The Isolation and Characterisation of murine CDP-diacylglycerol synthase genes, *Cds1* and *Cds2*.

Suzanne Louise Inglis

Department of Molecular Genetics, Institute of Ophthalmology, University College London, University of London

A thesis submitted to the University of London for the degree of Doctor of Philosophy (PhD) 2003
UMI Number: U602538

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.

UMI®
Dissertation Publishing

UMI U602538
Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code.

ProQuest
ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346
Abstract

Eye specific CDP-diacylglycerol synthase is a key regulator in the phototransduction pathway of the fruit fly *Drosophila melanogaster*. The enzyme is responsible for maintaining regeneration of phosphatidylinositol 4, 5 bisphosphate (PIP$_2$) which is required for normal vision. The *Drosophila eye-cds* mutants display light dependent retinal degeneration. Mammalian homologues of *Drosophila* mutants such as *eyeless* and *eyes absent* have been associated with mammalian disease. Based on these studies and the discovery of *CDS* expression in human retina, this gene would appear to be a good candidate for causing a retinal phenotype in mammals.

The aim of the work presented here has been to isolate and characterise the murine homologues of *eye-cds*. To this end both cDNA and genomic clones have been isolated using a combination of traditional molecular biology techniques and a bioinformatic approach. Using the identified clones, the structures of *Cds1* and *Cds2* have been confirmed. The clones were subsequently used to characterise the genes. In this thesis I have identified the expression pattern of both genes, showing that *Cds2* is ubiquitously expressed in both the embryonic and adult mouse whereas *Cds1* expression is more specific to the adult eye and brain. In addition, *Cds1* has been shown to be expressed only in the rod photoreceptor cells within the retina. *Cds1* and *Cds2* clones were also used to map the genes to chromosome 5E1 and 2G1 respectively. The remit of this thesis was further expanded to include the potential function of the murine genes. To this end, transgenic constructs of both genes were made for the rescue of the mutant *Drosophila* phenotype and a construct was prepared as a first step in the generation of a *Cds2* null mouse.

In summary, the study reported in this thesis has succeeded in isolating and characterising the murine *Cds1* and *Cds2* genes. The expression pattern of the two genes has been determined both in a large panel of tissues and within specific cell types. The chromosomal location of both genes has been determined and constructs made for use in developing transgenic mice and flies for the study of gene function.
DECLARATION

I declare that this thesis submitted for the Doctor of Philosophy is my own composition and that the data presented herein is my own original work, unless otherwise stated.

Suzanne Louise Inglis, B.Sc. (Hons).
Dedication

To Mum, Dad and Matt.
Acknowledgement

Thanks to the staff in the Department of Molecular Genetics at the Institute of Ophthalmology, University College London, for all their help and encouragement.

Thanks especially to my family, my extended family and my London urban family, your help, support and encouragement has been very greatly appreciated.

Most thanks go to Mum, Dad and Matt.
CONTENTS

ABREVIATIONS ..................................................................................................... 17

CHAPTER ONE - INTRODUCTION ....................................................................... 22

1.1 OCULAR STRUCTURE AND DEVELOPMENT ................................................. 22
  1.1.1 Drosophila ocular structure ................................................................... 22
  1.1.2 Drosophila ocular development ............................................................... 26
  1.1.3 Mammalian eye structure ..................................................................... 28
  1.1.4 Mammalian ocular development ............................................................ 37

1.2 PHOTOTRANSDUCTION .............................................................................. 40
  1.2.1 Drosophila phototransduction ................................................................ 40
  1.2.2 Mammalian phototransduction ............................................................... 50
  1.2.3 Comparison of Drosophila and mammalian phototransduction .......... 54
  1.2.4 Evidence for the presence of the Drosophila PIP\(_2\) pathway in mammals... 55
    1.2.4.1 rdgA .............................................................................................. 58
    1.2.4.2 rdgB .............................................................................................. 58
    1.2.4.3 rdgC .............................................................................................. 59
    1.2.4.4 norpA............................................................................................ 60
    1.2.4.5 eye-cds ...................................................................................... 61

1.3 RETINOPATHIES ...................................................................................... 63
  1.3.1 Retinitis Pigmentosa ............................................................................. 66
    1.3.1.1 Mode of inheritance ........................................................................ 66
    1.3.1.2 Animal models of Retinitis pigmentosa .......................................... 69
  1.3.2 Cone, cone-rod and macular dystrophy .................................................. 73
    1.3.2.1 Best's macular dystrophy ............................................................... 73
    1.3.2.2 Stargrdt's macular dystrophy ......................................................... 73
    1.3.2.3 Doyne's honeycomb macular dystrophy ......................................... 73
    1.3.2.4 Cone and cone-rod dystrophy ....................................................... 74
  1.3.3 Congenital stationary night blindness ..................................................... 75
    1.3.3.1 Description of CSNB ..................................................................... 75
    1.3.3.2 Genes and animal models associated with CSNB ......................... 75
CHAPTER TWO-MATERIALS AND METHODS .................................................98

2.1 MATERIALS .............................................................................................98
  2.1.1 Reagents ..............................................................................................98
  2.1.2 Vectors and Libraries .........................................................................98
  2.1.3 Bacterial strains and Competent cells .................................................98
  2.1.4 Kits ........................................................................................................98
  2.1.5 Oligonucleotide primers ......................................................................98

2.2 BUFFERS, SOLUTIONS AND MEDIA ......................................................99
  2.2.1 Media ....................................................................................................99
  2.2.2 Solutions and Buffers .........................................................................99
  2.2.3 In situ hybridisation buffers ................................................................101
  2.2.4 Protein buffers ...................................................................................102
  2.2.5 Gels .....................................................................................................104

2.3 METHODS .................................................................................................106
  2.3.1 DNA ....................................................................................................106
    2.3.1.1 Preparation of plasmid DNA (Miniprep) .......................................106
    2.3.1.2 Preparation of plasmid DNA (Maxiprep) ......................................106
    2.3.1.3 P1 Artificial Chromosome (PAC) Mini prep .................................107
    2.3.1.4 P1 Artificial Chromosome (PAC) Maxi prep ...............................107
    2.3.1.5 Phenol/Chloroform extraction ....................................................108
2.3.1.6 Genomic DNA preparation ............................................................... 108
2.3.1.7 Preparation of DNA from embryonic stem cells ............................. 109
2.3.1.8 Preparation of phage DNA ............................................................... 109
2.3.1.9 Restriction endonuclease digestions ................................................ 110
2.3.1.10 Agarose Gels and Visualisation ....................................................... 110
2.3.1.11 Southern blotting ............................................................................. 111
2.3.1.12 Cell preparations ............................................................................. 111
  2.3.1.12.1 Phage host cell preparation ........................................................ 111
  2.3.1.12.2 Preparation of DH5α cells ......................................................... 111
  2.3.1.12.3 Embryonic stem cell preparation ............................................... 111
2.3.1.13 Ligation of DNA fragments ............................................................. 112
2.3.1.14 Transformations ............................................................................... 113
2.3.1.15 Transfections of human embryonic kidney (HEK)-293T cells ......... 113
2.3.1.16 Polymerase chain reaction (PCR) ..................................................... 114
2.3.1.17 Colony PCR ...................................................................................... 114
  2.3.1.17.1 Cycle sequencing..................................................................... 115
  2.3.1.17.2 Direct sequencing .................................................................... 116
  2.3.1.17.3 Purification of extension products ............................................ 116
  2.3.1.17.4 Preparation of sequencing gels ................................................ 116
  2.3.1.17.5 Analysis of sequencing data .................................................... 117
2.3.1.18 Library screening .............................................................................. 117
  2.3.1.18.1 Phage libraries ........................................................................ 118
  2.3.1.18.2 Grided libraries ...................................................................... 118
2.3.1.19 Preparation of radioactive probes .................................................... 118
2.3.1.20 Hybridisation of filters .................................................................... 119
2.3.2 RNA ................................................................................................................ 119
  2.3.2.1 RNA extraction ............................................................................. 119
  2.3.2.2 mRNA preparation ........................................................................ 120
  2.3.2.3 RT-PCR ........................................................................................ 121
  2.3.2.4 Northern blotting ........................................................................ 121
  2.3.2.5 Riboprobe production ................................................................ 121
2.3.3 In situ hybridisation studies ................................................................. 123
2.3.3.1 *In situ* hybridisation ................................................................. 123
2.3.3.2 Fluorescent *in situ* hybridisation .................................................. 124
  2.3.3.2.1 Metaphase spreads .................................................................. 124
  2.3.3.2.2 Labelling and hybridisation ..................................................... 124
  2.3.3.2.3 Signal detection ....................................................................... 125
  2.3.3.2.4 Microscopy ............................................................................. 125

2.3.4 Protein ......................................................................................................... 125
  2.3.4.1 Tissue homogenisation .................................................................. 125
  2.3.4.2 Cell membrane preparation ........................................................... 125
  2.3.4.3 SDS-PAGE gels .......................................................................... 126
  2.3.4.4 Western blotting .......................................................................... 126
  2.3.4.5 Immunostaining of transfected cells .............................................. 127

CHAPTER THREE-RESULTS ........................................................................... 128

3.1 ISOLATION OF MURINE *Cds1* ............................................................... 128
  3.1.1 Identification of the cDNA clone 801-h9 ............................................ 128
  3.1.2 Identification of the cDNA clone 2279-d19 ....................................... 134
  3.1.3 Identification of the 5' end of mouse *Cds1* ....................................... 139
  3.1.4 Identification of murine *Cds1* 3'UTR ................................................. 145
  3.1.5 *Cds1* sequencing homology between species ..................................... 145
  3.1.6 Summary of section 3.1 ...................................................................... 148

3.2 ISOLATION OF MURINE *Cds2* ............................................................... 151
  3.2.1 Identification of cDNA clone 1500-M9 ............................................. 151
  3.2.2 Identification of cDNA clone 1247-c2 .............................................. 153
  3.2.3 *Cds2* sequence homology between species ........................................ 156
  3.2.4 Summary of section 3.2 ..................................................................... 156

3.3 GENOMIC STRUCTURE OF MURINE *Cds1* ........................................... 161
  3.3.1 Identification of a PAC clone representing *Cds1* ............................... 161
  3.3.2 Identification of *Cds1* intron/exon boundaries .................................... 161
  3.3.3 *Cds1* genomic structure ................................................................ 163
  3.3.4 Summary of section 3.3 ................................................................... 163

3.4 GENOMIC STRUCTURE OF MURINE *Cds2* .......................................... 168
  3.4.1 Identification of a genomic phage clone ............................................ 168
  3.4.2 Phage clone 20A ............................................................................. 168
  3.4.3 Identification of a PAC clone representing *Cds2* ............................... 172
3.4.4 Summary of section 3.4.1-3.4.3 ........................................................... 179
3.4.5 A comparison of the gene structure of Cds1 and Cds2 .................. 179
3.5 CHROMOSOMAL LOCALISATION OF Cds1 AND Cds2 .................. 185
  3.5.1 Bioinformatic mapping of Cds1 and Cds2 ................................. 185
  3.5.2 Previously mapped CDS genes ............................................. 185
  3.5.3 Fluorescence in situ hybridisation (FISH) mapping of Cds1 and Cds2 188
  3.5.4 Summary of section 3.5 ....................................................... 188
3.6 EXPRESSION STUDIES OF Cds1 AND Cds2 .................................. 193
  3.6.1 Mouse RNA MasterBlot™ ..................................................... 193
    3.6.1.1 Discussion points of section 3.6.1 ..................................... 193
  3.6.2 Northern blot analysis .......................................................... 196
    3.6.2.1 Preparation of Northern blot filter .................................. 196
    3.6.2.2 Northern blot probed with Cds1 ....................................... 198
    3.6.2.3 Northern blot probed with Cds2 ....................................... 198
    3.6.2.4 Discussion points of section 3.6.2 ................................... 201
  3.6.3 In situ hybridisation studies .................................................. 201
    3.6.3.1 Production of riboprobes ............................................... 201
    3.6.3.2 Location of gene transcripts ......................................... 201
    3.6.3.3 Discussion points of section 3.6.3 ................................... 214
  3.6.4 Rodless, Rodless-coneless and coneless .................................. 214
  3.6.5 Sub-cellular localisation of Cds2 ........................................... 218
    3.6.5.1 Production of the Cds2-myc tagged construct ...................... 218
    3.6.5.2 Transfection of cells for immunostaining .......................... 219
    3.6.5.3 Western blots .............................................................. 220
    3.6.5.4 Discussion points of section 3.6.4 ................................... 220
3.7 DROSOPHILA TRANSGENICS ....................................................... 222
  3.7.1 Cds rescue construct preparations ....................................... 222
  3.7.2 Mapping and expression studies of the transgenic Drosophila .... 227
  3.7.3 Discussion points of section 3.7 .......................................... 233
3.8 Cds2 TARGETTING CONSTRUCT ................................................. 236
  3.8.1 Construct vector pPNT ....................................................... 236
  3.8.2 Cds2 knockout construct production .................................... 236
    3.8.2.1 3kb BamHI construct fragment ..................................... 239
    3.8.2.2 1.4kb XhoI construct fragment ..................................... 239
3.8.3 Embryonic stem (ES) cell electroporation............................................242
3.8.4 Screening of potentially targeted ES cells............................................244
3.8.5 Discussion points of section 3.8...........................................................246
3.9 PATIENT SCREENING............................................................................251

CHAPTER FOUR-DISCUSSION.............................................................................254
4.1 RNA MASTERBLOTS™.............................................................................254
4.2 NORTHERN BLOTS.................................................................................257
4.3 IN SITU HYBRIDISATION........................................................................258
4.4 SUB-CELLULAR LOCALISATION OF Cds1 and Cds2 .............................259
4.5 MUTANT FLY RESCUE SYSTEM............................................................260
4.6 Cds2 KNOCKOUT MOUSE.................................................................261
4.7 CONCLUSIONS......................................................................................264

CHAPTER FIVE-FUTURE WORK........................................................................266
5.1 GENE STRUCTURE STUDIES.................................................................266
5.2 NORTHERN BLOT STUDIES...................................................................267
5.3 IN SITU HYBRIDISATION STUDIES......................................................267
5.4 PROTEIN STUDIES.................................................................................267
5.5 CDS MUTANT DROSOPHILA RESCUE STUDIES..................................268
5.6 KNOCKOUT MOUSE MODELS..............................................................269

BIBLIOGRAPHY...................................................................................................270
APPENDIX 1........................................................................................................300
LIST OF FIGURES

FIGURE 1.1 Drosophila compound eye.................................................................23
FIGURE 1.2 Adult ommatidium.................................................................25
FIGURE 1.3 Morphogenetic furrow.............................................................27
FIGURE 1.4 Differentiation of the rhabdomeres........................................29
FIGURE 1.5 Structure of the human eye......................................................28
FIGURE 1.6 A cross-section of the retina....................................................32
FIGURE 1.7 Rod and cone cells..................................................................34
FIGURE 1.8 A rhodopsin molecule in the plasma membrane..................36
FIGURE 1.9 Two structures of retinal.........................................................37
FIGURE 1.10 Development of the mammalian eye.................................39
FIGURE 1.11 Comparison of phototransduction in Drosophila and mammals..................................................41
FIGURE 1.12 Structure of an opsin molecule............................................42
FIGURE 1.13 Drosophila phototransduction cascade.............................44
FIGURE 1.14 Drosophila PIP_{2} regeneration cascade............................46
FIGURE 1.15 Signalplex..............................................................................48
FIGURE 1.16 Human phototransduction cascade.................................53
FIGURE 1.17 Mapped and cloned retinal disease graph.........................64
FIGURE 1.18 Inheritance Pedigree..............................................................67
FIGURE 1.19 P-elements.............................................................................88
FIGURE 1.20 Drosophila microinjections.....................................................90
FIGURE 1.21 Homologous recombination.................................................93
FIGURE 3.1 Gel photograph of 801-h9 insert fragment.........................129
FIGURE 3.2 Sequence generated from 801-h9........................................131
FIGURE 3.3 Alignment of amino acid sequences derived from human CDS1 and 801-h9...............................................................132
FIGURE 3.4 Restriction map of 801-h9........................................................133
FIGURE 3.5 Alignment of 2279-d19 and 801-h9 amino acid sequences..........................................................135
FIGURE 3.6 2279-d19 nucleotide and derived amino acid sequence......136
FIGURE 3.7 2279-d19 and human CDS1 nucleotide alignment..............137
FIGURE 3.8 2279-d19 and human CDS1 amino acid alignment............138
FIGURE 3.9 2279-d19 and rat Cds1 amino acid alignment.......................140
FIGURE 3.10 Sequence generated from the 5' end of $Cds1$.................................142
FIGURE 3.11 Summary of the generation of $Cds1$ sequence.................................143
FIGURE 3.12 Rat, human and mouse $Cds1$ amino acid alignment.........................144
FIGURE 3.13 Composite sequence of mouse $Cds1$.............................................146
FIGURE 3.14 $Cds1$ and $Drosophila$ cds amino acid alignment.............................147
FIGURE 3.15 Mouse $Cds1$ and $Drosophila$ hydrophobicity plots.........................149
FIGURE 3.16 Mouse $Cds1$ and human hydrophobicity plots.................................150
FIGURE 3.17 Amino acid alignment of 1500-b19 and human CDS2.......................152
FIGURE 3.18 Sequence derived from 1247-c2....................................................154
FIGURE 3.19 Restriction map of 1247-c2.............................................................155
FIGURE 3.20 Amino acid alignment of $Cds2$ and human CDS2..............................157
FIGURE 3.21 Amino acid alignment of $Drosophila$ cds and mouse $Cds2$.............158
FIGURE 3.22 Mouse $Cds2$ and $Drosophila$ hydrophobicity plots.......................159
FIGURE 3.23 Mouse $Cds2$ and human CDS2 hydrophobicity plots.......................160
FIGURE 3.24 Autoradiograph of the elimination of a PAC clone............................162
FIGURE 3.25 Gel photograph of the amplification of $Cds1$ introns.......................165
FIGURE 3.26 Genomic map of $Cds1$.................................................................167
FIGURE 3.27 Gel photograph of 20A restriction digests........................................169
FIGURE 3.28 Autoradiograph of 20A Southern blot..............................................171
FIGURE 3.29 Basic restriction map of 20A.........................................................174
FIGURE 3.30 Genomic structure of $Cds2$............................................................177
FIGURE 3.31 Identification of intron 1 in PAC clone 439-n24...............................178
FIGURE 3.32 Alignment of available cds amino acid sequences..........................180
FIGURE 3.33 Phylogenetic tree of available cds sequences....................................182
FIGURE 3.34 Comparison of the genomic structures of $Cds1$ and $Cds2$.............184
FIGURE 3.35a Synteny of mouse and human 4q21.............................................186
FIGURE 3.35b Synteny of mouse and human 20p13............................................187
FIGURE 3.36 $Cds1$ metaphase spreads...............................................................189
FIGURE 3.37 $Cds1$ chromosome pairs...............................................................190
FIGURE 3.38 $Cds2$ metaphase spreads..............................................................191
FIGURE 3.39 $Cds2$ chromosome pairs..............................................................192
FIGURE 3.40 RNA MasterBlots™......................................................................194
FIGURE 3.41 Photograph of RNA gel.................................................................197
FIGURE 3.42 Autoradiograph of $Cds1$ probed Northern blot..............................199
FIGURE 4.1a Alignment of Cds1 and *E.coli* cds amino acid sequences ...............255
FIGURE 4.1b Alignment of Cds2 and *E.coli* cds amino acid sequences ...............256
FIGURE 4.2 Diagram of Cre/lox site specific recombination system ......................263
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Cds gene annotation</td>
<td>21</td>
</tr>
<tr>
<td>1.1</td>
<td>Genes from the <em>Drosophila</em> phototransduction cascade</td>
<td>45</td>
</tr>
<tr>
<td>1.2</td>
<td>Human homologues of <em>Drosophila</em> genes from the phototransduction cascade</td>
<td>57</td>
</tr>
<tr>
<td>1.3</td>
<td>Human retinopathies</td>
<td>65</td>
</tr>
<tr>
<td>1.4</td>
<td>List of ABCR gene mutations</td>
<td>71</td>
</tr>
<tr>
<td>1.5</td>
<td>List of genes and loci for cone-rod dystrophy</td>
<td>74</td>
</tr>
<tr>
<td>1.6</td>
<td>URLs for common gene prediction packages</td>
<td>81</td>
</tr>
<tr>
<td>1.7</td>
<td>Description of the contents of gene prediction packages</td>
<td>83</td>
</tr>
<tr>
<td>3.1</td>
<td>801-h9 sequencing primers</td>
<td>130</td>
</tr>
<tr>
<td>3.2</td>
<td>2279-d19 sequencing primers</td>
<td>139</td>
</tr>
<tr>
<td>3.3</td>
<td><em>Cds1</em> 5' end amplification primers</td>
<td>141</td>
</tr>
<tr>
<td>3.4</td>
<td><em>Cds1</em> 5' end sequencing primers</td>
<td>141</td>
</tr>
<tr>
<td>3.5</td>
<td>1500-b19 sequencing primers</td>
<td>151</td>
</tr>
<tr>
<td>3.6</td>
<td>1247-c2 sequencing primers</td>
<td>153</td>
</tr>
<tr>
<td>3.7</td>
<td>Primers used to amplify <em>Cds1</em> introns</td>
<td>164</td>
</tr>
<tr>
<td>3.8</td>
<td><em>Cds1</em> splice junction information</td>
<td>166</td>
</tr>
<tr>
<td>3.9</td>
<td>Primers used to amplify <em>Cds2</em> introns</td>
<td>173</td>
</tr>
<tr>
<td>3.10</td>
<td><em>Cds2</em> splice junction information</td>
<td>175</td>
</tr>
<tr>
<td>3.11</td>
<td><em>A</em>$<em>{260}$/<em>A</em>$</em>{280}$ ratios of RNA samples used on the Northern blot</td>
<td>196</td>
</tr>
<tr>
<td>3.12</td>
<td>Tissue types used for <em>in situ</em> hybridisation analysis</td>
<td>214</td>
</tr>
<tr>
<td>3.13</td>
<td>Primers used to amplify myc-tagged fragments</td>
<td>219</td>
</tr>
<tr>
<td>3.14</td>
<td>Primers used to amplify rescue construct</td>
<td>224</td>
</tr>
<tr>
<td>3.15</td>
<td>Primers used to verify <em>Drosophila</em> construct</td>
<td>228</td>
</tr>
<tr>
<td>3.16</td>
<td>Primers to verify 3kb <em>BamHI</em> construct fragment orientation</td>
<td>239</td>
</tr>
<tr>
<td>3.17</td>
<td>Primers used to generate 1.4kb construct fragment</td>
<td>242</td>
</tr>
<tr>
<td>3.18</td>
<td>Primers to verify the <em>Cds2</em> knockout construct</td>
<td>242</td>
</tr>
<tr>
<td>3.19</td>
<td>Human CDS2 exon-intron information</td>
<td>252</td>
</tr>
<tr>
<td>3.20</td>
<td>Primer pairs used for CDS2 mutation analysis</td>
<td>253</td>
</tr>
<tr>
<td>4.1</td>
<td>Knockout construct vectors efficiency table</td>
<td>262</td>
</tr>
<tr>
<td>4.2</td>
<td>Tissue specific and inducible Cre-transgenic mice</td>
<td>265</td>
</tr>
</tbody>
</table>
ABBREVIATIONS

A  adenosine deoxyribonucleoside monophosphate residue
amp  ampicillin
AP  alkaline phosphatase
APS  ammonium persulphate
APS  aminopropyltriethoxysilane (see text)
Ax  optical absorbance at xnm
arr  arrestin
ATP  adenosine triphosphate
BCIP  bromo-chloro-indolyl-phosphate
bp  base pairs
BSA  bovine serum albumin
C  cytosine deoxyribonucleoside monophosphate residue
Ca2+  calcium ions
cAMP  3’ 5’ cyclic adenosine monophosphate
CaM  calmodulin
CDP  cytosine diphosphate
Cds  CDP-diacylglycerol synthase
cGMP  3’ 5’ cyclic guanosine monophosphate
cM  centimorgans
cDNA  complementary DNA
cfu  colony forming units
CHED  corneal hereditary endothelial dystrophy
cps  counts per second
Cre  cyclization recombination
Cys  cysteine residue
DAG  diacylglycerol
DAGK  diacylglycerol kinase
dATP  adenosine deoxyribonucleoside triphosphate residue
dCTP  cytosine deoxyribonucleoside triphosphate residue
dGTP  guanosine deoxyribonucleoside triphosphate residue
dTTP  thymine deoxyribonucleoside triphosphate residue
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPC</td>
<td>diethylpyrocarbonate</td>
</tr>
<tr>
<td>dH₂O</td>
<td>distilled, sterile water</td>
</tr>
<tr>
<td>DIG</td>
<td>digoxygenin</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulphoxide</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's modified eagle medium</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNase</td>
<td>deoxyribonuclease</td>
</tr>
<tr>
<td>d.p.c</td>
<td>days post coitus</td>
</tr>
<tr>
<td>EC</td>
<td>embryonic carcinoma</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminotetra-acetic acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>ethylene glycol-O, O'-bis (2-aminoethyl-N,N,N',N'-tetraacetic acid</td>
</tr>
<tr>
<td>en</td>
<td>engrailed</td>
</tr>
<tr>
<td>EOR</td>
<td>endoplasmic reticulum overload response</td>
</tr>
<tr>
<td>ES</td>
<td>embryonic stem</td>
</tr>
<tr>
<td>FCS</td>
<td>foetal calf serum</td>
</tr>
<tr>
<td>FISH</td>
<td>fluorescent <em>in situ</em> hybridisation</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>g</td>
<td>acceleration due to gravity (9.8 metres per second per second)</td>
</tr>
<tr>
<td>G</td>
<td>guanosine deoxyribonucleoside monophosphate residue</td>
</tr>
<tr>
<td>GCAP</td>
<td>guanylate cyclase activating protein</td>
</tr>
<tr>
<td>GDP</td>
<td>guanosine diphosphate</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine triphosphate</td>
</tr>
<tr>
<td>HEK</td>
<td>hamster embryonic kidney</td>
</tr>
<tr>
<td>HGMP-RC</td>
<td>human genome mapping project-resource centre</td>
</tr>
<tr>
<td>ina</td>
<td>inactivation no afterpotential</td>
</tr>
<tr>
<td>IP₃</td>
<td>inositol triphosphate</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropylthiogalactoside</td>
</tr>
<tr>
<td>K⁺</td>
<td>potassium ions</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodaltons</td>
</tr>
<tr>
<td>Klenow</td>
<td>large fragment of DNA polymerase I</td>
</tr>
<tr>
<td>λ</td>
<td>bacteriophage lambda</td>
</tr>
<tr>
<td>LIF</td>
<td>lipofectamine</td>
</tr>
</tbody>
</table>
TMED  (N,N,N',N')-tetramethylethylenediamine
trp transient receptor potential
trpl transient receptor potential-like
Tulp tubby-like protein
TUNEL TdT-mediated dUTP digoxigenin nick end labelling
U units
UPR unfolded protein response
UTR untranslated region
UV ultraviolet
V/cm volts per centimetre
v/v volume for volume
w/v weight for volume
w/w weight for weight
X-gal 5-bromo-4-chloro-3-indolyl-β-D-galactoside

<table>
<thead>
<tr>
<th>Gene annotation used</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>cds</td>
<td>Drosophila lower species</td>
</tr>
<tr>
<td>Cds</td>
<td>Rodent</td>
</tr>
<tr>
<td>CDS</td>
<td>Human</td>
</tr>
</tbody>
</table>

Table I. List of annotations used to describe the CDP-diacylglycerol synthase gene in various species.
The use of *Drosophila melanogaster* as a model system has proved to be an invaluable tool in trying to understand the complexities of human disease. Mammalian homologues of several *Drosophila* genes involved in the development of the visual system have been shown to be involved in human eye disease. The work presented in this thesis is a result of extending this model to the phototransduction pathway. The enzyme CDP-diacylglycerol synthase (cds), is required for the regeneration of phosphatidylinositol, 4, 5-bisphosphate (PIP$_2$). The hydrolysis of PIP$_2$ is required in the *Drosophila* phototransduction cascade to mobilise internal calcium stores which operate the ion-gated membrane channels (see section 1.2). eye-cds is a tissue-specific form of the enzyme in *Drosophila*. When *eye-cds* is mutated in flies it results in light-induced retinal degeneration. The work presented here is the analysis of two murine homologues of *eye-cds*, *Cds1* and *Cds2*.

This chapter is divided into three sections, the first is a description of the structure and development of the *Drosophila* and mammalian eye. The second section will discuss phototransduction and the third section is an overview of some of the techniques employed in this study.

### 1.1 Ocular Structure and Development

#### 1.1.1 *Drosophila* ocular structure

The fly ocular structure is very complex and highly organised. The adult compound eye (figure 1.1) is composed of eight hundred individual units, the ommatidia. Each ommatidium is positioned to capture light from distinct, but overlapping, visual fields. Each ommatidium is composed of twenty cells, which include, eight photoreceptor cells with twelve accessory bristle, lens secreting and pigment cells. This arrangement of cells means that each compound eye contains approximately sixteen thousand cells. The cells of the eye are all derived from a single monolayer of epithelium and the
Figure 1.1 Compound eye of *Drosophila melanogaster*

Shown is a scanning electronmicrograph of the compound eye of *Drosophila melanogaster*. The honeycomb pattern is formed from the 800 hexagonal ommatidia that are conical in shape resulting in the curvature seen in the adult eye. The bristles and the regular lenses which overlay the ommatidia can also be seen (Brennan and Moses, 2000).
columnar appearance is retained in the adult cells. The cells still contain their apical and basal connections but each ommatidium widens basally to apically to form a conical shaped unit that gives the curvature to the surface of the adult eye (Wolf and Ready, 1993).

The photopigments of each photoreceptor are embedded into the microvillar membrane stacks, called rhabdomeres, that extend from each photoreceptor cell into a central space in the ommatidium (figure 1.2). There are a total of eight rhabdomeres (R1-R8) arranged in a specific orientation within the ommatidium. The rhabdomeres of the outer photoreceptors, R1-R6, can be individually identified by the position they occupy in the irregular trapezoid shape that they form. R1-R6 span the full height of the ommatidium and contain a blue-sensitive opsin molecule (see section 1.1.3). These outer receptors provide 'high-sensitivity' light-capturing properties to the retina and have been likened to the vertebrate rod cells (Bowmaker et al., 1991; Zuker, 1996) (see section 1.1.3). The smaller, more central photoreceptors containing R7 and R8 have UV and blue-green sensitive opsin molecules respectively. These photoreceptors are similar to the vertebrate cone cells and confer a 'high-acuity' photoreceptive capacity (Zuker, 1996).

The axons from each of the eight photoreceptors of the ommatidium pass through the basement membrane and extend directly into the brain via the optic stalk for signal transduction. The projections from the ommatidia to the optic lobe of the brain are complex. Since each ommatidium is cone shaped, the rhabdomeres within each ommatidium are not parallel. Individual axons from each rhabdomere in the ommatidium have to separate and join with parallel axons from neighbouring ommatidia before projecting into the brain. This arrangement allows a high degree of sensitivity while maintaining visual acuity. The precise array of components must be maintained for the complex visual system of the fly to operate (Meinertzhagen and O'Neil, 1991; reviewed by Brennan and Moses, 2000).
Figure 1.2 Schematic representation of an adult ommatidium
A schematic representation of the adult *Drosophila* ommatidium, showing the arrangement of the rhabdomeres (R1-R8) and the various pigment cells taken at different cross-sections, indicated, in the ommatidium (adapted from Freeman *et al.*, 1997; Brennan and Moses, 2000).
1.1.2 Drosophila ocular development

*Drosophila melanogaster* belongs to the Endopterygota group of insects. This group progresses through three larval stages called instars before pupating to form the adult fly. The eye of the adult fly develops from an epithelial pouch, called an imaginal disc, consisting of approximately twenty cells derived from the optic primordium in the embryonic blastoderm. The disc grows through larval life having increased in cell number by one hundred fold by the third larval instar. Differentiation of the cells in the disc is initiated in the final larval instar and is completed during larval metamorphosis. At this stage the disc evaginates and fuses with other imaginal tissues (Haynie and Bryant, 1986). To form the adult eye structure, a small number of discrete units are formed and are used as a template to form identical repeats units which join to give the final eye structure. The template units in the eye are the ommatidia. The development of the imaginal disc towards the adult eye structure starts in the third larval instar at the posterior edge, and expands anteriorly following a morphogenetic furrow. This furrow, which is an indentation in the eye disc that marks the anterior boundary of neural patterning, sweeps anteriorly across the disc (figure 1.3). Ommatidial photoreceptor clusters are specified one column at a time, beginning at the posterior margin of the eye imaginal disc during initiation of differentiation. They emerge from the furrow at the rate of one column every two hours (Ready *et al.*, 1976). The morphogenetic furrow is caused by the co-ordinated apical-basal contractions of cells and moves in a wave-like motion over the disc in a period of approximately two days. Anterior to the furrow, the cells divide randomly, but at the furrow they synchronise and arrest in the G1 phase of mitosis. In, and posterior to the furrow, a subset of cells are determined as the initial five photoreceptors. The remaining cells undergo another round of co-ordinated mitosis several hours later, this new pool of cells provide the components required to complete the ommatidium. The remaining components include the final three photoreceptor cells, the four cone cells and, after pupation, the bristle and pigment cells (reviewed by Brennan and Moses, 2000).

As briefly mentioned above, the photoreceptor cell specific rhabdomeres differentiate sequentially. The founding photoreceptor cell in each ommatidium contains rhabdomere 8 (R8). R8 is required for the pair-wise recruitment of the other seven
Figure 1.3 *Drosophila* ocular morphogenetic furrow

(A) shows a late third-instar eye disc stained to reveal filamentous cytoplasmic actin. Anterior is to the right where cells are randomly dividing. The posterior shows a more organised cell pattern. The intense condensation of apical cell profiles, that indicate the morphogenetic furrow, is highlighted by arrows. (B) a late third-instar eye disc stained to show neural differentiation, this staining notes the maturity of the cells and the formation of the ommatidia behind the path of the morphogenetic furrow (Brennan and Moses, 2000).
cells. The assembly of the photoreceptors of the ommatidium is dependent on cell-cell interactions. R2 and R5 are the first pair of cells to be recruited by R8, followed by R3 and R4, then by R1 and R6. The final cell to be recruited is R7 followed by the cone and pigment cells (figure 1.4). The steps in retinal determination are tightly controlled, in terms of rate of expansion and differentiation, with synchronised cell cycles and recruitment of cell types to the ommatidium cluster, as this is essential to build the compound eye with such high complexity.

1.1.3 Mammalian eye structure
All mammals have simple eyes with a single optical system that forms a single image. The roughly spherical human eye is a good representative of a mammalian eye. The adult eye is a fluid filled chamber, about 24mm in diameter, contained within three coats, or layers, the sclera, the choroid and the retina.

Figure 1.5 Schematic representation of a human eye
A schematic representation of a cross-section of a human eye. The main structures of the eye are indicated. The sclera, choroid and retinal are represented by the beige, brown and yellow layers respectively.

The outermost coat, the sclera is a tough fibrous layer that protects the ocular tissues and provides structural support for the intraocular contents as well as forming
Figure 1.4 Differentiation of *Drosophila* rhabdomeres
A schematic representation of ommatidial assembly. Shown are the first steps of ommatidial assembly of photoreceptor cells starting with the photoreceptor containing rhabdomere 8 (R8), then the pair wise recruitment of the other cells which have been numbered. Cone cells are labelled C and primary pigment cells labelled P (adapted from Kumar and Moses, 1997).
attachments to the extraocular muscles. The sclera is opaque and constitutes the ‘white’ of the eye except at the front where it becomes transparent and forms the cornea. The cornea is responsible for most of the refraction of light entering the eye, thus the transparency of the cornea is of upmost importance. The transparency arises due to the regularity and smoothness of the covering epithelium, its avascularity and the regular arrangement of the extracellular and cellular components of the central corneal layer called the stroma. The state of hydration, metabolism and nutrition of the stromal elements controls the effectiveness of the cornea. The surface of the cornea, which provides a tough physical barrier to infection and damage, is kept moist and dust-free by secretions from the tear glands.

The choroid, the middle layer, is deeply pigmented with melanin. The pigmentation reduces the reflection of stray light within the eye. At the front of the eye the choroid coat forms the iris which is also pigmented. It is the iris that is responsible for ‘eye-colour’. There is an opening at the front of the choroid, the pupil, the size of which is variable and is under the control of the autonomic nervous system. The pupil can vary in size from 1-8mm. The functions of the pupil include the regulation of the amount of light entering the eye (it increases 16-fold on dilation from 2 to 8mm), and increases the depth of focus for near vision and minimises optical aberrations.

The lens is located just behind the iris, a third of the way between the front and back of the eye. It is held in position by suspensory ligaments, called zonules, which extend from an encircling ring of muscle. The lens is a highly organised system of specialised surface ectoderm-derived cells, which are highly ordered and tightly packed with structural proteins called crystallins. The lens constitutes an important component of the optical system of the eye. The lens can alter the refractive index of light entering the eye and focus it onto the retina. The lens has less refractive power than the cornea but due to the cilary muscles of the eye the lens can change shape and alter its refractive power. The lens and the iris divide the eye into two main chambers: the front chamber is filled with a water-like material, the aqueous humour, and the rear chamber is filled with a jelly-like material, the vitreous humour. These fluids transport oxygen and nutrients as well as removing waste products and helping to maintain the shape of the eye.
The retina is the innermost coat of the eye and consists of two primary layers, the inner neurosensory layer and the outer retinal pigment epithelium (RPE). The retina has two roles, firstly, to form a retinal image by detection of light and, secondly, to inform the brain about the features of the image that can be used to construct a mental picture of the external objects that are imaged on the retina. The retina is approximately 1250mm² and varies in thickness from 100μm at the periphery to 230μm at the optic nerve head. The neural retina is highly complex, it contains the photoreceptor cells, the rods and cones, as well as a variety of nerve cells. In all vertebrates there is an optical disk where the nerve fibres exit the eye via the optic nerve. This disk contains none of the normal retinal layers and is referred to as the ‘blind spot’. The peripheral retina is 100-140μm thick and is rich in rod type photoreceptor cells. The foveal pit is an area of the retina near the optic nerve where cone photoreceptors are concentrated at maximum density with exclusion of the rods, and arranged at their most efficient packing density, which is in a hexagonal mosaic. The whole foveal area is referred to as the macula of the human eye. The macula has a yellow hue due to the reflection from yellow screening pigments, the xanthophyll carotenoids, zeaxanthin and lutein present in the cone cell axons. The macula is thought to act as a short wavelength filter, in addition to that provided by the lens. As the fovea is the most essential part of the retina for human vision, protective mechanisms for avoiding bright light and especially ultraviolet irradiation damage are essential (Forrester et al., 1996)

The retina is traditionally shown to contain ten main layers, but these can be subdivided further (figure 1.6). The layer furthest from the incident light, the vascular choroid, supplies the outer retina with nutrients. Bruch’s membrane, a thin layer which delimits the retina proper, supports the basal surfaces of the cells of the retinal pigment epithelium (RPE). The RPE separates the retina from the choroid circulation and consists of a single continuous layer of roughly cuboidal cells that contain the pigment melanin. While it is not directly involved in any of the neural events in vision, the RPE is critical for the normal functioning of the photoreceptors. The RPE has multiple functions such as, maintaining the interphotoreceptor matrix, transporting water and metabolites, maintenance of the blood-retinal barrier, immunoregulation and free radical scavaging, but the most important role of the RPE is the maintenance of the physiology
Figure 1.6 Cross-section of the adult human retina
The section of the retina shown is taken from near the fovea illustrating the 10 main layers of the human retina. Overlaid is a schematic representation of the locations of the major classes of neurones found in the retina (adapted from Oyster, 1999).
of the photoreceptors. The RPE cells act as phagocytes to remove ‘spent’ photoreceptor tips in the diurnal process of photoreceptor renewal where the outer segments are shed. The RPE also participates in the recycling of 11-cis-retinal required for phototransduction.

There are two types of visual photoreceptors in the mammalian retina, the rod and cone cells, identified in accordance with the morphological appearance of their light sensitive tips. The human retina contains around 120 million rods and 6 million cones. Each cell type has the same basic structure: an outer segment linked via a connecting cilium to the inner segment containing the nucleus and finally the synaptic body, which allows the transmission of impulses (figure 1.7). The photoreceptor cells span four of the traditional retinal layers: their outer and inner segments lie in the photoreceptor layer; their nuclei form the outer nuclear layer, and their axons and terminal endings make up the outer plexiform layer (figure 1.6).

A photoreceptor cell is characterised by the photopigment that it contains. All the rods contain the same visual pigment, giving one type of rod photoreceptor. There are three types of cone photoreceptors, each ‘tuned’ to absorb light maximally from a portion of the spectrum of visible light due to the differing visual pigment; long-wavelength light (558nm) (red), middle-wavelength light (531nm) (green) and short-wavelength light (419nm) (blue) (Oprian et al., 1991; Merbs and Nathans, 1992).

The visual pigments are embedded in the membranes of the lamellae (flattened discs) of the outer segments of the photoreceptor cells. The discs are produced at the base of the outer segments and over the course of approximately ten days travel to the tips, which are enclosed by the apical microvilli of the RPE where spent tips are phagocytosed. The visual pigment molecules are the most abundant proteins within the outer segments, accounting for approximately 80% of membrane proteins present (Nathans, 1987). The density of the visual pigments in the discs of the outer segments is such that it is easy for even a single photon of light to stimulate a visual pigment molecule. The absorption of photons by these light sensitive pigments triggers a conformational change that stimulates the phototransduction cascade, which leads to a neural output from the photoreceptor cell which is modified and sent to the brain. The phototransduction cascade will be described later in section 1.2.
Figure 1.7 Structural comparison of rods and cones
A schematic diagram of a transverse section through a rod and cone photoreceptor cell. The morphological similarities and differences between the cells are shown (adapted from Ali and Klyne, 1985).
Each visual pigment consists of a transmembrane protein, opsin, coupled to a chromophore, retinal, a derivative of vitamin A. The opsins belong to the large family of seven transmembrane domain G-protein coupled receptors. These are integral membrane proteins that have an \( \alpha \)-helical conformation (Baldwin, 1993; Schertler et al., 1993; Alkorta and Du, 1994). The opsin molecules are characterised by seven hydrophobic transmembrane domains (helices I-VII) that are linked by extramembrane hydrophilic loops (Mollon, 1991; Khorana, 1992; Trumpp-Kallmeyer et al., 1992). The amino-terminus (N-terminus) and the carboxy-terminus (C-terminus) of the protein lie within the extracellular and cytoplasmic regions of the cell membrane respectively (figure 1.8). The chromophore molecule is located in the seventh helix of the opsin molecule and is connected by a covalent Schiff-base linkage to a lysine residue (Bownds, 1967; Wang et al., 1980; Findlay et al., 1981; Thomas and Stryer, 1982). The chromophore is usually the 11-\text{cis} isomer of vitamin A aldehyde, 11-\text{cis}-retinal (figure 1.9). Retinal is a long-chain molecule that can exist in two forms: a straight chain form called all-\text{trans}-retinal (figure 1.9) or a bent form, 11-\text{cis}-retinal. It is only the bent conformation, that can bind to the opsin molecule. The retinal is the portion of the photopigment first affected by light absorption and is considered to be the active part of the opsin-retinal complex.

