1 transduced bone marrow compared to MigR1 control mice (mean of 33.3±8.6 vs 26.6±3.3, n=3).
Figure 5.7A. B-cell Immunophenotyping of GFP+ve Cells Obtained From the Spleens of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow. GFP+ve cells were stained with phycoerythin (PE)-labelled CD19 and biotin (Bio)-labelled B220 mAbs, which fluoresce in the FL2 and FL1 channels respectively. Mice reconstituted with MigR1-ALDH1A1 transduced bone marrow demonstrate perturbations in B-cell development as evidenced by reduced numbers of CD19^B220^ B-cells, compared with control MigR1 mice in spleen (mean of 53±13.6% vs 21.4±3.8%, n=3). No significant difference in B-cell differentiation was detected in mice reconstituted with MigR1-FHL1 transduced bone marrow compared to MigR1 control mice (mean of 53±13.6% vs 45.9±15.9%, n=3).
Figure 5.7B. B-cell Immunophenotyping of GFP+ve Cells Obtained From the Bone Marrow of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow. GFP+ve cells were stained with phycoerythrin (PE)-labelled CD19 and biotin (Bio)-labelled B220 mAbs, which fluoresce in the FL2 and FL1 channels respectively. Mice reconstituted with MigR1-ALDH1A1 transduced bone marrow demonstrate perturbations in B-cell development as evidenced by reduced numbers of CD19+B220+B-cells, compared with control MigR1 mice in bone marrow (mean of 27.1±15.1% vs 5.1±0.9%, n=3). No significant difference in B-cell differentiation was detected in mice reconstituted with MigR1-FHL1 transduced bone marrow compared to MigR1 control mice (mean of 27.1±15.2% vs 19.4±7.7%, n=3).
control MigR1 mice (Figure 5.8), a reduction in the contribution of ALDH1A1-GFP-expressing cells towards the thymus compared to the bone marrow and spleen, was observed at three different timepoints post-reconstitution (Figure 5.5A, B, C). This is in keeping with a model whereby ALDH1A1 affects the commitment of HSCs to either the myeloid or lymphoid lineage. Alternatively, ALDH1A1 may inhibit thymic colonisation by T-cell precursors. In contrast, the percentage of MigR1-GFP and MigR1-FHL1/GFP-expressing cells was relatively consistent in all three tissues examined (Figure 5.5A, B, C). Nor was overexpression of FHL1 associated with a change in B-cell differentiation, in either spleen (mean of 53±13.6% vs 45.9±15.9%, n = 3; Figure 5.7A) or bone marrow compartments (mean of 27.1±15.2% vs 19.4±7.7%, n=3; Figure 5.7B). Similar to ALDH1A1, overexpression of FHL1 did not appear to affect T-cell development (CD4+ and CD8+; Figure 5.8).

Notably, in the case of ALDH1A1, there appeared to be a slight trans-effect, in which the effect of ALDH1A1 expression on the myeloid and B-cell compartments extended to neighbouring GFP/ALDH1A1-negative cells. This effect was most pronounced in the myeloid compartment, where the GFP/ALDH1A1-negative population was expanded in mice transduced with MSCV-ALDH1A1 transduced bone marrow compared to GFP-negative cells taken from MigR1 control mice (Figure 5.9).

Finally, the effect of ALDH1A1 and FHL1 overexpression on the maturational stage of differentiating erythroblasts was examined by staining for erythroid-specific CD71 and Ter119 cell surface markers. Erythropoiesis occurs in the order: proerythroblast, to basophilic and polychromatophilic erythroblasts, and then the orthochromatophilic erythroblast, giving rise to reticulocytes by enucleation (Fawcett, 1997). The transferrin receptor CD71 is expressed at high levels by early erythroid precursors and at the proerythroblast stage, but is downregulated during erythrocyte maturation (Ponka & Lok, 1999). Conversely, the cell surface erythroid specific Ter119 antigen is first expressed at the proerythroblast stage and is highly expressed at all later stages of differentiation (Kina et al., 2000). Overexpression of ALDH1A1 did not appear to have a significant effect on erythropoiesis as assessed by calculating the means for Ter119/CD71 double positive cells in either the spleen (mean of 27.5+9.1 vs 34.5+4.3%, [n=3], 5.10A) or bone marrow (mean of 22.8+2% vs 22.1+3.4%, [n=3];
Figure 5.8. T-Cell Immunophenotyping of GFP+ve Cells Obtained From the Thymus of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow. GFP+ve cells were stained with phycoerythrin (PE)-labelled CD4 and biotin (Bio)-labelled CD8 mAbs, which fluoresce in the FL2 and FL1 channels respectively. No significant difference in T-cell development was detected in mice reconstituted with MigR1-ALDH1A1 or MigR1-FHL1 transduced bone marrow compared to control MigR1 mice.
**Figure 5.9.** Trans-effect of *ALDH1A1* Expression on Myeloid Differentiation. GFP-ve cells obtained from the spleens of lethally irradiated mice 4 weeks post-reconstitution with either MigR1, MigR1-*ALDH1A1* or MigR1-*FHL1* transduced bone marrow were stained with phycoerythrin (PE)-labelled Gr1 and biotin (Bio)-labelled Mac Abs, which fluoresce in the FL2 and FL1 channels respectively. Mice reconstituted with MigR1-*ALDH1A1* transduced bone marrow displayed elevated numbers of mature (Mac1*Gr1*) myeloid cells compared to MigR1 control or MigR1-*FHL1* mice in the spleen.
Figure 5.10B). In the case of FHL1, however, there appeared to be a significant increase in erythropoiesis in the bone marrow (mean of 22.8±2% vs 41.4±6.1%, [n=3]; Figure 5.10B), however this effect was not observed in the spleen (mean of 27.5±9.1% vs 26.7±8.5%, [n=3]; Figure 5.10A). Although the bone marrow is the most important erythropoietic organ under resting or stimulated conditions, there appeared to be abnormally high numbers of Ter119/CD71 double positive erythroid cells in the spleen, suggesting that extramedullary haematopoiesis was taking place. This may reflect a state of anaemia, which is thought to trigger compensatory mechanisms resulting in the increased erythropoiesis, principally in the spleen (Kam et al., 1999), and indicates that haematopoietic homeostasis had not occurred 4 weeks post-transplant. Nevertheless, analysis at this timepoint has the advantage of enabling the acute effects of transgene expression to be identified.

In summary, the reconstitution of lethally irradiated recipients with primary murine bone marrow overexpressing the HOX11 target gene, ALDH1A1, resulted in enhanced myelopoiesis and a suppression of B-lymphopoiesis, whilst erythroid development appeared unaffected. Although ALDH1A1 did not appear to affect T-cell differentiation, overexpression of ALDH1A1 was associated with a quantitative decrease in cells contributing to the T-cell compartment, as measured by GFP expression. In contrast, the overexpression of a second HOX11 target, FHL1, was associated with an increase in erythropoiesis in the bone marrow, whereas myeloid, T-cell and B-cell compartments remained unaffected.
Figure 5.10A. Erythroid Immunophenotyping of GFP+ve Cells Obtained From the Spleens of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow. GFP+ve cells were stained with phycoerythrin (PE)-labelled CD71 and biotin (Bio)-labelled Ter119 mAbs, which fluoresce in the FL2 and FL1 channels respectively. Mice reconstituted with MigR1-ALDH1A1 demonstrated no change in the numbers of Ter119⁺CD71⁺ erythroid cells compared with control MigR1 mice in spleen (mean of 27.5±9.1% vs 34.5±4.2%, n=3). Similarly, mice reconstituted with MigR1-FHL1 demonstrated no change in the numbers of Ter119⁺CD71⁺ erythroid cells compared with control MigR1 mice in spleen (mean of 27.5±9.1% vs 26.7±8.5%, n=3).
Figure 5.10B. Erythroid Immunophenotyping of GFP+ve Cells Obtained From the Bone Marrow of Lethally Irradiated Mice 4 Weeks Post-Reconstitution With Either MigR1, MigR1-ALDH1A1 or MigR1-FHL1 Transduced Bone Marrow. GFP+ve cells were stained with phycoerythrin (PE)-labelled CD71 and biotin (Bio)-labelled Ter119 mAbs, which fluoresce in the FL2 and FL1 channels respectively. Mice reconstituted with MigR1-ALDH1A1 demonstrated no change in the numbers of Ter119^CD71^ erythroid cells compared with control MigR1 mice in the bone marrow (mean of 22.8±2% vs 22.1±3.4%, n=3). An increase in Ter119^CD71^ cells was observed for MigR1-FHL1 mice compared to MigR1 control mice (mean of 22.8±2% vs 41.4±6.1%, n=3).
5.4 DISCUSSION