In addition to the photoreceptors, the retina contains bipolar and ganglion cell neurones that relay impulses generated by light. The activity of these neurones is modulated by other cell types such as horizontal cells, amacrine cells and possibly by non-neuronal elements. Synaptic junctions that form between bipolar, amacrine and ganglion cells give rise to the inner plexiform layer of the retina. The ganglion cell layer of the retina, as the name suggests, contains the ganglion cell bodies and the axons of the ganglion cells that travel to the optic disk from the nerve fibre layer of the retina. The ganglion cell axons are the sole information channels from the eye to the brain. Ganglion cells remain active constantly, even in the dark they generate trains of action potentials and conduct them back to the brain along the optic nerve. Vision is based on the interpretation and modulation of these nerve impulses by the brain (reviewed by Forrester et al., 1996; Oyster, 1999).
Figure 1.8 Schematic representation of rhodopsin in the plasma membrane
Schematic illustration of the rhodopsin molecule situated in the plasma membrane of a photoreceptor outer segment disc. The illustration shows the site of retinal binding in the seventh transmembrane domain, indicated by a red arrow (adapted from Hargrave, 1996; rhodopsin molecule courtesy of S. Wilkie and D. Hunt).
1.1.4 Mammalian ocular development

The human eye is representative of most mammalian eyes, and for the purpose of this section will be used to describe the ocular developmental process in mammals. The mammalian eye is a projection from the brain, and is formed by evagination of the anterior neural tube. The walls of the optic evagination are continuous with the forebrain and consist of neural epithelium. As development progresses the evagination gets displaced laterally and becomes bulbous at the periphery while the walls towards the forebrain become narrowed to form the optic stalk. The optic stalk will eventually become the optic nerve. The bulbous portions are the optic vesicles; these face the lateral aspect of the embryonic head but will rotate with further development to ensure that the present rostral aspect will become the future medial aspect. Invagination of the optic vesicle occurs on the future medial aspect to complete the formation of the optic cup. The outer layer of neuroepithelium which lines the cup will form the pigmented

Figure 1.9 Molecular structure of the two retinal isoforms
Shown are the molecular structures of the bent 11-cis-retinal which covalently binds to the opsin molecule, and the straight all-trans-retinal (adapted from Montell, 1999).
epithelium of the retina, ciliary body and anterior iris epithelium. The inner lining layer will form the neural retina, non-pigmented epithelium of the ciliary body and the posterior iris epithelium. The primary vitreous and hyaloid vessel system is formed from invasion of vascular mesenchyme from the optic primordium. Growth of the optic cup brings ventromedial edges towards each other for fusion. The cup is completed by trapping hyaloid vessels in the stalk; these will become the future central vessels of the retina.

Through invagination, the neural retinal layers and the retinal pigmented epithelium (RPE) are formed and are continuous. The developmental margins are still apparent; and it is from these margins that the ciliary body and the iris will arise forming a pupillary margin. Figure 1.10 illustrates the development of the eye (reviewed by Forrester et al., 1996). The primitive neural retina has two layers; the outer nuclear zone and the inner acellular zone. The layers rest on their respective basal laminae, the inner zone is the inner limiting membrane and the outer is incorporated into Bruch’s membrane. Mitotic activity is greatest in the outer nuclear zone. The new cells then migrate to the inner zone to form the inner neuroblastic layer. The outer nuclear zone is now called the outer neuroblastic layer. Cells in the inner neuroblastic layer will later form ganglion cells, Müller cells and amacrine cells. The nerve fibre layer can be identified on the inner aspect of the neuroblastic layer due to the growth of ganglion cell axons converging at the optic stalk. The inner plexiform layer is formed from cellular processes intermingling. The new inner nuclear layer contains amacrine and Müller cell bodies. Shortly afterwards the bipolar and horizontal cells differentiate and migrate from the outer neuroblastic layer. The remaining components of the outer neuroblastic layer contain the cell bodies of the photoreceptors. The zone where the fibres from this layer intermingle with those of the inner nuclear layer constitute the new outer plexiform layer. The external-limiting boundary of the retina is identifiable in the earliest stages as rows of tight and gap junctions between adjacent neuroblasts.

Microglia invade the retina at four months gestation at around the same time as the synapses form in the cone pedicles. This is about one month before the same process occurs in the rod spherules. At four and a half months, the processes between the rod and cones can be identified and terminal expansion of the Müller cells beneath the inner limiting membrane mature around this time. At five months the photoreceptor outer
Figure 1.10 Schematic representation of the development of the mammalian eye

The retina is derived from the neural tube (1). The retina is formed during development of the embryo from optic vesicles out pouching from two sides of the developing neural tube (2). The primordial optic vesicles fold back in upon themselves to form the optic cup with the inside of the cup becoming the retina and the outside remaining a single monolayer of epithelium known as the retinal pigment epithelium (RPE) (3). Initially both walls of the optic cup are one cell thick, but the cells of the inner wall divide to form a neuroepithelial layer many cells thick which form the retina (taken from Forrester et al., 1996).
segments start to form and horizontal cells become distinguishable (reviewed by Forrester et al., 1996).

The RPE is first seen at 28 days when melanin can be identified. The cells form a mitotically-active ciliated epithelium. As melanogenesis progresses, the cilia disappear. The cells are uniformly sized and hexagonal in shape. Mitotic activity starts as early as 28 days after fertilisation but stops before birth. Growth then occurs by the enlargement of existing cells (reviewed by Forrester et al., 1996).

1.2 Phototransduction

The photoreceptors in both mammals and Drosophila are very specialised structures and although they are physically different, they both possess a large membranous surface area to facilitate ion transfer. Both the mammalian and Drosophila phototransduction cascades are initiated by a photon of light causing a conformational change in a rhodopsin molecule, which then activates a G-protein cascade. It is after this point that the two systems differ and follow two distinct pathways (Figure 1.11). In brief, in mammals a G-protein cascade causes hydrolysis of cGMP by phosphodiesterase that results in the closure of membrane channels. This is in contrast to Drosophila where the G-protein cascade results in the hydrolysis of phosphatidylinositol-4, 5, bisphosphate (PIP_2) generating inositol triphosphate (IP_3) and diacylglycerol (DAG). IP_3 starts a chain of events that leads to Ca^{2+} release which opens membrane channels. Although the end cycles differ and the resultant effect on the channels is converse, both systems use Ca^{2+} to cause deactivation and adaption. This section will describe in more detail both the Drosophila and mammalian cascades and the evidence that elements used in the Drosophila cascade are found in the mammalian retina.

1.2.1 Drosophila phototransduction

The Drosophila phototransduction pathway has been studied intensively but is still not fully understood. The G-protein cascade in this pathway is the fastest known, taking only a few tenths of a millisecond to occur. As discussed in section 1.1 the light sensitive molecule, rhodopsin, has two components, a seven α-helical transmembrane protein called an opsin, and a vitamin A derived chromophore which is covalently bound to the seventh helix of the opsin (figure 1.12). When a photon of light hits the rhodopsin molecule, a conformational change occurs in the chromophore. This event
Figure 1.11 Comparison of phototransduction in *Drosophila* and mammals
Shown is a very simple schematic representation of the similarities and differences between the *Drosophila* and human phototransduction cascades.
Figure 1.12 Structure of an opsin molecule
A secondary structure diagram of bovine Rho. Amino acid residues are depicted in single-letter code. The amino-terminal tail and extracellular domain is toward the top, and the carboxy-terminal tail and cytoplasmic domain is toward the bottom. Transmembrane α-helical segments (H1 to H7) and the cationic amphipathic helix H8 are shown in colored cylinders. An essential disulfide bond links Cys-110 and Cys-187. Cys-322 and Cys-323 are palmitoylated. Inset: structure of the 11-cis-retinylidene (RET) chromophore. Carbon atoms are numbered 1 through 20 (adapted from Menon et al., 2001)
results in the isomerisation of the 11-12 double bond in the 11-cis-retinal molecule which converts it into the all-trans-retinal form (figure 1.9) (Vogt and Kirchfeld, 1984; Tanimura et al., 1986). The all-trans conformation is active and is referred to as metarhodopsin. The active metarhodopsin molecule activates a heterotrimeric G-protein, G_αβγ. The inhibitory βγ subunits are removed and the GTP_α molecule becomes active. It is at this point that the Drosophila and mammalian cycles cease to follow similar pathways. The active GTP_α molecule couples to and activates a phospholipase C molecule (PLC_β). The PLC_β molecule is the effector molecule of the Drosophila pathway (Lee et al., 1990; Scott et al., 1995) (figure 1.13).

PIP_2 is a membrane lipid, which is hydrolysed by the PLC_β effector molecule into two intracellular messengers, inositol triphosphate (IP_3) and diacylglycerol (DAG). IP_3 mobilises Ca^{2+} from internal stores. DAG activates protein kinase C (PKC). This reaction leads to the opening of selective cation channels and the generation of depolarising receptor potentials (reviewed in Baylor, 1996; Zuker, 1996). It is thought that there may be up to fifty different gene products involved in this cascade (Smith et al., 1991; Zuker, 1992; Ranganathan et al., 1995).

Eye-specific isoforms of many of the elements involved in phototransduction have been identified. Table 1.1 lists known genes from the Drosophila phototransduction pathway with their functions. Examples include eye-specific PKC, which is required for deactivation of calcium feedback regulation (Ranganathan et al., 1991; Smith et al., 1991), and eye-specific CDP-diacylglycerol synthase (eye-cds) which catalyses the conversion of phosphatidic acid (PA) into CDP-DAG. CDP-DAG is the acceptor molecule for inositol head groups. Drosophila cds mutants undergo light-dependent retinal degeneration. The general points of the PLC signalling cascade are, (i) PIP_2 is synthesised on demand, (ii) PIP_2 pools for signalling are distinct from the general pool, (iii) PIP_2 availability is the rate-limiting step in regulating the output of the signalling pathway (Wu et al., 1995). Figure 1.14 schematically illustrates the PIP_2 regeneration cycle.

These initial steps in phototransduction are well established, however recent findings suggest that events that occur downstream are more complex than first thought (Tsunoda et al., 1997; Montell, 1998). The control mechanism of Drosophila
Phototransduction in *Drosophila* photoreceptors. Absorption of a photon causes a conformational change in the rhodopsin molecule (R) and activates its catalytic properties. Active metarhodopsin (M*) catalyses G-protein activation. Active G-protein catalyses the activation of *norpA*-encoded PLC which hydrolyses PIP$_2$ into IP$_3$ and DAG. Extracellular Na$^+$ and Ca$^{2+}$ enter the cell through the light-activated conductance and cause the depolarisation of the photoreceptor cells. Light-activated conductance is composed of at least two types of channels. The *trp* gene is required for a class of channel with high calcium permeability. DAG is thought to modulate a photoreceptor cell-specific PKC (encoded by *inaC*) that regulates deactivation and desensitisation of the light response. Metarhodopsin is inactivated via phosphorylation by rhodopsin kinase (RoK) and arrestin binding (encoded by the *arr1* and *arr2* genes (Mppp-Arr)). Inactive metarhodopsin is photoconverted back to rhodopsin and then dephosphorylated by the *rdgC*-encoded phosphatase (adapted from Zuker, 1996).
<table>
<thead>
<tr>
<th>Locus</th>
<th>Protein Product</th>
<th>Function</th>
<th>Retinal degeneration</th>
</tr>
</thead>
<tbody>
<tr>
<td>arrestin1</td>
<td>Arrestin</td>
<td>Rhodopsin inactivation</td>
<td>Not examined</td>
</tr>
<tr>
<td>arrestin2</td>
<td>Arrestin</td>
<td>Rhodopsin inactivation</td>
<td>Yes (light dependent)</td>
</tr>
<tr>
<td>cds</td>
<td>CDP-DAG synthase</td>
<td>Production of PIP$_2$ intermediate</td>
<td>Yes (light dependent)</td>
</tr>
<tr>
<td>$G_q \alpha$</td>
<td>$\alpha$-subunit of G-protein</td>
<td>Activation</td>
<td>Not examined</td>
</tr>
<tr>
<td>$G_q \beta$</td>
<td>$\beta$-subunit of G-protein</td>
<td>Activation/deactivation</td>
<td>Not examined</td>
</tr>
<tr>
<td>inaC</td>
<td>Protein kinase C</td>
<td>Deactivation/adaption</td>
<td>Not examined</td>
</tr>
<tr>
<td>inaD</td>
<td>PDZ containing protein</td>
<td>Activation/deactivation</td>
<td>Yes (light dependent)</td>
</tr>
<tr>
<td>ninaA</td>
<td>Cyclophilin homologue</td>
<td>Rhodopsin transport</td>
<td>Yes (light independent)</td>
</tr>
<tr>
<td>ninaC</td>
<td>Protein kinase</td>
<td>Deactivation/adaption</td>
<td>Yes (light dependent)</td>
</tr>
<tr>
<td>ninaE</td>
<td>Rhodopsin</td>
<td>Light receptors</td>
<td>Yes (light independent)</td>
</tr>
<tr>
<td>norpA</td>
<td>Phospholipase C</td>
<td>Activation</td>
<td>Yes (light dependent)</td>
</tr>
<tr>
<td>rdgA</td>
<td>DAG-kinase</td>
<td>Production of PIP$_2$ intermediate</td>
<td>Yes (light independent)</td>
</tr>
<tr>
<td>rdgB</td>
<td>PL transfer protein</td>
<td>Recovery from light stimulus</td>
<td>Yes (light dependent)</td>
</tr>
<tr>
<td>rdgC</td>
<td>Protein phosphatase</td>
<td>Rhodopsin phosphatase</td>
<td>Yes (light dependent)</td>
</tr>
<tr>
<td>trp</td>
<td>Cation channel</td>
<td>Light-responsive channel</td>
<td>Yes (light dependent)</td>
</tr>
<tr>
<td>trpl</td>
<td>Cation channel</td>
<td>Light-responsive channel</td>
<td>Not examined</td>
</tr>
</tbody>
</table>

Table 1.1 Genes from the *Drosophila* phototransduction cascade
Listed are the genes from the *Drosophila* phototransduction cascade with their protein product named, their predicted function and whether a mutation in that gene causes retinal degeneration to occur (taken from Montell, 1999).
Figure 1.14 *Drosophila* PIP$_2$ regeneration cascade

Upon activation phospholipase C hydrolyses phosphatidylinositol-4,5-bisphosphate (PIP$_2$) to yield inositol triphosphate (IP$_3$) and diacylglycerol (DAG). To regenerate PIP$_2$, DAG is phosphorylated by DAG kinase to phosphatidic acid (PA). PA is converted to CDP-diacylglycerol (CDP-DAG) by the enzyme CDP-diacylglycerol synthase (CDS). CDP-DAG is the activated donor of the phosphatidyl group to inositol. Phosphatidylinositol (PI) is phosphorylated by PI kinase to give phosphatidylinositol monophosphate (PIP) which is phosphorylated by PIP kinase to yield PIP$_2$. The genes encoding the enzymes of the pathway are shown in red. Broken arrows indicate a number of steps to reach the product (adapted from Wu *et al*., 1995)
phototransduction is still not fully understood, but recent findings suggest that the control is very complex and tightly regulated (Tsunoda et al., 1997; Montell, 1998).

Ca$^{2+}$ ions are required for the opening and closing of channels in the photoreceptor cells that govern phototransduction. The light dependent influx of these ions is highly transient and spatially localised to small regions within the rhabdomeres (Ranganathan et al., 1994). Since the Ca$^{2+}$ influx appears to be limited to small domains, and other signalling proteins require Ca$^{2+}$ as part of their negative feedback loop, a problem would arise if the upstream signalling proteins were spatially separated from the ion channels. Recent findings suggest that most proteins that function in Drosophila phototransduction are assembled into a supramolecule called a signalplex. The signalplex involves at least seven signalling proteins binding to a scaffold protein called INAD (reviewed in Montell, 1999).

INAD has five protein interaction domains referred to as PDZ domains, PDZ 1-5. PDZ is named after a domain common to PDS-95 (postsynaptic density protein), Drosophila disc large (dlg) and the tight junction protein ZO-1 (Shieh and Niemeyer, 1995; Huber et al., 1996b; Saras and Heldin, 1996). The proteins bound to the scaffold include rhodopsin (Rh1) (Xu et al., 1998a), PLC (Huber et al., 1996a; Chevesich et al., 1997), PKC (Huber et al., 1996b; Adamski et al., 1998; Xu et al., 1998b), neither inactivation nor afterpotential C (NINAC) (Wes et al., 1999), calmodulin (Xu et al., 1998a), and both light sensitive channels transient receptor potential (TRP) (Huber et al., 1996a; Shieh and Zhu, 1996), and transient receptor potential-like (TRPL) (Xu et al., 1998b). Figure 1.15 schematically illustrates the structure of the signalplex in the plasma membrane.

Mutations in INAD disrupt the rhabdomere specific localisation of TRP, PLC and PKC while spatial distribution of the remaining binding proteins are not influenced by the presence or absence of INAD (Chevesich et al., 1997; Shieh et al., 1997; Tsunoda, 1997; van Huizen et al., 1998). It seems that most if not all of the interacting proteins are complexed into a signal unit since Rh1 and PLC co-immunoprecipitate with TRP from wild type but not from inaD mutant photoreceptor cells (Chevesich et al., 1997). INAD appears to self assemble in vitro to form oligomers with a subunit composition of eight or greater (Xu et al., 1998a). The multimerisation is thought to occur under the
Figure 1.15 Schematic representation of the signalplex complex

Model of the signalplex. Shown is an INAD homomultimer consisting of several INAD monomers that interact through homo- and heterophilic interactions between PDZ3 and PDZ4. Each INAD monomer contains five PDZ domains indicated by numbered ovals. INAD binds to a minimum of seven target proteins: rhodopsin (Rh), PLC, PKC, NINAC, TRP, TRPL and calmodulin (CaM). With the exception of CaM, which binds to the linker region between PDZ1 and PDZ2, the remaining INAD targets bind to one or more PDZ domains (numbered 1-5). Because NINAC is an actin-binding protein, the actin-based cytoskeleton may be indirectly coupled to the signalplex (adapted from Montell, 1999)
control of PDZ3 and PDZ4. Homomultimerisation and target binding occurs simultaneously (Xu et al., 1998a). If this homomultimerisation happens in vivo a single signalplex could indirectly link many TRP and TRP/TRPL ion channels with a single activated rhodopsin. Amplification of the cascade may be mediated by activation of many ion channels through protein/protein interactions with the signalplex (reviewed in Montell, 1999).

The function of the signalplex within the phototransduction cascade is still not fully determined but may:
(i) promote termination of the photoresponse (Shieh and Zhu, 1996; Tsunoda, 1997).
(ii) facilitate rapid activation of phototransduction (Tsunoda, 1997).
(iii) amplify the phototransduction cascade (Scott and Zuker, 1998b).

Ca\(^{2+}\) has been implicated in the light response for many years (Ranganathan et al., 1991; Hardie, 1991; Hardie and Minke, 1994). Ca\(^{2+}\) levels have been shown to alter the kinetics of activation and deactivation of the photoresponse (Hardie, 1995) and are required for the maturation of light activated currents as well as maintaining a responsive state (Hardie et al., 1993; Hardie and Minke, 1992). The influx of Ca\(^{2+}\) and the generation of DAG from the breakdown of PIP\(_2\) activates eye-specific PKC. Eye-PKC is required for the return to resting levels of Ca\(^{2+}\) in the cells (Ranganathan et al., 1994). Active PKC then phosphorylates specific targets to mediate the termination of the light response by catalysing the inactivation of the active intermediates (reviewed in Montell, 1999).

Two channels are associated with calcium permeability in the photoreceptors, transient-receptor-potential (TRP) and transient-receptor-potential-like (TRPL). The TRP channel is responsible for most of the calcium permeability (Hardie and Minke, 1992). The trp mutant is defective in maintaining a light response (reviewed in Montell, 1999), and has lost the calcium entry pathways (Peretz et al., 1994; Hardie and Minke, 1992). TRP encodes a six transmembrane domain protein (Montell and Rubin, 1989; Wong et al., 1989). TRPL was isolated by Phillips et al. (1992) in a screen for retinal calmodulin-binding proteins. TRPL is >50% identical to TRP at the N-terminus. Both TRP and TRPL are highly expressed in the rhabdomeres and display weak homology to voltage-gated ion channels (reviewed in Montell 1999).
An indication that TRPL as well as TRP contributes to light dependent calcium influx is that *trpl* mutant photoreceptor cells exhibit changes in the relative permeabilities to various cations (Reuss *et al.*, 1997). *trp* and *trpl* double mutants are completely unresponsive to light (Niemeyer *et al.*, 1996; Scott *et al.*, 1997; Reuss *et al.*, 1997). Ca\(^{2+}\) influx via TRP channels mediates amplification of rapid response termination and light adaptation through multiple feedback targets (Scott and Zuker, 1997; Hardie and Minke, 1995). The Ca\(^{2+}\) influx has also been implicated in regulating PIP\(_2\) metabolism by inhibiting PLC and facilitating PIP\(_2\) regeneration (Hardie *et al.*, 2001). TRP has one calmodulin (CaM) binding site whereas TRPL has two. In TRPL these CaM binding sites seem to be involved in Ca\(^{2+}\) inactivation of the channel (Scott *et al.*, 1997). The TRP channel relies on an ion block which is voltage dependent. The block intensifies as the cell depolarises from -70mV to 0mV (Hardie and Mojet, 1995; reviewed by Hardie and Raghu, 2001).

When the proteins were expressed in 293T cells, TRP acts as a regulatory channel whereas TRPL is constitutively active. This constant activity in the TRPL channel is most probably due to the lack of an interacting regulatory protein. When the proteins were co-expressed *in vitro* they formed a store-operated channel combining both proteins. TRPL bound more readily to TRP than to itself suggesting this is what may occur *in vivo* (Xu *et al.*, 1997).

Wild type flies are thought to most likely form TRP homomultimers and TRP/TRPL heteromultimers (Xu *et al.*, 1997) but others argue that they do not present as separate channels (Reuss *et al.*, 1997; Scott and Zuker, 1998a). The TRP and TRPL channels are responsible for controlling the influx of calcium into the photoreceptor cells but the structure of the channel *in vivo* has not been determined (reviewed in Montell, 1999).

### 1.2.2 Mammalian phototransduction

The initial stages of mammalian phototransduction are very similar to those of *Drosophila* phototransduction described in section 1.2.1. As described in section 1.1, the photosensitive pigments are situated in the lamellae of the outer segments of the
photoreceptor cells. When a photon of light hits the rhodopsin molecule, a conformational change occurs in the chromophore changing the 11-cis-retinal to the all-trans form. The opsin molecule attached to the chromophore goes through a series of intermediates, one of which is the active form, metarhodopsin II. The conversion to metarhodopsin causes an increase in affinity for transducin, the photoreceptor specific G-protein coupled to rhodopsin.

Transducin is a heterotrimeric molecule consisting of three subunits ($T_\alpha$, $T_\beta$, $T_\gamma$) in rods and two ($T_\alpha$, $T_\alpha$, $T_\beta_\gamma$) in cones. In the inactive state, transducin is bound to a molecule of guanosine diphosphate (GDP). When metarhodopsin II binds to transducin, a molecule of guanosine triphosphate (GTP) is substituted for GDP on the transducin $\alpha$-subunit. This results in the dissociation of the metarhodopsin II-transducin-GTP complex to give, metarhodopsin, $T_\alpha$-GTP and $T_\beta_\gamma$. Metarhodopsin II can catalyse up to 500 such exchanges before it is inactivated. Inactivation occurs in two stages, firstly, the protein is inactivated by phosphorylation of a cluster of serine residues near the C-terminus (Ohguro et al., 1993). This phosphorylation event is mediated by rhodopsin kinase and protein kinase C (PKC) under the control of a Ca$^{2+}$ binding protein called recoverin (Udovichenko et al., 1997; Sanada et al., 1996). This reduces the ability of the active rhodopsin molecule to interact with transducin but does not completely quench it (Krupnik et al., 1997). Complete inactivation, and restoration of rhodopsin to its original light-responsive state, occurs upon capping of the phosphorylated molecule by retinal arrestin (Granzin 1998; Freedman et al., 1996; Hausdorff et al., 1990). The deactivated opsin spontaneously re-binds to 11-cis-retinal to form light responsive rhodopsin (Reviewed by Pepe, 1999).

The released active $T_\alpha$-GTP molecule modulates the activity of the effector molecule, cGMP phosphodiesterase (PDE). The cGMP-PDE molecule is tetrameric and consists of an alpha, beta and two inhibitory gamma subunits. The active $T_\alpha$-GTP subunit modulates the removal of the two inhibitory gamma subunits from PDE. When free from the inhibitory subunits, the catalytic $\alpha$ and $\beta$ PDE subunits can hydrolyse cGMP into GMP and H$^+$, a critical step in phototransduction.
Conversion of GTP to GDP deactivates the $T_\alpha$-GTP.PDE$_{\alpha\beta}$ complex and releases PDE$_{\alpha\beta}$ which reassociates with the PDE $\gamma$-subunits. The $T_\alpha$ subunit, once again bound to GDP, reassociates with the $T_\beta\gamma$ subunits to complete the cycle initiated by the absorption of a photon by rhodopsin. Figure 1.16 schematically illustrates the human phototransduction cascade in rod photoreceptors. Hydrolysis of cGMP results in the reduced levels of the molecule within the photoreceptor cell which, in turn, causes the ion-gated membrane channels to close and the receptors to hyperpolarise (reviewed by Lem, 1998; Müller, 1995).

$Ca^{2+}$ concentrations also change during phototransduction and are involved in the recovery of the dark state of photoreceptors. One mechanism is through the regulation of retinal guanylate cyclase (ret-GC1). Ret-GC1 is a retinal-specific membrane associated guanylate cyclase found in the outer segments of photoreceptor cells (Dizhoor et al., 1994). This reduced level of cGMP also results in the closure of cation channels. This means that there is an outflux of $Ca^{2+}$ ions and $Ca^{2+}$ can no longer enter the cell, thus resulting in a decrease of intracellular $Ca^{2+}$. The falling levels of $Ca^{2+}$ activates ret-GC1 via the $Ca^{2+}$ binding proteins guanylate cyclase activating proteins (GCAPs) (Palczewski et al., 1994; Gorczyca et al., 1995; Dizhoor et al., 1995) to produce cGMP (Koutalos and Yau, 1996). The cGMP-gated cation channel is thought to be responsive to the concentration of $Ca^{2+}$ and is modulated by calmodulin (Hsu and Molday, 1993).

The biochemical and physiological properties of the components of phototransduction in cones is essentially the same as described above for rods. Although functionally analogous, rods and cones use different protein isoforms (products of a distinct set of genes) in most of these steps (reviewed in Yau, 1994). For example, different numbers of isoforms of transducin subunits are present in rods and cones, this may explain some of the physiological differences between the two receptor types (Peng et al., 1992).
Figure 1.16 Human phototransduction cascade
Shown is a schematic diagram of the human phototransduction cascade adapted from Baehr et al., 1996. After a photon of light is absorbed by a rhodopsin molecule it undergoes a conformational change which induces activation of transducin. The α-subunit dissociates from transducin which leads to the removal of the inhibitory γ-subunits of phosphodiesterase which results in the hydrolysis of cGMP and the closure of membrane ion channels (adapted from Polans et al., 1996).
1.2.3 Comparison of *Drosophila* and mammalian phototransduction

The first two cascades in the phototransduction pathway of *Drosophila* and mammals are the same, a photon of light induces a conformational change in the rhodopsin molecule which then stimulates a G-protein coupled cascade, after this point they differ. This section will describe the similarities and differences between phototransduction in *Drosophila* and mammals.

The major difference between the two pathways are that in darkness mammals have a 'dark current' that leaves the inner segments of the photoreceptors as K⁺ ions and enters the outer segments as Na⁺ and Ca²⁺ ions. This movement of ions causes a hyperpolarisation event (Copenhagen, 1991; Detwiler, 1992; Lagnado, 1992). This is in contrast to *Drosophila* where light depolarises the photoreceptors by stimulating Na⁺ influx across the plasma membrane (Rayner *et al*., 1990).

The mechanism of both the mammalian and *Drosophila* primary photoevents are fundamentally similar with the conversion of rhodopsin into metarhodopsin by a conformational change in 11-cis-retinal. This conformational change event in both system activates a G-protein cascade (Rayner, 1990). Mammalian transducin is responsible for metarhodopsin II coupling to cGMP-PDE (Hurley, 1987). In *Drosophila*, light stimulates the Gₐₐ subunit of the heterotrimeric protein (Pottinger *et al*., 1991; Nobes *et al*., 1992; Ryba *et al*., 1993). The Gₐₐ subunit binding to rhodopsin occurs in the same ratio as transducin binding in mammals. The Ca²⁺ binding protein recoverin plays a key role in controlling the binding of transducin and hence phototransduction. This is mirrored by G₉₉ inhibitory subunits that control *Drosophila* phototransduction (Lagnado *et al*., 1992).

It is from this point that the cascades in *Drosophila* and mammals diverge. In mammals TαGTP stimulates PDEα and β catalytic subunits by binding inhibitory PDEγ subunits. PDEγ remains bound to Tα even after GTP has been hydrolysed. Reassociation of Tβγ with Tα releases PDEγ. Light stimulated hydrolysis of cGMP by PDE results in the closure of the plasma membrane cation channels (Fesenko *et al*., 1985; Yau *et al*., 1985). This is compared to *Drosophila* where the molecules inositol triphosphate (IP₃) and Ca²⁺ are the photoexcitation transmitters (Bloomquist *et al*., 1988). Multiple

54
channels with different ionic selectivities exist in the photoreceptor plasma membranes (Nagy, 1991). The Drosophila trp channel is responsible for light activated calcium conductance (Hardie et al., 1992). Ca$^{2+}$ and calmodulin then control a second channel, trpl, which is light sensitive and uses Ca$^{2+}$ release to open the channels (Phillips et al., 1992).

Both mammals and Drosophila inactivate metarhodopsin by using rhodopsin kinases (Inglese et al., 1993; Palczewski et al., 1991) and arrestins (Hargrave et al., 1992). Drosophila have an additional measure to convert metarhodopsin back to rhodopsin quickly and efficiently. Drosophila metarhodopsin is unusually thermostable and can be directly re-isomerised back from all-trans to the 11-cis conformation of rhodopsin by absorption of longer wavelength light. The conversion occurs using the ambient light diffusing through the Drosophila eye tissue (Hardie, 1986).

The changes in concentration of free Ca$^{2+}$ induced by light cause deactivation and adaption in both mammals and Drosophila. However they differ in that light lowers the concentration of Ca$^{2+}$ in mammalian photoreceptors but increases it in Drosophila photoreceptors (Rayner, 1990; Yau, 1991; Selinger et al., 1993).

So in summary, both the Drosophila and mammalian phototransduction cascades are initiated by a photon of light which causes a conformational change in a rhodopsin molecule, which in turn activates a G-protein cascade. The G-protein cascade stimulates different pathways in Drosophila and mammals. In mammals the activation of cGMP by phosphodiesterase results in the closure of membrane channels. This is in contrast to Drosophila where hydrolysis of PIP$_2$ generates IP$_3$ molecules and, along with Ca$^{2+}$ release, causes membrane channels to open (figure 1.11). Although the end cycles of both systems differ, components of the Drosophila final PIP$_2$ hydrolysis and regeneration cascades have been identified in the mammalian retina. Described in section 1.2.4.

1.2.4 Evidence for the presence of the Drosophila PIP$_2$ pathway in mammals
As described in section 1.2.3, phototransduction in mammals and Drosophila are very similar in the initial stages, the cascades only differ in their final stages. In mammals,
the final stage of phototransduction is the activation of phosphodiesterase subunits whereas in *Drosophila*, the final stage involves the hydrolysis of phosphatidylinositol 4, 5-bisphosphate (PIP$_2$) by phospholipase C$_\beta$ (PLC$_\beta$). Although mammals use a different final pathway to *Drosophila*, homologues of the components of the *Drosophila* PIP$_2$ regeneration pathway have been identified in the mammalian retina (table 1.2). It has been through the examination of naturally occurring *Drosophila* mutants that the fly visual cascade has been determined and this in turn has given some insight into mammalian ocular disease.

This section will describe the elements from the *Drosophila* pathway that have been found in the mammalian retina. These genes are good candidates for mammalian retinal diseases, since they are expressed in the mammalian retina and many have been mapped to chromosomal regions that are known to contain retinal disease loci.
<table>
<thead>
<tr>
<th>Drosophila Mutant</th>
<th>Gene Product</th>
<th>Human Homologue</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{norpA} ) (no receptor potential A)</td>
<td>PLC</td>
<td>PLC4</td>
</tr>
<tr>
<td>( \text{rdgA} ) (retinal degeneration A)</td>
<td>DAGK</td>
<td>DAGK1, DAGK3, DAGK4</td>
</tr>
<tr>
<td>( \text{rdgB} ) (retinal degeneration B)</td>
<td>PITP</td>
<td>PITPN1, NIR2, NIR3</td>
</tr>
<tr>
<td>( \text{rdgC} ) (retinal degeneration C)</td>
<td>PP</td>
<td>PPEF1, PPEF2, PP7</td>
</tr>
<tr>
<td>( \text{cds} ) (CDP-diacylglycerol synthase)</td>
<td>Eye-CDS</td>
<td>CDS1, CDS2</td>
</tr>
</tbody>
</table>

Table 1.2 Human homologues of *Drosophila* genes from phototransduction cascade

Listed are the genes from the *Drosophila* phototransduction cascade that are known to have human homologues which have been identified in the mammalian retina. The mammalian homologues are listed in the right hand column.
1.2.4.1 rdgA
Retinal degeneration A (rdgA) encodes the enzyme diacylglycerol (DAG) kinase used in the breakdown of DAG to phosphatidic acid (PA). This conversion is the initiation point of the regeneration cycle of PIP$_2$ used in phototransduction. The enzyme DAG kinase is found in the photoreceptor cells. Mutations in this gene lead to retinal degeneration in the fly (Masai et al., 1993). Four mammalian homologues of rdgA have been identified from human retinal and brain libraries (Pilz et al., 1995; Ding et al., 1998). They are called DAGK1, DAGK1 (DAGK2, which is the most homologous to Drosophila rdgA), DAGK3 and DAGK4, and the human chromosomal locations of each have been determined. The genes map to human chromosome 12q13.3 (Hart et al., 1994), 7q32.3-33 (Ding et al., 1998), 3q27-28 (Fitzgibbon et al., 1995) and 4p16.3 (Endele et al., 1996) respectively. The chromosomal localisation of DGK1 contains a locus for an inherited form of retinitis pigmentosa (RP), a disease involving the progressive degeneration of the retina. No mutations in the DGK1 gene have, thus far, been associated with human RP (Bowne et al., 2000). DAGK3 gene maps to a region that contains a locus for optic atrophy (OPA1). Mutation screening in a cohort of German OPA1 patients did not reveal any mutations in the DAGK3 gene (Stohr et al., 1999). Houssa et al. (1997) found high levels of DAGK4 expression in rat brain, specifically in the cerebellum and hippocampus. The human gene DAGK4 which was mapped to chromosome 4p16.3 by Endele et al. (1996), report that the gene which is expressed in the mammalian retina and is 90% identical to the rat homologue. To date this gene has not been associated with any mammalian ocular disease. DAGK1, which was cloned by Schaap et al. (1990), and mapped to chromosome 12q13.3 by Hart et al. (1994), has also not been associated with any mammalian ocular disease.

1.2.4.2 rdgB
The various phosphatidylinositol transfer protein (PITP), retinal degeneration B (rdgB), and amino-terminal domain interacting receptor (Nir) PITPs can be divided into two structural families. The small, soluble PITP isoforms contain only a phosphatidylinositol transfer domain and have been implicated in phosphoinositide signaling and vesicle trafficking. In contrast, the rdgB proteins, which include Nir2 and Nir3, contain an amino-terminal PITP-like domain, an acidic, Ca$^{2+}$-binding domain, six putative transmembrane domains, and a conserved carboxyl-terminal domain.
However, the full biological function of mammalian rdgB-like proteins are still unclear (Fullwood et al., 1999).

Flies with mutations in the rdgB gene undergo light dependent retinal degeneration as this protein plays a role in the recovery after phototransduction (Hotta and Benzer, 1970). The first human homologue of the rdgB gene to be described was PITPN1 which was cloned in 1994 (Dickeson et al.) and subsequently mapped to human chromosome 17q13.3 (Fitzgibbon et al., 1994). The human Nir family of PITPs were then isolated and mapped, human RDGB-alpha (now called Nir2) maps to chromosome 11q13, and human Nir3 (second homologue to be identified) which maps to chromosome 12, constitute a family of phosphatidylinositol transfer proteins. All three homologues are abundantly expressed in the brain and retina (Guo et al., 1997; Aikawa et al., 1997; Fullwood et al., 1999; Lev et al., 1999). Expression studies using in situ hybridisation on mouse embryonic sections showed that Nir2 was expressed in the neural retina and the central nervous system. The pattern of expression in the mouse follows the pattern of expression observed in Drosophila (Rubboli et al., 1997). The mouse Nir2 gene was used successfully to rescue the Drosophila rdgB mutant phenotype, a result that suggests there are similarities between the later steps of invertebrate and vertebrate phototransduction not previously appreciated (Chang et al., 1997). In summary, PITPN1 maps to 17p13, Nir2 to 11q13 and Nir3 to chromosome 12. The regions on 17p13 and 11q13 contain loci for several retinal dystrophies including cone-rod dystrophy 5 (CORD5), central areolar choroidal dystrophy (CADC), exudative vitreoretinopathy (EVR1) and neovascular inflammatory vitreoretinopathy (VRNI) making Nir1 and 2 excellent candidate genes for these disorders (Guo et al., 1997).

1.2.4.3 rdgC

The retinal degeneration C (rdgC) gene encodes a serine/threonine phosphatase. rdgC is involved in the final stages of returning the metarhodopsin to the resting rhodopsin conformation. Drosophila rdgC mutants undergo light dependent retinal degeneration (Steele and O'Tousa, 1990; Steele et al., 1992). The Drosophila rdgC gene is expressed in the visual system and mushroom bodies of the central brain (Montini et al., 1997). To date, three human homologues of rdgC have been identified (PPEF1, PPEF2 and PP7) and all are expressed in the mammalian retina (Sherman et al., 1997; Montini et al., 1997; Huang and Honkanen, 1998). Protein protease with EF-hand motifs (PPEF),
and protein phosphatase 7 (PP7) are from a family of protein phosphatases. **PPEF1** has been mapped to human chromosome Xp22.2-22.1 and is 61% similar at the amino acid level to *Drosophila rdgC*. The **PPEF1** gene has been shown to be expressed in the rod photoreceptors, somatosensory neurones and inner ear cells of the developing mouse. The chromosomal location to which the gene has been mapped has a locus for X-linked juvenile retinoschisis (RS), a progressive vitreoretinal degenerative disorder. However, no mutations in **PPEF1**, have been detected in RS patients (Montini *et al*., 1997; van de Vosse *et al*., 1997).

The gene, **PPEF2** has been mapped to human chromosome 4 (Sherman *et al*., 1997) and splice variants of the gene have been identified. The rat homologue of **PPEF2**, was shown to be expressed in the inner segments of the rod photoreceptors and in the pineal gland. Due to the location within the photoreceptors, this gene is probably not directly involved in phototransduction since the phototransduction cascade occurs in the outer segments of the photoreceptors. Of the two mammalian **PPEFs**, **PPEF2** is reported to be most like **rdgC** (Sherman *et al*., 1997).

The gene encoding protein phosphatase 7 (PP7) was identified from a human retinal cDNA library. PP7 shares the catalytic core domain of the protein phosphatase family but has unique N- and C-terminal regions. The unique C-terminal region of PP7 has multiple calcium binding sites and shares 42% identity with the *Drosophila* gene **rdgC**. PP7 is not ubiquitously expressed like other PPases but is found only in the retina (Huang and Honkanen, 1998).

### 1.2.4.4 norpA

The *no receptor potential A (norpA)* gene encodes PLCβ. The PLCβ protein catalyses the hydrolysis of PIP₂ to IP₃ and DAG. *Drosophila norpA* mutants develop light dependent retinal degeneration (Bloomquist *et al*., 1988). The homologous mammalian PLC₄ has been identified in the retina and the brain and mapped to human chromosome 20p12 (Alvarez *et al*., 1995). Immunohistochemical staining using bovine retina identified PLC₄ in the photoreceptors, bipolar cells, horizontal cells and ganglion cells. Studies using mice lacking PLC₄ showed that the mice had impaired visual processing abilities (Pak, 1995; Jiang *et al*., 1996; Miyata *et al*., 2001). To date, however, mutations in the PLC₄ protein have not been associated with any human disease.
1.2.4.5 eye-<i>cds</i>
The <i>eye-<i>cds</i></i> gene encodes the regulatory enzyme CDP-diacylglycerol synthase (<i>cds</i>) which catalyses the formation of CDP-diacylglycerol (CDP-DAG) from phosphatidic acid (PA) and is thought to be the rate limiting step for phosphoinositide synthesis (Wu <i>et al.</i>, 1995). PA is a key regulator of the amount of PIP$_2$ available for signalling (Monaco and Gershengorm, 1992; Heacock <i>et al.</i>, 1996; Weeks <i>et al.</i>, 1997). CDP-DAG is an essential precursor in the biosynthesis of phospholipids such as cardiolipin, phosphatidyglycerol and phosphatidylinositols (Bishop and Bell, 1988; Kent <i>et al.</i>, 1991), which are required for cellular localisation, regulation and signalling (Raetz and Dowhan, 1990; Cronan and Rock, 1996; Dowhan, 1997). CDP-DAG controls inositide triphosphate (IP$_3$) levels (Lavie and Agranoff, 1996), and therefore affects levels of 3’phosphoinositides, which are the intermediates used by multidomain proteins to recruit their host proteins to specific regions in cells, (Lavie <i>et al.</i>, 1996; Cullen <i>et al.</i>, 2001), and hence have influence over membrane traffiking (Shepherd <i>et al.</i>, 1996).

cds genes have been identified in many species but the first to be described was from <i>E. coli</i> (Langley and Kennedy, 1978). The identification of the <i>E. coli</i> gene was followed by the cloning of the <i>Drosophila melanogaster</i> (Wu <i>et al.</i>, 1995), yeast (<i>S. cerevisiae</i>) (Shen <i>et al.</i>, 1996) and the human <i>CDS</i> genes (Heacock <i>et al.</i>, 1996; Weeks <i>et al.</i>, 1997; Halford <i>et al.</i>, 1998; Volta <i>et al.</i>, 1999). Comparisons of <i>cds</i> sequences derived from the various species show that the genes share a 20% or greater sequence identity at the amino acid level. The lowest level of identity is between human and bacteria that show 20% identity (Ranganathan <i>et al.</i>, 1991) and the highest is human CDS2 and <i>Drosophila</i> <i>cds</i> with 64.5% identity (Halford <i>et al.</i>, 1998; Volta <i>et al.</i>, 1999). The most interesting tissue-specific isoform of <i>cds</i>, in relation to this project, is <i>Drosophila</i> eye-<i>cds</i> (Heacock <i>et al.</i>, 1996). As detailed above, this isoform is a key regulator of phototransduction. The eye-<i>cds</i> enzyme is required for the continuous supply of PIP$_2$, which is required to be regenerated for fly vision (Wu <i>et al.</i>, 1995). The eye-<i>cds</i> mutant flies are viable but cannot sustain a light activated current as PIP$_2$ is depleted. Light exposure results in retinal degeneration in the mutant flies (Wu <i>et al.</i>, 1995). Wu <i>et al.</i> (1995) showed that the <i>cds</i> protein is found in all eight rhabdomeres in <i>Drosophila</i>.