Elevated levels of ALDH1A1, a retinoic acid-synthesising enzyme, have been detected in most primitive human haematopoietic progenitor cells (Kastan et al., 1990, Jones et al., 1995, Storms et al., 1999), while ALDH1A1 has been reported to be downregulated in a number of tumour cell lines and in murine hepatocellular and lung tumours (Banfi et al., 1994, Dragani et al., 1996). Despite this, the role of ALDH1A1 in haematopoiesis and leukaemic transformation remains largely uncharacterised. Indeed, although there have been numerous studies investigating the effects of exogenously applied RA on haematopoietic differentiation under in vitro culture conditions (Gratas et al., 1993, Smeland et al., 1994, Douer et al., 2000, Purton et al., 2000), there have been no such studies on the effects of modulating RA availability in vivo, via the induction or repression of RA synthesising enzymes. Similarly, although the role of retinoic acid receptors (RARα, β, and γ and their isoforms) and retinoid X receptors (RXRα, β and γ) in haematopoiesis and associated malignancies involving dysregulation of specific target genes by dominant negative mutant forms of these receptors has been studied in great detail (Pandolfi et al., 1991, Pitha-Rowe et al., 2003), very little attention has been paid to the involvement of other aspects of retinoid metabolism in tumorigenesis. It is plausible that dysregulation of RA synthesis, which is tightly controlled by a select group of enzymes including ALDH1A1, as well as the closely related ALDH1A2 (RALDH2), may interfere with processes regulating cellular differentiation and proliferation, thereby contributing to the tumorigenic phenotype. In a separate role, ALDH1A1 may also affect the drug sensitivity of specific tumours to chemotherapeutic agents, since aldehyde dehydrogenases catalyse the oxidation of the drug metabolite aldophosphamid to the inactive form, carboxyphosphamid (Bunting et al., 1994).

In this thesis, ALDH1A1 was confirmed to be functionally linked to HOX11 as a subordinate target gene. We therefore assessed the in vivo effects of dysregulated ALDH1A1-expression on the differentiation of the various haematopoietic lineages at early progenitor and/or mature stages of development. This was achieved by transplanting lethally irradiated mice with bone marrow cells transduced with an ALDH1A1-expressing retrovirus (MigR1-ALDH1A1). The differentiation profiles of the myeloid, erythroid and lymphoid (T-cell and B-cell) lineages were then examined by
immunophenotypic analysis of cells taken from test (MigR1-ALDH1A1) and control (MigR1) mice. When analysed one month post-transplant, primary MigR1-ALDH1A1 recipients had a significantly higher number of myeloid cells in samples obtained from spleen and bone marrow, compared to control MigR1 animals (Figure 5.6A, B). Since the overall number of myeloid cells increased in both tissues, albeit by different magnitudes, it is possible that overexpression of ALDH1A1 induces the proliferation/differentiation of myeloid precursors. Presumably, given the known physiological function of this enzyme, this effect is mediated via altered RA synthesis. Indeed, the effects of RA on the granulocytic lineage are well documented, with studies by Douer and Koeffler (1982) revealing that the addition of supraphysiological concentrations of ATRA (micromolar range) to unseparated human bone marrow mononuclear cell cultures, increased the number of myeloid progenitor CFU-GM colonies induced by colony stimulating factors by approximately 2-fold. Later studies by the same group using CD34+ enriched bone marrow cells, confirmed that ATRA stimulates the growth of myeloid progenitors from less mature CFU-GMs in the presence of IL-3 and GM-CSF colony stimulating factors, and accelerates the terminal granulocytic maturation of these progenitors (Douer et al., 2000). Interestingly, ATRA was found to inhibit the growth of more mature G-CSF-induced myeloid progenitors from purified CD34+ cells stimulated with various cytokines, suggesting that the growth stimulating and differentiation properties of ATRA are targeted at less mature myeloid progenitors and decreases as cells begin to differentiate (Tohda et al., 1991, Smeland et al., 1994). Although the mechanism by which ATRA increases the number of early myeloid progenitors in culture is not known, the ability of ATRA to shift the differentiation of human foetal liver CD34+ cells from a mixed/erythroid/monocytic pathway towards the granulocytic lineage, raises the possibility that ATRA may redirect multi-lineage progenitors to alternative programmes of differentiation (Tocci et al., 1996).

The overexpression of ALDH1A1 in haematopoietic progenitors was also associated with reduced numbers of CD19/B220 double positive B-cells (~2.5-fold) in the spleen and bone marrow of MigR1-ALDH1A1 mice compared to MigR1 controls (Figure 5.7A, B). The inhibitory effect of constitutive ALDH1A1 expression observed in this study is supported by studies in which exogenous application of physiologic levels (10pmol-
10nM) of all-trans or 9-cis RA resulted in growth inhibition of human (CD19+IgM-) and murine (Lin-B220+) B-cell precursors in vitro, by more than 50% (Fahlman et al., 1995). The inhibition of B-cell proliferation by RA was not due to toxicity, since the effect was reversible after 24h, even after exposure to concentrations as high as 1uM. Moreover, Blomhoff et al., (1992) demonstrated that incubation of purified human B-cells with RA [30nM-3uM], led to growth inhibition of normal B-cells (as estimated by incorporation of [³H]-thymidine during DNA synthesis) without affecting cell viability. Analysis of cell cycle events known to occur at distinct points in the G1 phase of B-cell activation revealed that RA inhibits lymphoid cell-cycle progression at the mid-G1 phase (Blomhoff et al., 1992), by affecting components of the cell-cycle machinery at the transcriptional level, via nuclear RARs (Naderi & Blomhoff, 1999). Specifically, RA was shown to downregulate cyclins E and A, which are required for the activation of the late G1 phase cyclin dependent kinase, CDK2 and to induce the expression of the cyclin kinase inhibitor (CKI) p21Cipl (Naderi & Blomhoff, 1999). These events are likely to account for the growth inhibition by RA, since cyclin dependent kinases are key regulators of pRB phosphorylation, which is a critical event in G1/S progression. In addition, application of RA appeared to reduce the number of S.aureus plaque-forming cells, indicating that retinoids may affect normal B-cell differentiation as well as proliferation (Blomhoff et al., 1992). In a separate study, ATRA was shown to inhibit the in vitro differentiation of purified B-cells from patients with macroglobulinemia, which spontaneously differentiate in vitro to plasma cells (Levy et al., 1994). Taken together, these results indicate an important role for vitamin A and its biologically active derivatives in the regulation of B-cell lymphopoiesis.

As an alternative to the hypothesis that ALDH1A1 induces the expansion of myeloid progenitors and inhibits the differentiation and growth of B-cells, it is possible that ALDH1A1 may affect the commitment of haematopoietic stem cells to either the myeloid (CMP) or lymphoid (CLP) lineages, thereby accounting for an increase in the myeloid compartment at the expense of the lymphoid compartment. In support of this hypothesis, the percentage of GFP-expressing cells was reduced in cells taken from the thymus of MigR1-ALDH1A1 mice compared to spleen and bone marrow samples, suggesting that overexpression of ALDH1A1 may be associated with a quantitative reduction in the T-cell compartment. Such a role has previously been reported for the
stem cell leukaemia protein (SCL/TAL1), which, like ALDH1A1, is also expressed in HSCs, and is essential for haematopoietic cell differentiation into myeloid, erythroid and megakaryocyte lineages (Hoang et al., 1996, Elwood et al., 1998, Kunisato et al., 2004). Kunisato et al., (2004) performed HSC transplantation experiments using stem cells retrovirally transduced with either wild type SCL or a dominant negative form of SCL. The overexpression of wild type SCL was associated with an increase in myelopoiesis (Mac1+/Gr1+) and concomitant decrease in lymphopoiesis (B220+, Thy1.2+), whereas in contrast, overexpression of the mutant form of SCL, corresponding to a downregulation of SCL, resulted a shift from myelopoiesis to lymphopoiesis (Kunisato et al., 2004). These results suggested that SCL is capable of influencing the commitment of a HSC to either the myeloid or lymphoid lineage, by controlling the generation of CMPs and CLPs (Kunisato et al., 2004). In order to more accurately determine whether ALDH1A1 affects the commitment of HSCs to myeloid or lymphoid lineages or alternatively, affects the expansion of myeloid progenitors, the ability of ALDH1A1-expressing HSCs to generate T-cells, B cells and myeloid cells might be assessed in future work using a multilineage progenitor (MLP) assay. The MLP assay is a clonal assay system which employs a modified foetal thymic organ culture (FTOC) capable of supporting the development of T, B and myeloid cells (Kawamoto et al., 1997). Thus, the developmental potential of individual haematopoietic progenitors towards T, B or myeloid lineages, following the transduction of HSCs with ALDH1A1-expressing retrovirus may be assessed. Given that the overexpression of SCL and ALDH1A1 resulted in an expansion of the myeloid compartment and a decrease in B220+ cells, it is tempting to speculate that these genes may regulate haematopoiesis via a common mechanism. Intriguingly, SCL has been shown to regulate the transcription of a gene encoding a second RA-synthesising enzyme, ALDH1A2, better known as RALDH2 (Ono et al., 1998). It therefore seems plausible that overexpression of SCL, and hence RALDH2, phenocopies the effect of overexpressing ALDH1A1. Furthermore, given that RALDH2 has been shown to be a target of a complex involving SCL in T-ALL cell line, HPB-ALL, and that the status of ALDH1A1 as a target gene of T-cell oncprotein HOX11 has been confirmed, it is possible that both SCL and HOX11 may affect the expression of RA synthesising enzymes when dysregulated in T-ALL, converging on a central theme for haematopoietic perturbation via altered RA signalling.
Intriguingly, overexpression of Hoxa9 in murine bone marrow transplantation assays results in an increase in mature granulocytes (Mac1+) and a partial block in B-cell differentiation at the pre-B progenitor stage (Thorsteinsdottir et al., 2002). This is similar to the effect demonstrated for ALDH1A1 in our study. It is interesting to note that ALDH1A1, Hoxa9 (Lawrence et al., 1997) and SCL are all expressed in primitive HSCs, and that HOX11 has been cloned from HSC mRNA (Moretti et al., 1994), and that overexpression of these genes in haematopoietic progenitors results in perturbations in the myeloid and B-cell compartments. Moreover, both SCL and HOXA9 genes are involved in chromosomal translocations in human leukaemias (Begley & Green, 1999, Nakamura et al., 1996), suggesting that genes responsible for regulating the functions of HSCs may also facilitate leukaemic transformation.

As stated above, the effect on haematopoietic differentiation demonstrated by the enforced expression of ALDH1A1 was most likely to have occurred via altered RA synthesis. Some support for this notion was provided by an analysis of GFP negative (i.e. ALDH1A1 negative) cells from mice that had been reconstituted with MSCV-ALDH1A1 bone marrow. This revealed that, similar to the GFP positive cells, GFP/ALDH1A1 negative samples taken from the spleen demonstrated enhanced myelopoiesis as compared to the corresponding cells of control mice (Figure 5.9). Such a trans-effect may be explained by the fact that RA is a diffusible morphogen. This would mean that cells lacking the ALDH1A1 transgene could still be affected to some degree by the RA produced by neighbouring transgene-expressing cells.

The finding that constitutive expression of HOX11 in murine haematopoietic precursors results in the immortalisation of cells belonging to the myeloid lineage (Hawley et al., 1994) and the discovery in the present study that overexpression of ALDH1A1, a downstream target of HOX11 also affects the myeloid compartment, is intriguing and adds weight to the argument that ALDH1A1 may represent not only a physiological target in spleen development, but possibly an oncogenic target. Indeed, the well-documented role of retinoids in the development and disorders of haematopoietic cells makes ALDH1A1 a likely candidate for dysregulation, since the subversion of retinoid signalling pathways leading to the development of leukaemia is firmly established (reviewed by Oren et al., 2003). The question thus arises; if the transforming
capabilities of HOX11 derive from the ability of HOX11 to cause a block in myeloid differentiation, how is this reconciled with the stimulatory properties induced by overexpression of ALDH1A1? Although our results, in conjunction with studies investigating the effect of retinoids on normal haematopoiesis in vivo, suggest that RA stimulates granulopoiesis (van Bockstaele et al., 1993), there are also reports that RA inhibits the proliferation and differentiation of myeloid precursors (Bradley et al., 1983, Tohda et al., 1991). These studies highlight the pleiotrophic effects mediated by RA, which may be dependent on the maturational status of precursors present in the original population under investigation (Purton et al., 1999). The primary bone marrow cells utilised in these experiments represented multiple lineages at various stages of differentiation. As such, the effects of overexpressing ALDH1A1 may not be truly reflective of an in vivo situation whereby cells harbouring the HOX11 translocation resulting in altered ALDH1A1 expression, are blocked at a defined stage in development to yield a predictable phenotype. A detailed understanding of the effect of overexpressing ALDH1A1 in the various lineages at specific stages of development may yield insight into the potential for aberrant RA synthesis to cause differentiation or proliferative abnormalities associated with leukaemic transformation. Thus, in order to circumvent problems associated with using heterogenous bone marrow populations, future studies should involve transduction of ALDH1A1 containing retrovirus into purified populations of cells representing the different lineages/maturational stages in development, thereby yielding a clearer phenotype. Future work to test the hypothesis that HOX11 and ALDH1A1 are linked within a transcriptional hierarchy capable of perturbing haematopoiesis and possibly causing T-cell tumours, would ideally require a definitive murine model capable of recapitulating the tumorigenic capacity of HOX11 in T-ALL.