The <i>Drosophila</i> <i>cds</i> gene transcript is 2.5kb (Wu <i>et al.</i>, 1995) and the corresponding protein is 447 amino acids in length with an estimated molecular mass of 28kDa and is
localised to the photoreceptor neurones in both the compound eye and the ocelli. Two human homologues of the *Drosophila cds* gene, *CDS1* and *CDS2*, have been identified (Halford *et al.*, 1998; Volta *et al.*, 1999). Human *CDS1* has been mapped to chromosome 4q21 (Halford *et al.*, 1998; Volta *et al.*, 1999). The predicted *CDS1* protein is 444 amino acids in length with eight predicted membrane-spanning domains (predicted in Volta *et al.*, 1999, using the PHD computer analysis program). Human *CDS1* has been reported to show two transcripts on a Northern blot of bands of 3.9kb and 5.6kb (Lykidis *et al.*, 1995). The transcripts have been reported to be expressed at different levels in different tissues. Expression of *CDS1* has been reported in the inner segments of the photoreceptors (Volta *et al.*, 1999). Human *CDS2* has been mapped to chromosome 20p13 (Halford *et al.*, 1998). The predicted protein is 445 amino acids in length, and Volta *et al.* (1999) have predicted that there are nine transmembrane domains. Volta *et al.* (1999) mapped the mouse homologue of *CDS2* to chromosome 2, to a region that is syntenic with human chromosome 20. Kent *et al.* (1995) showed dual localisation of rat Cds in the microsomal and mitochondrial fractions from liver extracts. Saito *et al.* (1998) used fluorescence studies on rat tissues to show that the Cds protein is found in close association with the membranes of the endoplasmic reticulum, with mRNA expression observed in the cerebellar Purkinje cells, pineal body and inner segments of photoreceptor cells. To date, there have been no reports of mutations in either human *CDS* genes causing retinal diseases.

The *CDS* genes remain good candidates for retinal disease primarily due to their retinal expression patterns but also their mapped positions on human chromosomes. *CDS2* is a particularly good candidate in this respect with several ocular disease loci positioned around the same chromosomal location including, Hallervorden-Spatz syndrome and Corneal hereditary endothelial dystrophy 2 (CHED2).

Hallervorden-Spatz syndrome (Taylor, *et al.*, 1996), is a progressive disorder that affects muscle tone and voluntary movement and causes, mental deterioration, emaciation (due to feeding difficulties) and visual impairment. Death usually occurs before the age of 30 years. At post mortum iron deposits were found in the brain tissue (Elejalde *et al.*, 1979). Casteels *et al.*, (1994) described patients who developed visual impairments and bilateral optic atrophy before developing any of the other classical symptoms listed above. This study suggested that the patients should be studied.
ophthalmologically to exclude the coincidental occurrence of optic atrophy with typical Hallervorden-Spatz syndrome. Since the map location remains close to that of the CDS2 gene and the expression pattern of the CDS2 gene implies multiple tissue expression, this gene is a good candidate for this disease. CHED2 was mapped to chromosome 20p13 (Hand et al., 1999). CHED2 is the autosomal recessive form of the disease (CHED1 being the dominant form). Patients suffer from corneal clouding but a corneal transplant restores their vision. This disease is interesting again due to the map position and results reported in this thesis (see section 3.6.3) enhance this interest as the results show there is expression of murine Cds2 in the mouse cornea.

1.3 Retinopathies

The retinal dystrophies are a heterogeneous group of diseases in which the retina degenerates, leading to either partial or complete blindness. They are the major cause of incurable blindness in the western world. The inherited forms of retinal degeneration, including retinitis pigmentosa (RP) and various macular degenerations, affect approximately 1 in 3000 people but many more suffer from age-related macular degeneration in later life. The heterogeneous nature of these disorders was evident even before the advent of molecular genetics with wide variations in severity, mode of inheritance and phenotype. RP can have an autosomal dominant, autosomal recessive, X-linked or digenic mode of inheritance. With the advent of molecular genetic techniques the genes underlying these disorders started to be identified. However it was not until 1990 that Dryja and colleagues linked mutations in rhodopsin (RHO) in a panel of patients with autosomal dominant form of RP (Dryja et al., 1990). Since then, many other mutations in the RHO gene showing both autosomal dominant and autosomal recessive inheritance of RP and the dominant form of congenital stationary night blindness (CSNB) have been reported. Although mutations in RHO are the most common cause of adRP with an estimated prevalence varying between 25-50% (Inglehearn 1998, Sochocki et al., 2001) many other genes for RP, CSNB and other retinal disorders have been identified through positional cloning and candidate gene approaches (reviewed in Gregory-Evans and Bhattacharya 1998, Inglehearn 1998, Molday 1998, Rattner et al., 1999, Clark et al., 2000 and Farrar et al., 2002). Figure 1.17 shows an ideogram of the progress made in the identification of new loci and genes involved in inherited retinal diseases in the past 22 years (taken from RetNet: 63
http://www.sph.uth.tmc.edu/RetNet/). To date 134 loci accounting for different retinal disorders have been reported and 89 genes have so far been cloned. Many of the genes identified code for photoreceptor specific proteins, which is not surprising considering the complexity and the number of genes involved in the various steps of the phototransduction cascade (see section 1.2.2), and the extreme metabolic activity of the photoreceptor cells which makes them vulnerable to even minor changes in the genes expressed in or near them. In fact, the underlying assumption of the candidate gene approach used to clone the majority of these genes is that a mutation in a photoreceptor-specific gene known to be indispensable for photoreceptor function is likely to confer a phenotype of retinal degeneration or dysfunction. Table 1.3 list some of the genes identified as causing a retinal phenotype and lists the function, where known, of these genes.

**Figure 1.17** This graph shows the number of genes causing retinopathies which have been mapped but not cloned (blue) and mapped and cloned (red) from January 1980 to January 2003. The number of genes identified has increased rapid from the early 1990s. Graph taken from http://www.sph.uth.tmc.edu/RetNet/.
<table>
<thead>
<tr>
<th>Retinal Degeneration</th>
<th>Gene</th>
<th>Function</th>
<th>Animal Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP</td>
<td>Rhodopsin</td>
<td>Visual Transduction</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>α-subunit of rod cyclic GMP-PDE</td>
<td>Visual Transduction</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>β-subunit of rod cyclic GMP-PDE</td>
<td>Visual Transduction</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>α-subunit of rod cyclic GMP-gated channel</td>
<td>Visual Transduction</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>Arrestin</td>
<td>Visual Transduction</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>PRKC3</td>
<td>Visual Transduction</td>
<td>No</td>
</tr>
<tr>
<td>Oguuchi Disease</td>
<td>RHOK</td>
<td>Visual Transduction</td>
<td>No</td>
</tr>
<tr>
<td>CSNB</td>
<td>CACNA1F</td>
<td>Visual Transduction</td>
<td>No</td>
</tr>
<tr>
<td>CSNB</td>
<td>GNAT1 (transducin)</td>
<td>Visual Transduction</td>
<td>No</td>
</tr>
<tr>
<td>CSNB</td>
<td>NYX</td>
<td>Visual Transduction</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>RDS-peripherin</td>
<td>Photoreceptor Structure</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>ROM-1</td>
<td>Photoreceptor Structure</td>
<td>Yes</td>
</tr>
<tr>
<td>Usher Syndrome</td>
<td>Harmonin</td>
<td>Photoreceptor Structure</td>
<td>No</td>
</tr>
<tr>
<td>Usher Syndrome</td>
<td>Myosin VIIa</td>
<td>Photoreceptor Structure</td>
<td>Yes</td>
</tr>
<tr>
<td>Usher Syndrome</td>
<td>Usherin</td>
<td>Extracellular matrix</td>
<td>No</td>
</tr>
<tr>
<td>Usher Syndrome</td>
<td>Cadherin-like protein</td>
<td>Extracellular matrix</td>
<td>No</td>
</tr>
<tr>
<td>RP/LCA</td>
<td>CRB1</td>
<td>Retinal development</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>RP2</td>
<td>Protein folding</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>RPGR</td>
<td>Protein transport</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>CRBP</td>
<td>Retinoid cycle</td>
<td>No</td>
</tr>
<tr>
<td>RP/LCA</td>
<td>RPE65</td>
<td>Retinoid cycle</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>ABCA4</td>
<td>Retinoid cycle</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>RGR</td>
<td>Retinoid cycle</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>LRAT</td>
<td>Retinoid cycle</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>MERTK</td>
<td>Disc shedding</td>
<td>Yes</td>
</tr>
<tr>
<td>RP/LCA</td>
<td>CRX</td>
<td>Transcription Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>NRL</td>
<td>Transcription Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>RP/LCA</td>
<td>TULP1</td>
<td>Transcription Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>NR2E3</td>
<td>Transcription Factor</td>
<td>Yes</td>
</tr>
<tr>
<td>RP</td>
<td>PRPC8</td>
<td>Pre-mRNA slicing</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>PRP3</td>
<td>Pre-mRNA slicing</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>PRPF31</td>
<td>RNA processing</td>
<td>No</td>
</tr>
<tr>
<td>RP-deafness syndrome</td>
<td>MTTTS2</td>
<td>tRNA</td>
<td>No</td>
</tr>
<tr>
<td>RP</td>
<td>TTPA</td>
<td>Transport</td>
<td>Yes*</td>
</tr>
<tr>
<td>LCA</td>
<td>AIP1L1</td>
<td>Protein folding/trafficking</td>
<td>No</td>
</tr>
<tr>
<td>LCA</td>
<td>RPRGRIP</td>
<td>Photoreceptor cilia</td>
<td>No</td>
</tr>
<tr>
<td>LCA</td>
<td>RetGC-1</td>
<td>cGMP synthesis</td>
<td>Yes</td>
</tr>
<tr>
<td>LOA</td>
<td>LHON</td>
<td>Metabolism</td>
<td>No</td>
</tr>
<tr>
<td>STGD3</td>
<td>ELOVL4</td>
<td>Metabolism</td>
<td>No</td>
</tr>
<tr>
<td>DHRD</td>
<td>EFEMP1</td>
<td>Structural</td>
<td>No</td>
</tr>
<tr>
<td>LHON</td>
<td>OPA1</td>
<td>Structural</td>
<td>No</td>
</tr>
<tr>
<td>LHON</td>
<td>OPA3</td>
<td>To be established</td>
<td>No</td>
</tr>
<tr>
<td>BB</td>
<td>BBS2</td>
<td>To be established</td>
<td>No</td>
</tr>
<tr>
<td>BB</td>
<td>BBS4</td>
<td>To be established</td>
<td>No</td>
</tr>
<tr>
<td>BB</td>
<td>BBS6</td>
<td>To be established</td>
<td>No</td>
</tr>
<tr>
<td>Retinoschisis</td>
<td>RS1</td>
<td>To be established</td>
<td>Yes</td>
</tr>
<tr>
<td>RP/NFRCD</td>
<td>RIBP1</td>
<td>Vitamin A cycle</td>
<td>No</td>
</tr>
<tr>
<td>Stargardts</td>
<td>ABCA4</td>
<td>Transport</td>
<td>Yes</td>
</tr>
<tr>
<td>Cone-dystrophy</td>
<td>RetGC1</td>
<td>To be established</td>
<td>No</td>
</tr>
<tr>
<td>Cone dystrophy</td>
<td>GCAP1</td>
<td>Calcium binding</td>
<td>No</td>
</tr>
<tr>
<td>Fundus</td>
<td>RDH5</td>
<td>Vitamin A cycle</td>
<td>Yes</td>
</tr>
<tr>
<td>albinopunctatus (SNB)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMD</td>
<td>VMD2</td>
<td>Protein folding</td>
<td>No</td>
</tr>
</tbody>
</table>

**Table 1.3** A list of the different types of retinal degeneration, the corresponding mutated gene, the gene function, if known, and existence of an animal model. The animal models are not all rodents and are a mixture of laboratory generated, diet induced and naturally occurring (adapted from Farrar et al., 2002). RP-Retinitis Pigmentosa, CSNB-Congenital Stationary Night-blindness, LCA-Leber’s Congenital Amaurosis, LHON-Leber hereditary optic neuropathy, STGD-Stargardt Disease, DHRD-Doyne Honeycomb Retinal Dystrophy, BB- Bardet Beidl Syndrome, NFRCD-New Foundland Rod-Cone Dystrophy, BMD-Best’s Macular Dystrophy.
1.3.1 Retinitis Pigmentosa

Retinitis pigmentosa is the most common retinal generation that is clearly hereditary and is estimated to affect 1 in 5000 people worldwide (Weleber and Gregory-Evans 2001). RP is the result of the loss of both rod and cone photoreceptors throughout the retina. If the rods and cones are lost within the first two years of life or are already dead or non functional at birth the diagnosis of the patient is that of congenital retinal blindness, also known as Leber congenital amaurosis, especially if recessively inherited (Francois, 1968).

1.3.1.1 Mode of inheritance

RP is genetically heterogeneous. Only a few genes for retinal degenerations had been discovered by the 1990s, but this number dramatically increased to 75 genes identified and 45 additional genes assigned to chromosome locations but not identified, by January 2002. It has been estimated that the actual number of genes could be two or three times that currently identified (Wright and van Heyningen, 2001).

Most forms of RP are monogenic and come under the classical inheritance patterns of autosomal dominant, autosomal recessive, X-linked or mitochondrial (maternally inherited). However there have been some families which have exhibited more complex patterns of inheritance such as digenic diallelic inheritance, digenic triallelic inheritance, dominant RP with reduced penetrance, and uniparental disomies (Rivolta et al., 2002). These modes of inheritance are shown schematically in figure 1.18.

Digenic diallelic disease begins in the offspring of unaffected parents, each of whom carry one of the two mutations for a disease. Only offspring that inherit both mutant alleles (double heterozygotes) are affected. Therefore the first generation of affected individuals mimics a recessive disease. However like a dominant disease, the affected individuals can transmit both mutations and hence the disease to their offspring. The transmission ratio would only be 1 in 4, not 1 in 2 as seen in a dominant disease. This mode of inheritance had been seen in laboratory animals when specific mating is carried out, but the first reported human disease of this nature came from studies of the RDS and ROM1 genes in a handful of RP families. All the affected individuals were double heterozygotes for a null mutation in the ROM1 gene on chromosome 11q and a
Figure 1.18 Schematic pedigrees demonstrating dominant, dominant with reduced penetrance, digeneic diallelic, recessive, and digenic-triallelic disease. All of these inheritance patterns have been observed in families with RP. Affected individuals are represented by filled circles and squares. A diamond in an individual's symbol indicates that he or she is an unaffected carrier of a dominant RP allele. The genotypes appear under each family member's symbol. A, a_1, a_2, B and b all designate mutant alleles, with A, a_1, and a_2 alleles found at one locus and the B, and b alleles found at a second not linked to the first. + designates a wild type allele. In the family with autosomal dominant disease with reduced penetrance, +_1 and +_2 designates isoalleles, with the isoallele +_2 preventing the expression of RP in a patient with the mutant A allele in trans (a situation presumed to be present in an unaffected carrier of an RP11 mutation). Note that the digenic diallelic disease mimics a recessive pedigree in the first affected generation and mimics a dominant disease with a reduced transmission ratio in subsequent generations. Digenic triallelic disease differs from autosomal recessive disease by its lower recurrence risk (adapted from Rivolta et al., 2002).
missense mutation in the RDS gene on chromosome 6p (Kajiwara et al., 1994, Dryja et al., 1997). Relatives with only one of the mutations were unaffected. (Kedzierski et al., 2001). Both ROM1 and RDS genes have been shown to be expressed in the rims of photoreceptor cell discs, where they form homodimers and interact to form tetrameric complexes that are essential for the structure of the rod and cone outer segment discs and without which the photoreceptors degenerate (Sanyal et al., 1980, Molday et al., 1987, Travis et al., 1989, Bascom et al., 1992, Lefkowitz, 1992, Goldberg et al., 1995, Goldberg and Molday, 1996a, Goldberg and Molday, 1996b, Moritz and Molday, 1996, Kedzierski et al., 1999a, Kedzierski et al., 1999b).

Digenic triallelic inheritance is more often found in syndromic diseases such as Bardet-Biedl (section 1.3.5). In this mode of inheritance there are generally more than two loci for the disease. In the case of Bardet-Biedl syndrome six genes have been associated with the disease. Until recently it was thought that the disease was recessive and that only one of the loci would be involved (Beales et al., 1997, Katsanis et al., 2000, Mykytyn et al., 2001, Nishimura et al., 2001). However it has been discovered recently that affected individuals have mutations affecting both alleles at one locus and another mutation affecting one allele at a second locus (Katsanis et al., 2001).

Families with dominant RP with reduced penetrance all exhibit linkage between the responsible gene, RP11, and markers within chromosome 19q13.4 (Al-Maghtheh et al., 1994, Xu et al., 1995, McGee et al., 1997). These studies provided evidence that a single locus determines whether or not carriers of an RP11 mutation would develop RP and this penetrance locus was either closely linked to RP11 or it was RP11 itself (McGee et al., 1997). The evidence was generated from siblings who were both carriers but were affected and unaffected. The unaffected carriers all inherited the same wild-type RP11 allele from their noncarrier parent and the affected siblings all inherited the other noncarrier RP11 allele from the unaffected parent. So there are wild-type alleles (isoalleles) at the RP11 locus or a closely linked locus which by themselves produce no observed phenotype, but, when in trans with a dominant mutation in RP11, permit or inhibit the pathogenicity of that mutation. The RP11 gene has recently been shown to encode a ubiquitously expressed protein which forms part of the spliceosome, a factor called PRPF31 (Vithana et al., 2001). It is thought that the deficiency in the splicing
function under the conditions of high splicing demand to replenish the photoreceptor
discs may underlie the pathology of RP11 (Deery et al., 2002).

Uniparental disomies occur when, as a result of abnormal events during either meiosis,
fertilization, or early stages of embryogenesis, an individual inherits both copies of a
chromosomal pair from one parent (Engle, 1980). This can either be heterodisomy, if
both chromosomes are the same as the donor parent, or homodisomy if one of the
parental chromosomes has been duplicated. This mode of inheritance has been reported
in RP patients with mutations in the MERTK gene (the human orthologue of the RCS
rat retinal dystrophy gene) (Thompson et al., 2002), the RPE65 gene (Hamel et al., 1993,
Nicoletti et al., 1995) and for a patient with rod monochromacy but where the gene
remains as yet unidentified (Pentao et al., 1992, Gal et al., 2000, Thompson et al., 2000,
Thompson et al., 2002).

1.3.1.2 Animal models of RP

Table 1.3 lists the animal models, which are available for retinitis pimentosa and Leber
congenital amaurosis. Some of these models have been found to be naturally occurring
while others have been generated in the laboratory as a tool to help in the investigation
of the human disease.

Various rhodopsin (RHO) transgenic mice have been generated. Mutations in this gene
account for approximately 15% of all inherited human retinal degenerations. Olsson et
al. (1992) demonstrated that the overexpression of the normal opsin gene results in
photoreceptor cell degeneration. Humphries et al. (1997) generated mice carrying a
targeted disruption of the Rho gene. Rho-/- mice did not develop rod outer segments and
lost their photoreceptors within three months. Heterozygotes retained most of their inner
and outer segments but with a degree of structural disorganisation. A second targeted
disruption of Rho was carried out by Lem et al. (1999). In these mice, the rod outer
segments failed to develop and retinal degeneration quickly followed. Retinae with a
single normal opsin allele developed normally except that there was only half the
normal quantity of rhodopsin protein the outer segments. The photoreceptors of these
mice also degenerated. These models allow for a better understanding of the functional
and structural role of rhodopsin in the normal retina and the pathogenesis of retinal disease. Recently, Kijas *et al.* (2002) has identified a naturally occurring mutation in the Rho gene of English Mastiff dogs which gives a phenotype similar to that seen in human patients with a RHO mutation.

Mice with an inactivated Rpe65 gene, which is involved in the retinoid cycle, also lack the visual pigment rhodopsin. Redmond *et al.* (1998) showed that the outer segment discs of the rod photoreceptors are disorganised and that rod responses are abolished. A four-base-pair deletion in the RPE65 was found to occur naturally in a pedigree of Swedish Briard dogs. Initially this mutation was thought to cause an analogous disorder to that of human congenital stationary night blindness (Narfstrom *et al.*, 1989) but was later found to be progressive and was termed hereditary retinal dystrophy (Wrigstad *et al.*, 1994).

Another gene that is thought to play a role in the retinoid cycle is ATP-binding cassette transporter, subfamily A member 4 (ABCA4). This gene can also be called ABCR (retinal specific) (Allikmets *et al.*, 1997). The ABCR gene maps to chromosome 1 p22 (Azarian and Travis, 1997, Nasonkin *et al.*, 1998, Allikmets *et al.*, 1998). The gene spans approximately 150 kb and is comprised of at least 51 exons Gerber *et al.*, 1998). Mutations in this gene have been found in autosomal recessive Stargardt’s disease 1 patients (Allikmets *et al.*, 1997, Nasonkin *et al.*, 1998, Rozet *et al.*, 1998), age-related macular dystrophy (ARMD) patients (Allikmets *et al.*, 1997, De La Paz *et al.*, 1998, Bernstein *et al.*, 2002), Retinitis pigmentosa-19 patients (Martinez-Mir *et al.*, 1998) and Cone-rod dystrophy patients (Maugeri *et al.*, 2000). Table 1.4 lists the known disease causing mutations in the ABCA4/ABCR gene. A knockout mouse model was generated and characterised by Weng *et al.* (1999). The outcome of the investigation indicated that the retinal degeneration in this case was caused by ‘poisoning’ of the RPE by lipofuscin fluorophore deposits in the RPE as a result of delayed dark adaptation.

The rds and rom1 genes encode proteins required for photoreceptor structure. Mice with rds and rom1 genotypes similar to those found in the families with digenic diallelic inheritance (described above) have been found to have corresponding phenotypes
through the generation of rds and rom1 knockout mice (Clarke et al. 2000, Kedzierski et al. 2001).

<table>
<thead>
<tr>
<th>Disease</th>
<th>Disease Causing Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-related macular dystrophy 2</td>
<td>D2177N, G1961E</td>
</tr>
<tr>
<td>Retinitis Pigmentosa 19</td>
<td>1847delA</td>
</tr>
<tr>
<td>Cone-rod dystrophy 3</td>
<td>Splice site mutations IVS30 G-T, +1, IVS40, G-A, +5.</td>
</tr>
</tbody>
</table>

**Table 1.4** List of known mutations in the ABCR gene that result in a disease phenotype.

A naturally occurring RPGR mutation was first seen in Siberian husky dogs and the phenotype closely resembled X-linked RP in humans (Zeiss et al., 2000). Hong et al. (2001) then created an RPGR deficient mouse model to further examine the structure and function of this protein. In this model the cone cells exhibited ectopic localization of cone opsins in the cell bodies and synapses and the rod cells had a reduced level of rhodopsin. Subsequently both cell types degenerated. RPGR appears to have a role in protein transport and in the maintenance of photoreceptor viability.

The Royal College of Surgeons (RCS) rat is a classical model of recessively inherited retinal degeneration whereby the RPE fails to phagocytose the shed outer segments and the photoreceptor then die. D’Cruz et al. (2000) used positional cloning to identify the retinal dystrophy (rdy) locus in the rat. They found that a small deletion that disrupted the Mertk gene. With this conclusion, Vollrath et al. (2001) sought to determine whether
gene transfer of human MERTK into a RCS rat retina would correct the defect. They successfully showed that the defect could be corrected and hence photoreceptor degeneration in the rat avoided.

CRX is a transcription factor that is expressed in the photoreceptors of the retina and also in the pinealocytes of the pineal gland. Furukawa et al. (1999) generated mice with a targeted disruption of the Crx gene. This resulted in mice that did not develop photoreceptor outer segments and lacked rod and cone cell activity and had disrupted circadian rhythms. Expression levels of several other photoreceptor- and pineal-specific genes were also reduced in these mice.

Mears et al. (2001) generated mice with a deletion of the NRL gene. The null mutant mice had complete loss of rod function and supernormal cone function mediated by S cones. The photoreceptors in the mutant retina had a cone-like morphology and had abnormal outer segment discs. Analysis of the retina confirmed the apparent functional change of rods into S cones. The authors suggest that NRL acts as a molecular switch during rod-cell development by directly modulating rod-specific genes while simultaneously inhibiting the S-cone pathway through the activation of NR2E3. Haider et al. (2001) proposed that NR2E3 may function by regulating genes involved in cone cell proliferation, and mutations in this gene may lead to retinal dysplasia and degeneration by disrupting normal photoreceptor cell topography as well as cell-cell interactions. A splicing error in the NR2E3 gene is thought to be responsible for an autosomal recessive retinal degeneration of the rd7 mouse that was identified at the Jackson Laboratory (Akhmedov et al., 2000, Haider et al., 2001).

The retinal degeneration mouse (rd) is a naturally occurring mutant exhibiting mutations in the gene coding the β-subunit of the rod cyclic guanosine monophosphate (cGMP) phosphodiesterase (Bowes et al., 1990). The rods degenerate rapidly and are essentially lost by one month of age (Carter-Dawson et al., 1978). The retinal degeneration slow (rds) exhibit mutations in the gene coding for rd/peripherin (Travis et al., 1989). The rod degeneration is much slower than in the rd mouse with degeneration occurring over months rather than days (Sanyal et al., 1986).
1.3.2 Cone, cone-rod and macular dystrophy

Patients with macular dystrophy lose the rods and cones of the central retina (macula) while the peripheral photoreceptors are spared (Green and Enger, 1993). Cone-rod dystrophies show an initial loss of cones, followed by rods, and cone dystrophies are where photoreceptor loss is restricted to the cones.

1.3.2.1 Best’s macular dystrophy

In Best disease, a yellow mass like the shape of an egg yolk forms at the macula. To identify the gene that is mutant in Best macular dystrophy, Petrukhin et al. (1998) defined the minimum genetic region by recombination breakpoint analysis, and mapped to this region a novel retina-specific gene designated VMD2. Genetic mapping data lead to the identification of 5 independent disease-specific mutations, and expression studies, provided evidence that mutations within the VMD2 gene were a cause of the disorder. The 3’ untranslated region (UTR) of the gene contained a region of antisense complementary to the 3’ UTR of the ferritin heavy-chain gene, suggesting the possibility of antisense interaction between VMD2 and FTH1 transcripts. The gene is thought to encode a protein which plays a role in protein folding. Thus far no animal model has been produced for this disorder.

1.3.2.2 Stargardt’s Macular Dystrophy

Most instances of Stargardt’s disease shows an autosomal recessive mode of inheritance and is characterized by macular dystrophy with flecks (Cibis et al., 1980) as described in section 1.3.1.1. Stargardt disease 3 (STGD3) is a progressive form of the disease and has an autosomal dominant mode of inheritance (Stone et al., 1994). The STGD3 patients have normal vision until the age of 5 years. Central vision is affected between 5 and 23 years of age, this is followed by central atrophy and visual acuity decreasing to 20/200 or worse in all patients over the age of 31 years. This disease is as the result of mutations in the gene encoding elongation of very long chain fatty acids-4 (ELOVL4) (Griesinger et al., 2000, Bernstein et al., 2001, Zhang et al., 2001). As yet there are no animal models for this disease.

1.3.2.3 Doyne Honeycomb Retinal Dystrophy (DHRD)

DHRD is characterized by small round white spots in the posterior pole of the eye including the macula and the optic disc that appear in early adult life. These progress to
form a honeycomb-like pattern (Doyne, 1899). These spots form as the result of swelling of the inner part of Bruch’s membrane (Collins, 1913). These physical changes occur long before impairment of vision. The majority of cases are a result of a mutation in the gene encoding EGF-containing fibulin-like extracellular matrix protein-1 (EFEMP1) (Stone et al., 1999, Matsumoto and Traboulsi, 2001, Tarttelin et al., 2001).

1.3.2.4 Cone and Cone-Rod Dystrophy

Cone dystrophies often present with photophobia, reduced visual acuity and abnormal colour vision. COD1 occurs due to mutations in the RPGR gene (Seymour et al., 1998, Demirci et al., 2002, Yang et al., 2002). COD1 is also sometimes referred to as incomplete achromatopsia, this however should be viewed as a symptom and not a diagnosis due to the effects the disease has on colour vision. COD3 is as the result of mutations in the guanylate cyclase activator-1A (GUCA1A) gene (Payne et al., 1998, Wilkie et al., 2001). The initial symptoms of this disease are reduced visual acuity associated with loss of colour vision are often apparent between the 20 and 40 years of age. Changes in the RPE at the macula are often apparent prior to the visual loss. This is followed by central atrophy of the retina.

Cone-Rod dystrophy is characterised by early loss of visual acuity and colour discrimination associated with the loss of cone cells followed by nyctalopia and progressive peripheral field loss as the rods subsequently degenerate (Yagasaki and Jacobson, 1989, Moor, 1992). There are several known cone-rod dystrophy loci, CORD1, 2, 6, 7, 8 and 9, although some of the causative genes remain unknown. Table 1.5 lists the known loci and where known the causative gene.

<table>
<thead>
<tr>
<th>CORD</th>
<th>LOCI/GENE</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18q21.1-q21.3</td>
<td>Warburg et al., 1997</td>
</tr>
<tr>
<td>2</td>
<td>CRX 19q13</td>
<td>Freund et al., 1997</td>
</tr>
<tr>
<td>6</td>
<td>GUCY2D 17p12</td>
<td>Kelsell et al., 1998</td>
</tr>
<tr>
<td>7</td>
<td>6q</td>
<td>Kelsell et al., 1998</td>
</tr>
<tr>
<td>8</td>
<td>1q12-q24</td>
<td>Khaliq et al., 2000</td>
</tr>
<tr>
<td>9</td>
<td>8p11</td>
<td>Danciger et al., 2001</td>
</tr>
</tbody>
</table>
Table 1.5 List known genes and loci for Cone-Rod Dystrophy.

CORD2 is inherited in an autosomal dominant pattern and is caused by mutations in the CRX gene on chromosome 19q13. Mutations in the CRX gene cause adCRD either by haploinsufficiency or by a dominant negative effect (Furukawa et al., 1997, Papaioannou et al 1997, Freud et al., 1997). See above for details of the Crx knockout mouse phenotype.

Three different mutations in codon 838 of the GUCY2D gene which encodes retina guanylate cyclase 1 (retGCl) on chromosome 17p12-p13 have been linked to autosomal dominant cone-rod dystrophy at the CORD6 locus (Kelsell et al 1998, Perrault et al., 1998). The three different mutations confer differing severities of the disease (Wilkie et al 2000). As yet no animal models for CORD6 have been generated.

1.3.3 Congenital Stationary night blindness

1.3.3.1 Description of CSNB

Patients with congenital stationary night blindness usually have a full complement of rod and cone cells, but the rod photoreceptors are insensitive to dim light or require an abnormally extended time to adapt to dim light. In other patients there is a defect in the bipolar cell layer which prevents the correct transduction of the signals from the photoreceptors (Dryja, 2000).

1.3.3.2 Genes and animal models associated with CSNB

The X-linked NYX gene is responsible for congenital stationary night blindness type 1 (Pusch et al., 2000). The X-linked form is distinguished from the autosomal form by the association in the former case with myopia. In general, X-linked congenital stationary night blindness is a nonprogressive retinal disorder resulting from a presumptive defect of neurotransmission between the photoreceptors and the bipolar cells. Both ‘complete’ and ‘incomplete’ forms of the disease can be present. The complete form lacks rod function by ERG and dark adaptometry and is accompanied by refractive error ranging from mild to severe myopia. The incomplete type shows some rod function on scotopic testing and is accompanied by refraction ranging from moderate hyperopia to moderate myopia (Musarella et al., 1989).
Congenital stationary night blindness type 2 is caused but mutations in the X-linked CACNA1F (calcium channel, voltage dependent, alpha-1F subunit) gene involved in cation exchange in photoreceptors. The various mutations in this gene, all give the 'incomplete' type of the disease (Boycott et al., 2001). The third form of CSNB is caused by mutations in the alpha subunit of the rod transducin gene, GNAT. This causes an autosomal dominant form of the disease (Dryja et al., 1996).

1.3.4 Retinoschisis

1.3.4.1 Description of Retinoschisis
Retinoschisis is intraretinal splitting due to degeneration. The abnormality may not be clinically apparent until middle life. Affected males show cystic degeneration leading to split in the retina, detachment of the retina, and finally complete retinal atrophy with sclerosis of the choroid. Cystic maculopathy is sometimes the only finding in these patients. The basic lesion is cystic degeneration in the deep nerve layer. Visual handicap is usually mild up to age 40 to 50 years; thereafter, impairment of vision is slowly progressive. The lesions of the disease are thought to be related to defects in Müller cells (Sauer et al., 1997).

1.3.4.2 Genes and animal models associated with Retinoschisis
The disease-causing RsI gene was identified on the X-chromosome (Sauer et al., 1997). To gain insight into the function of the retinoschisin protein and its role in the cellular pathology of RS, Weber et al. (2002) generated knockout mice deficient in RsI. They showed that the pathologic changes in hemizygous mutant male mice were evenly distributed across the retina, apparently contrasting with the macular-dominated features in human. Similar functional anomalies in human and hemizygous mutant male mice, however, suggested that both conditions are a disease of the entire retina affecting the organization of the retinal cell layers as well as structural properties of the retinal synapse.
1.3.5 Retinal degeneration syndromes

1.3.5.1 Usher’s Syndrome

Gorlin et al. (1979) summarized the classification of Davenport and Omenn (1977) as follows: type I has profound congenital deafness with onset of RP by age 10; type II has moderate to severe congenital deafness with onset of RP in late teens; type III has RP first noted at puberty with progressive hearing loss; type IV is a possible X-linked form.

El-Amraoui et al. (1996) reported that the myosin VIIA gene is expressed in the cochlear sensory hair cells during mouse embryonic development and that myosin VIIA expression is restricted to sensory hair cells in the developing human optic vesicle (Pearsall et al., 2002). They noted that this expression pattern correlated to the vestibular and cochlear dysfunctions resulting in balance problems and hearing impairment observed in Usher type 1 patients. As yet a knockout model has not been made of the MyoVIIa gene.

The USH2A gene encodes a protein, designated usherin by Weston et al. (2000), which is responsible for the less severe form of Usher syndrome. The protein is thought to have a role in the extracellular matrix.

1.3.5.2 Bardet Biedl syndrome

Bardet Biedl syndrome is characterized by mental retardation, pigmentary retinopathy, polydactyly, obesity, and hypogenitalism (Bardet, 1920, Biedl, 1922) and a high frequency of renal abnormalities (Alton and McDonald, 1973).

Bardet-Biedl syndrome is a genetically heterogeneous disorder with linkage to 6 loci: BBS1 on 11q13; BBS2 on 16q21; BBS3 on 3p13; BBS4 on 15q22.3; BBS5 on 2q31; and BBS6 on 20p12. Mutations have been found in genes at several BBS loci: BBS1, BBS2, BBS4, and BBS6 (MKKS). Although BBS had been originally thought to be a recessive disorder, Katsanis et al. (2001) demonstrated that clinical manifestation of some forms of Bardet-Biedl syndrome requires recessive mutations in 1 of the 6 loci plus an additional mutation in a second locus. While Katsanis et al. (2001) called this 'triallelic inheritance,' Burghes et al. (2001) suggested the term 'recessive inheritance
with a modifier of penetrance.' Mykytyn et al. (2002) found no evidence of involvement of the common BBS1 mutation in triallelic inheritance. There are as of yet no animal models of Bardet Biedl syndrome.

### 1.4 The Human Genome Mapping Project

The ultimate goal of the Human Genome Project is to sequence the entire human genome and to outline the position of all genes by the year 2005. The Human Genome Project is the result of an international 15-year collaboration involving over 20 laboratories and hundreds of people around the world. The concept was born through scientific meetings held during the years 1984 to 1986 (Palca 1986). However the US National Research Council proposed a broader programme to create genetic, physical and sequence maps of the human genome in parallel with other species and deposit all of this information into publicly accessible databases. The US Department of Energy launched the initiative and was supported by the US National Institute of Health and many other Institutes around the world including the UK Medical Research Council.

After the construction of the genetic and physical maps of the human and mouse genomes (Donis-Keller et al. 1987, Hudson et al. 1995, Nusbaum et al. 1999), and the partial sequencing of the S. cerevisiae (yeast) and C. elegans (worm) genomes (Oliver et al. 1992, Wilson et al. 1994), a new, more rapid strategy for genome sequencing was developed. This strategy was initiated by Craig Venter who, once a NIH employee, did not agree with the slow strategy being employed to sequence the genome using a series of ‘genomic signposts’ to assemble the sequence generated. Instead he opted for a more rapid method of ‘shotgun’ sequencing manageable sized fragments of the genome performed on random genomic fragments, the data was then entered into a second phase whereby the gaps in the sequence were closed and any ambiguous sequence was resolved by direct analysis. These random fragments were conceived as a short cut to the finish line by Venter, but as an obstacle by the HGP consortium due to the possibility of producing large numbers of errors. Venter however took the idea to the private sector and to a company called Celera Genomics who took up the idea and ploughed the much-needed funds into the project. The sequences produced from Celera were not however made publicly accessible. The competition between the two groups meant that the Human Genome Project moved into full production in early 1999. Under
the new strategy, a draft sequence of more than 90% of the genome was anticipated by
the year 2001, (Collins et al. 1998). The goal was achieved with the completion of the
draft sequence of the human genome (International Human Genome Sequencing
Consortium 2001, Hattori et al. 2000, Dunham et al. 1999). The new goal is to have a
full, finished sequence by the year 2003, 2 years ahead of previous predictions.

The information obtained so far predicts that the Human Genome contains between
30,000 and 40,000 genes which is significantly less than previous estimates of 100,000
(Antequera and Bird 1994, Liang et al. 2000), and only approximately twice as many as
C. elegans (19000) or Drosophila (13600).

The sequence information generated by HGP and on a smaller scale by independent
scientists is deposited and maintained by GenBank. GenBank is housed at the NIH and
provides annotated DNA sequences. The volume of sequences collected here has
rapidly escalated, especially in the last three years, and as of August 2002 there were
approximately 18,197,000 sequences. GenBank is part of an international collaboration
including the DNA DataBank of Japan (DDBJ) and the European Molecular Biology
Laboratory (EMBL) with exchange of DNA data occurring daily between the groups.
Submission of sequence data to GenBank comes under three forms, ESTs, STSs and
HTGs.

Expressed sequence tags (ESTs) are short sequences of around 300-500bp in length that
have been generated from cDNAs. When produced by the International Collaboration
they are deposited into GenBank in large batches, however independent scientists can
register single sequences. ESTs represent a 'snapshot' of genes expressed in a given
tissue at a given developmental stage, some even provide a chromosomal location when
the mapping data is known. Most of the accumulated sequences are stored in the dbEST
division of GenBank database (Boguski et al. 1993). The IMAGE consortium who
deposit an array of cDNA sequences, cDNA libraries and expression and mapping data
from clones into GenBank make the clones they have generated available for academic
researchers making ESTs a very useful tool to get a ‘handle’ on a gene of interest
(Lennon et al. 1996). Caution has to be observed however since artefacts including
genomic contamination, sequencing errors, pseudogenes, multiple transcripts of the
same gene due to different tissue expression patterns and improper splicing have been
associated with deposited ESTs, so 'bench' verification of obtained clone sequence is still required (Wolfsberg and Landsman, 1997). Whatever the negative aspects of ESTs they have produced a resource whereby homologous genes within and between species can be identified quickly and obtained in clone form for further analysis.

Sequenced Tagged Sites (STSs) are also short (200-500bp) and are known to be unique to the genome, ie they can be specifically detected by PCR in the presence of all other genomic sequences. These tags can be used to define a specific position on a physical map. High Throughput Genomics (HTGs) allow the rapidly growing genomic data to be available to the scientific community even in the 'unfinished' forms. Unfinished HTG contigs contain sequence greater than 2kb generated from a single cosmid, BAC, YAC or PAC clone but the sequence may be unordered or contain sequence gaps. HTGs progress through three phases and it is only when in phase 3 that the sequence has been ordered and the gaps closed that the information is moved in its final form to another division of GenBank.

The publication of the draft sequence means that the project is now entering the final stage of assembling consensus maps of the human genome and entering the transition period from the era of physical mapping to that of functional genomics.

1.4.1 Bioinformatic tools for sequence data analysis

Part of the facilitation of the Human Genome Project was the 'development of computational tools such as algorithms, software and databases for the collection, interpretation and dissemination of the vast quantities of complex mapping and sequence data generated by the Human Genome Project' (Pearson and Soll, 1991). Since then several hundred programs using different algorithms have been designed to predict individual coding features within any genomic sequence, but none of these tools cover all aspects or is 100% accurate. Table 1.6 lists some of these most commonly used tools. However the most effective results to date have been produced by the new generation exon predictor programs such as GENSCAN, but many of the programs still
### Table 1.6

Listed are the URLs of some of the most commonly used gene prediction programs and packages.

<table>
<thead>
<tr>
<th>Database</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGMP</td>
<td><a href="http://www.hgmp.mrc.ac.uk">http://www.hgmp.mrc.ac.uk</a></td>
</tr>
<tr>
<td>BLAST</td>
<td><a href="http://www.ncbi.nlm.nih.gov/BLAST">http://www.ncbi.nlm.nih.gov/BLAST</a></td>
</tr>
<tr>
<td>GRAIL</td>
<td><a href="http://avalon.epm.ornl.gov/Grail-bin/EmptyGrailForm">http://avalon.epm.ornl.gov/Grail-bin/EmptyGrailForm</a></td>
</tr>
<tr>
<td>NIX</td>
<td><a href="http://www.hgmp.mrc.ac.uk/NIX">http://www.hgmp.mrc.ac.uk/NIX</a></td>
</tr>
<tr>
<td>RUMMAGE</td>
<td><a href="http://gen100.imb-jena.de/rummage">http://gen100.imb-jena.de/rummage</a></td>
</tr>
<tr>
<td>GenOtator</td>
<td><a href="http://www.hgmp.mrc.ac.uk/Registered/option/geneotator.html">http://www.hgmp.mrc.ac.uk/Registered/option/geneotator.html</a></td>
</tr>
<tr>
<td>EMBOSS</td>
<td><a href="http://www.hgmp.mrc.ac.uk/software/EMBOSS">http://www.hgmp.mrc.ac.uk/software/EMBOSS</a></td>
</tr>
<tr>
<td>dbEST</td>
<td><a href="http://www.ncbi.nlm.nih.gov/dbEST">http://www.ncbi.nlm.nih.gov/dbEST</a></td>
</tr>
<tr>
<td>Ensembl</td>
<td><a href="http://www.ensembl.org">http://www.ensembl.org</a></td>
</tr>
<tr>
<td>GenWise</td>
<td><a href="http://www.sanger.ac.uk">http://www.sanger.ac.uk</a></td>
</tr>
<tr>
<td>GeneScan</td>
<td><a href="http://bioweb.pasteur.fr/seqanal/interfaces/genescan.html">http://bioweb.pasteur.fr/seqanal/interfaces/genescan.html</a></td>
</tr>
<tr>
<td>SNP</td>
<td><a href="http://www.snp.cshl.org">http://www.snp.cshl.org</a></td>
</tr>
<tr>
<td>SWISS-PROT</td>
<td><a href="http://expasy.ch/sprot/">http://expasy.ch/sprot/</a></td>
</tr>
<tr>
<td>FEX,HEXON,Fgene</td>
<td><a href="http://genomic.sanger.ac.uk/gf/gf.html">http://genomic.sanger.ac.uk/gf/gf.html</a></td>
</tr>
<tr>
<td>GeneMark</td>
<td><a href="http://genemark.biology.gatech.edu/GeneMark">http://genemark.biology.gatech.edu/GeneMark</a></td>
</tr>
<tr>
<td>Genefinder, MZEF</td>
<td><a href="http://sciclio.cshl.org/genefinder">http://sciclio.cshl.org/genefinder</a></td>
</tr>
<tr>
<td>GetMaps</td>
<td><a href="http://telethon.bio.unipd.it/GATMaps/">http://telethon.bio.unipd.it/GATMaps/</a></td>
</tr>
<tr>
<td>Mouse genomic</td>
<td><a href="http://informatics.jax.org/informatics">http://informatics.jax.org/informatics</a></td>
</tr>
</tbody>
</table>

81
have difficulty in accurately predicting splice sites (Thanaraj, 2000). To minimise the errors generated by using these programs, it is advisable to analyse genes using a combination of two or more gene prediction packages. Packages include many individual programs to analyse different aspects of sequence characteristics. In the following paragraphs four of the common gene prediction packages, NIX, RUMMAGE, Genotator and EMBOSS, will have their advantages and disadvantages described.