Despite reports that physiological levels of ATRA are capable of stimulating the proliferation of human peripheral blood T-cells (Naderi & Blomhoff, 1999), no changes in the T-cell profile between MSCVMigR1 and MSCVMigR1-ALDH1A1 mice were observed in these studies. Since a progressive maturation of T-cells in both control and test mice sacrificed on 38 days post transplantation (Replicate 3; Figure 5.5C), compared to mice analysed on 33 days post transplantation (Replicate 1; 5.5A) was evident, an observable T-cell phenotype may have been masked by the incomplete
regeneration of the T-cell compartment following engraftment with transduced bone marrow. Indeed, thymuses extracted 33 days post transplantation, were visibly smaller than those taken 38 days post transplantation, suggesting that thymic reconstitution was not complete within the first set of mice. It is possible that the lack of a discernable T-cell phenotype in MSCVMigR1-ALDH1A1 mice may be due to the requirement for deregulation of a second gene. Alternatively, it is possible that the failure of ALDH1A1 to perturb T-cell development, reflects the requirement for specific timing of overexpression or the level of ALDH1A1 expression. Indeed, preliminary studies by Izon et al., (unpublished observations), suggest that low level overexpression of HOX11 alone in murine bone marrow, is capable of inducing T-cell leukaemias. The use of an inducible retroviral vector allowing for the temporal and quantitative control of HOX11/ALDH1A1 expression, would therefore be useful in future overexpression analyses for studying human leukaemias in murine models. Thus, although these studies demonstrate that the overexpression of ALDH1A1 in primary murine bone marrow has the ability to perturb normal haematopoietic differentiation in vivo, resulting in increased myeloid differentiation and inhibition of B-cell development, the oncogenic relevance of ALDH1A1 as a target gene of HOX11 in a T-cell background has yet to be established.

Although the over-expression of ALDH1A1 in primary murine bone marrow is associated with phenotypic changes in both myeloid and B-cell lineages, the exact mechanism/s underlying these changes remains unknown. Since RA serves as a ligand for inducible transcriptional regulators (RARs, RXRs)(Giguere et al., 1987), analysis of the changes in expression of the various receptor subtypes and their isoforms, which are often auto-regulated in order to amplify the RA response, following over-expression of ALDH1A1, may yield insights into transcriptional networks perturbed by aberrant RA synthesis. Indeed, there is precedence for altered retinoic acid nuclear receptor activities in the development of cancer (Kakizuka et al., 1991, Dejean et al. 1986). One particular study demonstrated that RARβ is inducible by RA in a hormone-dependent breast cancer cell line and is responsible for mediating the growth inhibitory effects of retinoids, partly by inducing apoptosis, such that loss of RARβ may facilitate cancer development (Liu et al., 1996). These studies highlight the fact that alterations in the cellular complement of retinoid receptors, as regulated by RA, can result in abnormal
cellular differentiation and proliferation associated with tumorigenesis. Presumably, the aberrant signalling effected by RA involves the ligand-induced regulation of specific target genes by RAR-RXR heterodimers, and as such these studies would optimally be coupled with an investigation into genes differentially expressed following changes in *ALDH1A1* expression. Given the precedence for RA to affect cell cycle machinery and the proliferative capacity of lymphoid cells, it would be of interest to analyse the effects of over-expressing *ALDH1A1* on intracellular levels of cyclins, CKIs and the phosphorylation status of pRB.

In this study, only the acute effects of transgene expression on haematopoietic differentiation were assessed one month post-transplantation. This approach was advantageous from the perspective that high levels of transgene (as estimated by GFP expression levels) were still observable in transduced cells. However, ideally, the functional phenotype of transduced cells should also be assessed following long-term reconstitution, to allow irradiated animals sufficient time to re-establish haematopoietic homeostasis and to observe chronic effects (*e.g.* development of leukaemias). In order to ensure long-term reconstitution, the repopulation of lethally irradiated recipients with sufficient numbers of virus transduced long-term self-renewing haematopoietic stem cells (LT-HSC) is crucial. Studies by Klug *et al.*, (2000) demonstrated that as few as 30 LT-HSC are required to establish long-term haematopoiesis from donor derived cells. However, standard bone marrow reconstitution experiments result in the injection of significantly less LT-HSCs, which may be detrimental to the long-term repopulating ability of these stem cells, since overt differentiation pressure is placed on these cells in order to reconstitute a lethally irradiated recipient (Hawley *et al.*, 1996). Transduction of FACS-purified, LT-HSC (Sca-1<sup>+</sup>, Lin<sup>-</sup>, Thy-1.1<sup>lo</sup>, c-kit<sup>+</sup>) with retroviral supernatants and sorting of cells for GFP expression prior to injection into irradiated mice may improve the success of future long-term transplant experiments. This will minimise the risk of potentially oncogenic integration events into differentiated haematopoietic cells, which do not contribute to long-term repopulation, but may obscure a transgene-mediated phenotype (Klug *et al.*, 2000). Alternatively, the use of a novel lentiviral vector with a hybrid MSCV/HIV LTR, capable of transducing human CD34+ cells with greater efficiency than an MSCV-based vector when assayed with normalised viral titres under the same transduction conditions may also improve the success of long term
transplant experiments. This is because HIV-based lentiviral vectors are capable of transducing non-dividing cells and are therefore more suited to slowly dividing haematopoietic stem cells (Choi et al., 2001). Although standard HIV vectors driven from an internal cytomegalovirus (CMV) promoter express transgenes 100-1000 fold less than MSCV based vectors, the hybrid vector containing the MSCV/HIV LTR expresses the transgene 10-100 fold higher than the original HIV-based vector and the researcher may choose to forfeit lower transgene expression for sustained expression (Choi et al., 2001).

In contrast to the effects observed by ALDH1A1, the overexpression of FHL1 in primary murine bone marrow was not associated with perturbations in either the myeloid or lymphoid lineages. The expression of FHL1 was, however, associated with a small, yet reproducible increase in the erythroid compartment when compared to the vector control. It is possible that FHL1 represents a physiological target gene of HOX11, and may not be associated with the tumorigenic function of HOX11. Intriguingly, however, FHL1 was shown to be expressed in 7 out of a panel of 12 T-ALL cell lines tested by Northern blot, which is noteworthy considering that FHL1 is normally expressed at low levels in lymphoid tissues (Greene et al., 1998, Greene et al., 1999). Moreover, there is precedence for the involvement of another family of LIM-domain proteins in T-ALL, namely, LMO1 and LMO2 (reviewed by Sanchez-Garcia & Rabbitts, 1993). Despite these interesting observations, the relationship between HOX11 and FHL1 in T-ALL remains unclear, and further work is required to establish the relevance of FHL1 as a target gene of HOX11, in either a physiological or oncogenic setting.

In summary, the work described in this Chapter was aimed at assessing whether ALDH1A1 and FHL1, two genes transcriptionally linked to HOX11 (Greene et al., 1998) could mimic the phenotypic effects previously observed for HOX11 on murine haematopoiesis (Hawley et al., 1994). Retroviral transduction of primary bone marrow cells coupled with in vivo reconstitution was chosen for this task, as it is a powerful method for assessing the effects of oncogenes on haematopoietic differentiation (Hawley et al., 1994). Although the data presented in this Chapter are only suggestive of ALDH1A1 being able to account for the ability of HOX11 to perturb haematopoiesis
and cause cellular immortalisation, the finding that enforced expression of \textit{ALDH1A1} in bone marrow cells produces phenotypic effects reminiscent of that previously reported for \textit{HOX11}, is an important first step in addressing this question.
Chapter 6

General Discussion

Despite the increasing body of evidence supporting the role of HOX11 as a transcription factor involved in the genesis of T-ALL, the mechanism(s) by which HOX11 induces T-cell tumours remain enigmatic (Kees et al., 2003). It is likely that the dysregulation of downstream target genes controlling processes of cellular differentiation and proliferation, accounts for the ability of HOX11 to potentiate cellular transformation and ultimately leukaemogenesis. As such, the identification and characterisation of relevant target genes is essential if our knowledge of HOX11 as a developmental regulator and a nuclear oncogene is to increase. Class I Aldehyde Dehydrogenase (Aldh1a1) and Four and a Half Lim Domain Protein I (Fhl1), were identified as genes transcriptionally upregulated by HOX11 in the murine NIH 3T3 fibroblast cell line (Greene et al., 1998). Aldh1a1 was also shown to be physiologically regulated by Hox11 in the developing spleen (Greene et al., 1998). Although HOX11 has been shown to affect the endogenous expression of ALDH1A1 in different cellular milieu (Greene et al., 2002, Greene, unpublished observations), the ability of HOX11 to specifically regulate the transcription of ALDH1A1 and FHL1 has not been demonstrated.