NIX is comprised of a collection of programs and is available on the UK HGMP website (see table 1.6 for URL). The programs include, GRAIL, TSSNW, GENSCAN, Fgene, FEX, HEXON, MZEF, GeneMark, Genefinder, HMMGene, Polyah, BLAST, Repeatmasker and tRNAscan (Table 1.7 shows comparison between packages and gives brief details of the programs). NIX main disadvantage is that all the parameters of the programs are preset and cannot be altered by the user. Up to 500kb of sequence can be analysed at any one time but for better resolution of the generated data, smaller fragments up to 20kb are more advisable. The analysis occurs on a unix server and the output data displayed on a web browser where it is stored for one month. Regions of interest can be enhanced using the web browser. Analysis of the input sequence occurs on both strands and the output data separated accordingly. The output data are grouped by program of the same function and are colour co-ordinated appropriately. For example the information is grouped as follows, CpG islands and promotor predictions, exons and gene predictions, BLAST searches against protein, EST, Unigene, mRNAs and other sequence datasets. Also identified are *E. coli* sequences, tRNA and vector sequences. The strength of the predictions is represented by the intensity of the coloured blocks, the raw data can be accessed by ‘clicking’ on these blocks (Reviewed Jones et al 2002).

RUMMAGE contains twenty-six programs including prediction of repeat motifs, exons and homology searches against a number of databases (see table 1.7). In contrast to the NIX parameters, RUMMAGE parameters can be altered but only by request to the program administrator. This package can analyse up to 1000kb, but shorter sequences give clearer results. Again the data are generated on unix and are visualised via a web browser. The output is similar to NIX but only the features present within the sequence are shown for any one stretch. Again both strands are analysed and regions of similar function are grouped together by colour. The selection of the blocks produces complex
<table>
<thead>
<tr>
<th>Feature</th>
<th>NIX</th>
<th>RUMMAGE</th>
<th>Genotator</th>
<th>EMBOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promoter prediction</td>
<td>GENESCAN</td>
<td>GENESCAN</td>
<td>NNPP</td>
<td>CpG plot</td>
</tr>
<tr>
<td></td>
<td>Fgene</td>
<td>Proscan promoter</td>
<td></td>
<td>newcpgseek</td>
</tr>
<tr>
<td></td>
<td>TSSW</td>
<td>Compile-Alu</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GRAIL CpG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon prediction</td>
<td>GRAIL</td>
<td>GRAIL</td>
<td>GRAIL</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>GENESCAN</td>
<td>GENESCAN</td>
<td>Genefinder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MZEF</td>
<td>Xpound</td>
<td>Xpound</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GeneMark</td>
<td>Exonsampler</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FEX</td>
<td>Compile-exon</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HEXON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Prediction</td>
<td>GENESCAN</td>
<td>GENESCAN</td>
<td>GeneMark</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Genefinder</td>
<td></td>
<td>Genefinder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GRAIL</td>
<td>Genesampler</td>
<td>Genie</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fgene</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poly A site prediction</td>
<td>GENESCAN</td>
<td>GENESCAN</td>
<td>PolyA</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PolyA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fgene</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GRAIL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORF prediction</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>EST searches</td>
<td>BLAST</td>
<td>BLAST</td>
<td>BLAST</td>
<td>Water</td>
</tr>
<tr>
<td>mRNA searches</td>
<td>BLAST</td>
<td>BLAST</td>
<td>No</td>
<td>Water</td>
</tr>
<tr>
<td>BLAST-protein db</td>
<td>Embl</td>
<td>Embl</td>
<td>Genpept</td>
<td>Genpept</td>
</tr>
<tr>
<td></td>
<td>Swissprot</td>
<td>PROSITE</td>
<td>PROSITE</td>
<td>OWL</td>
</tr>
<tr>
<td>STS identification</td>
<td>BLAST</td>
<td>BLAST</td>
<td>No</td>
<td>stsssearch</td>
</tr>
<tr>
<td>Repeats</td>
<td>Repeatmasker</td>
<td>Repeatmasker</td>
<td>BLAST</td>
<td>einverted</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Etandem</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>maskfeat</td>
</tr>
<tr>
<td>E.coli/vector sequences</td>
<td>BLAST</td>
<td>BLAST</td>
<td>No</td>
<td>vectorstrip</td>
</tr>
<tr>
<td>TRNA scan</td>
<td>TRNAscan-SE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

83
Table 1.7 Listed are some of the programs used in the packages described in the text to predict genes. The function has been indicated and those programs common to two or more packages are highlighted in bold (adapted from Jones et al., 2002).

tables of sequence data and hence this program's main disadvantage is that it is less user-friendly than NIX.

The Genotator package runs fewer programs than the previous two, but it includes the novel promoter program NNPP, the gene prediction program, Genie and a BLASTX search of GenBank (table 1.7). This package differs from the previous two by allowing all parameters to be altered by the user. The package can easily handle sequences of 25kb and is run in unix with the output displayed in X-windows. On the whole the data are viewed in a similar way to NIX and RUMMAGE. The raw EST data are shown in separate windows when the colour block is selected. Generally the output data are easier to extract in Genotator than they are in RUMMAGE.

Finally, the EMBOSS package contains a suite of over 100 applications (some are included in table 1.7), but exon and gene prediction programs are not available, however open reading frame detection programs are available. This program can be run in both unix and in windows HTML and the output data are given in the form of tables rather than graphically as in the previous packages. This makes comparison between analysis packages more difficult if pooling results.

On the whole these packages contain several common programs but some are unique, so running two or more can be beneficial. Caution must be taken when comparing packages, since some have the option of altering parameters and hence the results may be compromised. Probably the most commonly used package is NIX as it contains most of the required programs and it has the added bonus of being user-friendly.

The packages described above go a long way in predicting characteristics of the input sequence. However to predict a functional role of a gene, the information on expression patterns from the EST database can be examined. This is useful when trying to identify potential disease-causing genes in a specific tissue. Programs such as GetMaps can speed up this process as GetMaps interprets all the expression information that is stored.
in UniGene (a subdivision of GenBank). However resultant clones identified in this way should still be verified by Northern blot or in situ hybridisation using the tissue of interest.

Another useful site for exploring expression patterns in other species is the Mouse Genomic Informatics site (MGI). Using this site, the nineteen species listed, including cattle, dogs, cats, rats and mice, can be screened to see if the expression pattern of a gene of interest has been reported. This site also gives information on naturally occurring and generated animal models of genetic disease which may help outline the function of a particular gene (Lichanska and Simpson, 2002).

As more sequence information is generated and processed, database emphasis has been changing. There has become a need for the generation of techniques, databases and algorithms to deduce the function of genes and how they interact with each other. These databases and software programmes, such as SAGE, are starting to emerge with the introduction of tissue- and micro-array techniques. However a discussion of these databases is outwith the remit of this thesis.

In summary, from the advent of the Human Genome Mapping Project a vast quantity of sequencing data has been deposited into Genbank. This information however is useless without the facilities to analyse it. So under the current demand a plethora of resources have been developed (core resources listed in table 1.6). The software packages are at an advanced stage where the sequence data banks can be used to identify interesting genes based on homology (BLAST), possible gene features can be predicted (NIX), and the expression pattern derived (ESTs, MGI). This however is virtual, and although the development of ‘data-mining’ tools has cut down the laborious bench practice of library screening for initial gene identification, caution must be applied when interpreting computational results and the need for bench verification is still very much here.
1.5 Transgenic Techniques

1.5.1 *Drosophila melanogaster* as a genetic tool

*Drosophila* has been used as a model to understand genetic mechanisms for many years. Around the late 1960's *Drosophila* research was beginning to dwindle and interest was only associated with old-fashioned classical genetic, as most of the exciting new discoveries were being found using *Eschericia coli*. This was until P-elements were discovered in the *Drosophila* genome in the mid-1970s. Within the last twenty years these elements not only have become ubiquitous tools for *Drosophila* geneticists, but have also been used as a link between *Drosophila* and mammalian genetics.

1.5.1.1 *Drosophila* transposable P-elements

The first P-element found its way into the *Drosophila melanogaster* genome from another *Drosophila* species, *D. willistroni*, somewhere in Latin America. The sequences of the P-element derived from both species differ by only one base pair in nearly three thousand base pairs of sequence, which includes intronic sequence. *D. melanogaster* and *D. willistroni* diverged 60 million years ago, so this insertion most likely happened due to horizontal transmission from the species *D. willistroni* in recent years, mostly likely through the introduction of *D. melanogaster* from its native Western Africa to its now global colonies. Once in the genome, the P-elements made use of their new host's DNA repair mechanism to increase their copy number while transposing to new genomic positions. Within a few decades they spread worldwide to encompass nearly the entire species. Only the stocks maintained in laboratories from early *Drosophila* geneticists remained uninfected (reviewed by Engels, 1997).

P-elements are transposable DNA elements (one of a family of approximately thirty elements referred to as transposons), which are highly mobile in the germline of *Drosophila melanogaster*. The prototype of a P-element is a 2.9kb element from which other members have arisen from different internal deletion events. The elements are characterised by a perfect 31 base-pair inverted terminal repeat. This is the site of action for the transposase, the enzyme that cuts the chromosomal DNA where the P-element inserts. Transposase makes staggered cuts in the DNA and in a replicative process, a copy of the transposon inserts at the target site. The gaps are then filled using the host's DNA repair mechanism being sealed with DNA polymerase I and DNA
ligase, resulting in a duplication of the target site and formation of a new direct repeat sequence. Three long open-reading frames have been identified in the prototype DNA sequences. One of these open-reading frames encodes the transposase protein. The messenger RNA for this protein is produced by the removal of the introns from its primary transcript. The germline specificity of transposition is due to the fact that splicing out of one of these introns occurs in germ cells but not somatic cells (Laski et al., 1986). Some naturally occurring short P-elements are defective. They cannot encode functional transposase but are transposable when they are found in the nucleus of a cell which contains a non-defective P-element. The act of transposition has been referred to as illegitimate recombination because no homology between sequences is required and it is not site specific. The P-elements have been affectionately termed the Swiss army knife of Drosophila genetics as they are used as tools in many techniques. P-elements are now used in mutagenesis, transposon tagging, germline transformation and to provide the framework for the Drosophila genome project (reviewed by Engels, 1997).

1.5.1.2 Fly mutant rescue

In the early 1980s, Spradling and Rubin (1982) exploited the newly discovered P-elements (described in section 1.5.1.1). They used a naturally-occurring, short (1.2kb), internally deleted member of the P-element family. This defective P-element cannot encode any of the putative protein products of the 2.9kb prototype element (O'Hare and Rubin, 1983). The first experiments of this kind used the wild type rosy (ry) gene as the selectable marker for identification of insertion of target DNA (inserted into a prepared polylinker site) into the genome (figure 1.19). The ry gene encodes the enzyme xanthine dehydrogenase and conveys the wild type red eye colour, whereas the null mutant ry flies have brown eye colour making the gene easily identifiable. The construct was then co-injected with a plasmid containing a non-defective 2.9kb P-element into the posterior pole of a developing ry embryo.
**A. Structure of P-element derivative**

- **Polylinker**
- **31bp inverted repeat (IR)**
- **7.2kb \( r_y \) fragment**
  - Lacks own transposase gene

**B. Structure of helper P-element**

- **IR**
- **23bp of IR deleted therefore cannot be transposed**

*Figure 1.19 P-element derivatives as a vector system*

The figure shows (A) a short derivative of a P-element which lacks its own transposase gene. This derivative has been utilised by inserting a selectable marker gene \( (r_y) \) which confers eye colour and a polylinker site to insert gene specific DNA. In order for this construct to be inserted into the genome, it is co-injected into the posterior pole of an embryo (see figure 1.18) with (B) a helper P-element. The helper P-element encodes the transposase gene required for the insertion of (A) into the embryo’s genome, but lacks an entire repeat sequence at one end preventing its own transposition (adapted from Old and Primrose, 1994).
The injections into the posterior pole were performed at the blastoderm stage as the multinucleated embryo has not yet become partitioned into individual cells. The developing germ cells are at the posterior pole and are the first to be partitioned. Injection in this region increases the chance of the construct DNA being randomly inserted into one or more germ cells (figure 1.20). Ry progeny were recovered from 20-50% of injected embryos. This finding showed that foreign DNA could be delivered into the fly genome via vectors containing the components used by P-elements (reviewed in Primrose and Old, 1994). This technique was further developed to include the insertion of DNA from species other than Drosophila and the ‘tagging’ of the target DNA with reporter genes such as green fluorescent protein (GFP) which allows the detection of transgene expression using fluorescence microscopy.

Mammalian homologues of Drosophila genes have since been shown to be capable of insertion into the fly genome and can rescue mutant fly phenotypes and hence confirm that the genes confer a similar function in both species. An example of this is the rescue of the Drosophila rdgB (phenotypes detailed in section 1.2.4.2) mutant phenotype using the mouse homologue (Chang et al., 1997).

In this project, the CDP-diacylglycerol synthase genes Cds1 and Cds2 are tagged with GFP and cloned into P-element containing vector before being injected into fly embryos in the attempt to rescue the Drosophila eye-cds mutant phenotype (see section 3.7).

1.5.2 Gene targeting in the mouse
The previous section discussed the random insertion of foreign DNA into a genome, generally referred to as transgenics. In this section the more site-specific form of gene manipulation, gene targeting, will be discussed. This section will discuss the molecular basis of gene targeting events used in the preparation of a knockout mouse. Gene targeting relies on homologous recombination of DNA to introduce a change in an organism's DNA at a specific location within the genome.

The ability to insert foreign DNA into a host’s genome was first discovered through tissue culture experiments in the mid-1970’s. Jaenisch et al. (1976) demonstrated that a
Figure 1.20 Injection of transgenic DNA into embryo
Shown is a schematic representation of an early *Drosophila* embryo. Transgenic DNA is injected into the posterior pole just prior to pole cell formation and is incorporated into the germline (adapted from Old and Primrose, 1994).
murine retrovirus, Moloney murine leukaemia virus (MMLV), could infect mouse embryos and insert a proviral copy of DNA into the genome which was then transmitted as a Mendelian trait to the offspring. This however, did not imply that recombinant DNA, that is a DNA molecule formed by joining segments of DNA from different sources, could be integrated into the genome since the virus contained all the biological activity required to function and may seek out specific sights for integration. These experiments suggested that novel DNA could be tolerated by the developing mouse, even though it was under very specialised conditions (Mann et al., 1983; Jahner et al., 1985; Van der Putten et al., 1985). In early gene transfer experiments, selectable markers, which would result in the transformants being easily recognised, were transferred into cells (Maitland et al., 1977; Wigler et al., 1977). The frequency of success of insertions into cells was low when the genes were transfected into the cell. The success rate increased to 20% when the genes were injected directly into the cell nucleus (Graessman et al., 1979; Capechi, 1980).

However, probably the most important discovery in the history of germline transformation was from mammalian embryologists, who developed techniques to remove embryos, culture them briefly in vitro and then return them to foster mothers where the remainder of normal embryogenesis could occur. This opened the way for the production of chimaeric animals (reviewed by Palmiter and Brinster, 1986). It was through the study of the development of teratocarcinoma, which are tumours in the germ cells of the testis and ovaries many of which are benign, that led the way forward in transgenic studies. Malignant tumours in the germ cells are called embryonic carcinomas (EC). Mouse strain 129 has a high incidence of spontaneous embryonic carcinoma formation. EC cells are multipotent and can proliferate and differentiate into tissue and cell types representative of all the embryonic germ layers. The malignant cells are capable of transplantation due to the presence of the embryonic stem (ES) cell material in the tumours. It has been shown by transplantation of EC cells into blastocysts, which are then placed in a pseudopregnant mouse, a normal mouse is obtained and is composed of both cells from the host blastocyst and from the malignant carcinoma. This technique was exploited further to select mutant cells, while in culture, and transplant the mutants into the blastocyst. This resulted in a mouse carrying the mutant cells. If the mutant cells populate the germline it is possible to create a stock of mice with a lesion in every cell (Lehman, 1980; Brinster, 1993). This technique
progressed away from EC cells and instead healthy embryonic stem (ES) cells were used for manipulation of transgenic mice.

Transgenic technology takes advantage of four unique characteristics of the ES cell systems:
(i) they are derived from a single mouse embryo.
(ii) they can be maintained in culture.
(iii) they consist of pluripotent cells.
(iv) they are germline compatible (Nagy, 1997).

ES cells are derived from the inner cell mass of very early blastocysts. These cells express all the housekeeping genes as well as the essential transcription factor genes which are required to maintain the cells (Tada et al., 1993). These cells occur after the first differentiation in the cleaving mammalian embryo. The inner and outer cell masses separate, with the inner cell mass forming the first polarised extra embryonic epithelium. The inner cell mass is totipotent (Tarkowski, 1967). The 129Sv mouse strain was continued to be used to harvest the ES cells, due to the knowledge gathered from the EC cells. The preimplanted blastocyst is the optimum starting stage for manipulation (Wells et al., 1991). Genomic DNA is then introduced into the ES cells, usually by electroporation, and then the cells are implanted into the pseudopregnant mouse.

Gene targeting has many functional roles in genetics. By the manipulation of specific genes, their functional role can be examined. In this project a gene knockout construct (the principles of which are described later in this section) has been prepared to target the murine Cds2 gene in order to prevent the production of wildtype Cds2 protein. This technique relies on the genetic principle of homologous recombination. Homologous recombination relies on two regions of sequence identity to align and exchange genetic information during a crossover event (figure 1.21). In gene targeting studies, the principle of homologous recombination is used to manipulate an organism's DNA at a specific location. In knockout studies, a construct is prepared whereby two regions of gene specific, genomic DNA are separated by the insertion of a neomycin' gene. The role of the neomycin gene is two fold, firstly it disrupts the gene of interest and gives rise to the production of a truncated protein and secondly, the gene confers resistance to
Figure 1.21 Homologous recombination

Shown is a schematic representation of homologous recombination which is used in making transgenic models. A construct is designed using an expression vector which contains the genes neo' and thymidine kinase (t) which are employed as selectable markers for detecting targeted ES cells.
the antibiotic neomycin. The resistance to neomycin is one of two selectable markers used in this technique. The second selectable marker to be positioned in the targeting construct is thymidine kinase (tk) gene. The tk gene is sensitive to the antibiotic gancyclovir. The tk gene is positioned downstream of the 3’ region of specific gene sequence. If homologous recombination occurs effectively, the neomycin gene will be carried into the targeted embryonic stem (ES) cells but the tk gene will not (figure 1.21). Therefore correctly targeted cells will grow on medium containing both neomycin and gancyclovir and are easily identified. The linearised, cloned constructed DNA fragment with selectable markers are transferred into ES cells for homologous recombination (Thomas et al., 1986; Viville, 1997; Bishop, 1999).

Transgenic mice enable the analysis of gene function and provide models for gene therapy studies. Examples of successful gene targeting knockout mouse models, which give rise to retinal phenotypes, include the \( \gamma \)PDE knockout (Tsang et al., 1996), rhodopsin knockout models (Humphries et al., 1997, Lem et al., 1999), the retinol binding protein (RBP) knockout (Quadro et al., 1999), the Tulp1 knockout (Hagstrom et al., 1999), the arrestin knockout (Chen et al., 1999), GCAP knockout (Mendez et al., 2001), the retGC knockout (Yang et al., 1999) and the deletion of the rod transducin \( \alpha \)-subunit (Calvert et al., 2000).

Many mutations found in the rod opsin gene have been associated with RP, a progressive disease resulting in the loss of photoreceptor cells. In the rhodopsin\(^{-/} \) mice the gene has been disrupted in the first exon (Humphries et al., 1997; Lem et al., 1999). Later studies used transgenic lines expressing a rhodopsin triple mutant (V20G, P23H, P27L (GHL)) (Fredrick et al., 2001). The triple mutant mouse heterozygote shows slow progression of disease similar to human autosomal dominant RP (Frederick et al 2001) but the phenotype is further enhanced by crossing the mice with the rhodopsin null mutation. Both transgenic studies conclude that the absence of the rhodopsin gene results in the progressive loss of photoreceptor cells and thinning of all retinal layers. Complete loss of photoreceptor cells is seen by day 90 in the null mutant rhodopsin mice and by day 30 in the triple mutant/rhodopsin mice (Humphries et al., 1997; Toda et al., 1999; Hobson et al., 2000; Frederick et al., 2001).
Retinol binding protein (RBP) is the specific transport protein for retinol (vitamin A) in the circulation, its only known function being to deliver retinol to tissues. In tissues, retinol is converted to retinoic acid which regulates many hundreds of genes. In the eye, retinol is converted to 11-cis-retinal, the visual chromophore. The RBP knockout mouse has several phenotypes including reduced blood retinol levels and impaired retinal function in the first few months of life. The retinal function of the mice improved however when they were fed a vitamin A sufficient diet to the point that, at 5 months, they were practically indistinguishable from wild type mice. However, when deprived of vitamin A, vision remained abnormal and blood retinol became undetectable. RBP null mutants cannot mobilise internal stores of retinol so their vitamin A status is dependent on a regular dietary vitamin A intake. From the knockout of the RBP, it was concluded that RBP was required for normal vision in young animals and for retinal mobilisation in times of dietary vitamin A deficiency (Quadro et al., 1999).

Mutations in TULP1, a member of the Tubby family of genes, have been associated with autosomal recessive RP (Banerjee et al., 1998; Gu et al., 1998; Hagstrom et al., 1999). Tulpl protein is found exclusively in the photoreceptor inner segments and photoreceptor degeneration involving both the rods and cones cells is seen in the tulpl-f/+ knockout mice. A function of Tulpl is thought to be in maintenance and viability of rod and cone cell photoreceptors and the phenotype in the mice is consistent with the loss of polarised transport of nascent opsins to the outer segments (Hagstrom, 1999).

Calvert et al., (2000) deleted the gene for the rod transducin α-subunit in mice. As described in section 1.2, transducin is a heterotrimeric G-protein involved in coupling rhodopsin during phototransduction. The α-subunit is the active component of the molecule. Exons 3-6 were deleted an replaced by a neomycin cassette in the knockout construct. The results of this were that the hemizygous mice showed a small reduction in retinal transducin α-subunit content but retinal morphology and physiology of single rods are largely normal. In homozygous knockout mice, a mild retinal degeneration occurs with age. Rod-driven components were absent from electroretinograms but cone-driven components were retained. These results suggested that rod phototransduction is driven by a single type of transducin subunit.
As described in section 1.2, mammalian phototransduction follows from the photoisomerisation of a visual pigment to a G-protein mediated activation of a phosphodiesterase (PDE) hydrolysing cGMP. This results in the closure of ion-gated channels which causes a decrease in intracellular Ca^{2+} concentration, which acts through guanylyl cyclase activating proteins (GCAPs) to increase guanylyl cyclase activity. This increase in cyclase activity represents an important mechanism for photoreceptor adaption and recovery. Two eye-specific guanylyl cyclases have been identified in humans, RetGC1 and RetGC2 (Shyjan et al., 1992; Lowe et al., 1995) with homologues GC-E and GC-F in rats (Yang et al., 1995). When the GC-E (RetGC1) gene was disrupted, the rats displayed morphologically normal rods up to 1 year of age but the cone cells rapidly degenerated. The rods displayed a normal dark current despite the absence of GC-E with no compensatory increase in GC-F levels. The role of the GC-E gene has not yet been determined but it appears to be required for the maintenance of the normal retinal structure and function and seems critical for survival of cone cells (Yang et al., 1999).

Mendez et al. (2001) disrupted the expression of the GCAP genes to assess their function in the retina. The mutant mice appeared to have no Ca^{2+} regulation and had longer and slower rod responses to flashed light than wild type mice. Null mutant mice for both GCAPs had a GCAP2 transgene inserted and expressed in rods cells which restored maximum light induced guanylyl cyclase (GC) activity but did not restore normal flash response kinetics. From evidence obtained from the transgenic mice GCAPs seem to strongly regulate GC activity in rods and thereby extend the rod operating range, but GCAP1 and GCAP2 may subserve different specific function in mouse rods.

As described above and in previous sections, cGMP-PDE is a key regulator of vertebrate phototransduction. Tsang et al. (1996) disrupted the murine PDE\gamma gene, thus preventing the production of the two inhibitory \gamma subunits, which would normally bind to the active \alpha and \beta subunits. This disruption resulted in a rapid retinal degeneration which resembled human retinitis pigmentosa (RP), a common inherited degenerative retinal disease which progresses to blindness. Homozygous mutant mice displayed...
reduced levels of PDE activity instead of the predicted increase due to the lack of inhibition. In the mutant mice the α and β subunits formed but did not possess hydrolytic activity. Evidence from the mice therefore seems to suggest that the γ-PDE subunits are necessary for the survival of photoreceptors and PDE activity.

Arrestins are molecules used in the visual cascade to switch off activation of the G-protein cascade and terminate the signalling pathway that triggers cellular response by binding to G-protein coupled receptors. Chen et al. (1997) produced knockout models where one or both copies of the arrestin gene were disrupted. When one copy was removed, photoresponses were unaffected suggesting that the binding of arrestin is not a rate limiting step for the recovery of the rod photoresponse in humans while it is in Drosophila. When both copies of the gene were disrupted the flash response in the mice displayed a rapid partial recovery followed by a long final phase. This suggested that an arrestin-dependent mechanism initiates the quench of rhodopsin’s catalytic activity and that arrestin completes the quench. The evidence from the behaviour of the mutant mice suggests that arrestin is required for normal signal transduction but does not participate directly in light adaptation.

Mouse knockout models are useful tools in identifying the possible functional role of a gene or protein in an in vivo situation. However, although these models are useful tools in their own right, caution should be applied since mice are obviously not identical to humans and although the functional role in the mouse has been deduced, this does not mean the functional role in man will always be the same.
CHAPTER TWO

MATERIALS AND METHODS

2.1 Materials

2.1.1 Reagents
All reagents were purchased from Sigma, Promega, GibcoBRL (now trading as Invitrogen), Boehringer Mannheim (now trading as Roche Diagnostics) or Amersham Pharmacia Biotech unless otherwise stated; Oxoid supplied the broth stocks.

2.1.2 Vectors and Libraries
pBlueScript was supplied by Stratagene, and pGEM-T Easy by Promega. The targeting vector, pPNT was a kind gift from Victor Tybulewicz.
All gridded libraries were obtained from the Human Genome Mapping Project Resource Centre (HGMP-RC), Hinxton, UK (http://www.hgmp.mrc.ac.uk). The 129λ mouse spleen genomic library in λFIXII vector was purchased from Stratagene.

2.1.3 Bacterial strains and competent cells
DH5α competent cells were purchased from GibcoBRL. XL1 Blue cells were purchased from Stratagene and Promega supplied JM109 cells for cloning into pGEM T-easy.

2.1.4 Kits
Kits for the preparation of DNA, gel purification and PCR product purification were supplied by Qiagen. The oligolabelling kit and the cDNA synthesis kits were supplied by Amersham Pharmacia Biotech.

2.1.5 Oligonucleotide Primers
All primers used for polymerase chain reaction (PCR) and for sequencing were supplied by Sigma-Genosys Ltd.
2.2 Buffers, Solutions and Media

2.2.1 Media

NZY:
- For 1 litre
  - 21 g NZY mix
  - 2 NaOH pellets

L-Broth (LB):
- 10 g Tryptone
- 5 g Yeast extract
- 10 g NaCl

2YT:
- 16 g Tryptone
- 10 g Yeast extract
- 5 g NaCl

For plates:
- 15 g agar per litre

For top agarose:
- 6 g agarose per litre

2.2.2 Solutions & Buffers

SM Buffer:
- 100 mM NaCl
- 10 mM MgSO₄·7H₂O
- 50 mM 1 M Tris Cl (pH 7.5)
- 0.01% gelatin (w/v)

20x SSC:
- 3 M NaCl
- 0.3 M Sodium Citrate

Denaturing Solution:
- 0.5 M NaOH
- 1.5 M NaCl
Neutralising Solution: 1.5M NaCl
1M Tris Base
1mM EDTA (pH 7.2)

100x Denhardt’s: 2% Ficoll (w/v)
2% Polyvinyl pyrrolidine (PVP) (w/v)
2% Bovine Serum Albumin (w/v)

50x TAE: 9M Tris Base
57.1g glacial acetic acid
0.5M EDTA (pH8.0)

Homogenization Buffer: 1M NaCl
0.5M EDTA

Lysis Buffer: 1M NaCl
0.5M EDTA
20% SDS (w/v)

10x 3-[N-Morpholino]propanesulfonic acid (MOPS): 0.2M MOPS
50mM NaOAc
10mM EDTA
pH to 7 with NaOH

Transformation Buffer I: 1M KAc
1M RbCl
0.1M CaCl₂
1M MnCl₂
35% Glycerol (w/v)

Transformation buffer II: 100mM MOPS
0.1M CaCl₂
1M RbCl
35% Glycerol (w/v)
Pre/Hybridisation Buffer: 50mM Phosphate Buffer (pH 6.8)
4x SSC
75μg/ml Salmon Sperm DNA
5x Denhardt's Solution
10% Dextran Sulphate (w/v)
0.3% SDS (w/v)
0.15% Sodium Pyrophosphate (w/v)

P1 solution: 15mM Tris-Cl pH8
10mM EDTA pH8
100μg/ml RNase A

P2 solution: 0.2M NaOH
1% SDS (w/v)

P3 solution: 3M KOAc pH 5.5

Nuclease solution (10mls): 50mg DNase I
50mg RNase A
10ml 50% glycerol (v/v)
30mM sodium acetate (pH 6.8)

### 2.2.3 In situ Hybridisation Buffers

**DIG 1 Buffer:**
100mM Tris Base
100mM Tris HCl
150mM NaCl
pH to 7.6 with NaOH

**DIG 3 Buffer:**
100mM Tris Base
100mM NaCl
50mM MgCl₂.6H₂O
pH to 9.5 with HCl
Nitro Blue Tetrazolium (NBT) (Sigma) (100μl):
7.5mg NBT
70% dimethyl formamide

Bromo-Chloro-Indolyl-Phosphate (BCIP) (Sigma) (100μl):
5mg BCIP
100% dimethyl formamide

Levamisole (Sigma):
2.3mg in 10μl DIG 3 buffer

Pre Hybridisation Buffer:
10% 20x SSC
50% Formamide (v/v)
in dH₂O

Hybridisation Buffer:
5M Tris pH 7.5
100x Denhardt’s
20xSSC
0.5% SDS (v/v)
50% Formamide (v/v)
10% Dextran Sulphate (v/v)
75μg/ml Sheared Salmon Sperm DNA
in dH₂O

2.2.4 Protein Buffers

Electrophoresis Buffer (pH 8.3):
0.1M glycine
50mM Tris
0.1% SDS (w/v)

Transfer Buffer:
25mM Tris
190mM Glycine
20% Methanol (v/v)
Blocking Buffer:  
50mM Tris  
150mM NaCl  
5% Skimmed Milk (w/v)  
0.05% Tween-20 (v/v) pH 7.4

Tris buffered saline (TBS):  
10mM Tris  
0.9% NaCl (w/v)  
0.05% Tween-20 pH 7.4

Detection Buffer:  
100mM Tris, pH 9.5  
100mM NaCl  
50mM MgCl₂

Substrate Buffer:  
0.04mM NBT  
0.04mM BCIP  
in detection buffer

Quenching Buffer:  
10mM Tris, pH7.5  
1mM EDTA, pH8.0  
150mM NaCl

Coomassie Blue Stain:  
2% Coomassie Brilliant Blue R250 (w/v)  
50% methanol (v/v)  
10% acetic acid (v/v)

De-stain 1:  
50% methanol (v/v)  
10% acetic acid (v/v)

De-stain 2:  
5% methanol (v/v)  
7% acetic acid (v/v)

Anti-crack solution:  
30% methanol (v/v)  
3% glycerol (v/v)
Homogenisation Buffer:
- 5mM Tris
- 2mM EDTA
- 2mM EGTA
- 2% SDS (w/v) in dH₂O

Sample Buffer:
- 125mM Tris-HCl (pH 6.8)
- 4% SDS (w/v)
- 20% Glycerol (v/v)
- 10% 2-mercaptoethanol (v/v)
- 0.2% Bromophenol blue in dH₂O (w/v)

2.2.5 Gels

RNA gel:
- 30ml dH₂O DEPC
- 0.6g agarose
- 8ml 5x MOPS Buffer
- 2ml formaldehyde

Preparation of RNA samples for gel electrophoresis:
- 10μl formamide
- 3.5μl formaldehyde
- 2μl 5x MOPS Running Buffer
- 1μl RNA sample
- 2μl loading dye
- 5μl 1μg/ml Ethidium Bromide.

Gel Loading Dye:
- 25% Ficoll (w/v)
- 0.25% Xylene cyanol (w/v)
- 0.25% Bromophenol blue (w/v)
RNA loading buffer: 750μl formamide
240μl formaldehde
100μl DEPC H₂O
100μl glycerol
80μl 10% bromophenol blue
150μl 10x MOPS buffer.

ABI loading dye: 25mM EDTA (pH8.0)
Blue Dextran (50mg/ml)
Deionised formamide
5:1 formamide to EDTA/blue dextran

ABI gel: 0.04g Ammonium persulphate (APS)
40ml Sequagel-6 (National Diagnostics)
10ml Sequagel complete

SDS-PAGE Resolving gel: 30% Acrylamide mix
1.5M Tris (pH8.8)
10% SDS (w/v)
10% APS (w/v)
0.05% TEMED (v/v)

SDE-PAGE Stacking gel: 30% Acrylamide
1M Tris (pH6.8)
10% SDS(w/v)
10% APS(w/v)
0.05% TEMED(w/v)
2.3 Methods

2.3.1 DNA

2.3.1.1 Preparation of plasmid DNA (Miniprep)
Plasmid DNA was prepared by the alkaline lysis method (Bimboim and Doly, 1979) using a Qiagen miniprep spin kit according to manufacturer's instructions. 10mls of selective media was inoculated with a single bacterial colony and incubated with shaking at 37°C overnight. 1.5ml of the culture was spun in an eppendorf tube in a microcentrifuge at 13,000g for 3 minutes. The supernatant was removed to leave the cell pellet. The cell pellet was resuspended in 200μl of P1 solution and inverted several times to mix. 200μl of P2 solution was added and inverted to mix. 300μl of neutralising P3 solution was added immediately and the tube inverted to mix. The mixture was spun at 13,000g for 10 minutes, resulting in the separation of cellular debris and the DNA-containing aqueous solution. A pre-equilibrated Qiagen miniprep spin column was placed inside an eppendorf tube and the solution containing the DNA applied to the column. The column was spun at 13,000g for 1 minute. The eluted solution was discarded. The column resin was washed with 700μl of ethanol wash buffer and spun twice at 13,000g for 1 minute, with the eluted solution being discarded between spins. The column was then transferred to a clean eppendorf tube and the DNA was eluted with 50μl of dH2O applied to the centre of the resin. This was incubated at room temperature for 1 minute and then spun at 13,000g for 1 minute to collect the plasmid DNA.

2.3.1.2 Preparation of plasmid DNA (Maxiprep)
Up to 50μg of plasmid DNA were prepared using a Qiagen Maxi prep Qiafilter kit. 250ml of selective media was inoculated with 250μl of a 10ml overnight culture (described in 2.3.1.1) and grown with shaking at 37°C overnight. The cells were pelleted by centrifugation at 10,000g for 10 minutes at 4°C. The media was discarded and the cell pellet was resuspended in 10ml of solution P1. 10ml of solution P2 was added and the tube inverted to mix. The mixture was then incubated at room temperature for 5 minutes. 10ml of chilled P3 solution was added to the mixture and again inverted to mix. The mixture was applied to a Qiafilter column and incubated without the plunger for 10 minutes at room temperature before the mixture was applied to a pre-equilibrated Qiatip by pressure through the Qiafilter to remove any cellular
debris. The column was washed with 60ml of wash buffer. The DNA was eluted from
the column using 15ml of elution buffer. The DNA was precipitated from the eluate by
the addition of 10.5ml of isopropanol. The solutions were mixed then immediately spun
at 10,000g for 30 minutes at 4°C. The isopropanol was removed and the pellet was
washed by adding 5ml of 70% ethanol. This was spun at 10,000g for 10 minutes at 4°C.
The ethanol was removed and the pellet allowed to air dry briefly before being
resuspended in 100μl of dH2O.

2.3.1.3 PI Artificial Chromosome (PAC) Mini Prep
10ml of 2YT, supplemented with 25μg/ml kanamycin, was inoculated with a single
PAC colony and grown overnight at 37°C. 1.3ml of this culture was then transferred to
an eppendorf tube and spun in a microfuge for 3 minutes. The supernatant was
removed and another 1.3mls of culture added and spun again. The supernatant was
removed again and the pellet resuspended by vortexing in 300μl of P1 solution. 300μl
of P2 was added and the tube inverted several times to mix. It was then left to stand at
room temperature for 5 minutes. 300μl of P3 solution was added and mixed. The tubes
were then incubated on ice for 10 minutes. The samples were spun for 10 minutes and
the supernatant transferred to a tube containing 800μl of isopropanol at room
temperature. The tubes were incubated at -20°C for 20 minutes before being spun for
20 minutes at 4°C. The excess isopropanol was removed and 500μl of 70% ethanol was
overlaid. The tubes were spun for 10 minutes. Excess 70% ethanol was removed and
the pellet resuspended in 50μl of dH2O.

2.3.1.4 PI Artificial Chromosome (PAC) Maxi Prep
250ml of 2YT with 25μg/ml kanamycin was inoculated with 2.5ml of an overnight
culture (described 2.3.1.3) and grown at 37°C overnight in a shaking incubator. The
culture was then decanted into six 50ml Falcon tubes (Greiner, UK) and spun at
3000rpm for 10 minutes. The supernatant was removed and each cell pellet was
resuspended in 10ml of P1 solution. The cells were then lysed with 10ml of P2 solution
and left to stand on ice for 5 minutes. 10ml P3 solution was added and stirred. The
tubes were incubated on ice for 10 minutes before being spun down for 10 minutes.
The supernatant from the six tubes was decanted onto one Qiagen column pre-
equilibrated with buffer QBT. Polymer wool was used as a pre-filter to prevent the
column from becoming blocked. The column was then washed with 2 volumes of
Qiagen wash buffer. The DNA was eluted using 15ml of elution buffer. 10ml of isopropanol at room temperature was added to precipitate the DNA. The DNA was pelleted by spinning at 10,000rpm at 4°C for 30 minutes. The DNA pellet was rinsed in 70% ethanol and spun again. Excess 70% ethanol was removed and the pellet was left to air-dry briefly before being resuspended in 200μl dH2O.

2.3.1.5 Phenol/Chloroform Extraction
The addition of phenol to an aqueous solution of nucleic acids helps remove any contaminating proteins. This purification process is enhanced by the addition of chloroform/isoamyl alcohol (CHCl3/IAA in a ratio of 24:1) which denatures proteins and enhances the separation of the organic and inorganic phases. The extraction was carried out by adding 1 volume of phenol, vortexing, then spinning at 13000rpm for 5 minutes. The top aqueous layer was removed and transferred to a clean tube. 0.5 volumes of phenol and 0.5 volumes of CHCl3/IAA are added, vortexed, and spun down as before and the aqueous layer removed and transferred to a clean tube. 1 volume CHCl3/IAA was added and vortexed. The tube was spun again as before, and the top layer removed and transferred to a clean tube. The DNA was precipitated by the addition of 1ml of 100% ethanol and incubated overnight at -20°C (or -80°C for 45 minutes). The tube was spun at 13000rpm for 30 minutes at 4°C. Excess ethanol was removed and the pellet was then washed with 100μl of 70% ethanol. The excess 70% ethanol was then removed, the pellet was allowed to air dry briefly and was then resuspended in an appropriate volume of dH2O, and stored at -20°C.

2.3.1.6 Genomic DNA Preparation
Genomic DNA was prepared from various mouse tissues. Initially the weight of the tissue to be used was determined. 5ml of homogenization buffer was added per gram of tissue. The tissue was homogenized using a small Ultra-turrax probe (Janke & Kunkel-IKA-Labortechnik) for 30 seconds. The resulting homogenate was poured through gauze into a 30ml centrifuge tube. The cells were pelleted by spinning the tube at 10,000rpm at 4°C for 30 minutes. The cells were then resuspended in 1.5ml of ice cold 3M sodium perchlorate (NaClO4). 3ml lysis buffer was added and the tube was left on a rotating wheel at room temperature for 45 minutes. An equal volume of CHCl3:IAA (5:1) was then added to the sample and mixed for 10 minutes before being spun for 10 minutes at 10,000rpm. The aqueous layer was then removed and transferred into a clean tube and the last step was repeated. The DNA was then precipitated by adding 0.1
volume of 3M sodium acetate (pH 4.8) and 3 volumes of absolute ethanol. The DNA was spooled onto a glass rod, it was washed in 70% ethanol before being left to air dry. The DNA was resuspended in 1.5ml of TE (pH 8). SDS to a final concentration of 0.5% (37.5μl of a 20% stock) and proteinase K to a final concentration of 500μg/ml (15μl of a 50mg/ml stock) were added and incubated for 60 minutes at 50°C. The sample was phenol/chloroform extracted (see 2.3.1.7), reprecipitated, spooled onto a glass rod, washed and dried as before, then resuspended in an appropriate volume of TE and stored at 4°C.

2.3.1.7 Preparation of DNA from Embryonic Stem cells.
All embryonic stem (ES) cell work was performed by Dr Kim Wells at the Gene Targeting Unit, Charing Cross Hospital, London.

ES cells for DNA preparation were grown to confluency until they acidified the culture media every day for several days (see section 2.3.12.3). The cells were then washed in PBS and resuspended in 0.5ml of 50mM Tris pH8, 200mM NaCl, 100mM EDTA and 100μg/ml proteinase K and incubated overnight at 37°C. An equal volume of isopropanol was added and the tube shaken gently until the DNA was visible. The DNA was spooled out onto a glass rod, rinsed in 70% ethanol for 15 minutes and redissolved in 100μl of TE.

2.3.1.8 Preparation of Phage DNA
Single phage plaques (see section 2.3.1.20.1) were extracted using a Pasteur pipette and put into 1ml of SM buffer then allowed to elute for 2 hours. 50μl of this phage stock was added to 500μl of host cells (prepared as detailed in section 2.3.1.21.1) and incubated for 30 minutes at 37°C. The pre-adsorbed phage were then added to 37ml of NZY broth in a 250ml flask and incubated overnight at 37°C with gentle shaking. The mixture was then transferred to a centrifuge tube and 100μl of chloroform and 370μl of nuclease solution added. This mixture was incubated at 37°C for 30 minutes. 2.1g NaCl was then added and dissolved gently. After spinning at 7000rpm for 20 minutes at 4°C the supernatant was transferred to clean tubes containing PEG. The PEG was dissolved and the tubes incubated on ice for 60 minutes. They were then spun at 7000rpm for 20 minutes at 4°C. The supernatant was poured off and the phage pellet resuspended in 500μl SM, this was allowed to mix for 30 minutes. One chloroform extraction was carried out (see section 2.3.1.5). The aqueous phase was transferred into
clean eppendorf tubes and 20μl of 0.5M EDTA, 10μl 10% SDS and 10μl proteinase K (2.5mg/ml) were added. The tubes were incubated at 65°C for 30 minutes. One phenol and one chloroform extraction was carried out (see section 2.3.1.5). 170μl of 6M ammonium acetate was added to the aqueous phase and 700μl of isopropanol was added to precipitate the DNA. The phage DNA was then hooked onto a glass rod, washed in 70% ethanol and allowed to air dry briefly before being resuspended in 500μl TE.

2.3.1.9 Restriction Endonuclease Digestions
Restriction endonuclease digests were carried out using the reaction buffer supplied by the enzyme manufacturer. 1-2 U of enzyme was used to digest 1μg of DNA, however the volume of enzyme added did not exceed 1/10 of the total volume. The digests were incubated at the optimum temperature recommended by the manufacturer for 1-3 hours, except for genomic DNA digests which were incubated for 6-8 hours.

2.3.1.10 Agarose Gels and Visualisation
Agarose gel electrophoresis was used to separate DNA fragments, the size of the DNA fragments to be separated determined the percentage of the agarose (w/v) used. The appropriate amount of agarose was dissolved in 1x TAE buffer by heating in a microwave oven. When the solution had cooled to below 60°C ethidium bromide was added to a final concentration of 0.5μg/ml. The solution was then poured into a gel former with an appropriate comb. The gel was allowed to set completely and was placed in a horizontal gel tank and covered in 1x TAE buffer, the comb was then removed. The samples to be run had 1/10 volume of loading dye added and were loaded into the slots. A commercial 1kb DNA ladder (Invitrogen, UK) was also loaded for size determination. The DNA was electrophoresed horizontally along a voltage gradient of 1-5 V/cm until the required separation had been achieved. During electrophoresis, the ethidium bromide in the gel becomes intercalated with the DNA, this enables the DNA to be visualised on a UV transilluminator (302nm). The gel was either photographed using a Polaroid MP-4+ camera (Kodak) with a Wratten gelatin filter attached or with Gene Genius Bio Imaging system (Syntoptic Ltd) using Syngene imaging software (Synoptic Ltd).

2.3.1.11 Southern blotting
Southern blotting allows the transfer of DNA fragments from an agarose gel onto a nitrocellulose or nylon filter (Southern, 1975). DNA fragments were separated using
agarose gel electrophoresis (described in 2.3.1.10) and then transferred to a nylon membrane (Hybond-N) (Amersham Pharmacia Biotech, UK) using the technique described by Southern (1975). Briefly the DNA on the gel was denatured and neutralised before being blotted onto a nylon membrane by capillary action. The blotting step was performed overnight in 20xSSC. The position of the wells was marked onto the filter before removal from the gel. The filter was then washed in 3xSSC, air dried and baked for 2 hours at 80°C to bind the DNA to the filter.

2.3.1.12 Cell preparations

2.3.1.12.1 Phage host cell preparation
Host cells were prepared by inoculating 10ml of NZY broth, to give a final concentration of, with 0.2% maltose (w/v) and 10mM MgSO₄ with a single colony of the host strain required and incubated overnight at 37°C. The cells were pelleted by centrifugation at 3000rpm for 5 minutes and resuspended in 0.5 volumes of 10mM MgSO₄.

2.3.1.12.2 Preparation of competent DH5α cells
10ml of LB was inoculated with a single DH5α colony and grown overnight at 37°C. 50ml of LB was then inoculated with 0.5ml of the overnight culture. The culture was grown with shaking at 37°C for 60 minutes or until the O.D₅₅₀nm was 0.45 - 0.55. The culture was chilled on ice for 10 minutes and then spun down at 4000rpm at 4°C for 10 minutes to pellet the cells. The cell pellet was resuspended in 20ml of Transformation Buffer (TfB) I and incubated on ice for a further 5 minutes and spun as before. The pellet was resuspended in 2ml TfB II and incubated on ice for 10 minutes. The competent cells were stored at -80°C in 200µl aliquots.

2.3.12.3 Embryonic Stem (ES) cell preparation
R1 ES cells (129J background, kind gift of Andreas Nagy, University of Toronto) were grown on mitomycin C treated STO SNL2 fibroblast feeder layers (ATCC CRL-2225). These feeder cells were transfected with a construct expressing LIF and are G418 resistant. Feeders were grown in DMEM containing Glutamax-1, 10% FCS, 20 µg/ml Gentamycin (all supplied by Gibco BRL now Invitrogen). Feeders were routinely passaged, and split 1:5 every 2 days over several weeks. A new vial was thawed to provide feeders at the point of picking the selected ES colonies to avoid problems with senescence over multiple passages. When cultures reached confluence the medium was
replaced with medium containing 10 μg/ml mitomycin C (Sigma) and incubated for 2-3 hours. The media was removed and the cells washed twice with PBS, once with PBS/EDTA (0.02% EDTA final) and trypsinised with 1ml of 0.1% trypsin in PBS/EDTA per 90mm dish for a maximum of 5 minutes. The action of trypsin was stopped by the addition of 10ml of normal growth media, cells were pelleted, resuspended in a small volume, counted and plated on gelatin coated dishes at a density of 3 x 10^6/90mm dish (gelatin 0.1% final concentration in sterile water). Cells were prepared the day before thawing a vial of ES cells.

The ES cells were grown in DMEM/F12 media (including Glutamax-1) supplemented with 1% NEAA (non-essential amino acids), 0.1mM ME, 1000u/ml Penicillin/Streptamycin or 20 μg/ml Gentamycin and 15% GlobePharm FCS (batch tested for ES cell growth). During the initial rounds no additional LIF was provided but for the dose response experiment and the final targeting attempt, LIF (ESGRO™ Chemicon) was added to a final concentration of 1000U/ml. Early passage ES cells (p15) were thawed and plated on mitomycin C treated feeders. The cells were fed every day and passaged 1:3 2 days after plating and again two days later to give 9 plates. Two days later the cells were fed early in the day and were later passaged as follows:

ES cell cultures were washed with PBS then PBS/EDTA and trypsinised with 3ml of standard trypsin solution for 5 minutes in the incubator. The cells were pipetted up and down to ensure a single cell suspension, 7ml of growth media added and the cells were mixed again by pipetting. The cells were then left for 30 minutes in the incubator for the fibroblast feeder population to adhere back to the plate. The less adherent ES cells were collected carefully from the plate. The cells were pelleted and resuspended in 5 ml of ice cold PBS and an aliquot counted. Cells were diluted to a concentration of 1 x 10^7/ml in PBS or as required.

2.3.1.13 Ligation of DNA fragments
Ligation reactions allow a specific fragment of DNA to be inserted into a vector so that the DNA can then be manipulated. The target DNA to be ligated was excised from the vector it was in by the appropriate restriction enzyme(s) or generated by the polymerase chain reaction (PCR) (section 2.3.1.17). The vector for subcloning was digested with the appropriate enzyme(s). If only one enzyme was used for digestion alkaline
phosphatase (AP) was added to prevent re-ligation of the vector. AP catalyses the hydrolysis of the 5' phosphate group at the 5' phosphorylated termini of the vector. Both vector and target DNA were then phenol/chloroform extracted (detailed in section 2.3.1.5), precipitated and resuspended in water at concentrations of 50ng/μl and 10ng/μl respectively. To optimise the ligation reactions a number of different ratios of vector : insert were set up. In addition to the target DNA and the vector, 1/10 volume of 10x ligation buffer and 1 U T4 ligase were added in a final volume of 20μl. Appropriate control reactions were also set up. Reactions were incubated overnight at 14°C.

2.3.1.14 Transformations
Transformation reactions allow fragments of DNA contained in a vector to be stably inserted into a bacterial cell, the cells can then be propagated which allows the production of large quantities of the specific DNA fragment to be generated. Competent cells were either DH5α (GibcoBRL), subcloning efficiency prepared cells (see section 2.3.1.12.2) or JM109 (Promega) depending on the vector used. 5μl of the ligation reaction was mixed with 50μl of competent cells and incubated on ice for 30 minutes. The reactions were heat shocked at either, 37°C for 20 seconds if DH5α cells were used or at 42°C for 45 seconds if JM109 cells were used, then returned to ice for two minutes. 950μl of SOC broth (Invitrogen, UK) at room temperature was added and allowed to incubate at 37°C for 60 minutes, shaking. 200μl of the mixture was then plated onto selective 2YT agar plates. With vectors where blue/white (white coloration indicating the presence of successfully cloned insert DNA) selection was possible 100μl IPTG (24mg/ml) and 50μl of X-GAL (10mg/ml) were spread onto each plate prior to plating the transformed cells. The plates were incubated at 37°C overnight. Colonies were minipreped as in 2.3.1.1.

2.3.1.15 Transfections of Human embryonic kidney (HEK)-293T Cells
Previously frozen stocks of HEK-293T cells were thawed from liquid nitrogen. After two passages, cells were split onto 90mm petri dishes and grown to 40-50% confluency. 5μg of expression construct per plate was added to 10ml of DMEM (Sigma-D5796). Lipofectamine transfection reagent (Invitrogen) (40μl per plate) was then added and the plates incubated at room temperature for 45 minutes. The media was removed and the cells were washed with 4ml of media per plate. Cells were incubated at 37°C for 24
hours. The media was then changed on the cells by adding 5-10ml of media containing 10% FCS, 1% antibiotics (penicillin/streptomycin) and 1% antimycotics was added and incubated for a further 24 hours at 37°C.

The media was removed and the cells were washed in 1xPBS (Oxoid). The cells were scraped off in to PBS and spun at 2000rpm for 10 minutes. The supernatant was removed and the cell pellet was washed again in PBS, and spun as before. The cell pellet was stored at -80°C.

2.3.1.16 Polymerase Chain Reaction (PCR)
The polymerase chain reaction (PCR) allows the amplification of specific DNA sequences, using two oligonucleotide primers which hybridise to opposite strands and flank the target region of DNA. PCR reactions were performed in a final volume of 50μl, using 1x PCR buffer (16mM (NH₄)₂SO₄, 67mM Tris-Cl (pH 8.3), 0.01% Tween-20), 10pmoles of each oligonucleotide and 50-100ng of genomic DNA or 0.1-1ng cloned DNA as template. The constituents were mixed in a 0.5ml eppendorf tube and 0.2U Taq polymerase (Bioline) was added. The DNA was denatured by incubation at 96°C for 3 minutes, the reaction was then allowed to proceed in a DNA thermal cycler programmed for the required conditions. Standard reactions generally involved 35 cycles of denaturation at 96°C for 30 seconds, annealing at x°C for 30 seconds and extension at 72°C for 1 minute. The appropriate annealing temperature (x°C) was estimated from the Tₘ of the oligonucleotide primer pair, which is calculated as:

\[ Tₘ = (2°C \times \text{no. A + Ts}) + (4°C \times \text{no. G + Cs}) \]

and was 6°C below the Tₘ. Followed by a final extension step at 72°C for 5 minutes.

2.3.1.17 Colony PCR
Colony PCR was used to screen many colonies generated in a transformation reaction (see section 2.3.1.14) for a specific DNA fragment using the selected colony as the template. Colonies picked from transformation plates were first restreaked onto a gridded selective agar plate then the loop used in the streaking process was used to inoculate a 50μl PCR reaction. PCR was performed as described in 2.3.1.16.
2.3.18 Sequencing
Sequencing reactions were performed using the ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready reaction kit (PE Applied Biosystems) with AmpliTaq® DNA polymerase, FS. AmpliTaq® is a variant of Taq DNA polymerase, with point mutation resulting in less discrimination against nucleotides, and the elimination of 5'→3' nuclease activity. The BigDye™ terminator ready reaction mix contains AmpliTaq® and also BigDye™ terminators labelled with novel, high-sensitivity dyes. These contain a fluorescein donor dye linked to a dichlorohodamine acceptor dye. The donor dye has an excitation spectra corresponding to the argon ion laser of the ABI sequencer. However, the 4-dideoxynucleotide terminators are each labelled with a slightly differing acceptor dye, with corresponding different emission spectra, allowing differentiation between nucleotides. The terminators are labelled as follows:

<table>
<thead>
<tr>
<th>Terminator</th>
<th>Acceptor Dye</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Dye</td>
<td>dichloro [R6G]</td>
</tr>
<tr>
<td>C-Dye</td>
<td>dichloro [R110]</td>
</tr>
<tr>
<td>G-Dye</td>
<td>dichloro [ROX]</td>
</tr>
<tr>
<td>T-Dye</td>
<td>dichloro [TAMRA]</td>
</tr>
</tbody>
</table>

The BigDye™ ready reaction mix also contains deoxynucleoside triphosphates (dATP, dCTP, dITP, dUTP), MgCl₂ and Tris-HCl buffer (pH 9.0).

2.3.18.1 Cycle Sequencing
This protocol was used to prepare a successfully cloned fragment for sequencing. The following were added to a 0.2ml microcentrifuge tube:

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABI BigDye terminator ready reaction mix</td>
<td>4µl</td>
</tr>
<tr>
<td>Primer</td>
<td>1.6pmols</td>
</tr>
<tr>
<td>DNA template</td>
<td>0.2-3µg</td>
</tr>
<tr>
<td>dH₂O</td>
<td>to 10µl</td>
</tr>
</tbody>
</table>

The reaction was then mixed, and the following cycling conditions used (using a Perkin Elmer Cetus system 9600 thermal cycler): 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes, with a final holding step of 4°C.
2.3.1.18.2 Direct sequencing
If a single band, with no obvious non-specific bands, was obtained by PCR the band was either excised from the agarose gel and eluted using a Wizard™ column or the PCR was used directly. In this case, the cycle sequencing reaction was set up as described in section 2.3.1.18.1, but typically using 4µl of BigDye, 1.6pmol primer (0.5µl) and 5.5µl of eluted PCR product as template.

2.3.1.18.3 Purification of extension products
Following the labelling reaction, the labelled template was purified prior to sequencing. Firstly, the labelling mix was transferred to a 0.5ml eppendorf tube and 30µl of dH2O and 60µl of 96% ethanol were added. The reaction was vortexed briefly and then incubated at room temperature for 15 minutes. Following this incubation, the sample was spun at 13,000rpm for 20 minutes. The supernatant was removed by aspiration, leaving the pellet of labelled DNA. 200µl of 70% ethanol was then added to wash the pellet and centrifuged at 13,000rpm for 10 minutes. Finally all the ethanol was removed by aspiration. The tubes containing labelled product were then left to dry on the benchtop for 1-2 hours. When dry 3µl of ABI loading dye was added to the sample and vortexed to mix. Immediately prior to loading the sample onto the sequencing gel, the sample was denatured by heating to 95°C for 3 minutes then quenched on ice.

2.3.1.18.4 Preparation of sequencing gels
Automated sequencing was conducted using a 373a DNA sequencer (ABI) using 6% acrylamide gels. The glass gel plates were thoroughly cleaned with tap water then rinsed with deionised water prior to assembly with 1mm spacers. The gel mix consisted of 40ml of Sequagel™ sequencing gel solution with 6ml of Sequagel™ complete buffer reagent. 0.04g of ammonium persulphate (APS) was then added and dissolved to induce gel polymerisation. The gel mixture was then injected between the glass plates using a 50ml syringe, with tapping of the plates to encourage gel movement if slow and to prevent the formation of air bubbles, finally a flat comb was placed at the gel front. The gel was left to polymerise for at least one hour before the plates were washed again first with tap then deionised water, and assembled into the sequencer for a pre-scan.

The rest of the components of the sequencer were assembled according to the manufacturer's instructions, and 1x TBE buffer was added to both top and bottom tanks. Following a pre-run of 15-60 minutes, a 48 or 64 well comb was inserted into the gel and the wells were washed out to remove any traces of urea. 1.5µl of denatured,
labelled sequencing sample (prepared as in 2.3.1.18.3) was loaded per lane. The 373a DNA sequencer was run in accordance with the manufacturer's instructions, typically electrophoresing for 12 hours and yielding roughly 300-700bp read, depending on the sample purity and structure.

2.3.1.18.5 Analysis of sequencing data
Data from the sequencer was stored in raw form as an electropherogram, showing the dye emissions as peaks, and then in a translated form as the actual sequence read by the operating program. Electropherograms were based upon automatic tracking of the sample, but it was often necessary to track each lane by eye then re-extract the data to optimise the sequencing read. When clean unambiguous sequence was produced base-calling was usually accurate allowing analysis of the translated data. However, in some cases, particularly when faced with difficult templates, when reads extended beyond 400 bases (when longer extension products are scarce) or after runs of identical bases, analysis of the electropherogram to confirm the translated sequence was necessary.

2.3.1.19 Library screening
2.3.1.19.1 Phage libraries
All cDNA libraries were plated and screened in the same way, with only the bacterial host strain being different. Serial dilutions of the library were prepared in SM buffer. These were then mixed with 500µl of the prepared host cells (described 2.3.1.12.1) and allowed to adsorb for 30 minutes at 37°C. 10ml aliquots of NZY top agarose were prepared and kept molten at 50°C. An aliquot of top agarose was mixed quickly with the cell/phage mixture and poured onto an NZY agar plate. The top agarose was allowed to harden before the plates were incubated overnight at 37°C. The number of plaques formed were counted and the titre of the library calculated.

A library screen usually consisted of 20 plates each with 20,000 - 25,000 plaque forming units (pfu). The plates were incubated overnight at 37°C and then placed at 4°C for 30-60 minutes to ensure the top agarose did not lift off in later steps. Duplicate replica filters were made from each of the master plates by placing a Hybond-N filter onto the plate, the first filter was left on the plate for 3 minutes and the second for 5 minutes, orientation marks were made with a 21 gauge needle. The DNA was released by the method of Benton and Davis (1977). The filter was carefully peeled from the
plate and placed DNA side up into a tray containing 3MM filter paper soaked in
denaturing solution. After 5 minutes the filter was placed on 3MM soaked in
neutralising solution, after a further 5 minutes the filter was placed in another tray
containing neutralising solution for 3 minutes. The filter was rinsed in 3xSSC and then
allowed to air dry. The whole process was then repeated for the second duplicate filter.
The same orientation marks were used. The filters were baked for 2 hours at 80°C to
bind the DNA to the filters. The filters were prehybridised and hybridised as described
in section 2.3.1.21.

2.3.1.19.2 Gridded libraries
Gridded libraries obtained from HGMP-RC were screened by pre-wetting the filters in
3xSSC before being covered with hybridisation solution in a hybridisation bottle and
hybridised as described in 2.3.1.21.

2.3.1.20 Preparation of radioactive probes
DNA probes were prepared by excising the appropriate band from a low melting point
(LMP) agarose gel with a scalpel blade on an UV transilluminator. The LMP agarose
gel slice was processed in one of two ways. The gel slice was either placed in a
Wizard™ column (Promega) in an eppendorf tube and spun for 3 minutes at 5,000g or
the band was weighed and 3ml water per gram of agarose added. The solution was
heated to 65°C to melt the agarose, vortexed then stored at –20°C until needed.

The DNA extracted from the agarose gel was first denatured at 95°C for 7 minutes, then
incubated at 37°C for 5 minutes. The DNA was randomly radiolabelled using the
RadPrime DNA labelling system (Invitrogen). 25-100ng of DNA was added to 20μl of
2.5x random primer solution, 1μl of each of 500μM dATP, dGTP, dTTP, 5μl α-32P
dCTP and make up to a final volume of 49μl with distilled sterile water. This mixture
was centrifuged briefly then 1μl Klenow was added and the reaction was incubated at
37°C for 10 minutes. Finally 5μl of Stop buffer was added to terminate the labelling
reaction. The reaction was then loaded onto a ProbeQuant™ G-50 Micro column
(Amersham Pharmacia Biotech) and spun for 2 minutes at 3,000g to remove any
unincorporated nucleotides. The labelled probe was denatured for 5 minutes at 95°C
and then quenched on ice for 2 minutes before being added to hybridisation buffer.
2.3.1.21 Hybridisation of Filters
Filters were pre-hybridised in hybridisation solution in petri dishes or hybridisation bottles with shaking or rotating for 2 hours or overnight. Both pre-hybridisation and hybridisation were carried out at 65°C unless otherwise stated. The probe (prepared as in 2.3.1.20) was added to fresh hybridisation solution. The pre-hybridisation solution was removed from the filters and the hybridisation solution containing the probe was added. Hybridisation was performed overnight with shaking or in a rotating oven.

Following hybridisation, the filters were washed at varying stringencies depending on the probe used. Generally filters were first washed in a solution of 6xSSC, 0.1% (w/v) SDS for 30 minutes at 65°C, then subsequently in decreasing salt concentrations, but not less than 0.1xSSC 0.1% (w/v) SDS, until monitoring with a hand-held Geiger counter (Mini Instruments Ltd, UK) indicated an appropriate level of radioactivity (usually 5-10 cps).

Washed filters were mounted onto intensifying screens under clingfilm and exposed to X-ray film in autoradiographic cassettes, with a second screen, usually overnight at -80°C. The exposed film was developed in an automated developing machine using Fuji reagents.

2.3.2 RNA
All reagents in this section are prepared with 0.1% diethylpyrocarbonate (DEPC) solutions. 0.1% DEPC is added to distilled water and all autoclavable solutions, the solutions are then incubated overnight in a fumehood to allow the DEPC to deactivate any contaminating RNAases. The solutions are autoclaved to remove the DEPC. A glass and plasticware used in this section was washed with DEPC-treated water before and after use.

2.3.2.1 RNA extraction
RNA was extracted from tissue using TRIzol™ reagent (Invitrogen) according to the manufacturer’s instructions. 1ml of TRIzol was added per 50mg of tissue and then homogenized using a small Ultra-turrax probe (Janke & Kunkel-IKA-Labortechnik) for 30 seconds. The homogenate was incubated at room temperature for 5 minutes. 0.2ml chloroform was then added per 1ml of TRIzol before being mixed and centrifuged at
12,000g for 15 minutes at 4°C. The upper aqueous phase was transferred to a clean eppendorf. The RNA was precipitated from the aqueous layer by the addition of 0.5ml of isopropanol. The reaction was incubated at room temperature for 10 minutes before being spun at 12,000g for a further 10 minutes. The supernatant was removed and the RNA pellet was washed once with 75% ethanol. The pellet was allowed to air dry and was then dissolved in an appropriate volume of DEPC-treated water before stored at -80°C until required.

2.3.2.2 mRNA preparation
mRNA was prepared from total RNA using Oligotex Spin Columns (Qiagen). 1mg of total RNA was combined with the appropriate volume of water, 2x binding buffer and oligotex according to the manufacturers instructions depending on the concentration of RNA used. This was incubated for 3 minutes at 65°C, then at room temperature for 10 minutes. The column was centrifuged for 2 minutes and the supernatant removed. The pellet was resuspended in 400μl of wash buffer OW2 and pipetted on to a spin column. The column was spun at 13,000rpm for 30 seconds. 400μl of OW2 wash buffer was applied to the column and then the column was centrifuged for a further 30 seconds at 13,000rpm. The RNA was eluted from the spin column by the addition of 50μl of elution buffer preheated to 70°C, the column was spun for 30 seconds at 13,000rpm and the eluted RNA stored at -80°C until required.

First strand cDNA synthesis allows the generation of cDNA from a purified mRNA population. First strand cDNA synthesis was performed using the You-Prime First-Strand Beads Ready To Go kit (Amersham Pharmacia Biotech, UK). For this procedure oligo-dT primers were used. Oligo-dT primers bind to the polyA tail of mRNA, and this allows the generation of cDNA from all the mRNA population in a sample. Gene specific primers complementary to the mRNA sequence or random primers could have also been used. The You-Prime First-Strand Beads Ready To Go kit provides beads containing all the reagents for the production of first strand cDNA, except primers. The beads contain buffer, dNTPs, murine reverse transcriptase, RNA guard and RNase/DNase-free BSA. 0.02μg-2μg of mRNA (see above) or 0.1μg-5μg of total RNA (as prepared in section 2.3.2.1) was used in the reaction. The RNA sample was made up to a volume of 29μl with DEPC treated water. The sample was then heated at 65°C for 10 minutes, then chilled on ice for 2 minutes. This solution was transferred to the
reaction tube containing the beads. 1μl of 100nM of an oligo-dT primer was added. The reaction was incubated at room temperature for 1 minute, and then mixed by vortexing. The reaction was then incubated at 37°C for 60 minutes. The product was stored at -20°C until required.

2.3.2.3 RT-PCR
First strand cDNA was synthesised as detailed in section 2.3.2.2. The cDNAs were used for PCR with gene specific primers as detailed in section 2.3.1.17.

2.3.2.4 Northern Blotting
Northern blotting allows the transfer of RNA from a gel onto a nylon membrane. 10-15μg of total RNA (1 volume) was combined with 2 volumes of RNA sample buffer and heated to 65°C for 15 minutes then quenched on ice before 1/10 volume of RNA loading buffer was added. The samples were electrophoresed in a DEPC-treated environment, this was to avoid the degradation of the RNA samples by RNases. The gel was blotted overnight on to a Hybond-N filter. After blotting the filter was soaked in 2xSSC for 5-10 minutes then air dried on 3MM paper before baking for 1 hour at 80°C. It was then placed in a UV Stratalinker (Amersham Pharmacia Biotech) for 7 seconds which crosslinks the RNA to the filter. The filter was washed in prehybridisation buffer for 2 hours at 42°C. DNA probes were prepared as detailed in section 2.3.1.20. Freshly prepared probes were added to hybridisation buffer and applied to the filter overnight at 42°C. The radioactive hybridisation buffer was removed from the filter. The filter was washed twice in 1xSSC, 0.1% SDS for 30 minutes at 42°C then twice in 0.05xSSC, 0.1% SDS for 30 minutes at 42°C before being exposed to X-ray film. All hybridisation and washing was performed in a rotisserie oven.

2.3.2.5 Riboprobe production
The production of riboprobes allows the detection of specific mRNA in tissues. The DNA to be transcribed is cloned into the polylinker of a vector containing promoters for SP6 (9or T3) and T7 RNA polymerases adjacent to the polylinker. After linearisation of the template DNA at a suitable site, the RNA polymerases are used to produce "run-off" transcripts. DIG-UTP is used as a substrate and incorporated into the transcript. Every 20-25th nucleotide of the newly synthesised RNA is a DIG-UTP. Since the nucleotide concentration does not become limiting in the standard transcription assay, a large
amount of labelled RNA can be generated. Digoxigenin (DIG)-labelled probes were synthesised using 1μg of purified linearised DNA, DIG-labelling mix, 10x Transcription buffer, RNase inhibitor, RNA polymerase in a final volume of 20μl. This was incubated at 37°C for 2 hours. After incubation 1U of RNase-free DNase I was added and incubated for a further 15 minutes at 37°C to degrade the DNA template. 70μl of DEPC (0.5% v/v) H₂O, 10μl 3M NaOAc and 300μl absolute ethanol were added and stored at -20°C overnight to precipitate the RNA. The RNA was pelleted and then resuspended in 100μl DEPC (0.5% v/v) H₂O.

The digoxigenin (DIG) system is a synthetic molecule based on the steroid hapten digoxigenin. Digoxigenin is the so-called aglucon of the steroid digoxin, which occurs in certain digitalis plants. Aglucon means that the three sugar moieties at the C3 OH group which are part of the digoxin molecule are lacking. There are two features of the digoxigenin molecule which make it exceptionally valuable for establishing detection systems. First, if digoxigenin is coupled as a hapten to a suitable carrier molecule high affinity antibodies can be easily generated, e.g. in sheep. Since digoxin occurs exclusively in digitalis plants there are no endogenous background problems with these antibodies as in the case of other haptens, such as biotin. Second, digoxigenin can be coupled to nucleotides like dUTP or UTP and incorporated into nucleic acids using generally available polymerases like Klenow polymerase, Taq polymerase, or RNA polymerases. The probes thus generated can be used in standard blotting and hybridization procedures including in situ hybridization and detected with anti-digoxigenin conjugates. The most frequently used conjugates are fluorescent or alkaline phosphatase labelled antibodies. If phosphatase labelled antibodies are employed hybridized probes can be detected either by chemiluminescence or a colour reaction. DIG labelled probes that hybridised to a target sequence are detected with an alkaline phosphatase labelled anti-DIG antibody. If the blot is incubated with suitable reagents, such as NBT and BCIP, phosphatase activity is detected by a colour reaction.

In order to ensure the riboprobes were labelled a dot-blot was performed. This technique involves the immunodetection of the DIG-labelled riboprobes with anti-digoxigenin Fab fragments conjugated to alkaline phosphatase. The bound antibody conjugate is visualised by colour detection using substrates NBT and BCIP. 1μl of
riboprobe was dotted onto a Hybond-N+ charged membrane and UV cross-linked for 12 seconds. The filter was washed for 30 minutes in DIG2 buffer. 1μl of anti-DIG AP was diluted in 5ml of DIG2 buffer. The filter was washed in this solution for 45 minutes, subsequently washed twice in DIG1 buffer for 5 minutes, then for 5 minutes in DIG3 buffer. Finally the filter was washed for one hour in a "light-tight" box in DIG3/BCIP/NBT (10mls/34μl/44μl).

2.3.3 In situ Hybridisation Studies

2.3.3.1 In situ Hybridisation

In situ hybridisation is used to detect the location of specific gene transcripts in tissue sections. 5μm thick formalin fixed whole mouse eye sections were adhered to APS (aminopropyltriethoxysilane) (Sigma) treated slides. The sections were dewaxed in fresh xylene for 10-15 minutes then taken through a graded series of ethanol from 100-30%, two minutes in each, before being washed in DEPC-treated PBS. The slides were then immersed in 0.2M HCl for 15 minutes. They were then rinsed in DEPC-treated PBS before being washed in 0.3% (w/v) Triton X-100 for 15 minutes on a shaker at room temperature. The slides were washed again in PBS for 10 minutes before being incubated with 100μg/ml proteinase kinase (Proteinase K). The sections were left covered in proteinase K solution for 10-30 minutes in a moist chamber at 37°C. The slides were rinsed again in PBS (DEPC) then post-fixed in 2% (w/v) paraformaldehyde solution for 5 minutes at room temperature. The slides were washed again in PBS for 10 minutes and were then covered in pre-hybridisation solution and incubated for 1-2 hours at 37°C. The riboprobe (see section 2.3.2.5) was adjusted to the required dilution (1:50-1:100) using hybridisation buffer. 25-50μl of probe was dotted onto a cover slip of an appropriate size. The excess pre-hybridisation solution was removed and the sections were mounted using the coverslip dotted with probe. The slides were heated to 70°C for 2 minutes, the temperature was lowered to 55°C and the slides were hybridised overnight in the moist chamber. The moist chamber prevents the sections from drying out by maintaining a constant temperature in a humid environment.

The cover slips were removed by washing the slides twice in 2xSSC for 20 minutes each. The solution was changed to 0.1xSSC at room temperature for 10 minutes, then 0.1xSSC at 50°C for 30 minutes then 0.1xSSC for 10 minutes at room temperature. The
slides were rinsed in DIG1 buffer for 5 minutes at room temperature. Anti-DIG/AP in 1:2000 DIG1 buffer (containing 10% normal swine serum) was applied for at least 2 hours at room temperature in a moist chamber. The excess antibody was washed off in DIG1 buffer twice for 15 minutes each. The slides were then rinsed in DIG3 buffer for 5 minutes. The slides were finally immersed in NBT/BCIP/levamisole for between 30 minutes to 3 hours or until a colour change was observed, this was checked microscopically every 30 minutes. When the reaction was complete, the slides were washed in DEPC-treated dH2O then mounted.

2.3.3.2 Fluorescent in situ Hybridisation (FISH)
All FISH mapping in this project was performed by Ruby Banerjee at the MRC Cytogenetics Unit, Harwell, UK. The protocol used is briefly outlined below.

2.3.3.2.1 Metaphase spreads
Mouse metaphase chromosome spreads were made from spleen cultures. Intact spleens were dissected and collected into cold RPMI 1640 medium, cells were grown in complete RPMI 1640 medium supplemented with glutamine, penicillin and streptomycin, foetal calf serum, concanavalin A and lipopolysaccharide. The cells were allowed to incubate in a 5% CO2 incubator for 48 hours and then harvested the following day. Before harvesting, cells were incubated for a further 30 minutes in colcemid solution (final concentration of 0.05μg/ml). Chromosome spreads were made using standard procedures, treatment in hypotonic (KCl) and subsequent fixation (methanol:glacial acetic acid - 3:1) followed by air drying. Slides were examined using phase contrast microscopy and stored at 4°C with a desiccant. Prior to application of the probes, slides were treated with RNase and then dehydrated through an alcohol series (70%, 90%, 100%) and the chromosomes were denatured in 70% formamide in 2xSSC for 2 minutes at 72°C, then passed through the ethanol series at 4°C a second time and air dried.

2.3.3.2.2 Labelling and Hybridisation
Probes were labelled by nick-translation with Biotin-16dUTP (Boehringer Mannheim) using the Bio Nick Kit (Invitrogen) at 16°C for 90 minutes. 400ng of the product were ethanol precipitated along with Cot-1 DNA and salmon sperm DNA and the pellet was resuspended in hybridisation mix containing 50% formamide and 10% dextran sulphate. It was then denatured at 75°C for 8 minutes followed by pre-annealing at 37°C for 30
minutes. 12μl of the hybridisation mix was applied to each denatured slide. Slides were incubated overnight in a humid chamber at 37°C.

2.3.3.2 Signal detection
Post hybridisation washes were performed in three changes of 50% formamide, 2xSSC and 4xSSC, 0.05% Tween 20 at 42°C followed by another wash in 4xSSC, 0.05% Tween 20 at room temperature. Slides were incubated in a Blocking Reagent (Roche) for 30 minutes. For signal detection, slides were incubated at 37°C with Avidin FITC (Vector Labs) for 25 minutes then washed with 4xSSC Tween 20 three times. Incubation at 37°C was repeated with Biotinylated Anti Avidin and Avidin FITC again with washes in 4xSSC Tween 20 in between. In this case PAC 439 has been detected with Avidin FITC and PAC 572 with Avidin Texas Red. Finally, the slides were embedded in Vectashield mounting medium (Vector Labs) containing 1μg/ml DAPI.

2.3.3.2.4 Microscopy
FISH images were generated with an epifluorescence microscope equipped with a cooled CCD camera (Photometrics, Sensys) and computer-controlled filter wheels with excitation and emission filters for visualization in DAPI, FITC and Texas Red.

2.3.4 Protein
2.3.4.1 Tissue Homogenisation
To generate protein samples from *Drosophila* heads, the flies were decapitated and a 20% homogenate made in homogenisation buffer. Samples were homogenised and the supernatants collected. Homogenates were diluted 1:1 with sample buffer and denatured under reducing conditions by boiling at 100°C for 5 minutes in a dry block. Volumes of 100μl were stored at -20°C until required.

2.3.4.2 Cell Membrane Preparation
The cell pellet (see section 2.3.1.15) was resuspended in cold homogenisation buffer (1ml per dish). The cells were lysed on ice for 30 minutes and by 5 passes through a 0.5mm needle, spun down at 2000rpm for 5 minutes at 4°C. The supernatant was transferred to a new tube, 50μl 5M NaCl was added and the tube inverted to mix. This was spun at 13000rpm for 30 minutes at 4°C. The supernatant was removed and the membrane pellet was resuspended in 50μl of PBS/dish of cells. 1 volume of 4x Sample
Buffer was added and proteins were denatured by heating to 100°C for 10 minutes before being loaded onto an SDS-PAGE gel or stored at -20°C until needed.

2.3.4.3 SDS-PAGE gels
Gels were cast using equipment purchased from BIO-RAD Laboratories, Inc. The cast dimensions were 8 cm x 5 cm, with a thickness of 0.75 cm. The 12% resolving gel was overlaid with a 6% stacking gel layer into which the samples were loaded.

Each well was loaded with up to 20 µl of denatured sample. To determine the molecular weights of sample proteins, pre-stained molecular weight markers, Rainbow™ coloured protein markers (molecular weight range 10-250 kDa) (Amersham Pharmacia, Biotech), were diluted 1:1 with sample buffer and were also loaded onto the gel. Proteins were separated by electrophoresis. Samples were subjected to SDS-PAGE for 1.5 hours at 100 V.

Protein bands separated in the gels were visualised with Coomassie blue stain. Excess stain was removed using de-stain 1 and de-stain 2. Gels were placed in anti-crack solution and dried in a Gel Air Dryer (BIO-RAD Laboratories).

2.3.4.4 Western Blotting
Western blotting is the transfer of proteins separated by SDS-PAGE (as described in 2.3.4.2-3) onto nitrocellulose membranes. The gel was placed on a nitrocellulose membrane and the gel and membrane were sandwiched between a stack of Whatman 3MM paper pre-soaked in transfer buffer. The entire sandwich was secured in a plastic frame and placed in the transfer apparatus (Trans-blot® SD Semi Dry Transfer cell, BIO-RAD Laboratories) according to the manufacturer's instructions. Proteins were transferred from the gel onto the nitrocellulose membrane by electrophoresis at 20 volts for 1 hour in transfer buffer. The gel was stained, as described in 2.3.4.3, to determine the amount of protein left on the gel after blotting. The membrane was then blocked at room temperature for 30 minutes shaking then 30 minutes not shaking in blocking buffer. Primary antibody was added to the membrane, diluted in blocking solution and incubated overnight at 4°C with shaking.
The membrane was washed 5 times in Tris buffered saline (TBS) and incubated for 1 hour at room temperature with secondary antibody. Unbound secondary antibodies were removed by washing in TBS. Antibody binding was detected by nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) (alkaline phosphatase) as chromogens. The membrane was washed in TBS, equilibrated in detection buffer and exposed to substrate solution. The colour reaction was stopped by washing the membrane in quenching buffer. Blots were air-dried and the molecular weights of the protein bands were determined by comparison with Rainbow™ molecular weight markers.

2.3.4.5 Immunostaining of transfected cells
293T cells were plated on coverslips a day prior to transfection (see section 2.3.1.15). 5µg of the DNA construct of interest was mixed with 40µl of Lipofectamine Transfection reagent (Gibco, UK) and incubated for 45 minutes at room temperature. Cells were washed in DMEM and 1ml of the DNA/transfection reagent was added per coverslip and incubated at 37°C overnight. Cells were fed the following day with DMEM containing 10% foetal calf serum and 0.1% antibiotic/antimycotics. Cells were prepared for immunostaining by washing twice in PBS prior to being fixed in 4% (v/v) formaldehyde in PBS for 10 minutes at room temperature. Cells were washed in PBS and permeabilised in 0.2% (v/v) Triton X-100 in PBS for 10 minutes at room temperature. Cells were subsequently washed in PBS and blocked with 5% (v/v) foetal calf serum in PBS for 30 minutes at room temperature. Cells were washed in PBS and incubated with a 1:500 dilution of anti-myc antibody (Santa Cruz Biotechnology, Inc) in PBS. Cells were washed six times in PBS prior to being incubated with the secondary antibody. An anti-mouse FITC conjugated secondary antibody was used to visualise primary antibody binding. Anti-mouse FITC (Pierce, UK) was used at a 1:50 dilution and incubated at room temperature for 30 minutes. Cells were washed 6 times in PBS before being mounted on glass slides. Immunolabelling was visualised using a fluorescence microscope.
CHAPTER THREE

RESULTS

The work presented in this thesis describes the isolation and characterisation of the murine genes CDP-diacylglycerol synthase 1 (Cds1) and 2 (Cds2). These genes are the murine homologues of the Drosophila eye-cds gene. The Drosophila eye-cds gene causes retinal degeneration when mutated (this is described in detail in section 1.2). The mammalian homologues of eye-cds were identified from mammalian retinal samples, thus the Cds genes are good candidate genes for mammalian retinal disease phenotypes. The data gathered forms two main sections. The first section includes the identification, characterisation and expression of the two murine Cds genes and the second section uses the data generated in the first section to define possible gene functions through transgenic studies.

3.1 Isolation of murine Cds1

3.1.1 Identification of cDNA clone 801-h9

Two human homologues of the Drosophila eye-cds gene, CDS1 and CDS2 have previously been described (Heacock et al., 1996; Weeks et al., 1997; Halford et al., 1998; Volta et al., 1999). The human CDS1 nucleotide sequence (GenBank accession number: XM_034414 (Weeks et al., 1997) was used to search the Expressed Sequence Tag (EST) database using the BLAST-N algorithm (http://www.ncbi.nlm.nih.gov/BLAST) to identify homologous murine cDNA clones. This resulted in the identification of one murine cDNA clone (GenBank accession number: w30593, I.M.A.G.E number: 349880). The clone was obtained from the Human Genome Mapping Project Resource Centre (HGMP-RC) Hinxton, UK (http://www.hgmp.mrc.ac.uk) and was given the laboratory number 801-h9. The clone originated from a Soares mouse p3NMF19.5 library and had been cloned into the NotI and EcoRI restriction sites of a pT7T3D vector. DNA was prepared from the clone (see section 2.3.1.1-2) and was digested with the enzymes NotI and EcoRI to release the 1.3kb insert and the 3kb pT7T3 vector (figure 3.1) (see section 2.3.1.9). The ends of the clone were sequenced
Figure 3.1 Insert from Cds1 cDNA clone 801-h9
801-h9 was digested with the restriction enzymes EcoRI and NotI. This generated two fragments, the insert of ~1.3kb and the 3kb vector.
using vector primers M13F and M13R (see section 2.3.1.18). This confirmed that the clone was the expected \textit{Cds1} homologue and the entire insert was then sequenced on both strands using gene specific primers (see table 3.1 for primer sequences).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence (5'\textendash3')</th>
<th>Sequence Start Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>801R2</td>
<td>AACCTGGCCAGTGAGGG</td>
<td>1097</td>
</tr>
<tr>
<td>801F2</td>
<td>GGTAGAGGGCGAAGGATATC</td>
<td>19</td>
</tr>
<tr>
<td>801R3</td>
<td>GTCTACCTAAATCATGG</td>
<td>1114</td>
</tr>
<tr>
<td>801F3</td>
<td>CTTGAAAGGTGTAGCTG</td>
<td>817</td>
</tr>
<tr>
<td>801R4</td>
<td>CATTATCCCGCAGTGCGAGG</td>
<td>636</td>
</tr>
<tr>
<td>801R5</td>
<td>ATGTGACTGAGTGACGGTTA</td>
<td>132</td>
</tr>
<tr>
<td>801F5</td>
<td>AATGCACTGGAGGCACCTTG</td>
<td>1203</td>
</tr>
</tbody>
</table>

Table 3.1 Primers used to generate the 3' sequence of \textit{Cds1} using cDNA clone 801-h9.

The nucleotide sequence of 801-h9 and the derived amino acid sequence of the longest open reading frame are shown in figure 3.2 (the primer sequences listed in table 3.1 are highlighted in red). Aligning the derived amino acid sequence of 801-h9 and that of human \textit{CDS1} (figure 3.3), using the standard alignment algorithm in GeneWorks 2.5.1N (supplied by IntelliGenetics), it can be seen that the 5' end of the murine \textit{Cds1} gene is not represented in 801-h9. The human \textit{CDS1} and mouse 801-h9 nucleotide sequences are 85% identical and the amino acid sequences are 96.7% identical.