The first aim of this study, therefore, was to investigate the transcriptional regulation of ALDH1A1 and FHL1 by HOX11, by identifying HOX11 responsive elements within the 5’ regulatory regions of these genes. Such information would not only serve to validate the status of ALDH1A1 and FHL1 as target genes of HOX11, providing an insight into the complex gene networks controlled by HOX11, but would also increase our knowledge of the transcriptional properties of HOX11 and the mechanisms of gene regulation utilised by this T-cell oncprotein. A complete understanding of the role of HOX11 and downstream target genes in normal blood cell development is also paramount to our understanding of the functions disrupted by HOX11 in T-ALL. A second major objective of this study, therefore, was to assess the ability of ALDH1A1 and FHL1 to perturb haematopoiesis, in order to determine whether either of these
target genes might account for the ability of HOX11 to predispose cells to leukaemic transformation.

The identification of ALDH1A1 as a putative target gene of HOX11 is of interest, given the role of ALDH1A1 in regulating the delicate balance of retinoic acid, a potent modulator of cell growth, differentiation and apoptosis, within the cell. The ability of HOX11 to modulate ALDH1A1 promoter activity was assessed by luciferase reporter assays, following transient introduction of nested deletions of the ALDH1A1 promoter (ranging from -2159 to +1/+42bp) and a HOX11 expression vector into either the T-ALL cell line, PER-117, or the human erythroleukaemic cell line, HEL (Chapter 2). In the case of PER-117, HOX11 repressed transcription for promoter constructs ranging between -2159 to -91/+42bp. Deletion of a functional CCAAT box (-74/-70bp) resulted not only in a massive decline of promoter activity, but was also associated with a loss of HOX11-mediated repression. Given that the CCAAT box is the primary cis-regulatory element within this region, we hypothesised that the observed repression of ALDH1A1 by HOX11 occurred through the CCAAT box, via either direct (involving HOX11 itself) or indirect (involving transcription factors whose expression is affected by HOX11 higher in a transcriptional hierarchy) mechanisms. This may occur by a passive mechanism, involving the competition for transcription factor binding sites (occlusion), or through active mechanisms, via the specific inhibition of the formation of a functional preinitiation complex, inhibition of adjacently bound transcriptional activators (quenching), or recruitment of repressor complexes including histone deacetylases (HDACs), which modify local chromatin structure causing a reduction in gene expression.

Given that the expression of HOX11 in PER-117 was associated with the loss of a protein binding complex at the CCAAT box (complex A) and at the putative GATA box (-34/-28bp)(complexes I and J), and that inclusion of HOX11 antibody was associated with a quantitative increase in complex I, it was hypothesised that HOX11 directly represses transcription by interfering with trans-acting factor(s) at the CCAAT box (passive repression), and with components of the basal transcriptional machinery at the GATA box (active repression). In support of a model whereby HOX11 displaces CCAAT box binding transcription factors, a DNA-binding site selection assay
performed by Zhang et al., (1999), using in vitro translated HOX11, enriched a consensus sequence (TGGCANNNGCCAA) for the CCAAT-box binding transcription factor (CTF). Since no sequences corresponding to a consensus HOX11 binding site were obtained, this group suggested that HOX11 physically interacted with CTF in the binding assay, and later demonstrated that HOX11 and CTF1 did indeed associate in vitro and in vivo (Zhang et al., 1999). However, it is also possible that this site represents a novel HOX11 recognition sequence, in which case HOX11 may mediate repression by occlusion of CTF1 (or other CCAAT box binding factors) at the CCAAT box. The direct interaction of HOX11 with members of the basal transcriptional machinery has also been documented, with Owens et al., (2003) demonstrating the ability of HOX11 to repress transcription from the SV40, HSVtk, CMV and neu promoters by interacting with the RNA polymerase II holoenzyme. In accordance with this, evidence was provided to suggest that HOX11 physically interacts with TFIIB, a member of the basal transcriptional machinery, that functions in the correct positioning of RNA polymerase II at promoter DNA (Chapter 4, Heidari, unpublished observations). Consistent with our mutant (HOX11ΔH3) studies, Owens et al., (2003) also demonstrated that the DNA binding helix of HOX11 was required for transcriptional repression. Since homeodomain point mutations previously shown to abrogate or alter DNA binding were capable of repressing transcription as effectively as wild type HOX11, however, Owens et al., (2003) postulated that repression was mediated via protein-protein interactions as opposed to a mechanism requiring sequence specific DNA binding. In support of their finding, the interaction between HOX11 and TFIIB was significantly reduced upon deletion of the DNA binding helix in our in vitro GST-pulldown assays (Chapter 4). Future studies to address the role of the HOX11 homeodomain in transcriptional repression, specifically the DNA binding helix 3, may clarify the exact requirements for HOX11-mediated repression. Thus, HOX11 appears to repress ALDH1A1 transcription in PER-117 cells by inhibiting/enhancing the formation of specific complexes at the CCAAT and GATA boxes resident within the proximal promoter of this gene. Although the homeodomain recognition helix is required for this activity, it is unlikely to be due to the need for HOX11 to bind DNA since no new complexes were formed in the presence of HOX11. Rather, Helix 3 appeared to be required for protein-protein interactions, as was the case for TFIIB. The likelihood that TFIIB assembles at the GATA box is strengthened by the presence of a
TFIIB recognition element (BRE)(SSRCGCC) immediately in front of this site (Lagrange et al., 1998).

BRE Consensus  S S R C G C C T A T A
ALDH1A1         C C G T G C A G A T A

Figure 6.1. The Human ALDH1A1 Promoter Contains a Sequence Resembling a TFIIB Recognition Element (BRE). The putative TFIIB recognition element lies immediately in front of the TATA/GATA box, denoted by green letters. Nucleotide matches are denoted by red letters.

The ability of HOX11 to regulate ALDH1A1 promoter activity in the HEL cell line was also assessed by transient luciferase reporter assays. In contrast to the repression observed for PER-117, HOX11 weakly activated the ALDH1A1 promoter within +1/+42bp, in the HEL cell line. Given that HOX11 was also capable of activating transcription from the empty pGL3Basic luciferase vector, which has been specifically engineered to prevent read-through of transcription from cryptic elements within the vector backbone, we hypothesised that the activation function of HOX11 may occur independently of DNA binding, again possibly via protein-protein interactions with members of the basal transcriptional machinery. This activation was sustained for promoter constructs up to -91/+42bp, however the addition of 55bp of 5′ promoter sequence (-146/+42bp) resulted in a decline in promoter activity. This suggested that the region between -146 to -91bp binds a factor capable of exerting a negative regulatory influence on HOX11.

These studies highlight the potential for HOX11 to act as both a transcriptional repressor and a transcriptional activator, in a manner apparently dependent on both the cell-type specific transcription factors available. HOX11 may also be subject to different post-translational modifications (e.g. phosphorylation) depending on the cell background, thereby accounting for the differential action observed between PER-117 and HEL. Intriguingly, Western blot analyses of HOX11-transfected cell lysates to verify the expression of full length HOX11 in transient transfection assays identified an additional band of lower mobility than the major HOX11 band in HEL cells. This additional band was not present in HOX11 negative cell lysates nor in the PER-117 lysates. Future experiments to identify this band, which might represent a post-translationally modified of HOX11, may reveal a mechanism for the dual functionality
demonstrated by HOX11 in transient assays. It is well documented, for example, that phosphorylation status may affect the transcriptional properties of regulatory proteins (Yokoyama et al., 2002). The use of a HOX11 mutant lacking the third helix of the homeodomain revealed further differences between the activation and repressor functions of HOX11; in that repression was dependent on the DNA binding helix, whereas activation occurred independently of this domain. Moreover, deletion of this region switched HOX11 from being a repressor into a potent activator of transcription. Taken together, these results indicate that transcriptional repression may occur via a mechanism requiring DNA binding, whereas activation may occur in a DNA-binding independent manner, possibly via protein-protein interactions. Alternatively, an intact homeodomain may be required for interactions with specific cofactors, as appears to be the case for TFIIB.