GeneWorks was used to predict the restriction map of 801-h9 (figure 3.4) this was then verified by restriction digests. The map was prepared in order to identify restriction sites that could later be used to generate probes.
**Figure 3.2 Nucleotide and derived amino acid sequence of 801-h9**

The 1308bp nucleic acid sequence generated from cDNA clone 801-h9. The stop codon is highlighted in blue at position 828bp. The oligonucleotide primers used to generate the sequence are listed in table 3.1 and highlighted above in red.
Figure 3.3 Amino acid alignment of human CDS1 and 801-h9
This figure shows the amino acid sequence alignment of the mouse clone 801-h9 and the human CDS1 sequence. The regions of complete identity are shown boxed and shaded and regions of similarity are just shaded. 801-h9 aligns with the 3' end of the human sequence. This alignment shows that 801-h9 does not contain the 5' end of the mouse gene.
Figure 3.4 Schematic restriction map diagram of 801-h9

A simple restriction map of clone 801-h9 is shown with vector sites shown in black text and the restriction sites contained within the clone shown in red. The nucleotide sequence generated from 801-h9 is shown beneath the map with the actual restriction sites highlighted in red. Pstl- CTGCAG, BamHI- GGATCC.
3.1.2 Identification of cDNA clone 2279-d19

To identify a clone containing the 5' sequence of *Cds1*, the nucleotide sequence of 801-h9 was used to re-search the EST database. This resulted in the identification of a further clone (GenBank accession number: AI596413, I.M.A.G.E number: 917346) which was obtained from HGPM-RC and given the laboratory number 2279-d19. The clone originates from a Stratagene mouse adult testis library and is cloned into the *EcoRI* and *XhoI* sites of a pBluescriptSK- vector. DNA was prepared from the clone, as before. The DNA was digested with *EcoRI* and *XhoI* to release the insert. This produced three fragments of 3kb, 1.8kb and 950bp. The DNA was digested with *XhoI* only which linearised the clone and with *EcoRI* only which produced two fragments, 4.8kb and 950bp. This suggests that a co-ligation may have occurred or that the clone contains an internal *EcoRI* restriction site. Clone 2279-d19 was sequenced initially with vector primers M13F and M13R. The sequence from the 3' end of 2279-d19 was aligned with clone 801-h9 (figure 3.5). This shows that the 3' end of 2279-d19 is 100% identical to clone 801-h9, however the 5' end of clone 2279-d19 extends greatly in the 5' direction from 801-h9. The remainder of the insert from clone 2279-d19 was then sequenced with specific primers (table 3.2). The nucleotide sequence generated from clone 2279-d19 and the derived amino acid sequence is shown in figure 3.6 with the internal *EcoRI* site highlighted in red, and the sequencing primers highlighted in blue. The sequence showed that the 5'→3' adapter sequence (GGCACGAGAAACCCT) was present after the *EcoRI* site, making the presence of a co-ligation the most likely explanation for the location of the *EcoRI* site. The sequence preceeding the sequence representing *Cds1* in clone 2279-d19, appears to be running in the opposite direction (3'→5'), and contains a poly (A) tail at the start of the clone. The nucleotide sequence generated from clone 2279-d19 was aligned, as before, with the nucleotide sequences of human *CDS1* (figure 3.7) this alignment showed that the 5' end of 2279-d19 did not match the human sequence. The first 950bp of clone 2279-d19 was used to screen the EST database to determine what the co-ligation sequence represented. The 5' end of clone 2279-d19 did not align strongly with anything in the database. The first 950bp of sequence from clone 2279-d19 were removed and the longest open reading frame of the remaining 1.8kb of sequence was derived. The derived amino acid sequence from 2279-d19 was aligned with the human CDS1 amino acid sequence (figure 3.8). The alignment shows that the amino acid sequence derived from clone 2279-d19 does not align with the 5' end of the human sequence. This alignment shows
Figure 3.5 Amino acid alignment of 801-h9 and 2279-d19
The derived amino acid sequence from 2279-d19 was aligned with the derived amino acid sequence of 801-h9. The amino acid sequence of 2279-d19 was derived from the second reading frame as this gave the best alignment with 801-h9. Regions of identity are boxed and shaded, the alignment shows that the sequences are 100% identical at the 3' end.
Figure 3.6 Clone 2279 nucleotide sequence

The sequence of the clone 2279-dl9 which contains most of the mouse Cds1 gene is shown. The co-ligation is highlighted in red with the 5'-*3' adapter sequence highlighted in purple and sequencing primers (table 3.2) highlighted in blue.
Figure 3.7 2279-d19 and human CDS1 nucleotide alignment
As can be seen by the grey, boxed areas, the central sequence from clone 2279-d19 aligns with the nucleotide sequence of human CDS1. The 5' of clone 2279-d19 does not contain sequence homologous to the CDS1 gene. The 3' ends of both sequences represent the 3'UTR and are not expected to be highly homologous.
Figure 3.8 Human CDS1 and 2279-d19 amino acid alignment

The figure shows the alignment between the amino acid sequence of human CDS1 and the amino acid sequence derived from clone 2279-d19. The sequences are 96.6% identical over the aligned regions and 88.5% identical over the full length. Identical residues are boxed and shaded and similar residues are just shaded. The clone 2279-d19 is lacking the first nineteen amino acids if the start sites are maintained by both the mouse and human, this results was verified (see figure 3.10) by the rat sequence since the rat sequence is 19 amino acids longer than the mouse sequence derived from clone 2279-d19.
that 2279-d19 does not contain the translation start site and hence does not represent the entire Cds1 gene.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5'...</th>
<th>Sequence Start Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>801R5</td>
<td>AATGCACCTGGAGGCACCTTG</td>
<td>2647</td>
</tr>
<tr>
<td>2279F1</td>
<td>GTGAGTCCTCTAGCATGAGA</td>
<td>210</td>
</tr>
<tr>
<td>2279R1</td>
<td>AGGCTCAAGGCAAGTAGTG</td>
<td>571</td>
</tr>
<tr>
<td>CDS1exon1R</td>
<td>GTGCTCTCGTTTTCGTGGTC</td>
<td>996</td>
</tr>
<tr>
<td>MCDS1F4</td>
<td>CTTAATGTGAGTCTCCTAGC</td>
<td>205</td>
</tr>
<tr>
<td>MCDS1R4</td>
<td>AATTCCGACGAGAAACCT</td>
<td>935</td>
</tr>
<tr>
<td>CDS1intron1F</td>
<td>GACCACGAAACCGAGAGCAC</td>
<td>977</td>
</tr>
<tr>
<td>CDS1intron1R</td>
<td>TGGCATCGAGATCTCCATAC</td>
<td>1046</td>
</tr>
<tr>
<td>MCDS1R3</td>
<td>GAATCCACCAGTCTCCATAC</td>
<td>1154</td>
</tr>
<tr>
<td>CDS1intron7F</td>
<td>TTGGGAAGGATTCATTGGTG</td>
<td>1717</td>
</tr>
<tr>
<td>CDS1intron9F</td>
<td>TTGCCCTTTCAACGTTGCG</td>
<td>1959</td>
</tr>
<tr>
<td>MCDS1R1</td>
<td>CAAACATGCAGAAACCCTG</td>
<td>1484</td>
</tr>
<tr>
<td>MCDS1R2</td>
<td>GTGCGTCCATGCGAACATAT</td>
<td>1543</td>
</tr>
</tbody>
</table>

**Table 3.2** Primer names, sequences and the start position in the sequence generated from cDNA clone 2279-d19.

### 3.1.3 Identification of the 5' end of mouse Cds1

The Cds1 sequence generated from murine clone 2279-d19 was used to re-screen the EST database for a cDNA clone containing the 5' end of the mouse Cds1 gene. The search did not identify a mouse clone but did identify a rat cDNA Cds clone (GenBank accession number: NM_031242) with a full-length coding region, when compared to the human sequence, this provided evidence that the mouse gene was not complete. The amino acid sequences derived from clone 2279-d19 and rat Cds1 were aligned (figure 3.9) and are 99% identical. As a mouse clone containing the 5' end of the Cds1 gene had not been identified, the 5' end of the gene was amplified using polymerase chain reaction (PCR) (see section 2.3.1.16). Adult mouse eye and brain cDNA (see section 2.3.2.3) was used as template.
3.9 Amino acid alignment of rat Cds1 and clone 2279-d19
The figure shows the alignment between the amino acid sequence of rat Cds1 and the amino acid sequence derived from clone 2279-d19. The sequences are 98.8% identical over the aligning regions and 90% identical over the full-length. A primer was designed to the start site of rat and another was designed to the mouse nucleotide sequence. The primer pair was used to amplify the 5' end of the mouse gene from mouse cDNA.
In order to amplify the 5' end of the murine *Cds1* gene, the rat nucleotide sequence was used to design an oligonucleotide primer that contained the start site of the rat gene (rCDS1F). This was used with a second mouse specific primer, (CDS1intron2R) (for sequences see table 3.3). The PCR reaction resulted in the amplification of a 250bp fragment. The 250bp PCR fragment was directly sequenced (see section 2.3.1.18) using the primers in table 3.4. The nucleotide and derived amino acid sequences that were generated from the PCR fragment are shown in figure 3.10. The sequence from the PCR fragment was aligned with the sequence generated from clone 2279-d19 to generate the full-length amino acid sequence for mouse Cds1 (figure 3.11 summarises the generation of the coding region of mouse *Cds1*). The full-length human, murine and rat Cds1 amino acid sequences were aligned and are shown in figure 3.12.

### Table 3.3

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5'</th>
<th>Primer Start Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>RatCDS1F</td>
<td>TGACATGCTGGAGCTGCGGC</td>
<td>1(Rat)</td>
</tr>
<tr>
<td>CDS1intron2R</td>
<td>TCAGGAAGAAGAGGGAGATC</td>
<td>335</td>
</tr>
</tbody>
</table>

Table 3.3 Primer names, sequences and position on the sequence derived from cDNA clone 2279 (where relevant) that were used to amplify the 5' end of mouse *Cds1* from eye and brain cDNA.

### Table 3.4

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5'</th>
<th>Primer Start Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>RatCDS1F</td>
<td>TGACATGCTGGAGCTGCGGC</td>
<td>1(Rat)</td>
</tr>
<tr>
<td>CDS1intron2R</td>
<td>TCAGGAAGAAGAGGGAGATC</td>
<td>335</td>
</tr>
<tr>
<td>CDS1intron1R</td>
<td>TGGCATCGAGATCTCCATAC</td>
<td>186</td>
</tr>
</tbody>
</table>

Table 3.4 Primer names, sequences and position on the sequence derived from cDNA clone 2279 that were used to sequence the 5' end of mouse *Cds1* that was amplified from eye and brain cDNA.
Figure 3.10 Nucleotide sequence of 5' Cds1 fragment
The nucleotide and derived amino acid sequence generated from the fragment shown in figure 3.1.9. The start site is highlighted in red.
Figure 3.11 Summary of Cds1 sequence generation
Shown is a schematic diagram of the generation of the mouse Cds1 sequence. The central line (red) represents the whole coding region and the 3'UTR of the gene. Shown are clones 801-h9 (dark blue), 2279-d19 (purple) (dotted line represents the rest of the clone not representing Cds1), and the PCR product (green) from which the coding sequence was generated. The three murine ESTs which were identified from the BLAST database (pale blue) using the 3'UTR sequence from rat are shown beneath the central line. The start location in the full-length sequence is marked at the 5' end of each clone in bp.
<table>
<thead>
<tr>
<th>hCDS1</th>
<th>ratCDS1</th>
<th>mCDS1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLELRHRGCPGAP</td>
<td>MLELRHRGCPGAP</td>
<td>MLELRHRGCPGAP</td>
</tr>
<tr>
<td>7E</td>
<td>/ S</td>
<td>/ S</td>
</tr>
<tr>
<td>ratCDS1</td>
<td>MLELRHRGCPGAP</td>
<td>MLELRHRGCPGAP</td>
</tr>
<tr>
<td>J</td>
<td>CPGP</td>
<td>CPGP</td>
</tr>
<tr>
<td>J</td>
<td>T</td>
<td>G</td>
</tr>
<tr>
<td>J</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>mCDS1</td>
<td>MLELRHRGCPGAP</td>
<td>MLELRHRGCPGAP</td>
</tr>
<tr>
<td>J</td>
<td>CPGP</td>
<td>CPGP</td>
</tr>
<tr>
<td>J</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>J</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>133</td>
<td>133</td>
<td>133</td>
</tr>
</tbody>
</table>

**Figure 3.12 Amino acid alignment of sequences representing Cds1, derived from rat, human and mouse.**

The figure shows the alignment of the derived amino acid sequences from the rat, human and mouse Cds1. The areas boxed and shaded are identical residues and similar residues are just shaded. The human and mouse sequences share 96% identity, the human and rat sequences share 95% identity and the mouse and rat sequences share 98% identity.
The mouse and human CDS1 genes are 96% identical and the mouse and rat genes are 99% identical at the amino acid level.

### 3.1.4 Identification of murine Cds1 3' UTR

The first murine Cds1 cDNA clone, 801-h9, to be identified appeared to contain a poly A-tail although there was no obvious poly-adenylation signal present. Analysis of the human and rat sequences from the database, which were being used to find the 5' end of mouse Cds1, showed that the other two species had much longer 3' untranslated regions (UTRs) than the sequence derived from 801-h9. To verify if the mouse sequence was genuinely shorter than that of the human or rat sequence, the sequence representing the 3'UTR of the rat was used to search the mouse-EST database using BLAST-N. This search identified three overlapping clones (GenBank accession numbers: AI314024 (I.M.A.G.E: 1920907), AI316951 (I.M.A.G.E: 1920907), AK014670). Using the sequence data from the database more of the 3'UTR of the mouse Cds1 gene was generated. The composite nucleotide sequence with derived amino acid sequence of murine Cds1 is shown in figure 3.13. The generated sequence is of similar length to that of rat and human. The rat 3' UTR is 2345bp in length and is 68% identical to the 2303bp 3'UTR of the mouse gene at the nucleotide level.

### 3.1.5 Cds1 sequence homology between species

The mouse Cds1 gene was identified as a homologue of Drosophila eye-cds using the human CDS1 sequence. In order to establish the identity between the newly generated mouse Cds1 coding region and that of the Drosophila gene eye-cds (GenBank accession number: P 56079), which was used initially to identify the human CDS1 gene, an alignment of the amino acid sequences was performed (figure 3.14). The mouse Cds1 and Drosophila eye-cds are 56% identical at the amino acid level.

In order to predict if the Drosophila, mouse and human CDS1 proteins all have the same transmembrane structure, the amino acid sequences from each species were used to computer generate hydrophobicity plots using the computer programme TMPred (http://www.ch.embnet.org/software/TMPRED). The TMPred programme makes predictions of membrane-spanning regions based on statistical analysis of Tmbase, a database of naturally occurring transmembrane proteins.
ATGCTGGGACGCTGCGCCACCGGGGGCTGCCCCGGCCCTGGGGGAGCGGGGGCGCCGCCCCTCCGCGAGGGAGAGGCGGCCGGCGGTGACMLELRHRGGCPGPGGAGAPPLREGEAAGGD
90

CACGAAACCGAGAGCACCAGCGACAAAGAAACAGATATTGATGACAGGTATGGAGATCTCGATGCCAGAGGCGATTCCGACGTGCCTGAG 1 8 0
HETESTSDKETDIDDRYGDLDARGDSDVPE
450

GTCCCACCGTCCTCAGACAGGACCCCCGAGATTCTCAAGAAAGCCCTGTCCGGACTATCTTCAAGATGGAAGAACTGGTGGATTCGCGGG 2 7 0
VPPSSDRTPEILKKALSGLSSRWKNWWIRG
540

ATCCTCACCCTCACCATGATCTCCCTCTTCTTCCTGATTATCTATATGGGGTCCTTCATGCTGATGCTTCTGGTTCTGGGCATCCAAGTG 3 6 0
ILTLTMISLFFLIIYMGSFMLMLLVLGIQV
1 0 6

AAGTGCTTCCATGAGATCATCACCATCGGGTACCGGGTCTACCACTCCTACGACCTCCCATGGTTTAGGACACTAAGTTGGTACTTCCTC 4 5 0
KCFHEIITIGYRVYHSYDLPWFRTLSWYFL
2 4 0

CTGTGTGTGAACTACTTCTTCTATGGAGAGACGGTGGCAGATTACTTCGCCACGTTTGTTCAGAGGGAGGAGCAGCTGCAGTTCCTCATT 5 4 0
LCVNYVYETGTYTADFPATFVYQREEQLQFL
3 3 0

GTCCCACCGTCCTCAGACAGGACCCCCGAGATTCTCAAGAAAGCCCTGTCCGGACTATCTTCAAGATGGAAGAACTGGTGGATTCGCGGG 2 7 0
VPPSSDRTPEILKKALSGLSSRWKNWWIRG
540

ATCCTCACCCTCACCATGATCTCCCTCTTCTTCCTGATTATCTATATGGGGTCCTTCATGCTGATGCTTCTGGTTCTGGGCATCCAAGTG 3 6 0
ILTLTMISLFFLIIYMGSFMLMLLVLGIQV
1 0 6

AAGTGCTTCCATGAGATCATCACCATCGGGTACCGGGTCTACCACTCCTACGACCTCCCATGGTTTAGGACACTAAGTTGGTACTTCCTC 4 5 0
KCFHEIITIGYRVYHSYDLPWFRTLSWYFL
1 0 6

CTGTGTGTGAACTACTTCTTCTATGGAGAGACGGTGGCAGATTACTTCGCCACGTTTGTTCAGAGGGAGGAGCAGCTGCAGTTCCTCATT 5 4 0
LCVNYVYETGTYTADFPATFVYQREEQLQFL
3 3 0

Figure 3.13 Full length mouse Cds1 sequence
Shown is the known full-length sequence of mouse Cds1. To date no 5'UTR sequence has been identified. The start and stop codons are highlighted and underlined in red, the derived amino acid sequence is shown in blue and the 3'UTR excluding polyA+ tail (poly A+ signal highlighted and underlined in purple) is shown from position 1386bp. The 3'UTR was generated in full from mouse sequences homologous to the 3'UTR of rat and human.

146
<table>
<thead>
<tr>
<th>Dros cds</th>
<th>Mouse Cds1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEERRRGEDE-----------------</td>
<td>EANKRNSAADSD---------EKFVDELAKN</td>
</tr>
<tr>
<td>57</td>
<td></td>
</tr>
</tbody>
</table>
| UPDGLMTPEI--------------PS--| VXRIHSLWYFFLLSNYFFYGELVLDFVYVINRNLFL | 117
| 117                        |                                 |
| KCFIEIIG------KCFIEI--- | VYHSLWYFFLLSNYFFYGELVTAFYTFVQREHI   |
| 177                        |                                 |
| TYHFLSFALY---TYHFLS---   | FYWVFVLSLVKKVYRIQFNFAWTHVLTVOSLLIONFEGII |
| 237                        |                                 |
| WFVVSVMVCNDVAY-----WFVS-- | GFGFRTPLKSLPKTWEGFGGATFIFLYSVICN  |
| 297                        |                                 |
| YQYFCPQVEQGRMTSCQP---LYPFIVHSI--- | GIGK-TELNLYPFIVHSISLS |
| 353                        |                                 |
| IFSS1GPSGFFASAAGKRAFKIKDI| DIPHHGIMDFDCFLMATFVNYI SFFHT  |
| 413                        |                                 |
| PGGKLLQIVLYFLPOQYIELDKONG- | 447                           |

**Figure 3.14 Drosophila cds and mouse Cds1 amino acid alignment**

The amino acid alignment shows that the sequences are 56% identical at the amino acid level. The highly homologous regions highlighted in red correspond to peaks V and VI on the hydrophobicity plots shown in figure 3.26, which predict transmembrane regions. Regions of residue identity are boxed and shaded, while similar residues are shaded.
The prediction is made using a combination of several weight-matrices for scoring (Kyte and Doolittle, 1982; Hofman and Stoffel, 1993). The plots that were generated to compare mouse and *Drosophila* are shown in figure 3.15 and the plots comparing mouse and human Cds1 shown in figure 3.16. The plots predict that the *Drosophila*, mouse and human have seven transmembrane domains.

### 3.1.6 Summary of section 3.1

In summary, the mouse *Cds1* gene is 3686bp in length (excluding the 5' UTR), with a derived amino acid length of 461 amino acids and the protein has been predicted to form seven transmembrane domains. The gene sequence was generated by the identification of clone 801-h9 and 2279-d19 and by sequencing a PCR fragment amplified from mouse cDNA using primers designed to the start codon of rat *Cds1*. The 3' UTR was generated bioinformatically using the BLAST-N algorithm (http://www.ncbi.nlm.nih.gov/). The amino acid sequence generated for mouse Cds1 is 98% identical to rat, 96% identical to human and 56% identical to *Drosophila*. The experiments performed in this section were done in order to gather information about the mouse *Cds1* gene. This was then used in section 3.3 to generate the genomic structure, in section 3.6 to examine the expression patterns of the gene and in section 3.7 in the preparation of a *Drosophila* rescue construct.
Figure 3.15 Mouse and Drosophila cds1 hydrophobicity plots
The hydropobicity plots which were generated using TMPRED online software from the amino acid sequences of mouse Cds1 and Drosophila cds are shown. The package predicts seven transmembrane domains for both mouse Cds1 (A) and Drosophila cds (B). The transmembrane regions are numbered in red I-VII, and correspond to the amino acid region numbered along the base axis. Peaks V and VI correspond to regions of high sequence identity (see figure 3.16)
Figure 3.16 Mouse and human Cds1 hydrophobicity plots

The hydophobicity plots which were generated using TMPRED online software from the amino acid sequences of human CDS1 and mouse Cds1 are shown. The package predicts an almost identical pattern of seven transmembrane regions for both human (A) and mouse Cds1 (B). The transmembrane domains are numbered I-VII in red.
3.2 Isolation of Murine Cds2

3.2.1 Identification of cDNA clone 1500-b19
To isolate murine Cds2, a similar approach was taken as described in section 3.1 for Cds1. The nucleotide sequence of the previously identified human CDS2 (GenBank accession number: AF069532 (Halford et al., 1998)) was used to search the EST database to identify murine clones. This search identified one clone (GenBank accession number: AA170489, I.M.A.G.E number: 618162) which was obtained from HGMP-RC and given the laboratory number 1500-b19. This clone originated from a Soares mouse 3NbMS spleen library and was cloned into the NotI and EcoRI sites of a pT7T3D vector. DNA was prepared from the clone (see section 2.3.1.1-2). The insert, which was 660bp in length, was released from the vector by restriction digestion using enzymes EcoRI and NotI (see section 2.3.1.9). 1500-b19 was initially sequenced with vector primers M13F and M13R (see section 2.3.1.18) in order to verify that it was correct and then fully sequenced with specific primers (table 3.5).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5'[13]</th>
<th>Sequence Start Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>1500F1</td>
<td>GCCTTGGAGAGATGAGTAGA</td>
<td>544</td>
</tr>
<tr>
<td>1500R1</td>
<td>ATCTAAGTCACAGGCAGCTGG</td>
<td>623</td>
</tr>
<tr>
<td>1500F2</td>
<td>GGCACCTTTGTCAATGTGT</td>
<td>388</td>
</tr>
<tr>
<td>1500R2</td>
<td>GTACCCAGACATCACAT</td>
<td>57</td>
</tr>
<tr>
<td>1500R3</td>
<td>CCACAGTGAAAGCTGGT</td>
<td>109</td>
</tr>
</tbody>
</table>

Table 3.5 Primers used to sequence clone 1500-b19

The derived amino acid sequence for clone 1500-b19 was aligned with the known human CDS2 amino acid sequence (figure 3.17). The alignment in figure 3.20 shows that clone 1500-b19 represents the 3' end of murine Cds2 and that the 5' end is not contained within 1500-b19. The amino acid sequences of the two species differ in length by 12 amino acids at the 3' end. Clone 1500-b19 is 99.4% identical to the 3' end of human CDS2 at the amino acid level.
**Figure 3.17 Amino acid alignment of 1500-b19 and human CDS2**

Shown is the amino acid alignment of clone 1500-b19 and human CDS2 sequences. The human sequence differs from 1500-b19 at the 3' end, this sequence was verified (see figure 3.24). Residues which are identical are boxed and shaded whereas residues which are similar are only shaded. This alignment shows that a clone containing the 5' mouse Cds2 gene is required.
3.2.2 Identification of cDNA clone 1247-c2
To find a clone which contained the 5' end of mouse Cds2, the nucleotide sequence of clone 1500-b19 was used to re-screen the EST database to identify overlapping clones. This resulted in the identification of a further clone (GenBank accession number: AI427541, I.M.A.G.E number: 521017) which originated from a Stratagene adult mouse diaphragm (# 937303) library constructed in the EcoRI and NotI restriction sites of a pBluescript SK- vector. The clone was obtained from HGMP-RC and given the laboratory number 1247-c2. DNA was prepared from clone 1247-c2, as before, and the 1.5kb insert was released by restriction digestion with the enzymes EcoRI and XhoI.
The clone was sequenced initially using vector specific primers M13F and M13R and then with sequence specific primers (table 3.6). The 1556bp nucleotide sequence generated from clone 1247-c2 with derived amino acid sequence is shown in figure 3.18. In figure 3.18 the start and stop codons of Cds2 are highlighted and underlined in red, and the primer sequences (table 3.6) are highlighted in purple. The sequences derived from clone 1247-c2 and 1500-b19 align with 100% identity at the 3' end.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5' 3'</th>
<th>Primer Start Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>1500R2</td>
<td>GTACCCAGACATCAGCAT</td>
<td>975</td>
</tr>
<tr>
<td>1247F1</td>
<td>GTAACGCAGTCACACCTTTGT</td>
<td>744</td>
</tr>
<tr>
<td>1500R3</td>
<td>CCACTGACAGCTGTTGGTG</td>
<td>1030</td>
</tr>
<tr>
<td>1247R1</td>
<td>GCAGAATCTCTGTCTGATAC</td>
<td>660</td>
</tr>
</tbody>
</table>

Table 3.6 Primer names which were used to generate the sequence of clone 1247-c2.

GeneWorks was used to predict the restriction map of 1247-c2 (figure 3.19) which was then verified by restriction digests using the enzymes BamHI, HindIII, PstI and EcoRI in single, double and triple digests. The map was prepared in order to identify restriction sites which could later be used to generate probes.
Figure 3.18 Sequence generated from clone 1247-c2
The figures shows the nucleotide and derived amino acid sequence from clone 1247-c2 which represents Cds2. The start and stop codons are highlighted and underlined in red and the primer sequences of the primers (table 3.6) which were used to sequence the clone are shown in purple.
Figure 3.19 Restriction map of clone 1247
The figure shows a schematic restriction map of the *Cds2* cDNA clone 1247 and a gel photograph of a series of restriction digests. The clone was digested using *BamHI, HindIII, EcoRI* and *PstI* restriction enzymes. Some of the reactions have not gone to completion, but these bands are easily identified. The schematic map was compiled using the gel results then verified by sequencing data.

Key: H- *HindIII*, B- *BamHI*, E- *EcoRI*, P- *PstI*, x4 - *HindIII, BamHI, EcoRI*, and *PstI*.
3.2.3 Cds2 sequence homology between species
In order to examine the amino acid sequence identity between human and mouse Cds2 an alignment of the sequences were performed (figure 3.20). The human and mouse Cds2 sequences show a 96% sequence identity at the amino acid level. In section 3.1 an alignment of amino acid sequences from Drosophila eye-cds (GenBank accession number: P56076) and mouse Cds1 shared an identity of 56%. The same alignment was performed with the Drosophila eye-cds amino acid sequence but this time with the amino acid sequence representing Cds2 (figure 3.21). The identity between Drosophila eye-cds and mouse Cds2 is 57% at the amino acid level.

In order to predict the number of transmembrane domains each protein has all three, human CDS2, mouse Cds2 and Drosophila eye-cds, amino acid sequences were entered into the transmembrane-spanning region predictive programme TMPred (see section 3.1 for programme details). The resultant hydrophobicity plots for mouse Cds2 and Drosophila eye-cds are shown in figure 3.22 and the plots for human CDS2 and mouse Cds2 are shown in figure 3.23. The plots predict that the Drosophila, mouse and human have seven transmembrane domains.

3.2.4 Summary of section 3.2
In summary, the murine Cds2 gene is at least 1556bp in length (with only limited 5' and 3' UTR having been identified to date), with a derived amino acid length of 432 and predicted protein structure containing seven transmembrane domains. The sequence was generated using two cDNA clones, 1500-b19 and 1247-c2. The experiments performed in this section were done in order to gather information about the mouse Cds2 gene. The information generated was then used in section 3.4 to generate the genomic structure, in section 3.6 to examine the expression patterns of the gene and in section 3.7 in the preparation of a Drosophila rescue construct.
Figure 3.20 Amino acid alignment of human and mouse Cds2

An amino acid alignment of mouse Cds2 and human CDS2 sequences. The boxed and shaded areas indicate identical residues whereas similar residues are only shaded. The sequences are 96% identical.
Figure 3.21 *Drosophila* cds and mouse Cds2 amino acid alignment

Shown is the amino acid alignment of *Drosophila* cds and mouse Cds2. Residues which are identical are boxed and shaded whereas residues which are similar are just shaded. The sequences are 57% identical. Two regions of high sequence homology are highlighted in red. These regions correspond to predicted transmembrane domains shown in figure 3.26.
Figure 3.22 *Drosophila* and mouse Cds2 hydrophobicity plots
Shown above are the hydrophobicity plots that the online software program TMPRED plotted from the amino acid sequences of mouse Cds2 and *Drosophila* cds. The mouse (A) and the *Drosophila* (B) plots show seven predicted transmembrane domains numbered I-VII in red.
Figure 3.23 Human and mouse Cds2 hydrophobicity plots
Indicated above are the hydrophobicity plots that the online software program TMPRED plotted from the amino acid sequences of mouse Cds2 and human CDS2. The mouse (A) and the human (B) plots show seven predicted transmembrane domains (I-VII), the pattern shown is very similar between the two species.
3.3 Genomic structure of murine Cds1

3.3.1 Identification of a PAC clone representing Cds1
A genomic clone of mouse Cds1 was required for gene mapping studies (see section 3.5) and to determine the genomic structure of the gene. In order to identify a genomic clone a mouse P1 artificial chromosome (PAC) library (HGMP-RC) was screened (see section 2.3.1.19.2). The library was screened by hybridising the gridded filters with a probe made from the entire 1.3kb insert from the cDNA clone 801-h9, which represents Cds1 that had been randomly $^{32}$P-radiolabelled (see section 2.3.1.20). The filters were hybridised overnight (see sections 2.3.1.21) then washed at a stringency of 0.1X SSC, 0.1% SDS. The filters were exposed to autoradiography film (Kodak) at -80°C for 72 hours. Three positive clones were identified, 379-p6, 572-a24 and 620-k8 and obtained from HGMP-RC. DNA was prepared from each clone (see section 2.3.1.3-4) and the concentration estimated by spectrophotometry. The clones were verified on a Southern blot, 10µg of DNA from each PAC clone was digested with EcoRI, electrophoresed through an agarose gel (see section 2.3.1.10) and blotted onto Hybond-N (as 2.3.1.11). The blot was hybridised with the 1.3kb probe originally used to screen the PAC library filters. The autoradiography results (figure 3.24) show that the probe did not hybridise to clone 620-k8 and hence 620-k8 was deemed to be a false positive or alternatively the wrong clone had been sent from HGMP-RC. As there were two verified positive clones, 620-k8 was not pursued further. Clone 572-a24 was used for all further analysis.

3.3.2 Identification of Cds1 intron/exon boundaries
Previously a human PAC clone which contained human CDS1 (GenBank accession number: AC022739) had been identified by screening the high throughput genome sequence (htgs) database at NCBI (http://www.ncbi.nlm.nih.gov/BLAST) (S. Halford, personal communication) with the human CDS1 nucleotide sequence. The sequence information available from the human PAC clone indicated the positions of the intron splice sites of human CDSL. However there was no mouse PAC clone containing sequence information about murine Cds1 available in the htgs database so the information available from the human PAC clone was used. As described in section 3.1 the amino acid sequences of human and mouse Cds1 are 96% identical. Using an
Figure 3.24 Southern blot of potential PAC clones
The PAC clones which were identified from the HGMP-RC PAC library were digested with enzymes BamHI, EcoRI and HindIII. A Southern blot filter was made from the products. The filter was probed with the 801 probe. Shown is the autoradiograph eliminating PAC clone 620-k8 as a false positive. The two other clones appear to be very similar.
alignment of the two amino acid sequences derived from the cDNA sequence of both genes, the human intron splice sites were identified and hence used to represent the provisional placement of the mouse intron splice sites. In order to verify the position of the mouse intron splice sites and deduce the intron sizes, Cds1 specific oligonucleotide primers were designed approximately 50bp up- and downstream of the provisional splice sites (table 3.7) using mouse Cds1 nucleotide sequence.

3.3.3 Cds1 genomic structure
To deduce the size of the mouse introns and then verify the position of the splice boundaries, 10ng of PAC clone 572-a24 DNA was used as template in PCR reactions (see section 2.3.1.16). The resultant fragments indicated the size of the introns (figure 3.25). The PCR fragments were cloned into pGEM T-Easy vector (see section 2.3.1.13-14). The ends of the cloned intronic fragments were sequenced with M13F and M13R vector specific primers (see section 2.3.1.18). Some of the fragments proved difficult to clone and were sequenced directly. Some of the fragments proved difficult to clone and to sequence, but the intron sizes and the available sequence data confirming the position of the intron splice sites derived from the PAC clone are shown table 3.8. The Cds1 structural information gathered from PAC clone 572-a24 is shown schematically in figure 3.26.

3.3.4 Summary of section 3.3
In summary the murine Cds1 gene has twelve exons and genomically spans approximately 31kb. The experiments performed in this section generated a PAC clone representing murine Cds1 that was used to map the chromosomal location of the gene by fluorescent in situ hybridisation (FISH) in section 3.5. The information and resources reported in this section were essential, with the view to making a construct for the production of a knockout mouse.
<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5'→3'</th>
<th>Primer Start Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDS1intron1F</td>
<td>GACCACGAAACCGAGAGCAC</td>
<td>36</td>
</tr>
<tr>
<td>CDS1intron1R</td>
<td>TGGCATCGAGATCTCTCCATAC</td>
<td>106</td>
</tr>
<tr>
<td>CDS1intron2F</td>
<td>GTATGGGATCTCGATGCGGA</td>
<td>86</td>
</tr>
<tr>
<td>CDS1intron2R</td>
<td>TCAAGGAGAGAGGGAGATC</td>
<td>256</td>
</tr>
<tr>
<td>CDS1intron3F</td>
<td>GATCTCCCTCTCTCTCTCTCTGA</td>
<td>236</td>
</tr>
<tr>
<td>CDS1intron3R</td>
<td>TAAACCAGGGAGGTCGTA</td>
<td>376</td>
</tr>
<tr>
<td>CDS1intron4F</td>
<td>ATATCCCTTCGCCCTCTACCTGG</td>
<td>505</td>
</tr>
<tr>
<td>CDS1intron4R</td>
<td>TGCTTCCTTCACCAGGCCTCAG</td>
<td>566</td>
</tr>
<tr>
<td>CDS1intron5F</td>
<td>CTGAGCCTGGAAGAGAAGCA</td>
<td>547</td>
</tr>
<tr>
<td>CDS1intron5R</td>
<td>GGAATGAAGAGATGTGACTGA</td>
<td>646</td>
</tr>
<tr>
<td>CDS1intron6F</td>
<td>TCAGTCACATCTGTACCTCC</td>
<td>627</td>
</tr>
<tr>
<td>CDS1intron6R</td>
<td>GATGTCATAGATCATGACGC</td>
<td>711</td>
</tr>
<tr>
<td>CDS1intron7F</td>
<td>GCGTCTATCTGCAATGACATC</td>
<td>692</td>
</tr>
<tr>
<td>CDS1intron7R</td>
<td>CACCAATGAATGTTCTCCCAA</td>
<td>796</td>
</tr>
<tr>
<td>CDS1intron8F</td>
<td>TGGGAAGGATTCATTGGTG</td>
<td>777</td>
</tr>
<tr>
<td>CDS1intron8R</td>
<td>TCCACGGGCCACACAAAGTA</td>
<td>875</td>
</tr>
<tr>
<td>CDS1intron9F</td>
<td>TACTTTGTGTGCCCCTGGGA</td>
<td>856</td>
</tr>
<tr>
<td>CDS1intron9R</td>
<td>CGCGAAGGTTGAAAGGGCAA</td>
<td>1038</td>
</tr>
<tr>
<td>CDS1intron10F</td>
<td>TTGCCCCTTCACGTTGC</td>
<td>1019</td>
</tr>
<tr>
<td>CDS1intron10R</td>
<td>AACCTGTCATATCCTGCCG</td>
<td>1148</td>
</tr>
<tr>
<td>CDS1intron11F</td>
<td>GCCGGGATAATGGACAGTTT</td>
<td>1128</td>
</tr>
<tr>
<td>CDS1intron11R</td>
<td>AAGCACTAACAGCTGCTGG</td>
<td>1218</td>
</tr>
</tbody>
</table>

**Table 3.7 Primers used to amplify Cds1 introns**

The table lists the name, sequence and position of the primer in relation to the sequence generated from clone 2279-d19 homologous to human CDS1. The primer pairs were used to amplify the introns of Cds1 using PAC clone 572-a24.
Figure 3.25 Photograph of \textit{Cds1} introns

Shown are the PCR fragments representing the \textit{Cds1} intron sizes. The fragment for 6 (1.6kb) is missing from this photograph. Lane 1-10 represent introns, 1 (4kb), 2 (2kb), 3 (5kb), 4 (3.9kb), 5 (1.8kb), 7 (2.5kb), 8 (2kb), 9 (1.3kb), 10 (2kb), 11 (3.8kb) respectively.
### Table 3.8 Cds1 Splice junction information

The table lists the sizes of the exons and the introns of mouse *Cds1* to date. The table also list some of the sequence generated around the exon splice sites. The information was gathered using PCR and sequencing. The primers used to generate this data are listed in table 3.7.

<table>
<thead>
<tr>
<th>Exon</th>
<th>EXON SIZE (bp)</th>
<th>5’ splice donor</th>
<th>Intron</th>
<th>3’ splice acceptor</th>
<th>Intron Size (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>126</td>
<td>TATCTTCAAG</td>
<td>2</td>
<td>...</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>97</td>
<td>GATGCTTCTGAaacgcaagtattttagt...</td>
<td>3</td>
<td>...</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>238</td>
<td>TACCTGGCA</td>
<td>4</td>
<td>...</td>
<td>3.9</td>
</tr>
<tr>
<td>4</td>
<td>59</td>
<td>GTTCTATG</td>
<td>5</td>
<td>...tttgggcatataacagTTGATG</td>
<td>1.8</td>
</tr>
<tr>
<td>5</td>
<td>83</td>
<td>GCATGATATGtaaggactatgtgtgatg...</td>
<td>6</td>
<td>...ttcctgctaatttttaagTTTCTTATGT</td>
<td>1.6</td>
</tr>
<tr>
<td>6</td>
<td>88</td>
<td>TCTAATTTaactacaggctgcagcattg...</td>
<td>7</td>
<td>...caagaactcttaattaagCTGTTCTTATTA</td>
<td>2.5</td>
</tr>
<tr>
<td>7</td>
<td>70</td>
<td>GGATTCATTGtagtttagaggatttagtt...</td>
<td>8</td>
<td>...agtcatatagtttagcattagCTGCTTATGT</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>164</td>
<td>ACTGTGAGCTaagcaagagcccgaga...</td>
<td>9</td>
<td>...ccactgctcctctacagCTGATCCCT</td>
<td>1.3</td>
</tr>
<tr>
<td>9</td>
<td>108</td>
<td>CAAAATCAAGtttgctgtttgtgcattttctc...</td>
<td>10</td>
<td>...cacttgctctttgctgcagGATTTCCTTCA</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>104</td>
<td>GCTTCATCGaggcagcggagtttagcagaa...</td>
<td>11</td>
<td>...</td>
<td>3.88</td>
</tr>
<tr>
<td>11</td>
<td>128</td>
<td>TTGAAGGTGTAAG....</td>
<td>12</td>
<td>3’UTR...</td>
<td></td>
</tr>
</tbody>
</table>

1.6

2.5

3.88
Figure 3.26 Schematic representation of the known Cds1 genomic restriction map
Shown is a schematic illustration of mouse Cds1 restriction map to date. This was generated by a combination of restriction digests and sequencing. The intron sizes are shown in red.
3.4 Genomic structure of murine Cds2

3.4.1 Identification of a genomic phage clone
To determine the genomic structure and chromosomal location (see section 3.5) of Cds2 a genomic clone containing the gene was required. A phage library (Stratagene) constructed in the λFIX®II vector using 8 week male spleens from mouse strain 129/SvJ was screened (see section 2.3.1.19.1). The library filters were hybridised (see section 2.3.1.21) with the entire insert from the cDNA clone 1500-b19, which had been randomly 32P-radiolabelled (see section 2.3.1.20). This resulted in the identification of four primary positives. A further round of screening resulted in only two positives. The two positive clones were given the laboratory numbers 1A and 20A and were digested with the restriction enzyme NotI (see section 2.3.1.9) and the inserts cloned into pBluescriptSK+ plasmid vectors for easier manipulation (see section 2.3.1.13-14). Restriction digest analysis indicated that 20A was 1kb longer than 1A and was used for further analysis.

3.4.2 Phage clone 20A
The insert from clone 20A is 18kb in length. A restriction map of 20A was generated with the restriction enzymes BamHI, EcoRI and HindIII in a combination of single, double and triple digests (figure 3.27). The clone appears to cut most efficiently with enzyme HindIII, so a double digest with HindIII and NotI was performed to determine which fragment contained the vector. As seen in figure 3.27 (lane 6), when 20A in pBluescript SK+ is digested with only HindIII six fragments are produced, 9kb, 6kb, 4.5kb, 300bp, 250bp and 200bp but when the clone 20A is digested with HindIII and NotI, seven fragments are produced, 9kb, 6kb, 3kb (vector), 1.5kb and the same three small fragments. This data positioned the 1.5kb NotI/HindIII fragment at the 5' end of the clone by using the vetor NotI site as an anchor point. Since the NotI/HindIII fragments were digesting most efficiently the six insert fragments were subcloned into pBluescriptSK+ vectors for further analysis. The cloned fragments generated from 20A were sequenced initially using vector primers M13F and M13R, the small fragments did not sequence well and no sequence information was obtained. Using the sequencing data derived from the larger fragments, two of the fragments contained sequence identical to that derived from cDNA clone 1247-c2 and one fragment (9kb) did not produce any sequence recognisable as Cds2.
The clone 20A was restriction digested with various enzymes. The reactions are contained within the following lanes. Lane 1, *EcoRI* gives three bands of ~20kb, 2kb and 800bp. Lane 2, *EcoRI/BamHI* double digest gives eight bands ranging from 5kb to 200bp. Lane 3, *BamHI* gives seven bands ranging from 5kb to 200bp. Lane 4, *EcoRI* digest. Lane 5, *EcoRI/HindIII* digest gives seven bands ranging from 6kb to 200bp. Lane 6, *HindIII* digest giving six bands from 9kb to 200bp (small bands are barely visible in photo but were clearly visible on the gel) these fragments were subcloned for further analysis. Lane 7, *HindIII* digest. Lane 8, *HindIII/BamHI* digest giving seven bands ranging from 3.2kb to 200bp. Lane 9, *BamHI* digest. Lane 10, triple digest of all enzymes. This gel was used in forming the restriction map of 20A which can be seen in figure 3.28.
Sequence data derived from the 5’end of the 1.5kb *NotI/HindIII* fragment had ~150bp of unknown sequence before sequence from 510bp into clone 1247-c2 sequence was recognised, thus 20A was deemed not to contain all of *Cds2*. The first 90bp of the sequence derived from the 3’end of the 1.5kb fragment was not recognised then the sequence aligned with sequence from ~800bp into the cDNA clone 1247-c2. The sequence data from the 5’end of the 6kb *HindIII* subcloned fragment was not recognised as aligning with the *Cds2* cDNA sequence, but the sequence data from the 3’ end of the fragment aligned with the 3’ end of cDNA clone 1247-c2. With no information available from the small fragments, provisional data suggested that all the *Cds2* genomic information was contained within the 1.5kb and 6kb fragments which were provisionally positioned adjacent at the 5’ end of the restriction map of 20A. The positions of the 1.5kb and 6kb fragments were subsequently verified by fully sequencing intron 6. No *HindIII* sites were found within the intron that would indicate the presence of any of the small *HindIII* fragments between the 1.5kb and 6kb fragments. Initially the 9kb fragment was positioned at the 3’ end of clone 20A with the three small fragments positioned centrally after the 6kb fragment. The *EcoRI* fragments were positioned onto the map of 20A in a similar way. Since there are only three *EcoRI* fragment sized, 18kb, 2kb and 800bp (figure 3.27 lane1), the large 18kb fragment is the only fragment that could contain the 3kb vector fragment. The first *EcoRI* site was positioned on the map 15kb from the 5’ end of the clone using the vector *EcoRI* as an anchor point. The positions of the other two *EcoRI* fragments were determined later. To determine the position of the small *HindIII* fragments in the map of 20A, Southern blot panels (see section 2.3.1.11) were prepared from 20A DNA restriction digested with enzymes *BamHI*, *EcoRI* and *HindIII*. These panels were hybridised with probes made from the cloned inserts of the small *HindIII* fragments randomly $^{32}$P-radiolabelled. The results (figure 3.28) show that the 300bp and 250bp *HindIII* fragments both hybridise to the 15kb *EcoRI* fragment which are positioned adjacent to the 6kb fragment. The 200bp fragment probe appears to hybridise to the 2kb *EcoRI* fragment although it does not appear to hybridise to itself. This may indicate that the probe does not truly represent the 200bp *HindIII* fragment as first thought, and the position of the fragment, which has been positioned at the very 3’ end of clone 20A, remains tentative. The result using the 200bp probe positions the two remaining
Figure 3.28 Autoradiographs determining the position of the small HindIII fragments in 20A

Shown is the 20A digest gel (A), which was Southern blotted to provide the filter for hybridisation. The filter was probed with each of the small HindIII fragments from 20A in order to position them onto a restriction map of the 20A clone. (B) 300bp fragment that hybridises to the 18kb EcoRI fragment, 5kb BamHI fragment and a 300bp HindIII fragment. (C) 250bp fragment that hybridises as the 300bp fragment except to the 250bp HindIII fragment. (D) A new blot had to be prepared for the final fragment, 200bp fragment which hybridises to the 2kb EcoRI fragment and the 2.8 and 3.2kb BamHI fragments although there is no self hybridisation (hence the results are still tentative). It is apparent that the 200bp fragment is hybridising to a completely different set of fragments than the previous probes. The deduced map is shown in figure 3.29.
EcoRI fragments. The order is 800bp fragment followed by the 2kb fragment at the 3' end. The results from this section conclude that Cds2 genomic information is contained within the 1.5kb and 6kb HindIII fragments at the 5' end of clone 20A.