Notably, transient transfections assays to assess ALDH1A1 promoter activity in PER-117 and HEL cell lines consistently yielded an effect opposite to that of the endogenous ALDH1A1 gene. For example ALDH1A1 expression is upregulated by HOX11 at the level of PCR in the PER-117 T-cell line, however luciferase reporter assays indicated that HOX11 represses the proximal ALDH1A1 promoter when analysed in isolation. Conversely, ALDH1A1 expression is downregulated by HOX11 in HEL cell line, as assessed by Northern blot, contrasting with the activation observed in luciferase reporter assays. The reasons for this discrepancy are likely due to the inherent limitations associated with transient luciferase reporter assays. Specifically, the introduction of multiple copies of the promoter sequence under investigation (ALDH1A1) and protein expression vector (HOX11, HOX11ΔH3) far exceeds the in vivo cellular contextual requirements for endogenous gene expression. This is likely to affect the delicate ratio of cofactors within the cell, which in turn may affect specific DNA-binding capabilities and interactions with members of the basal transcriptional machinery (Mercola et al., 1985). In addition, distal regulatory elements not included within the luciferase reporter constructs, may be required to recapitulate the effect of HOX11 on endogenous gene expression. Finally, it should be noted that the episomal nature of the plasmid reporter constructs mean that the promoter sequences under investigation are constrained within an artificial configuration that may not be subject to chromatin specific mechanisms of gene regulation (Smith & Hager, 1997). As such, an extension of these promoter studies
might involve the analysis of specific promoter constructs in a stable transfection system, which would circumvent artefacts associated with artificial configuration and high copy numbers (Smith & Hager, 1997). There are, however, still limitations associated with stable transfection assays, particularly those involving the integration site of the reporter gene construct, which may affect the accessibility of the promoter sequence to transcription factors required for normal gene regulation. As such, a combination of both types of assay may be necessary to definitively assess regions of the ALDH1A1 promoter important for cell-type specific regulation and responsiveness to HOX11.

The role of HOX11 in T-ALL, via the deregulation of ALDH1A1 remains unclear. The repression of Aldh1a1 by Hox11 in splenic precursor cells appears requisite for normal splenogenesis, and it has been postulated that the inappropriate expression of Aldh1a1 in the spleen primordium of Hox11 null mice may contribute to splenic involution, possibly via pathway of programmed cell death, since the ability of retinoids to inhibit cell survival by inducing apoptosis has been previously demonstrated (Greene et al., 1998). More recent studies by Kanzler and Dear (2001), however, suggest that splenic involution observed in Hox11+/− mice is not due to apoptosis, but rather an arrest in spleen development, with splenic precursors contributing to surrounding tissues, and in this context it is also possible that changes in cellular levels of RA, a crucial morphogen, may be responsible for altered cell fate. In the context of T-cell development, which involves the positive selection and ‘default cell death pathway’ of T-cell clones in the thymus, RA may act as either an inducer or an inhibitor of TCR/CD3-mediated apoptosis, depending on the levels of 9-cis RA, which neutralises the inhibitory effect of RARα on activation-induced apoptosis (Iwata et al., 1992, Szondy et al., 1997). Since the conversion of all-trans RA to 9-cis RA only occurs at high cellular concentrations of RA, the dysregulation of ALDH1A1, hence perturbation in RA levels in developing T-cells, may therefore affect T-lineage apoptotic programmes or differentiation programmes, thereby predisposing cells to leukaemic transformation.

With respect to the link between changes in ALDH1A1 expression and cellular immortalisation associated with leukaemic transformation, Owens et al., (2003) suggest that since there are crucial differences between domains of HOX11 required for the
activation of Aldh1a1 in NIH 3T3 cells and those required for the immortalisation function of HOX11, Aldh1a1 is unlikely to represent a target gene involved in leukaemogenesis. Considering that HOX11 has been shown to differentially affect the expression of endogenous levels of ALDHI1A1 in the HEL cell line (downregulation) as compared to the PER-117 cell line (upregulation), and given the differential requirement for the DNA binding helix for activation and repression activities, it is likely that HOX11 utilises distinct mechanism(s) to regulate ALDHI1A1 expression. Therefore the regions of HOX11 involved in Aldh1a1 upregulation in NIH 3T3 cells are possibly different to those involved in transcriptional repression demonstrated in the PER-117 model, neither of which necessarily reflect a situation whereby HOX11 is deregulated in leukaemic T-cell precursors. The relationship between HOX11 overexpression and ALDHI1A1 gene expression in relation to T-ALL remains unclear, however. Although ALDHI1A1 upregulation was not observed in a small panel of T-ALL cell lines expressing HOX11, at the level of Northern blot (Greene et al., 1998), it remains unclear whether HOX11 activates or represses ALDHI1A1 in a tumorigenic setting, and by what magnitude. Moreover, evolution of in vitro cultured cell lines may complicate gene expression studies. Future work using primary patient T-ALL tumour specimens should help to clarify whether HOX11 or ALDHI1A1 expression patterns are correlated.

The second putative target investigated in this study, FHL1, previously known as SLIM1 (Striated Muscle LIM protein 1), encodes a LIM domain protein required for muscle differentiation and function (Brown et al., 1999). Following the identification of the transcriptional start site of FHL1, the effect of HOX11 on the 5’ regulatory region was undertaken using transient luciferase reporter assays. These studies indicated that HOX11 was capable of repressing FHL1 promoter activity in the PER-117 cell line, supporting the finding that FHL1 is indeed regulated by HOX11. Given the presence of multiple Sp1 binding sites in the proximal region of the promoter (-928/+73bp), it is possible that HOX11 may interfere with Sp1-mediated transactivation. Precedent for homeodomain proteins to suppress Sp-1 dependent transcription has been demonstrated for transcription factors Msx-1 (Shetty et al., 1999) and Hoxa11 (Suzuki et al., 2003), however further experiments to determine the specific mechanism(s) by which HOX11 represses FHL1 transcription are required.
With the studies described above indicating that both ALDH1A1 and FHL1 are regulated by HOX11, the next step was to determine whether a link exists between these target genes and leukaemic transformation. Initial investigations to evaluate the ability of ALDH1A1 and FHL1 to perturb murine haematopoiesis were therefore undertaken (Chapter 5). The reconstitution of lethally irradiated recipients with primary murine bone marrow overexpressing ALDH1A1 resulted in a significant increase in myelopoiesis and a suppression of B and T-lymphopoiesis, compared to control mice. This may reflect the ability of RA to stimulate the growth and differentiation of myeloid progenitors from less mature CFU-GMs (Douer et al., 2000) or the inhibitory effect of RA on B-cell precursors (Fahlman et al., 1995). Alternatively, the observed increase in the myeloid compartment and decrease in the lymphoid compartment in MigR1-ALDH1A1 mice may reflect the ability of RA to affect the commitment of HSCs to either the myeloid (CMP) or lymphoid (CLP) lineage. Indeed, such a role has previously been described for the Stem Cell Leukaemia (SCL/TAL1) protein, which has been shown to regulate the transcription of ALDH1A2/RALDH2, a second RA-synthesising enzyme. This highlights the possibility of a central theme for haematopoietic perturbation via altered RA synthesis. It was also noteworthy that the effect of ALDH1A1 overexpression on the myeloid compartment was reminiscent of the abnormal myeloid cell differentiation observed following the overexpression of HOX11 in primary murine bone marrow (Hawley et al., 1994). However, further investigations are required to establish whether dysregulated retinoic acid metabolism following the ectopic expression of HOX11 is capable of predisposing T-cells to malignant transformation. Unlike ALDH1A1, the overexpression of FHL1 was not associated with perturbations in myelopoiesis or lymphopoiesis, suggesting that FHL1 may represent a physiological, as opposed to an oncogenic target gene of HOX11.