3.4.3 Identification of a PAC clone representing Cds2

In a similar way as was described in section 3.3, a human PAC clone had been identified (GenBank accession number: AC016073) and the intron/exon boundaries determined for human CDS2 (S. Halford, personal communication). Using the information about the human boundary sites the provisional mouse intron boundary sites were positioned on the cDNA sequence. Sequence specific primers were designed approximately 50bp up- and downstream (table 3.9) using the provisional intron boundary sites. The primers listed in table 3.9 were used in PCR reactions (see section 2.3.1.16) to amplify the available introns using clone 20A derived DNA as template. This provisional intron splice site information was used to deduce that the 1.5kb NotI/HindIII cloned fragment contained 150bp of the 3' end of intron 3 to the first 90bp of intron 7. The information gathered from these PCR reactions found that the 6kb HindIII fragment contains the remainder of intron 7 to the start of the 3' UTR of Cds2 (figure 3.29). Sequencing of 20A using primers in table 3.9 verified the position of the intron splice sites, and the information obtained was used to compile part of table 3.10. Table 3.10 shows a large number of non-canonical splice sites. This may be due to the fragments which were analysed being generated via PCR. PCR errors leading to changes in the sequence of the boundaries may have produced this feature. These sites should therefore be verified using cloned DNA or various PCR products amplified on different occasions and both strands sequenced. Using the restriction information the remaining sites were positioned on the 5' end of 20A containing the genomic structural information of Cds2. This information confirmed that a more 5' genomic clone was required to deduce the remainder of the Cds2 genomic structure.

To obtain a genomic clone containing the 5' end of Cds2 a gridded PAC library (HGMP-RC) was screened (see section 2.3.1.19.2). The filters were hybridised overnight with a probe generated from the entire insert from the 1.5kb NotI/HindIII fragment, which was subcloned from 20A. This identified two positive clones, 439-n24 and 542-j8, which were obtained from HGMP-RC. DNA was prepared from both clones (see section 2.3.1.3-4) and the concentration determined by spectrophotometry.
10µg of each DNA was digested using the restriction enzyme EcoRI and a Southern blot was prepared. The Southern blot filter was probed overnight with the same 1.5kb fragment as used to screen the PAC library. The results showed that both clones were true positives and 439-n24 was used in further studies.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5' → 3'</th>
<th>Primer start position</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDS2 INTRON1F</td>
<td>CAGCATGACCAGAACTAGG</td>
<td>122</td>
</tr>
<tr>
<td>CDS2 INTRON1R</td>
<td>TGTTTCGCCTCGGCTCTC</td>
<td>246</td>
</tr>
<tr>
<td>MEX2F</td>
<td>GAATCAGAGTCCGAGCAAAA</td>
<td>180</td>
</tr>
<tr>
<td>MEX2R</td>
<td>TCATCAAACCACATTGGTCC</td>
<td>409</td>
</tr>
<tr>
<td>CDS2 INTRON2F</td>
<td>CTTTGCCATGATCGCGTT</td>
<td>349</td>
</tr>
<tr>
<td>CDS2 INTRON2R</td>
<td>CAGGTCTAGAGGAGTGGTAT</td>
<td>489</td>
</tr>
<tr>
<td>CDS2 INTRON3F</td>
<td>CAGATTAAGTGTTCATG</td>
<td>426</td>
</tr>
<tr>
<td>CDS2 INTRON3R</td>
<td>GACCAGATGAAGAAGTAG</td>
<td>582</td>
</tr>
<tr>
<td>CDS2 INTRON5F</td>
<td>GTCCAGAGACAGGAGCTT</td>
<td>579</td>
</tr>
<tr>
<td>CDS2 INTRON5R</td>
<td>CATGTAGAAGTCACAGCC</td>
<td>711</td>
</tr>
<tr>
<td>CDS2 INTRON6F</td>
<td>GATTCCTGCATGTGTGTAC</td>
<td>653</td>
</tr>
<tr>
<td>CDS2 INTRON6R</td>
<td>GATTCCTGCATGTGTGTAC</td>
<td>787</td>
</tr>
<tr>
<td>1500F3</td>
<td>GTAACGCAGTCACACCTTGT</td>
<td>744</td>
</tr>
<tr>
<td>MEX4R</td>
<td>CGTCTAGTCGCTGCTCTACC</td>
<td>825</td>
</tr>
<tr>
<td>CDS2 INTRON7F</td>
<td>CAAAGAAGACCTGGGAAGGC</td>
<td>889</td>
</tr>
<tr>
<td>1500R3</td>
<td>CCAAGAGACCTGGGAAGGC</td>
<td>1030</td>
</tr>
<tr>
<td>MEX8F</td>
<td>ACCAACAGCTTCACTGGGAC</td>
<td>1013</td>
</tr>
<tr>
<td>MEX9R</td>
<td>ATGAGCGACCGCAAGTGGGA</td>
<td>1168</td>
</tr>
<tr>
<td>MEX9F</td>
<td>TCCACCTTGTCCTGCTCTT</td>
<td>1149</td>
</tr>
<tr>
<td>MEX10R</td>
<td>CTGGCGATGTACACATTGAC</td>
<td>1319</td>
</tr>
<tr>
<td>1500F2</td>
<td>GCCCACCTTTGTCATGTGT</td>
<td>1289</td>
</tr>
<tr>
<td>1500R1</td>
<td>ATCTAAGTCACACAGCTGG</td>
<td>1544</td>
</tr>
</tbody>
</table>

Table 3.9 Primers used to amplify Cds2 introns
Listed are the primer name, sequences and the start site for the primers used to amplify the Cds2 introns. The primers were designed to the nucleotide sequence of 1247-c2.
Figure 3.29 Restriction map of 20A

Shown is the basic restriction map of clone 20A. The blots shown in figure 3.32 positioned the 250bp and 300bp HindIII fragments adjacent to the 6kb HindIII fragment. The blots were inconclusive for the third small HindIII fragment. Shown on the map in red are two possible positions of the 200bp HindIII fragment. The position of the EcoRI fragments were identified using the vector sites (shown in blue) as anchor points. The blot in figure 3.32 probed with the 200bp HindIII fragment implies the position of the 2kb and 800bp EcoRI sites and the positions are shown above. The size of the fragments are indicated on the map.
<table>
<thead>
<tr>
<th>Exon</th>
<th>EXON SIZE (bp)</th>
<th>5’ splice donor</th>
<th>Intron</th>
<th>3’ splice acceptor</th>
<th>Intron Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>176</td>
<td>GGAGGACAAGxxxxxxG....</td>
<td>1</td>
<td>...GAAATCAGAGTxxxxxx</td>
<td>1100</td>
</tr>
<tr>
<td>2a</td>
<td>139</td>
<td>CAAGGTGGGAacctgtagacacccca....</td>
<td>2a</td>
<td>gcccttcacactgtctcGAACTGGTGG</td>
<td>1200</td>
</tr>
<tr>
<td>2b</td>
<td>95</td>
<td>GATGATGATTTtttcttcattatcc...</td>
<td>2b</td>
<td>gatgcctctgtttttctcagGTATGTGTG</td>
<td>1800</td>
</tr>
<tr>
<td>3</td>
<td>98</td>
<td>CCCTCAGCTGxxxxxxG....</td>
<td>3</td>
<td>atgtcttctcgctttcagGTACTTCTCTG</td>
<td>225</td>
</tr>
<tr>
<td>4</td>
<td>140</td>
<td>TATCTGACAGgcagttacaggtcctc...</td>
<td>4</td>
<td>ttatttcttcacactagGATTCTGCAT</td>
<td>750</td>
</tr>
<tr>
<td>5</td>
<td>59</td>
<td>GGTCTACATGgtaagggagctgtccagt...</td>
<td>5</td>
<td>tcatttcgtctacatgctagTTTGCTGGA</td>
<td>620</td>
</tr>
<tr>
<td>6</td>
<td>83</td>
<td>GAATGATATGgtaagggagcgacagaagtttt...</td>
<td>6</td>
<td>actctttttttgcttcagTTTTATTGTT</td>
<td>820</td>
</tr>
<tr>
<td>7</td>
<td>88</td>
<td>ACTCATTAAGgtaagggagcgacccgctg...</td>
<td>7</td>
<td>atttgttcttcacatgctagCTCTCTCCA</td>
<td>1400</td>
</tr>
<tr>
<td>8</td>
<td>69</td>
<td>TGGCCTTTTGTgtaggtgtacgtcccat...</td>
<td>8</td>
<td>tctgtcttctcgctttagCTCTCTATG</td>
<td>620</td>
</tr>
<tr>
<td>9</td>
<td>153</td>
<td>CATTTGCTGGgtgctacccctccgtag...</td>
<td>9</td>
<td>...AAAACGGTGA....</td>
<td>1470</td>
</tr>
<tr>
<td>10</td>
<td>120</td>
<td>TAAAATCAAAGgtatgcgaactgctgccac...</td>
<td>10</td>
<td>...GATTTCCTCCA</td>
<td>400</td>
</tr>
<tr>
<td>11</td>
<td>104</td>
<td>GCTTTATCAGattatgtgaggcaactactct...</td>
<td>11</td>
<td>...GGTTCCAAAC</td>
<td>200</td>
</tr>
<tr>
<td>12</td>
<td>137</td>
<td>CACCTAACTGA.....3’ UTR</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.10 Known Cds2 splice junction information

The table lists the sizes of the exons and the introns of mouse Cds2 to date. The table also list some of the sequence generated around the exon splice sites. The information was obtained using PCR and sequencing. The primers used to generate this data are listed in table 3.9.
The remaining introns that could not be amplified from 20A were amplified using the primers listed in table 3.9 and 439-n24 as template. During the course of amplifying the remaining introns it was found that exon 2 had a 1kb insert in the centre, so the nomenclature of exon 2a and 2b was adopted. The size of the PCR fragments generated in these reactions determined the size of the introns in Cds2 (figure3.30). However the primer pair designed to amplify intron1 failed to amplify even after repeated attempts using varying condition within the PCR reaction. To identify if PAC clone 439-n24 contained exon1, a Southern blot was prepared from 439-n24 PAC DNA restriction digested with a combination of single and double enzymes including BamHI, EcoRI and HindIII (figure 3.31c). Two probes were generated from clone 1247-c2 (see section 3.2), these probes were generated in order to determine if exon 1 is contained in 439-n24 and to possibly deduce the size of intron 1. The first probe generated from 1247-c2 was prepared by restriction digest, using the enzymes EcoRI (artifical vector site) and BspEI (position 170bp, just 5' of exon 1/exon 2a boundary). This probe represented exon 1 (figure 3.31a). The second probe was generated by PCR, in this reaction exon 2a and 2b cDNA was amplified using primers Mex2F and Mex2R (table 3.9) (figure 3.31e). The results when the Southern blot was probed with exon 1 (figure 3.31b) showed that the probe hybridised to a 5kb EcoRI, EcoRI/BamHI, EcoRI/HindIII and triple digest fragments as well as an 8kb HindIII and HindIII/BamHI fragment. The filter was stripped before probing with the exon 2a/2b probe (figure 3.31d). This probe hybridises to a 17kb EcoRI, a 12kb HindIII, HindIII/EcoRI and a 7kb EcoRI/BamHI, BamHI and triple digest bands. These results indicated that clone 439-n24 contains both exon 1 and exon 2a/2b but they are not sufficiently close to be amplified using standard PCR and the results do not allow the intron size to be determined.

In an attempt to determine the size of intron 1 and thereby complete the structural information, the human PAC clone (GenBank accession number: AC01673), which was used to identify the intron splice sites was re-examined. The sizes of the human CDS2 introns are comparable to those identified for the mouse. Human CDS2 intron 1 is 43kb in length which is beyond the range of standard PCR amplification. The size of the mouse Cds2 intron 1 is yet to be determined.
Figure 3.30 Schematic representation of the genomic structure of mouse Cds2

Shown is a schematic representation of the mouse Cds2 genomic structure. The exons are shown in red and the exon and intron sizes are shown in blue. The sizes of the introns were determined by PCR using either clone 20A or PAC clone 439-n24 as template. The area of the gene contained within the 1.5 and 6kb HindIII fragments generated from clone 20A is indicated. The primers used to amplify the introns are listed in table 3.9. Full listings of the sizes of the introns and exons are given in table 3.10.
Figure 3.31 Southern blot confirming the presence of exon 1 in PAC clone 439-n24
The five panels represent (A) the 170bp exon1 probe generated by restriction digest using enzymes EcoRI and BspEI on clone 1247. (B) The autoradiograph of the bands to which the exon1 probe hybridised, 5kb EcoRI, EcoRI/BamHI, EcoRI/Hind III and the triple digest as well as the 8kb Hind III and Hind III/BamHI band. (C) The 439-n24 PAC ethidium stained digest gel. The enzymes used were, EcoRI (1), EcoRI/BamHI (2), BamHI (3), Hind III/BamHI (4), Hind III (5), Hind III/EcoRI (6) and Hind III/BamHI/EcoRI (7). The lanes were smeary due to possible bacterial contamination. (D) The autoradiograph of the bands to which the exon 2a/2b probe hybridised, a 17kb EcoRI, 12kb Hind III, Hind III/EcoRI and 7kb EcoRI/BamHI, BamHI and Hind III/BamHI/EcoRI fragments. (E) The gel photo of the 224bp probe representing Cds2 exon 2a/2b which was generated by PCR using primers Mex2F and Mex2R and clone 1247-c2 as template. This showed that although both exons are contained within 439-n24 PAC clone they are too distant to be amplified by PCR.
3.4.4 Summary of section 3.4.1-3.4.3
In summary the murine Cds2 gene has thirteen exons and genomically spans approximately 11kb excluding intron 1. The experiments performed in this section generated a phage clone and a PAC clone representing murine Cds2. The PAC clone was subsequently used to map the chromosomal location of the gene by fluorescent in situ hybridisation (FISH) in section 3.5. The information gathered about clone 20A was used in section 3.8 for the generation of a construct for the production of a knockout mouse.

3.4.5 A comparison of the gene structures of Cds1 and Cds2
In sections 3.1 through to section 3.4 the isolation and characterisation of the murine CDP-diacylglycerol synthase genes were reported. The isolation of the murine genes came about through database screening using CDS nucleotide sequences from other species. The first cds gene sequence to be used in this manner was the Drosophila cds sequences for the identification of the human CDS genes (Halford et al., 1998). The human sequences were then used to isolate the murine sequences. Since this project has commenced, several other types of CDS genes have become available through the EST database. The sequences include; rat Cds1 (NP_112521) and Cds2 (NP446095), human Cds1 (XP003308) and Cds2 (CAB89582), Drosophila cds (AAF50483), C.elegans cds (NP_501297), S. cerevisiae cds (NP009585) and S. pombe cds (CAB99396). Figure 3.32 shows a ClustralX alignment of the derived amino acid sequences of these genes and it can be seen clearly that certain regions of these different genes share very high levels of homology. Using the amino acids sequences shown in the alignments, a phylogenetic tree was generated by the neighbour joining method (Saitou., et al., 1986) (figure 3.33). The figure shows an evolutionary pattern that is expected with the tree rooted using the E.coli sequence.

An interesting point is that the Cds2 amino acid sequence shows marginally more identity to Drosophila cds (57%) than does mouse Cds1 (56%). This has previously also been reported to be the case for human CDS2 (Volta, et al., 1999). Although the expression patterns reported in section 3.6.3 and also by Volta et al. (1999) (discussed in section 4.3), suggests that Cds1 is more likely to function in a role similar to Drosophila eye-cds than Cds2.
Figure 3.32 Cds alignment

Shown is the alignment of mouse Cds1 amino acid sequence with cds sequences from \textit{S. pombe}, \textit{S. Cerevisiae}, \textit{C. elegans}, \textit{Drosophila}, rat Cds1, rat Cds2, mouse Cds2, human CDS1 and human CDS2. The red dots indicate residues identical to the mouse Cds1 sequence, blue dashes represent gaps made in the sequences to ensure the best alignment, and residues which are not identical to mouse Cds1 are represented by the appropriate letter. The alignment was produced using software ClustalW. As can be seen the sequences are highly homologous not only between the different species but also between the two genes.
Figure 3.33 Phylogenetic tree

Shown above is the phylogenetic tree derived from the amino acid sequences of rat Cds1 and Cds2, mouse Cds1 and Cds2, human CDS1 and CDS2, and *Drosophila, C. elegans, S. cerevisiae, S. pombe* and *E. coli* Cds. The tree is rooted using *E. coli*. The bootstrap values are shown and the distance of the lines corresponds to the number of substitutes in the sequences of each species relative to each other. The tree produces the expected pattern of evolution.
When the genomic structure of both mouse *Cds* genes are compared (figure 3.34) it is apparent that the exonic sizes of the genes are comparable whereas the intronic sizes vary greatly. It will be interesting to compare the gene structure again when intron 1 of *Cds2* has been identified and to analyse the sequences derived from the introns of both genes in an attempt to isolate regions which may represent promoters or enhancers of the genes. Even when all this data has been collected the true function role of the murine genes will only be elucidated by the generation of knockout mouse models for each gene. The preparation of a knockout construct for *Cds2* is described in section 3.8.
Figure 3.34 Comparison of genomic structure of *Cds1* and *Cds2*
Shown above are schematic representations of the *Cds1* and *Cds2* genes. They are drawn to the same scale, and it can be seen that *Cds2* is just slightly more than half the genomic size of *Cds1* (excluding intron 1). A proper comparison can only be made when intron 1 has been sized and the full lengths of the 5' and 3' UTRs of each gene determined.
3.5 Chromosomal Localisation of Cds1 and Cds2

3.5.1 Bioinformatic mapping of Cds1 and Cds2

The provisional mapping of mouse genes which have previously mapped human homologues, can be performed bioinformatically at the NCBI web site (see chapter 3.1 for web address). NCBI has a web page called the Human-mouse homology map, which compiles syntenic regions between human and mouse chromosomes. The maps are constructed by integrating homologues collected by the mouse genome database (Jackson Laboratory) with putative homologues identified by sequence homology. The sequence-based pairs are identified by a megaBLAST using non-EST mRNAs. Two megaBLAST analyses are performed, one in which human mRNA is the query and one in which the mouse mRNA is the query. Only pairs exhibiting reciprocal best megaBLAST scores are included in the analysis. Pairs for which identifying a reciprocal best hit is not ambiguous, or pairs identified in the megaBLAST that are not conflicting with the collected mouse homology information are reviewed. Human-mouse homologue pairs that meet the following criteria are used to calculate conserved synteny bins, (i) unambiguous location of the human mRNA on the human genome assembly, (ii) mouse homologues that are on the mouse genome database genetic map.

Conserved synteny bins are established using human sequence as the master, and looking for breaks in the mouse map. A conserved segment must include at least two loci from the same region of the mouse genome. Once the conserved synteny bins are established, cytogenetically mapped loci are integrated where possible. Pairs for which a human position can be determined, but for which there is no known mapping data are virtually mapped based on the homology bins. All these mapping assignments should be considered highly speculative in the absence of experimental data.

3.5.2 Previously mapped CDS genes

Both human CDS genes have been mapped using GeneBridge4 radiation hybrid mapping panel and fluorescence in situ hybridisation (FISH) to chromosomes 4q21 (CDS1) and to 20p13 (CDS2) (Halford et al., 1998; Volta et al., 1999). Volta et al. (1999) also described the mapping of mouse Cds2 to mouse chromosome 2. On examination of the syntenic maps, human chromosome 20 shares exclusive homology to mouse chromosome 2 (figure 3.35a). Human chromosome 4q21 shares homology
<table>
<thead>
<tr>
<th>Cytogen Position</th>
<th>Human symbol</th>
<th>Mouse Chr</th>
<th>Mouse symbol</th>
<th>cM Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>4q21</td>
<td>PRDM8</td>
<td>5</td>
<td>4930565F05Rik</td>
<td></td>
</tr>
<tr>
<td>4q21</td>
<td>FGF5</td>
<td>5</td>
<td>Fgf5</td>
<td>55</td>
</tr>
<tr>
<td>4q13.1-q21.1</td>
<td>PRKG2</td>
<td>5</td>
<td>Prkg2</td>
<td>53</td>
</tr>
<tr>
<td>4q13-q21</td>
<td>HNRPDLC</td>
<td>5</td>
<td>Hnrdcl</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>KIAA0905</td>
<td>5</td>
<td>1810024J13Rik</td>
<td></td>
</tr>
<tr>
<td>4q21.21-q21.23</td>
<td>MRPS18-1</td>
<td>5</td>
<td>1110037D14Rik</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>KIAA0993</td>
<td>5</td>
<td>2610509D04Rik</td>
<td></td>
</tr>
<tr>
<td>4q22.1-q23</td>
<td>MAPK10</td>
<td>5</td>
<td>Mapk10</td>
<td></td>
</tr>
<tr>
<td>4q21.3</td>
<td>PTPN13</td>
<td>5</td>
<td>Ptpn13</td>
<td>56</td>
</tr>
<tr>
<td>4q21</td>
<td>MLLT2</td>
<td>5</td>
<td>Mllt2h</td>
<td>56</td>
</tr>
<tr>
<td>4q21.3</td>
<td>DSPP</td>
<td>5</td>
<td>Dsppl</td>
<td>56</td>
</tr>
<tr>
<td>4q21</td>
<td>DMP1</td>
<td>5</td>
<td>Dmp1</td>
<td>56</td>
</tr>
<tr>
<td>4q21-q25</td>
<td>IBSP</td>
<td>5</td>
<td>Ibsp</td>
<td>56</td>
</tr>
<tr>
<td>4q21-q25</td>
<td>SPP1</td>
<td>5</td>
<td>Spp1</td>
<td>56</td>
</tr>
<tr>
<td>4q21-q23</td>
<td>PKD2</td>
<td>5</td>
<td>Pkd2</td>
<td>55</td>
</tr>
<tr>
<td>4q22</td>
<td>ABCG2</td>
<td>6</td>
<td>Abcg2</td>
<td>28.5</td>
</tr>
<tr>
<td>4q21</td>
<td>HERC3</td>
<td>6</td>
<td>5730409F18Rik</td>
<td></td>
</tr>
<tr>
<td>4q21</td>
<td>SNCA</td>
<td>6</td>
<td>Snca</td>
<td>29</td>
</tr>
<tr>
<td>4q21.3</td>
<td>HPSE *</td>
<td></td>
<td>Hpa</td>
<td></td>
</tr>
<tr>
<td>4q21</td>
<td>IGJ *</td>
<td></td>
<td>Igj</td>
<td></td>
</tr>
<tr>
<td>4q21</td>
<td>AMBN *</td>
<td>5</td>
<td>Ambn</td>
<td>46</td>
</tr>
<tr>
<td>4q21</td>
<td>SCYB1O *</td>
<td>5</td>
<td>Adh5</td>
<td></td>
</tr>
<tr>
<td>4q21-q25</td>
<td>ADH5</td>
<td>3</td>
<td>1700029C17Rik</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>KIAA0994</td>
<td>3</td>
<td>Eif4e</td>
<td>25</td>
</tr>
<tr>
<td>4q21-q25</td>
<td>EIF4E</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4q21-q24</td>
<td>PPP3CA</td>
<td>3</td>
<td>Ppp3ca</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.35a Schematic representation of human and mouse chromosome synteny**

Shown is a portion of the human chromosome 4q21 map displayed on the NCBI web site synteny with regions from mouse chromosomes. As can be seen human chromosome 4q21 has syntenic regions on mouse chromosomes 5, 6, 3, and 12. Some of the known gene positions are shown on this map with their distance in centimorgans (cM).
<table>
<thead>
<tr>
<th>Cytogenic pos</th>
<th>Human symbol</th>
<th>Mouse Chr</th>
<th>Mouse symbol</th>
<th>cM position</th>
</tr>
</thead>
<tbody>
<tr>
<td>20p13</td>
<td>SOX22</td>
<td></td>
<td>Sox12</td>
<td></td>
</tr>
<tr>
<td>20p12.2-20p13XAP4</td>
<td></td>
<td></td>
<td>Ubce7ip3-pending</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>CSNK2A1</td>
<td></td>
<td>Csnk2a1</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>TCF15</td>
<td></td>
<td>Tcf15</td>
<td>83.9</td>
</tr>
<tr>
<td>20p13</td>
<td>MGC4675</td>
<td>2</td>
<td>Fkbp1a</td>
<td></td>
</tr>
<tr>
<td>20q11.2</td>
<td>TGM3</td>
<td>2</td>
<td>Pdyn</td>
<td>73.3</td>
</tr>
<tr>
<td>20p13</td>
<td>SNRPB</td>
<td>2</td>
<td>Snrpb</td>
<td>73</td>
</tr>
<tr>
<td>20p13</td>
<td>FKBP1A</td>
<td></td>
<td>FKBP1A</td>
<td></td>
</tr>
<tr>
<td>20p12-p12</td>
<td>PDYN</td>
<td></td>
<td>PDYN</td>
<td></td>
</tr>
<tr>
<td>20p12.3-p13</td>
<td>CPX-1</td>
<td></td>
<td>CPX-1</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>OXT</td>
<td>2</td>
<td>Oxt</td>
<td>73.5</td>
</tr>
<tr>
<td>20p13</td>
<td>AVP</td>
<td>2</td>
<td>Avp</td>
<td>73.2</td>
</tr>
<tr>
<td>20p13</td>
<td>MGC2592</td>
<td>2</td>
<td>MGC2592</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>ITPA</td>
<td>2</td>
<td>ITPA</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>ATRN</td>
<td>2</td>
<td>ATRN</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>SN</td>
<td>2</td>
<td>SN</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>CENPB</td>
<td>2</td>
<td>CENPB</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>CDC25B</td>
<td>2</td>
<td>CDC25B</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>FLJ11729</td>
<td>2</td>
<td>FLJ11729</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>ADRA1D</td>
<td>2</td>
<td>ADRA1D</td>
<td></td>
</tr>
<tr>
<td>20pter-p12</td>
<td>PCNA</td>
<td>2</td>
<td>PCNA</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>PTPNS1 *</td>
<td>2</td>
<td>PTPNS1</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>PTPRA *</td>
<td>2</td>
<td>PTPRA</td>
<td></td>
</tr>
<tr>
<td>20p13</td>
<td>ANGPT4 *</td>
<td>2</td>
<td>ANGPT4</td>
<td></td>
</tr>
<tr>
<td>20q13</td>
<td>KIAA1434</td>
<td>2</td>
<td>KIAA1434</td>
<td></td>
</tr>
<tr>
<td>20q13</td>
<td>VAPB *</td>
<td></td>
<td>VAPB</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.35b Schematic representation of human and mouse chromosome synteny**

Shown is a portion of the human chromosome 20p13 map displayed on the NCBI web site synteny with regions from mouse chromosomes. As can be seen human chromosome 20p13 has syntenic regions only on mouse chromosomes. Some of the known gene positions are shown on this map with their distance in centimorgans (cM).
with a variety of mouse chromosomes, predominately regions of mouse chromosome 5 but also regions from chromosomes 3, 6 and 12 (figure 3.35b).

3.5.3 Fluorescence in situ Hybridisation (FISH) mapping of Cds1 and Cds2
The identification of the chromosomal locations of both mouse Cds genes were carried out using FISH mapping. The FISH mapping was performed in the MRC Cytogentic unit at Harwell by Ruby Banerjee (see section 2.3.3.2).

The mouse PAC clones (see chapters 3.3 and 3.4 for PAC information) were used as probes to identify the genes. PAC clone 572-a24, representing Cds1, was labelled by Nick translation with biotin-16dUTP and used to probe a metaphase spread of mouse chromosomes, hybridisation was detected with Avidin Texas Red fluorescence (figure 3.36). The mouse Cds1 gene was localised to the mouse chromosome 5 band E3. Figure 3.37 illustrates the position of the gene on four individual metaphase spreads, all pairs show hybridisation to 5 E3, this is then represented on a schematic diagram of mouse chromosome 5.

Mouse Cds2 was mapped in the same way using PAC clone 439-n24 labelled by nick translation with biotin 16dUTP and detected with avidin FITC fluorescence. This experiment was performed to verify the results reported in Volta et al. (1999). Cds2 was assigned to mouse chromosome 2 band G1. Figure 3.38 illustrates the metaphase spread with the gene location highlighted by green fluorescence. Figure 3.39 shows the position of the gene on a schematic diagram of mouse chromosome 2 G1.

3.5.4 Summary of section 3.5
The results from this section show the mapping of mouse Cds1 to chromosome 5 E3 and mouse Cds2 to chromosome 2 G1. These chromosomal results were predicted bioinformatically and Cds2 mapping confirms results by Volta et al. (1999). The FISH mapping in this section confirms a previous result and positioned the genes to a more precise location on the mouse chromosomes.
Figure 3.36 Metaphase spread probed with Cds1 PAC clone
This fluorescent in situ hybridisation panel shows the chromosomal location of mouse Cds1 on a metaphase spread. The metaphases were probed with Pac clone 572-n24 labelled with Flurophore Avidin Texas Red. Mouse Cds1 mapping to chromosome 5E1 is indicated by the white arrows.
Figure 3.37 Schematic representation of the position of Cds1 on mouse chromosome 5
Shown is a schematic diagram of mouse chromosome 5. Also shown are four individual metaphase spreads each show hybridisation to chromosome 5 band E3.
Figure 3.38 Metaphase spread probed with Cds2 PAC clone
This fluorescent in situ hybridisation panel shows the chromosomal location of mouse Cds2. The chromosomes were probed with PAC clone 439-n24 labelled with the fluorophore FITC. Mouse Cds2 maps to mouse chromosome 2G1 as indicated by the white arrows.
Figure 3.39 Schematic representation of the position of $Cds2$ on the mouse chromosome 2
Shown is a schematic diagram of mouse chromosome 2. Also shown are four individual metaphase spreads each show hybridisation to chromosome 2 band G1.
3.6 Expression studies of Cds1 and Cds2

3.6.1 Mouse RNA MasterBlot™

A mouse multiple tissue Masterblot™ (Clontech) was used to determine general sites of Cds1 and Cds2 gene expression. This commercially available blot contains polyA+ RNA samples from various mouse tissues. Figure 3.40b shows the key to the gridded filter, which indicates the positions of the various tissues and control samples. In addition to the various adult tissues, the blot has four embryonic tissue samples from 7 days post coitus (d.p.c) to 17 d.p.c. cDNA probes were generated (see section 2.3.1.20) using the entire inserts from the Cds1 clone 801-b9, (see section 3.1 for clone 801 information), and the Cds2 clone 1247-c2 (see section 3.2 for clone 1247 information). The probes were randomly radio-labelled with 32P α-dCTP. The radio-labelled probes were used in separate hybridisation reactions with the blot as per the manufacturer’s instructions, the blots were washed to 0.1xSSC, 0.1% SDS before exposing the probed filters to autoradiographic film for 48 hours at -80°C. After the first probe was used the blot was stripped (as per the manufacturer’s instructions), then hybridised with the second probe under the same conditions.

The expression patterns of the genes can be seen in figures 3.40a and 3.40c. The expression of Cds1 is limited to the adult brain, eye, smooth muscle and testis samples, with very weak expression in the kidney and no expression observed in any of the embryonic tissue samples (figure 3.40a). The expression pattern observed when the blot was hybridised with the Cds2 probe was very different to the pattern observed with the Cds1 probe. When the blot was hybridised with the Cds2 probe (figure 3.40c) every adult tissue on the blot showed expression (refer to the key in figure 3.40b for the list of tissues present on the blot). Expression of the gene in the brain sample showed the highest intensity. In addition to expression observed in all adult tissues, Cds2 expression was observed in all embryonic stages.

3.6.1.1 Discussion points of section 3.6.1

There are several negative controls on the blot for the purpose of reducing the level of non-specific background, which is crucial for achieving high sensitivity. The negative controls include, yeast total RNA, yeast tRNA, E.coli rRNA, E.coli DNA (for the detection of non-specific hybridisation), poly r(A) (detection of oligo (dT) sequences),
Figure 3.40 RNA Masterblots probed with *Cds1* and *Cds2*

The blots shown above are RNA Masterblots (Clontech) hybridised as with the whole inserts from the cDNA clones 801 (*Cds1*) and 1247 (*Cds2*). As can be seen in blot (a), probed with 801, the areas of expression are limited to the eye, brain, smooth muscle and testis, with weak expression in the kidney sample (barely visible in photograph but was apparent on autoradiograph film). No expression was observed in the embryonic tissues. Blot (c) was probed with 1247, the probe hybridised to every sample on the filter, with the highest level of expression observed in the brain sample. Expression could be seen from the earliest embryonic sample, implying that the gene may be developmentally important.
C₀t-1 DNA (detection of repetitive sequences) and two mouse genomic DNA controls (detection of repetitive sequences, if the gene is highly abundant or if the gene belongs to a multi-gene family).

Neither gene probe hybridised with the yeast, *E.coli* rRNA, poly r(A) or C₀t-1 DNA controls. Both probes hybridised to the *E.coli* DNA control sample. The manufacturer’s notes on this control state: ‘In some cases, you may observe a signal for *E.coli* DNA, but not the other three controls. This may be explained by the presence of *E.coli* sequences that are homologous to mammalian mRNA. If your probe hybridises to the dot containing *E.coli* DNA, but not yeast total RNA, yeast tRNA, *E.coli* rRNA, you should verify that your gene has homology to bacterial sequences.’ In section 1.2.4.5 the *cds* gene was described, the first *cds* gene identified was the *E.coli cds* gene (Langley and Kennedy, 1978). This finding will be discussed further in chapter 4.

The probes also hybridised to a second control on the panel, the mouse genomic DNA control. This control is used to detect repetitive sequences, highly abundant genes or genes which belong to multi-gene families. The detection of repetitive sequences could be eliminated as a reason for hybridisation since neither probe hybridised with the C₀t-1 DNA sample which is also present on the blot for the detection of repetitive sequences. The hybridisation with the genomic DNA control could be for either one of the other two reasons given above. The *Cds2* gene was shown on the blot to be present in every tissue and therefore occurs abundantly, and *Cds1* and *Cds2* form a family of genes of which other members have yet to be identified. The final reason may again be genomic DNA contamination in the probes from the laboratory environment. However, even with the hybridisation to two of the control samples, the pattern of expression of the genes on the blot is still significant.

The different patterns of expression suggest that *Cds1* and *Cds2* may have different functional roles. When the blot was hybridised with the *Cds2* probe, all adult tissues and embryonic tissues showed expression. This result suggests that the *Cds2* gene is involved in mouse development, and may be a ubiquitously expressed gene encoding a protein used in general signalling pathways requiring the regeneration of PIP₂ to occur. The expression pattern observed when the blot was hybridised with the *Cds1* probe, is limited to adult eye, brain, smooth muscle, kidney and testis tissues, which suggests that
this gene may have a specific function in certain tissues of the adult. The results will be discussed further in chapter 4.

### 3.6.2 Northern blot analysis

#### 3.6.2.1 Preparation of Northern blot filter

None of the cDNA clones identified for the two *Cds* genes are full-length since neither contains a poly-adenylation signal and the start codon was also absent in the cDNA clones identified for *Cds1*. In order to determine the full transcript length of *Cds1* and *Cds2*, a Northern blot was prepared (see section 2.3.2.4). Total RNA was prepared (see section 2.3.2.1) from the following adult mouse tissues; eye, brain, liver, spleen and kidney. $A_{260}/A_{280}$ ratio calculations were performed on the samples to assess the concentration and quality of the RNA where a ratio of 1.8 is the ideal. The rather lower ratios shown in table 3.11 indicate that some protein was present in the samples. The liver sample shows the least purity of the samples and the effects of this can be seen in the blot figures where there is the identifiable presence of high molecular weight proteins in the samples. 10μg of RNA from each tissue was used in the preparation of the blot. The agarose gel photo of the RNA samples is shown in figure 3.41. The blot was hybridised at 42°C (see section 2.3.1.21) with $^{32}$P α-dCTP radiolabelled probes generated using the whole insert from cDNA clones 801-b9 (*Cds1*) and 1247-c2 (*Cds2*) (see section 2.3.1.20).

<table>
<thead>
<tr>
<th>Tissue</th>
<th>$A_{260}$</th>
<th>$A_{280}$</th>
<th>Concentration (μg/μl)</th>
<th>$A_{260}/A_{280}$ ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain</td>
<td>0.714</td>
<td>0.488</td>
<td>2.86</td>
<td>1.5</td>
</tr>
<tr>
<td>Eye</td>
<td>0.082</td>
<td>0.55</td>
<td>0.328</td>
<td>1.5</td>
</tr>
<tr>
<td>Spleen</td>
<td>1.46</td>
<td>0.977</td>
<td>5.84</td>
<td>1.5</td>
</tr>
<tr>
<td>Kidney</td>
<td>0.676</td>
<td>0.460</td>
<td>2.7</td>
<td>1.4</td>
</tr>
<tr>
<td>Liver</td>
<td>0.447</td>
<td>0.312</td>
<td>8.94</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**Table 3.11** Shown are the $A_{260}/A_{280}$ ratios of the RNA samples generated from various mouse tissues.
Figure 3.41 Total RNA gel
Shown is the ethidium bromide stained gel photograph of the mouse RNA samples prior to Northern blotting. ~10μg each of eye, brain, liver, spleen and kidney tissues were used. The liver sample appears to be slightly degraded and showed least purity in the A260/A280 ratios (table 3.11). The ribosomal RNA bands at 3.9kb and 1.9kb are indicated.
3.6.2.2 Northern blot probed with \textit{Cds1}

The results of the Northern blot hybridised with the \textit{Cds1} probe are shown in figure 3.42. The blot shows that the \textit{Cds1} probe hybridises to two fragments, one \(~4.1\text{kb}\) and one \(~2\text{kb}\). These fragment sizes are very close to the sizes of the ribosomal bands, but on close inspection, the position of the bands in this blot are not identical to those which would be present if cross-hybridisation to ribosomal RNA had occurred and these bands are not present in all lanes. As far as can be determined therefore, the probe is not hybridising to ribosomal RNA. This result confirms what was previously known, that the cDNA clone 801-b9 (1.3kb in length) which represents the 3' end of \textit{Cds1} does not contain the full-length transcript of the gene. In section 3.1, the coding sequence of the \textit{Cds1} gene along with sequence from the EST database representing the 3'UTR of mouse \textit{Cds1} was compiled to form a 3.7kb transcript (figure 3.12) minus sequence information for the 5'UTR. This sequence information suggests that the larger 4.1kb fragment on the Northern blot represents the transcript size hence implying that the 5' UTR, which to date is not known, is likely to be approximately 400bp in length. The \textit{Cds1} Northern blot also confirms results from the Masterblot, that \textit{Cds1} is expressed in the eye, brain and kidney.

3.6.2.3 Northern blot probed with \textit{Cds2}

The results of the Northern blot hybridised with the \textit{Cds2} probe is shown in figure 3.43. In all lanes two fragments of \(~2.1\text{kb}\) and \(~4\text{kb}\) were observed. Ribosomal RNA cross-hybridisation was eliminated since, as before, on close inspection of the blots the ribosomal and \textit{Cds1} bands differ in size. In the lanes containing spleen and kidney RNA a third band of approximately 2.9kb was observed. In the brain and eye RNA sample lanes a band of \(~6\text{kb}\) was also identified; this band seems too large to be the transcript size and showed weaker intensity levels of hybridisation than the other two smaller bands on the blot, indicating it may represent cross-hybridisation to another transcript. These results confirm that the cDNA clone representing \textit{Cds2} (1.6kb in length) does not contain the full gene transcript. It is most likely that the 4kb band represents the full transcript length and that the 2.9kb band in the spleen and kidney samples is a splice variant isoform of the gene that has a specific role in certain tissues. The possible role of the splice variants will be discussed in chapter 4.
Figure 3.42 Cds1 probed Northern blot
Shown is the autoradiograph result of the Northern blot filter probed with the 801-b9 (Cds1) probe. The probe hybridises to two bands, one approximately 4.1kb and one approximately 2.1kb. The sequence reported in chapter 3.3 implies that the size of the transcript must be represented by the upper 4.1kb band.
Figure 3.43 Cds2 probed Northern blot
Shown is the Northern blot filter (from figure 3.6.3.1) probed with the 1247-c2 (Cds2) probe. The probe hybridises to three bands, one approximately 6kb, one approximately 4.1kb and one approximately 2kb in the eye and the brain but hybridises to a third band in the spleen and kidney of approximately 2.9kb. It is possible that the lower band is the transcript size and that a splice variant occurs in other tissues to give the 2.9kb band.
3.6.2.4 Discussion points of section 3.6.2
Although the two genes produce bands of relatively similar sizes, cross-hybridisation between the genes can be eliminated because the Cds1 probe does not hybridise to all samples on the blot. In addition, the same probes were used on the Northern blots as were used on the RNA MasterBlots™ (figure 3.40) where no cross-hybridisation occurred.

The Northern blots were washed at a much higher stringency than the Masterblots (0.01XSSC, 0.1% SDS compared to 0.1XSSC, 0.1%SDS) since there was much higher background on the Northern blots, this higher stringency has also reduced the likelihood of cross-hybridisation of genes occurring. Absence of cross-hybridisation was also noted in the in situ hybridisation studies (see section 3.6.3) where the genes displayed very different expression patterns. A full discussion on the results generated in this section can be found in chapter 4.