To date, attempts to recapitulate the tumorigenic activity of HOX11 in murine T-cells have been unsuccessful. It is possible that HOX11 requires the deregulation of a second gene, perhaps one encoding a protein partner, since the transforming function of HOX11 most likely involves the collaboration with protein cofactors to modulate oncogenically relevant target genes. The finding that HOX11 is capable of associating in vitro with the TALE homeodomain family members, MEIS1 and MEIS2A, provides a starting point for investigating mechanisms of HOX11-mediated transcriptional
regulation. In addition, the novel discovery that HOX11 homodimerises in vitro, provides an alternative mechanism by which HOX11 may regulate gene transcription. For example, HOX11 may act similarly to the Drosophila transcription factor Kruppel, which at low concentration, exists as a monomer with activator properties, whereas at higher concentrations, Kruppel forms homodimers with repressor function (Sauer & Jackle, 1993). The development of a murine T-cell model that would enable the coexpression of HOX11 with several candidate protein partners, or the overexpression of putative downstream target genes, is therefore necessary to identify which of these candidates is oncogenically relevant in HOX11-mediated tumorigenesis. The generation of inducible T-cell transgenic mice would provide a valuable in vivo model to more accurately determine which candidate protein partners or target genes may account for the oncogenic effect of HOX11. Alternatively, retroviral foetal thymic organ culture (FTOC), which involves the in vitro culture of foetal thymuses, which have been reconstituted with foetal liver derived haematopoietic stem cells retrovirally transduced with the gene(s) of interest and subsequent analyses of T-cell development by flow cytometry, may be employed. This technique has the advantage of speed (high throughput) and is less expensive than current transgenic options.

Conversely, the failure of HOX11 to cause T-cell tumours in murine models involving the overexpression of HOX11 in bone marrow (Hawley et al., 1994) and transgenic mice where HOX11 was targeted to the thymus using CD2 regulatory elements (Greene & Rabbitts, unpublished observations), may reflect the requirement for specific expression parameters not achieved in these models. This may relate to the timing of overexpression or the level of HOX11 expression. Indeed, preliminary bone marrow transplant experiments by Izon et al., (unpublished observations), suggest that low level overexpression of HOX11 alone in murine bone marrow is capable of inducing T-cell leukaemias. This is in keeping with the high proportion of T-ALL specimens that have been reported to express HOX11 at a low level (Kees et al., 2003), although an explanation is still required for the high levels of HOX11 expression observed in those tumours with a translocation involving the HOX11 locus. The requirement for low-level expression for the induction of relevant human leukaemias in mice has also been documented for other oncogenes including PML-RARα and TEL/ABL (reviewed by Ren, 2004). The use of an inducible retroviral vector allowing for the temporal and
quantitative control of gene expression, would therefore be useful in future overexpression analyses for studying human leukaemias in murine models.

The identification of relevant target genes is crucial if our understanding of HOX11 as a regulator of normal developmental processes including spleen organogenesis, and as a nuclear oncogene, implied by its overexpression in a significant proportion of patients with T-ALL and ability to immortalise haematopoietic progenitors in vitro, is to increase. To this end, the use of screening techniques such as cDNA representational difference analysis (RDA), has aided the identification of novel candidate target genes (Greene et al., 1998). More recently, the use of conventional cDNA microarray analysis to assess the network of genes affected by transcription factors deregulated in the leukaemic transformation process, has aided the classification of various leukaemia subtypes, in addition to providing an insight into the common pathways disrupted in many leukaemias (Ferrando & Look, 2003). Such studies do not enrich for direct targets of the overexpressed transcription factor, however, and the altered gene expression profiles often represent global cytoskeletal and metabolic alterations in the cell’s physiology that provide little insight into the original pathways disrupted in hierarchical order. Although the functional promoter analyses performed in this study confirmed the status of ALDH1A1 and FHL1 as genes transcriptionally regulated by HOX11, it is not known whether this is a result of direct interactions between HOX11 and cis-regulatory elements/trans-acting proteins within the ALDH1A1 promoter, or an indirect effect of HOX11 on genes higher in a transcriptional hierarchy. Given that ALDH1A1 expression is modulated by HOX11 in numerous cellular systems ranging from murine spleen to immortalised cell lines of both erythroid and T-cell origins, it is likely that ALDH1A1 represents a direct target of HOX11, however additional experiments, for example, chromatin immunoprecipitation (ChIP), which involves PCR amplification of specific regulatory regions bound by the transcription factor of interest, should clarify the status of ALDH1A1 as a direct target gene of HOX11.

Future studies of HOX11 function are likely to benefit from the identification and characterisation of direct targets, which provide a necessary platform upon which to study specific cis-regulatory elements and transcriptional complexes involving protein-protein interactions between HOX11 and cofactors. Such studies will provide necessary insights into HOX11 function in either normal or leukaemic settings. Although
ALDH1A1 may represent one such target, ChIP-CHIP, which incorporates current technologies of Affymetrix arrays in combination with ChIP, may be used to isolate novel, direct target genes of HOX11. This method involves chromatin immunoprecipitation of HOX11-DNA complexes and hybridisation of labelled, immunopurified DNA to a CpG island microarray CHIP. The advantage of this approach is that it can be used to identify HOX11-regulated promoters that harbour both consensus and non-consensus binding sites, therefore providing an unbiased approach in the elucidation of new, direct targets. The success of this technique, which is designed to increase the speed and number of significant targets isolated, has been demonstrated in the yeast system using yeast genomic arrays and now, with the development of mammalian CpG island arrays, should provide a high throughput method for the identification of in vivo target promoters (Ren et al. 2000, Cross et al., 1994).

In conclusion, the ability of HOX11 to repress the proximal promoters of both ALDH1A1 and FHL1 in PER-117 cells provides further evidence that these genes represent bona fide target genes of HOX11. In the case of ALDH1A1, HOX11 was shown to repress ALDH1A1 promoter activity, either directly or indirectly via interactions with activation complexes at a crucial cis-acting element, the CCAAT box (-74/-70bp). The expression of HOX11 was also associated with the disruption of specific complexes at the GATA box (-34/-29bp). Given the ability of HOX11 to interact with TFIIB and the presence of a TFIIB recognition element immediately 5’ upstream of the GATA box, it is possible that HOX11 represses transcription by interfering with members of a preinitiation complex on the ALDH1A1 promoter. Transcriptional repression and protein-protein interactions with TFIIB were dependent on homeodomain binding helix 3, and deletion of this region changed HOX11 from a repressor to a potent activator of transcription, suggesting that HOX11 requires DNA binding for repression or that this region is important for repressor-specific interactions.

In contrast to the repressive effect of HOX11 on the ALDH1A1 promoter in PER-117, HOX11 activated transcription from ALDH1A1 promoter constructs in addition to the empty pGL3Basic reporter vector in a DNA-binding independent manner in the HEL cell line. Thus, HOX11 is capable of acting as both an activator and a repressor of transcription, via distinct mechanisms, in a cell-type specific manner. Although the ability of HOX11 to predispose T-cells to malignant transformation by disrupting the balance of RA synthesis within the cell remains unclear, given the ability of ALDH1A1
to affect various aspects of murine haematopoiesis, it remains formally possible that \textit{ALDH1A1} represents an oncogenic target of HOX11. Conversely, although the \textit{FHL1} promoter appears to be transcriptionally regulated by HOX11, overexpression of \textit{FHL1} in murine bone marrow was not associated with perturbations in murine blood cell development, suggesting that \textit{FHL1} may represent a physiological target gene of HOX11.
REFERENCES


82. Duester, G. 1996. Involvement of Alcohol Dehydrogenase, Short Chain Dehydrogenase/Reductase, Aldehyde Dehydrogenase, and Cytochrome P450 in
the Control of Retinoid Signalling by Activation of Retinoic Acid Synthesis. Biochemistry \textbf{35}:12221-27.


133. Heller, H., Bengal, E. 1998. TFIIID (TBP) Stabilizes the Binding of MyoD to its DNA Site at the Promoter and MyoD Facilitates the Association of TFIIIB With the Preinitiation Complex. Nucleic Acids Research 26:2112-19.


311. Russo, J., A. Barnes, K. Berger, J. Desgrosellier, J. Henderson, A. Kanters, and L. Merkov. 2002. 4-(N,N-dipropylamino)benzaldehyde inhibits the
oxidation of all-trans retinal to all-trans retinoic acid by ALDH1A1, but not the differentiation of HL-60 promyelocytic leukemia cells exposed to all-trans retinal. BMC Pharmacol 2:4.


downstream of c-myc in a chromosomal 8;14 translocation in a cell line from a human T-cell leukemia. Proc Natl Acad Sci U S A 83:3439-43.


377. Tu, Y., F. Li, S. Goicoechea, and C. Wu. 1999. The LIM-only protein PINCH directly interacts with integrin-linked kinase and is recruited to integrin-rich sites in spreading cells. Mol Cell Biol 19:2425-34.
380. van den Berg, H., Zsiros, J., Veneberg, A., Schutten, N.J., Kroes, W., Slater, R.M., Behrendt, H. 1998. Favorable Outcome After 1-Year Treatment of
Chilhood T-Cell Lymphoma/T-Cell Acute Lymphoblastic Leukemia. Medical and Pediatric Oncology 30:46-51.


### PCR Programmes

<table>
<thead>
<tr>
<th>Programme</th>
<th>Cycles</th>
<th>Denaturation</th>
<th>Annealing</th>
<th>Elongation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extend 72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 2min</td>
<td></td>
<td>72°C, 1min</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 2min</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 2:40</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 3min</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 3:20</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 3:40</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 4min</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 4:20</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 5min</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 5:20</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 5:40</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 6min</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 6:20</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 6:40</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 7min</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 7:20</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 7:40</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 8min</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 8:20</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 10s</td>
<td>60°C, 30s</td>
<td>72°C, 8:40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5’RACE</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>94°C, 2min</td>
<td></td>
<td>72°C, 1min</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>94°C, 30s</td>
<td>52°C, 30s</td>
<td>72°C, 1min</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94°C, 30s</td>
<td>52°C, 30s</td>
<td>72°C, 7min</td>
</tr>
</tbody>
</table>

**5’RACE** - denaturation temperature 97°C

<table>
<thead>
<tr>
<th>SEQ1</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25</td>
<td>95°C, 30s</td>
<td>55°C, 36s</td>
<td>72°C, 84s</td>
</tr>
<tr>
<td>SEQKIM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>97°C, 1:40s</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>97°C, 40s</td>
<td>55°C, 1min</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>72°C, 1min</td>
<td>72°C, 7min</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TD1</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>94°C, 1:30s</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94°C, 30s</td>
<td>63°C, 1min</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94°C, 30s</td>
<td>62°C, 1min</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94°C, 30s</td>
<td>61°C, 1min</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94°C, 30s</td>
<td>60°C, 1min</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94°C, 30s</td>
<td>59°C, 1min</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94°C, 30s</td>
<td>58°C, 1min</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94°C, 30s</td>
<td>57°C, 1min</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94°C, 30s</td>
<td>56°C, 1min</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>94°C, 30s</td>
<td>55°C, 1min</td>
</tr>
</tbody>
</table>

**TD2** - denaturation temperature 97°C, final extension 72°C, 10min

320
Vector Constructs Maps

pCR®2.1-TOPO® Cloning Vector

pCR®2.1-TOPO®
3.9 kb

pCI-neo Mammalian Expression Vector.

(Genbank®/EMBL Accession Number: U47120)
pGL3Basic Luciferase Reporter Vector

(Genbank®/EMBL Accession Number: U47295)

pSV-β-Galactosidase Control Vector

(Genbank®/EMBL Accession Number: X65335)
### Monoclonal Antibodies Used For Flow Cytometric Analysis.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Tissue</th>
<th>Antibody</th>
<th>Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T-cell</strong></td>
<td>Thymus</td>
<td>CD4-APC</td>
<td>CD4 is a differentiation antigen expressed on most thymocytes and a subpopulation of mature T lymphocytes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8-PerCp</td>
<td>The CD8α chains form heterodimers with the CD8β chain on mature T lymphocytes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCRβi-Pe</td>
<td>The TCRβi chain is a component of the TCR receptor complex on T lymphocytes</td>
</tr>
<tr>
<td><strong>B-cell</strong></td>
<td>BM, Spleen</td>
<td>B220-Bio</td>
<td>B220 is a member of the Protein Tyrosine Phosphatase (PTP) family and is expressed on B lymphocytes from pro-B through mature and activated B-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD19-Pe</td>
<td>CD19 is a transmembrane glycoprotein and is expressed from pro-B cell through to mature B-cell stages; also reported to be more restricted to the B-cell lineage than B220</td>
</tr>
<tr>
<td><strong>Myeloid</strong></td>
<td>BM, Spleen</td>
<td>Mac1-Bio</td>
<td>Mac1 is an adhesion glycoprotein (165kDa) expressed on activated lymphocytes, monocytes, granulocytes and a subset of NK cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gr1-Pe</td>
<td>Gr1 is a myeloid differentiation antigen expressed on the monocyte lineage in the bone marrow and neutrophils eosinophils and monocytes in the periphery</td>
</tr>
<tr>
<td><strong>Erythroid</strong></td>
<td>BM, Spleen</td>
<td>Ter119-Bio</td>
<td>TER119 is expressed on erythroid cells from the early erythroblast through to mature erythroid stages in embryonic yolk sac, fetal liver, adult bone marrow, adult peripheral blood and adult lymphoid organs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD71-Pe</td>
<td>CD71 is a transferrin receptor. Although it is not erythroid specific, it is expressed at high levels in early erythroid precursors and levels decrease with maturation</td>
</tr>
<tr>
<td><strong>Haematopoietic Stem Cell</strong></td>
<td>BM, Spleen</td>
<td>Sca1-Pe</td>
<td>Sca1 is a phosphatidylinositol anchored protein (18kDa) expressed on the multipotent HSC found in adult bone marrow and fetal liver</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c-kit-APC</td>
<td>c-Kit is a trans-membrane tyrosine kinase receptor expressed on HSC, progenitors committed to the myeloid/erythroid lineages and precursors of T and B lymphocytes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LinNeg-Bio</td>
<td>A cocktail of monoclonal antibodies to exclude lineages with the following surface cell markers, CD3, CD4, CD8, Mac1, Ter119, B220.</td>
</tr>
</tbody>
</table>

**Streptavidin-Allophycocyanin (SA-APC)**  
**Streptavidin-PhycoerythrinCy5 (SA-PeCy5)**