3.6.3 In situ hybridisation studies

3.6.3.1 Production of riboprobes
To identify specific sites of Cds1 and Cds2 expression in the adult eye, the specific pattern of transcription was established by in situ hybridisation (see section 2.3.3.1) of gene specific digoxigenin-labelled riboprobes to 5μm thick sections of whole mouse eyes. Sense and anti-sense riboprobes were generated (see section 2.3.2.5). The sense probe shares the same sequence as the transcript and should not bind to the gene transcripts in the sections. The sense probes are used as the negative control in the experiments. The anti-sense probe is the complementary sequence to the gene transcript and should therefore hybridise to the gene transcript. Six riboprobes were generated in total. A sense and an anti-sense probe were generated for each of the following genes, Cds1 (cDNA clone 801-b19 was used as template), Cds2 (cDNA clone 1247-c2 was used as template) and human rod opsin. None of the sections were counter-stained after probing as this increases the chance of masking faint positive signals.

3.6.3.2 Location of gene transcripts
All sections used were whole adult eyes, the entire structure was examined for areas of positivity. Only regions displaying hybridisation to specific gene transcripts are shown from each section. All figures in this section show a Haematoxylin and Eosin stained section presenting the region of the eye being examined to define the various layers.
To test the methodology the rod opsin probes were used as positive controls. The rod opsin gene transcript is known to be expressed in the outer nuclear layer of the retina. Figure 3.44 shows the results from the control experiments. The anti-sense probe detects expression in the outer nuclear layer of the retina with all other layers of the retina being negative.

The results from using the Cds1 riboprobes can be seen in figure 3.45. Expression was only observed in the retina. On probing with the anti-sense probe, expression of the Cds1 gene transcript is observed only in the inner segments of the photoreceptor cells.

In contrast to the previous probes used, Cds2 gene transcript expression is not restricted to the retina. The Cds2 anti-sense probe detected expression in the cornea (figure 3.46) and cilary body (figure 3.47) as well as in the retina (figure 3.48). When the sections were probed with the Cds2 anti-sense probe, expression was observed in the epithelial and endothelial layers only with no expression detected in the stroma of the cornea (figure 3.46). Expression was observed in the cilary body (figure 3.47) and in the ganglion cell layer and the inner nuclear layer of the retina (figure 3.48).

The masterblot (figure 3.40) showed that Cds2 is expressed in all tissues present on the blot and that Cds1 is expressed only in the brain, eye and smooth muscle. To further examine these results an in situ hybridisation study was performed using a panel of mouse tissues. The tissues that were used in the panel are listed in table 3.12.

Using in situ hybridisation on paraformaldehyde fixed, paraffin embedded sections, a high level of Cds2 expression was found in various areas of the brain. This confirms the masterblot results for the brain since this tissue showed the highest level of expression on the blot. A very distinctive pattern was observed in the brain tissue with expression in the Purkinje cells of the cerebellum, Ammon’s horn of the hippocampus, the temporal cortex except for layer 1, the anterior hypothalamus, and the lateral and medulary reticular nuclei of the spinal cord (figures 3.49 A-E).

In addition to expression in the brain, Cds2 expression was seen in the sebaceous glands and epithelium of ear skin, the pancreas (3.50 A and B), the spleen, the chronic inflammatory cells of the core villi of the small intestine (3.51 A and B), the semi-
Figure 3.44 *Rod ops in* transcript expression in the adult mouse retina

All sections used for *in situ* hybridisation are adult mouse whole eye sections which have been formaldehyde fixed and paraffin embedded. Section (a) shows a Haematoxylin and Eosin stained section, section (b) was probed with the anti-sense positive control which was human rod opsin and section (c) shows a opsin sense probed section. The red arrow highlights the outer nuclear layer of the retina where immunoreactivity can be seen. All other layers of the retina are negative.
Figure 3.45 Cds1 expression in mouse retinal sections

Mouse adult retina sections stained with Haematoxylin and Eosin (a), a Cds1 anti-sense probed section (the lower region of the section has folded over) (b) and a Cds1 sense probed section (c). As demonstrated by the red arrow, immunoreactivity can be seen only in the inner segments of the photoreceptors.
Figure 3.46 Cds2 expression in adult mouse cornea
This figure shows the cornea stained with Haematoxylin and Eosin (a), probed with the anti-sense probe (b) and probed with the sense probe for Cds2 (c). As the red arrows indicate, areas of immunoreactivity can be seen in endothelium and the epithelium, but the stroma is completely negative.
Figure 3.47 Cds2 expression in the ciliary body

Section (a) shows an Haematoxylin and Eosin stained section of the ciliary body of an adult mouse eye. Section (b) shows immunoreactivity in the ciliary body (indicated with red arrows) when probed with the Cds2 anti-sense riboprobe. Section (c) was probed with the sense riboprobe and shows no positivity.
Figure 3.48 Cds2 expression in mouse retinal sections
Mouse adult retinal sections. Stained with Haematoxylin and Eosin (a), a Cds2 anti-sense probed section (b) and a Cds2 sense probed section (c). The anti-sense probed section shows expression in the ganglion cell layer and the inner nuclear layer of the retina. These positive regions are highlighted with red arrows.
Figure 3.49 *Cds2* expression in neural tissues
A panel of tissues were examined using sense and anti-sense *Cds2* probes. Various regions of positivity were found in the brain. Shown above are these positive regions along with H & E stained equivalent sections. Positivity was found in the cortex (A), cerebellum (B), hippocampus (C), the hypothalamus (D) and the spinal cord (E).
Figure 3.50 *Cds2* expression in ear skin and pancreas.

*Cds2* probed tissues showed positive expression in the skin (A) and the pancreas (B). This figure show the antisense probed tissue, a H & E stained equivalent section and the sense probed tissues. Positivity can be found in the epithelium (red arrow) and sebaceous glands (green arrow) of the skin will a general expression is seen in the pancreas.
Figure 3.51 *Cds2* expression in the spleen and small intestine. 
*Cds2* positivity was detected in the spleen (A) and small intestine (B). This figure shows the antisense probed sections, an H & E equivalent and the negative sense probed section.
nipherous tubular epithelium of the testis, and the lymph nodes (figure 3.52 A and B). However the in situ hybridisation studies did not show expression in the lung, liver, heart, kidney, thymus or muscle as they did on the blot. This may be due to lower levels of Cds2 mRNA on the tissue sections compared with the concentrated dot on the blot membrane. This interpretation is consistent with the observation that some of the samples that do not give expression in the tissue panel are expressed at a much lower level on the blot.

The role of Cds2 is still elusive, but the pattern of expression in the non-neural tissues may suggest a role in secretory cells. However the best way to define the function of the Cds2 gene is to generate a knockout model (see discussion).

The same panel of tissues was probed with Cds1. Like the masterblot, the results were more specific than the pattern seen in Cds2. Positivity was only seen in the cortex (except layer 1), the spinal cord, the lymph nodes and the testis (figure 3.53 A-D). The role of Cds1 is still unknown, but a knockout model would again be informative. However from the masterblot and from the in situ experiments, it appears that the role of Cds1 is more restricted than that of Cds2. A list of all tissues tested appears in table 3.12

The most likely explanation for the discrepancies between the results of the in situ hybridisation experiments and that of the masterblot is that there is a greater concentration of RNA on the blot than found in any given section of a tissue, hence weak signals may not have been detected using in situ hybridisation. Another explanation may be that Cds2 is not constantly expressed in all tissues and was not expressed in some of the tissues at the time of death of the animal used for the in situ hybridisation experiments. The conclusion that can be drawn from these experiments is that Cds2 is more widely expressed than Cds1, confirming the results of the masterblot.
Figure 3.52 Cds2 expression in testis and lymph nodes
The final areas of Cds2 positivity were the testis (A) and the lymph nodes (B). Shown in this figure are the antisense probed tissues, the equivalent H & E stained section and the negative sense probed tissues.
Figure 3.53 Cds1 expression in the brain, lymph nodes and testis
A panel of tissues were probed with Cds1 sense and anti-sense probes, only the cortex (A), the spinal cord (B), the lymph nodes (C) and the testis (D) produced positive results, these are accompanied above by H & E views of the positive tissues.
Table 3.12 Presence (+) and absence (-) of Cds1 and Cds2 expression in a range of tissues.

3.6.3.3 Discussion points of section 3.6.3
The results from this section show that the transcripts from \textit{Cds1} and \textit{Cds2} are expressed in very different discrete locations within the eye. \textit{Cds1} is confined to the retina whereas \textit{Cds2} is found in various ocular structures. Further discussion of the location of these transcripts can be found in chapter 4.

3.6.4 Rodless, coneless and rodless coneless mice
\textit{rd} or \textit{retinal degeneration} mice were first identified in the wild in Switzerland, England and France and subsequently in several laboratory strains (Pittler and Baehr., 1991). Mice homozygous for the \textit{rd} mutation display hereditary retinal degeneration and the classic \textit{rd} lines have subsequently served as models for human retinitis pigmentosa. In the affected retinas the rod photoreceptor cells begin to degenerate at postnatal day 8 and by four weeks there are no rod cells present. The degeneration is preceded by accumulation of cyclic GMP (Carter-Dawson \textit{et al.}, 1978, Bowes., \textit{et al} 1990) in the retina and is correlated with the deficient activity of the rod photoreceptor cGMP-phosphodiesterase. The degeneration of the rod cells is followed at a slower rate by the loss of cone cells. However, there are still 1.5-5% of the original population of cone cells present in 18 month old \textit{rd} mice (Carter-Dawson \textit{et al.}, 1978). For the analysis of
the function of the retina after the removal of cone cell function, Wang et al. (1992) generated a transgenic mouse (cl) by introducing of a construct consisting of a portion of the human red cone opsin promoter attached to an attenuated diphtheria toxin gene. These mice have the normal number of rod cells but lose all green cone cells and greater than 95% of UV cone cells. A study by Lucas et al. (1999) generated a strain of rodless, coneless mice by mating cl mice with homozygous rd mice. This generated a strain of mice lacking all photoreceptor cells.

With the evidence that the Cds1 transcripts were present in the photoreceptor layer of the retina, further investigations were carried out to establish if the transcripts were present in rod cells or cone cells or both. Initially this was investigated by RT-PCR using the primers Cds1 intron8F and Cds1 intron9R and Mex8F and Mex10R (as controls) with retinal cDNA templates from rd/rd, cl and rd/rd cl mice (kindly gifted by Professor Russell Foster, Imperial College London). Since the photoreceptor layer of the mouse retina is composed predominately of rod cells, the rd/rd retina is very degenerate due to the loss of these cells (figure 3.55). However it is possible that there are residual cone cells present in the rd/rd retina. The cl mouse retina is not as degenerated, since the predominately rod scaffolding structure of the photoreceptor layer is in place to maintain the retina. The rd/rd cl mouse retina does not contain any photoreceptor cells and is as degenerate as the rd/rd mouse. The results show that the Cds1 transcripts were only present in the cl cDNA, indicating that Cds1 is only expressed in rod cells (figure 3.54).

Unfortunately tissue was not available from cl or rd/rd cl mice since the mutant mouse lines have not been maintained by Professor Foster and sourcing the tissue from other laboratories was not possible. The in situ study was restricted therefore to rd/rd retinal sections (figure 3.55). Since the mutant mice are pigmented, age and sex matched pigmented wild-type mice were used as controls for this experiment (figure 3.55). The in situ experiments show that no expression can be detected in the cone cells of the rd/rd retina. This result confirms the RT-PCR result shown in figure 3.54.
Figure 3.54 The RT-PCR results shown confirm that Cds1 is only expressed in the rod photoreceptors. Lanes 1-6 Cds1, 1 – marker, 2- water, 3- WT, 4- rodless, 5 – coneless, 6- rodless/coneless. Yellow arrow indicates the Cds1 band of 300bp. Lanes 7-12 Cds2 following the same order. Green arrow indicates the Cds2 band of 250bp.

In conclusion therefore, the expression of Cds1 only in rod photoreceptors may imply that Cds1 has a role in phototransduction in dim light conditions, but this finding would have to be confirmed by the generation of a knockout mouse model for Cds1. Nevertheless, the results from this section has strengthened the case that a mutation in Cds1 may be responsible for some forms of RP and this should be investigated further using panels of RP patients where the causative gene is still unknown (see discussion).
**Figure 3.55 Comparison of *Cds1* transcript expression in wild type and *Rodless* retinas**

Section A. was a wildtype mouse adult retina probed with the anti-sense *Cds1* probe, section B. shows the sense probed wildtype retina C. shows a haematoxylin and eosin stained section. D. was a *Rodless* adult retina probed with the anti-sense *Cds1* probe, section E. shows and enlargement (40x) of the area of interest and section F. shows a haematoxylin and eosin stained *Rodless* section.
3.6.5 Sub-cellular localisation of Cds2

To identify the sub-cellular localisation of a protein in mammalian cells in the absence of a specific antibody, epitope tags can be used. This technique allows a small (approximately 10 amino acids) tag to be cloned in frame with the gene of interest. The tagged gene is then expressed and can be detected using an antibody to the tag. At the time this study was performed the whole coding sequence of Cds1 had not been deduced so only Cds2 was tagged. Cds2 protein was tagged with the epitope c-myc. C-myc tag is a ten amino acid (EQKLISEEDL) synthetic peptide designed against residues 408-439 of the human c-myc gene product. The Cds2-myc tagged construct was cloned into a mammalian expression vector and transfected into mammalian human embryonic kidney (HEK) 293T cells in culture. Gene expression was evaluated by immunohistochemistry and Western blotting.

3.6.5.1 Production of the Cds2-myc tagged construct

The Cds2-myc-tagged construct was amplified by PCR (see section 2.3.1.16). In order to amplify the construct, four primers were designed (see table 3.13) to the coding sequence of Cds2 deduced from the cDNA clone 1247-c2 (see section 3.2). Two of the primers were designed to the exact sequence of the gene. The primer MCds2For was designed at the start codon and the primer MCds2Rev was designed to the stop codon. The second set of primers, were designed to the same sequence but had the myc-tag nucleotide sequence added to the end of the Cds2 gene sequence. The primers were designed in this way to allow two constructs to be generated, one with the myc-tag at the N-terminus of the gene and the other with the tag at the C-terminus. The reason for generating two constructs with tags at either terminus is that some epitope tags can interfere with the folding and transportation of the protein. With tags at either terminus the results of the experiments can be compared to ensure that the presence of the tag has no effect on the localisation of the proteins.

10ng of clone 1247-c2 cDNA was used as template in the PCR reactions (see section 2.3.1.16), with the non-tagged primers MCds2For and Mcds2Rev to amplify the coding region of Cds2. This product was diluted and used as template for the next round of amplification. Two reactions were performed with primers MCds2F and MCds2C-myc in one reaction and MCds2Rev and MCds2N-myc in the other.
The products with a tag at either the C-terminal or N-terminal were subcloned into pGEMT-easy vector (Promega) (see section 2.3.1.13-14) and maxi prep stocks were generated (see section 2.3.1.2).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5'[\text{\textit{\textbf{I}}}]3'</th>
<th>Primer Start position</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCds2For</td>
<td>ggcgcaagcttATGACCGAACTACGGCAG</td>
<td>128</td>
</tr>
<tr>
<td>MCds2N-myc</td>
<td>ggcgcaagcttATGGAACAGAAGCTTATT TCCGAAGAGGATCTGAAGATGACCGAACT ACGGCAG</td>
<td>128</td>
</tr>
<tr>
<td>MCds2Rev</td>
<td>ggcgccggaattcTCAGTTAGGTGAGACTTGAAAG</td>
<td>1426</td>
</tr>
<tr>
<td>MCds2C-myc</td>
<td>ggcgccggaattcCTACTTCAGATCCTCTT CGGAATAAAGCTTCTGTTCCATTTCAGTT AGGTAAGACTTGAAAG</td>
<td>1426</td>
</tr>
</tbody>
</table>

**Table 3.13** The primer names and sequences used to amplify *Cds2* incorporating myc-tags (regions underlined) from 1247-c2.

The products were excised from the pGEM vector with restriction enzymes *EcoRI* and *HindIII* (see section 2.3.1.9). The fragments were subcloned into the *EcoRI* and *HindIII* sites of expression vector pcDNA3 (Invitrogen). The clones were sequenced (see section 2.3.1.18) to ensure that they were in frame and did not contain any PCR amplification errors.

### 3.6.4.2 Transfection of cells for immunolocalisation

HEK-293T cells were plated onto coverslips and transfected with 5μg of the myc-tagged constructs described above (see section 2.3.4.5). The transfected cells were incubated at 37°C overnight. The cells were incubated in a 1:500 dilution (in PBS) of anti-myc-primary antibody (Santa Cruz Biotechnology, Inc.) followed by an anti-mouse FITC conjugated secondary antibody (Pierce, UK) which detects the primary antibody used and can be used for visualisation using fluorescence microscopy. However, on examination under a fluorescent microscope, no signal was seen and the cells appeared to be apoptotic. To ensure that transfection had occurred successfully, fresh cells had...
3.6.5.3 Western blots
Since the immunostaining experiments did not yield any positive results, Western blots were prepared (see section 2.3.4.2-4) to assess if any protein was being produced by the transfected cells.

HEK-293T cells were grown on petri dishes until they reached 40-50% confluency. 5μg of construct (see section 3.6.4.1) was then transfected into the cells. The cells were incubated for 48 hours at 37°C before being scraped from the plates and pelleted. The pellets were stored at -80°C until required (see section 2.3.1.15). The cell pellets were used to make cell membrane preparations (see section 2.3.4.2). The cell membrane preps were electrophoresed through an SDS-PAGE gel (see section 2.3.4.3) and electroblotted onto nitrocellulose membranes (see section 2.3.4.4). The Western blots were incubated with commercial anti-C-myc mouse monoclonal antibody (Santa Cruz Biotechnology, Inc). Control cell preparations, from non-transfected cells were used as a negative control. The results of the Western blots show a series of non-specific bands in both sample lanes (figure 3.56) with no clear differences between the transfected and non-transfected cells. This suggests that either the commercial antibody was not specific enough to detect the c-myc tag or that no tagged protein was being produced.

3.6.5.4 Discussion point of section 3.6.5
No tagged protein appears to be produced by the cells and the cells which were transfected and immunostained appeared to be apoptotic, this could be due to the cells relaying a stress response which triggers apoptosis. At this stage it can only be postulated that the gene product may be toxic to the cells. These results will be discussed further in chapter 4.
Figure 3.56 Western blotting to detect myc-tagged protein expression
Filter (a) is a control with only the secondary antibody applied, filter (b) has had no antibodies applied and filter (c) has had both primary and secondary antibodies applied. As can be seen there is little difference between filters (a) and (c), this was probably due to lack of specificity of the anti-C-myc antibody or the absence of any tagged protein.
3.7 Drosophila Transgenics

This experiment was performed to elucidate which, if either, mouse Cds gene would successfully rescue the Drosophila eye-cds mutant phenotype. Drosophila with mutations in the eye-cds gene experience light dependent retinal degeneration (the cds gene is described in section 1.2.4.5). In a previous study another Drosophila mutant which experiences light dependent retinal degeneration, rdgC, has been successfully rescued using the murine homologue (Chang et al., 1997). The rescue experiment had two main sections, firstly, the production of the rescue construct and then the mapping and expression study of the transgenes in the flies.

3.7.1 Cds rescue construct preparations

The rescue constructs have two components, the Cds coding sequence and green fluorescent protein (GFP) tag which have to be fused together and inserted into a Drosophila expression vector. The vector has P-element ends, P5' and P3' which when coinjected with an intact P-element allows the transgene to insert into the fly genome (for a full description of P-elements see section 1.3.1). The vector also contains the white+ gene which acts as a marker for detection of the transgene. The fly strain into which the construct is injected is a laboratory strain which have white eyes compared to wild type flies which have red eyes, when the transgene is successfully incorporated the transgenic flies display a yellow eye phenotype due to the incorporation of the white+ gene. The GFP tag is used to identify the sites of Cds-tagged protein expression in the transgenic flies using fluorescence microscopy.

The primers shown in table 3.14 were used to amplify (see section 2.3.1.16) the Cds gene fragments using 10ng of the cDNA clones 2279-d19 (Cds1) and 1247-c2 (Cds2) as template. The GFP fragment was amplified from the fly vector expression vector, pUAST-GFP (Brand and Perrimon, 1993)(supplied by Dr Raghu Padinjat, Cambridge University fly laboratory) using primers GFPC and GFPD (table 3.14). The vector pUAST is an inducible expression system that utilises the yeast transcription factor GAL4 as a promoter (figure 3.57).
Figure 3.57  Schematic representation of vector pP[UAST]

pUAST consists of five tandemly arrayed optimized GAL4 binding sites (red) followed by the hsp70 TATA box and transcriptional start (blue), a polylinker (green) containing unique restriction sites for EcoRI, BgII, NotI, XhoI, KpnI and XbaI and the SV40 small t intron and polyadenylation site. These features are included in a P-element vector (pCaSpeR3) containing the P element ends (P3' and P5') and the white gene which acts as a marker for successful incorporation into the Drosophila genome (Brand and Perimon, 1993).
Table 3.14 Lists the primers used to amplify the components for the Drosophila rescue constructs. The primers are written in the 5’-3’ orientation.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5'→3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cds1intron1F</td>
<td>GACCACGAAACCGAGAGCAC</td>
</tr>
<tr>
<td>Cds1B</td>
<td>CTTTACTCCTATGAGCCCATC</td>
</tr>
<tr>
<td>Cds2intron1F</td>
<td>CAGCATGACCGAACTACG</td>
</tr>
<tr>
<td>Cds2B</td>
<td>CTTTACTCCTATGAGCCATGT</td>
</tr>
<tr>
<td>GFPC</td>
<td>ATGGCTCATAGGAAGAGAAGAA</td>
</tr>
<tr>
<td>GFPD</td>
<td>CATATGGTTGTATAGTTCATC</td>
</tr>
</tbody>
</table>

Three gene fragments were amplified. The fragment bands were excised from the agarose gel (see section 2.3.1.10) and spun through Wizard™ (Promega) purification columns in a microcentrifuge at 5000rpm for 2 minutes. The concentration of the purified DNAs were measured by spectrophotometry, then used in fusion reactions to join the fragments (figure 3.58 schematically summarises the fusion reaction). The 3’ end of the primers Cds1B and Cds2B were designed to contain complementary sequence to the 5’ end of the GFPC primer. This overhanging sequence was used in fusion reactions to join the two fragments. A total of 250ng of purified DNA was used in each fusion reaction, the ratios of Cds gene fragments to GFP fragment were calculated based on the relative sizes of each and their purified band concentrations. To perform the fusion reaction, the correct ratios of DNA fragments were set up then heated at 65°C for 10 minutes. The reactions were then cooled slowly in a water bath to 37°C over 20 minutes. At 37°C 10x Reaction buffer 1 (Promega), Klenow (GibcoBRL) and dNTPs (Promega) were added. The reaction was then incubated at 37°C for two hours to allow the fragments to fuse completely.

The product from each fusion reaction was used as a PCR template to amplify the construct using primers Cds1intron1F or Cds2intron1F and GFPD (table 3.14). The fragments that were generated are shown in figure 3.59. The amplified fragments were excised from the agarose gel (see section 2.3.1.11) and spin-column purified as before.
Figure 3.58  

**Cds and GFP fusion reaction**

Shown is a schematic representation of the fusion reaction performed to generate *Cds2* tagged with *GFP*. (A) Shows the initial PCR amplification reaction using primers Cds2intron1F with Cds2B and GFPC with GFPD. This results in two fragments being generated, which are shown in (B). (B) A *Cds2* fragment with an overhanging complementary sequence to GFP and a GFP fragment. (C) shows the fusion reaction where the products from (B) are heated then cooled slowly. The product from this reaction is shown in (D). (D) shows the amplification of the fusion product using primers Cds2intron1F and GFPD. This PCR reaction results in a Cds2/GFP fusion fragment, which is shown in (E).
Figure 3.59 GFP fusion PCR fragments
Shown are the fusion fragments generated in the fusion reaction performed using the fragments shown in figure 3.54. The fusion fragment is approximately 2kb and will be cloned into the *Drosophila* expression vector pUAST.
The purified fusion product was cloned into the vector pGEM T-Easy (see section 2.3.1.13-14) and DNA prepared (see section 2.3.1.1). Using the restriction enzyme EcoRI (see section 2.3.1.9) the fusion fragments were excised from the vector. The EcoRI Cds-GFP fusion fragments were then subcloned into the EcoRI site of the \textit{Drosophila} expression vector pUAST (supplied by Dr Raghu Padinjat, Cambridge University fly laboratory). As described above the pUAST vector contains the \textit{white}\textsuperscript{+} gene, conferring eye colour which will be used to detect flies containing the transgene. The vector has a GAL4 promoter, which can be induced by mating the transgenic \textit{Cds} flies with GAL4 expressing flies to initiate transgene expression. The vector pUAST does not contain the LacZ gene and hence the colonies are not colour selectable. As the pUAST vector is not colour selectable, colony PCR (see section 2.3.1.17) was performed on all colonies using the primers Cds1intron1F or Cds2intron1F and GFPD. Clones identified as containing inserts were sequenced to check the orientation of the insert. The primers used in the sequencing reactions are listed in table 3.15.

After verifying the orientation of the inserted fragments, the correct clones were then sequenced with primers 1500F2 and CDS1INTRON11F, which verified the fusion region of the fragment and ensured that reading frame was continuous. The entire sequence of each of the constructs was verified using automated sequencing to ensure amplification errors had not occurred. Maxi preps of the positive clones were prepared (see section 2.3.1.2).

\textbf{3.7.2 Mapping and expression studies of the transgenic \textit{Drosophila}}

I performed all fly mapping and expression experiments in the Fly Laboratory at the University of Cambridge under the guidance of Dr Raghu Padinjat. The Maxi preps of construct circular DNA (see section 3.7.1) were microinjected into the \textit{Drosophila} embryos. The circular construct DNAs were co-injected into the posterior pole of 300 \textit{w\textsuperscript{1118}} (white eyed) \textit{Drosophila} embryos with a second plasmid carrying the transposase enzyme (Rubin and Spradling, 1983) (figure 1.21). The \textit{w\textsuperscript{1118}} (white eyed) flies are a transgenic laboratory fly line which have white eye colour instead of the wild type red eye colour. The construct was injected into the posterior pole of the embryos, as this is where the developing germ cells are situated, thereby increasing the chance of the
construct being incorporated into the fly germline. The transgenic DNA will randomly incorporate into developing germ cell DNA (described in section 1.3.1).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5' → 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cds2intron1R</td>
<td>TGTTTCGGCTCGGCTCTC</td>
</tr>
<tr>
<td>Cds2intron2R</td>
<td>CAGGTCGTAAGGAGTGGTAT</td>
</tr>
<tr>
<td>Cds2intron1F</td>
<td>CAGCATGACCGAACTACG</td>
</tr>
<tr>
<td>Cds2intron6F</td>
<td>GATTCTGCATGTTGTAC</td>
</tr>
<tr>
<td>Cds2B</td>
<td>CTGTATGTTGTATAGTTCAT</td>
</tr>
<tr>
<td>GFPD</td>
<td>CATATGTTGTATAGTTCAT</td>
</tr>
<tr>
<td>1500F2</td>
<td>GGCCACTTTTGTCAATGTGT</td>
</tr>
<tr>
<td>Cds1intron1R</td>
<td>TGCCATCGAGATCTCCATAC</td>
</tr>
<tr>
<td>Cds1intron1F</td>
<td>GACCACGAAACCGAGACAC</td>
</tr>
<tr>
<td>Cds1intron8F</td>
<td>TTGGGAAGGATTCACTTGGTG</td>
</tr>
<tr>
<td>Cds1B</td>
<td>CTTACTCCTATGAGCCATGTTAGGTGAGACTTG</td>
</tr>
<tr>
<td>Cds1intron11F</td>
<td>GGCGGGATAATGGACAGGT</td>
</tr>
</tbody>
</table>

Table 3.15 Lists the primers used to verify the orientation of the fusion fragment with the fly pUAST vector. The first three primers were used to verify Cds2 and the others were used to verify Cds1.

The emerging adults had white eye colour and were separated by sex while still virgins. The potentially transgenic (Cds) flies were back-crossed with w^1118^ virgin Drosophila. The resulting progeny were selected as carrying the Cds transgene by eye colour. The transgenic progeny from the first cross, had eye colour ranging from pale yellow to dark orange due to the introduction of the mini-white^+^ gene that was inserted as part of the pUAST vector.

Subsequent fly standard crosses were performed to map the location of the inserted transgenes. The mapping crosses were performed with fly lines that have been developed with dominant gene marker phenotype (balancer chromosomes) (chart of crosses shown in figure 3.60). This resulted in stable transgenic Drosophila lines where
<table>
<thead>
<tr>
<th>Sex</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td>CyO</td>
<td>MKRS</td>
</tr>
<tr>
<td>w</td>
<td>If</td>
<td>TM6Tb</td>
</tr>
</tbody>
</table>

**X**

- \( P_{w^+} P_{w^+} P_{w^+} \)
- \( Y \ P_{w^+} P_{w^+} \)

**Y**

- \( P_{w^+} P_{w^+} P_{w^+} \)
- \( w \ CyoOTM6Tb \)

**W**

- \( w \ CyO MKRS \)
- \( Y \ If \ TM6Tb \)

**Pick**

- \( P_{w^+} \)

**(3rd chromosome line)**

- \( Y \ Cyo/If and P_{w^+} \)

**(2nd chromosome line)**

- \( Y \ If \ TM6Tb \)
- \( or \ MKRS \)

**Pick**

- \( TM6Tb/MKRS \)
- \( Tb, Sb \)

**W**

- \( W; DB \)

**(X chromosome lines)**

- \( P_{w^+} Cyo MKRS \)
- \( Y \ If \ or \ TM6Tb \)

- \( P_{w^+} Cyo MKRS \)
- \( Y \ If \ x \ TM6Tb \)

- \( W; DB \)

---

**Figure 3.60 Chart of mapping crosses** - see next page for legend.
Figure 3.60 Chart of mapping crosses

Shown are the various mapping crosses performed in order to identify the chromosomal location of the integrated transgene. Each of the horizontal lines represents a Drosophila chromosome, sex chromosome (X/Y), 2 and 3. Chromosome 4 is not analysed, because it is so small homologous recombination is infrequent and hence the transgene is unlikely to have integrated. The symbols represent the various dominant and recessive mutations, which produce various phenotypes in the crosses and allow the site of integration to be determined. Curly of Oster (CyO) a dominant marker representing an inversion on chromosome 2, which gives the curly wing phenotype. Enhancer of forked (If) is the recessive chromosome 2 marker and give a phenotype of abnormal bristles. MKRS represents a transposition on chromosome 3; the dominant marker is stubble (Sb) which produces flies with shortened thickened bristles, small wings and twisted legs. Third multiple six (TM6) is an inversion of chromosome 3, the dominant marker for this is tubby (Tb) where the flies display shorter thicker bodies. Male and female balancers are slightly different; they only have dominant markers.

the transgene is inserted onto only one chromosome. Sufficient lines were set up to have a transgene insert on each chromosome. Balancer chromosomes have multiple inversions which, prevent the transmission of recombinant chromosomes to the progeny, interfere with homologous pairing and reduce the frequency of crossing over along most of that chromosome. These characteristics allow the determination of the site of integration of the transgene.

The transgenic fly lines were then crossed with Rh1-GAL4 flies that induce GAL4 expression and hence result in the expression of the transgene and GFP. The rhodopsin 1 (Rh1) gene drives expression in rhabdomeres. The successful progeny should express the transgene and fluoresce in regions of the deep pseudopupil in live flies (figure 3.61).

When the desired cross was performed to induce the Cds transgene expression the progeny were examined using fluorescence microscopy. No fluorescence was detected in any of the transgenic Drosophila lines (figure 3.61). One possibility for the absence of fluorescence is that the level of expression of the transgene may have been too low to be detected by light microscopy.

To analyse if there was low level expression that could not be detected using the deep pseudopupil fluorescence a Western blot was prepared (see section 2.3.3.4) using protein samples generated from transgenic fly heads (see section 2.3.4.1). The Western blots were incubated with GFP-antiserum (figure 3.62). No specific bands were detected and the Cds protein samples were indistinguishable from the control sample.
Figure 3.61 Fluorescence microscopy of live *Drosophila* eyes

Shown in figure (A) is a positive control fly (courtesy of Raghu Padinjat) showing fluorescence in rhabdomeres 1-6 (deep pseudopupil) the fluorescent region is indicated by a red arrow. Shown in figure, (B) is an example of the observations made on examination of the *Cds* transgenic flies where no fluorescence was present. The eye region is indicted with a red arrow.
**Figure 3.62 Western blots of transgenic fly head extracts**

Protein extracts were generated from *Cds* transgenic fly heads and were blotted with a Potassium channel protein extract (KRIII) which was used as a control (~160kDa) (courtesy of Dr R Padinjat). The blots were incubated with anti-GFP primary antibody and then with anti-rabbit horseradish peroxidase (HRP) conjugated secondary antibody. As can be seen from the blot only non-specific bands were produced in the Cds samples. The conclusions were that no protein was being expressed or that the level of expression was too low to be detected by immunoblotting.
Since there was lack of protein expression, evidence for RNA expression was determined. RNA was extracted (see section 2.3.2.1-2) from the progeny of rhodopsin driven GAL4 expressing flies crossed with the Cds transgenic flies (Rh1-GAL4*UAS-cds) and RT-PCR experiments were performed (see section 2.3.2.3) (figure 3.63). Primers that were used were Cds1intron1F with Cds1intron8R and Cds2intron1F with Mex7R (see tables 3.7 and 3.9 for sequence information) as well as primers to amplify the GFP tag (primers GFPC and GFPD, see table 3.14). Expression of both genes and GFP was observed. These experiments show that the gene and the GFP tag are present and being transcribed in the transgenic flies but the protein is not being expressed.

Since low-level expression in adult flies can be difficult to observe, expression of the proteins were examined in the transgenic embryos since this is where the proteins are usually most highly expressed. The engrailed-GAL4 (en-GAL4) promoter was used for this experiment instead of the Rh1-GAL4 promoter. The en-GAL4 is one of the strongest drivers available for this type of experiment. The driver uses the engrailed gene, which defines segmentation in the developing embryo so if the transgenic proteins are being expressed they will follow the segmentation pattern of expression shown by characteristic stripes. The results of this experiment are shown in figure 3.64. The positive control (courtesy of Dr Raghu Padinjat) shows expression, but no expression is observed in either of the Cds transgene embryos.

### 3.7.3 Discussion points of section 3.7

Experiments to rescue the cds mutant fly phenotype were not pursued since there was no evidence, from the results reported in this section, to suggest expression of the Cds proteins. The conclusions that can be drawn from these experiments are that the transgenic fly lines do not express protein in the tissues checked or that the level of expression was too low to be able to detect by the procedures employed above. However the RNA transcribed from the transgenes seems to be present in the transgenic flies. There are various reasons as to why no protein was detected in these experiments and these will be discussed in detail in chapter 4.
Figure 3.63 RT-PCR experiments using transgenic fly RNA extracts
RT-PCR was performed using RNA extracted from Cds transgenic flies as template. These experiments were performed to check that transgenic RNA was being produced. From the photographs shown above, it can be seen that RNA from Cdl and Cds2 is present and gives a band of around 800bp. The third photograph shows that the GFP tag is also present in the transgenic DNA (using primers GAL41/GAL42 (provided by Dr R Padinijat) and is also approximately 800bp. The relevant controls were also performed and shown in the photographs.
Figure 3.64 Detection of transgene in embryos using the en-GAL4 promoter

(A) positive control transgenic fly embryo where expression can be observed in the position where the protein encoded by the *engrailed* gene is expressed (indicated by the arrows). (B) shows the *Cds1* transgenic fly embryo resulting from a cross with the flies expressing en-GAL4 promoter, no expression of the transgenic protein is observed. (C) shows the *Cds2* transgenic fly, crossed in the same way as (B), again no protein expression is observed.
3.8 Cds2 Targeting Construct

3.8.1 Construct vector pPNT

The expression vector, designed for construct production (Tybulewicz et al., 1991), pPNT was used in this experiment (see figure 3.65). The vector pPNT contains several unique restriction sites, which are utilised for inserting specific genomic gene fragments into the vector. These gene fragments are inserted around a neomycin cassette using the unique vector sites. The role of the neomycin cassette in the construct vector is two fold; it disrupts the regions of specific gene fragments and it conveys resistance to the antibiotic neomycin. Resistance to neomycin is one of the two selectable markers used for positively identifying targeted embryonic stem (ES) cell colonies. The second selectable marker is the thymidine kinase gene, also present in the pPNT vector, and causes susceptibility to the antibiotic gancyclovir. When successful homologous recombination (described in section 1.3.2) occurs between the ES cell DNA and the transfected construct DNA the position of the vector's selectable marker genes is such that the neomycin cassette is carried into the targeted ES cell DNA with the gene specific fragments surrounding it and the thymidine kinase gene excluded. This allows positively targeted cells to grow in media containing both neomycin and gancyclovir.

3.8.2 Cds2 knockout construct production

The data that was gathered in sections 3.2 and 3.4, was used to form the knockout mouse construct. The clone 20A (see section 3.4) was used in the production of the construct. As described in section 3.4, 20A was digested using NotI and HindIII restriction enzymes (see section 2.3.1.9) and the resulting fragments were subcloned into pBluescript SK+ vectors (see section 2.3.1.13-14). The adjacently positioned 1.5kb and 6kb HindIII fragments contained all the Cds2 genomic information from clone 20A. The restriction map deduced from data gathered in sections 3.2 and 3.4 is shown in figure 3.66. It was decided that exon 6 of Cds2 would be targeted, with the neomycin cassette replacing half of exon 6 and approximately 90bp of intron 6. Using the map of pPNT (figure 3.65) the unique XhoI and BamHI situated either side of the neomycin cassette sites were selected for insertion of Cds2 genomic fragments.
Figure 3.65  pPNT vector
The figure shows a schematic representation of the pPNT vector used to form the Cds2 knockout construct. The restriction sites that were known to be unique are highlighted with stars. The unique BamHI and XhoI sites were used to insert the gene specific fragments.
Figure 3.66 Wild type and targeted Cds2 structures

This figure shows a schematic representation of the genomic structure wild type Cds2 compared to the structure which would be seen in targeted ES cells, where a neomycin cassette is inserted into exon 6. The 750bp PstI/BamHI probe, which was used in the detection of correctly targeted ES cells, is shown as a pale blue box. The important restriction sites are indicated, B-BamHI, H-HindIII, P-PstI and E-EcoRI. The BamHI and EcoRI sites after the neomycin cassette in the recombinant structure are sites acquired from the insertion of the 3kb BamHI fragment and are derived from pBluescript vector.
3.8.2.1 3kb BamHI construct fragment

The 6kb HindIII fragment containing most of intron 6 to 12 was used to produce the 3kb construct fragment (figure 3.67). The 6kb HindIII fragment derived from 20A was digested using BamHI. The pBluescript vector (in which the 6kb fragment was contained), BamHI site was used and formed the 3kb fragment with the central BamHI site situated in intron 8. The 3kb BamHI fragment which was produced, was excised from the agarose gel (see section 2.3.1.10) and purified using a Wizard™ spin column (Promega). The fragment was then cloned into the BamHI site in the pPNT vector. To ensure the fragment was inserted into the vector in the correct orientation, it was sequenced (see section 2.3.1.18) using the primers listed in table 3.16.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5’[1]3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cds2int7F</td>
<td>CCAGGTAACACTAATATTGG</td>
</tr>
<tr>
<td>CDS2INTRON3R</td>
<td>CAGATTAAGTGTTCATG</td>
</tr>
<tr>
<td>1500F3</td>
<td>GTAACGCAGTCACACCTTGT</td>
</tr>
</tbody>
</table>

Table 3.16 The primer name and sequence used to verify the orientation of the gene fragments within the pPNT expression vector.

3.8.2.2 1.4kb Xhol construct fragment

The second fragment to be inserted into the pPNT construct vector was a 1.4kb Xhol fragment generated from the 1.5kb NotI/HindIII fragment derived from 20A. The 1.5kb NotI/HindIII fragment contains 150bp of intron 3 – 5’ end of intron 6 (figure 3.68). Since the restriction map generated for the 1.5kb NotI/HindIII fragment did not provide any suitable restriction sites for cloning into the pPNT vector, primers with Xhol tags (listed in table 3.17) were designed against sequence derived from intron 3 and exon 6. The primers were used to amplify a 1.4kb Xhol fragment using the 1.5kb NotI/HindIII fragment as template (see section 2.3.1.16). The PCR product was cloned into a pGEM T-Easy vector and the insert excised using the restriction Xhol to ensure the tags were in place. The fragment was excised from an agarose gel and purified as before. The purified Xhol fragment was then subcloned into the unique Xhol site in the pPNT vector which already contained the 3kb BamHI fragment.
Figure 3.67 The generation of the 3' Cds2 construct fragment
Shown is a schematic illustration of the 3kb BamHI fragment that was used as the 3' construct gene fragment. The fragment was derived from the 6k HindIII clone and utilises one of the vector restriction sites to create the fragment of interest. This fragment was then cloned into the pPNT expression vector.
Figure 3.68 The generation of the 5’ Cds2 construct fragment

Shown is a schematic illustration of the 1.4kb fragment that was used as the 5’ construct gene fragment. The fragment was amplified from the 1.5k HindIII clone using the primers Con1F and Con6R. This fragment was then cloned into the pPNT expression vector.
To ensure that the XhoI fragment had been inserted in the correct orientation and the subsequent sequence was not in frame preventing normal protein production, the construct was sequenced using the primers listed in table 3.18. A summary of the production is schematically illustrated in figure 3.69. A Maxi prep was prepared (see section 2.3.1.2) for the construct and was linearised using the unique NotI site in the pPNT vector.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5’[3’]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDS2ConF1</td>
<td>GGCCTCGAGG AGCTC AATT AACC</td>
</tr>
<tr>
<td>CDS2Con6R1</td>
<td>CGCCCTCGAGCAAGTTAGGATAACAAGGT</td>
</tr>
</tbody>
</table>

**Table 3.17** XhoI tagged primers used to generate the 5’ gene fragment for the Cds2 knockout mouse construct.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5’[3’]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int5F</td>
<td>GCATGCATTGTCCAIGTCTG</td>
</tr>
<tr>
<td>SIF3</td>
<td>GTAACGCAGTCACACCTTGT</td>
</tr>
<tr>
<td>ConFor</td>
<td>CAGGACGTGACAAAT</td>
</tr>
<tr>
<td>Con2R</td>
<td>ATTTGTCACGTCTGCACGA</td>
</tr>
</tbody>
</table>

**Table 3.18** Primers used to verify orientation and to verify that there is a stop codon present to prematurely terminate the translation of the gene.

### 3.8.3 Embryonic stem (ES) cell electroporation

The linearised construct DNA was given to Dr Kim Wells of the Transgenic unit, Imperial School of Medicine, Charing Cross hospital, for electroporation into ES cells (see section 2.3.1.12.3). All ES cell work was performed by Dr K Wells.

1x10⁷ cells in a volume of 0.6ml cold PBS were electroporated with either 25μg or 6μg of linearised CDS2 plasmid at 250volts 500μF in a Biorad Gene Pulser (0.4mm cuvette).
1. Restriction digest of the 6kb *Hind*III fragment, which was derived from 20A and subcloned into pBS vector, using *Bam*HI generates a 3kb *Bam*HI fragment, which incorporates the vector *Bam*HI and *Eco*RI sites.

2. The 3kb *Bam*HI fragment was cloned into the unique *Bam*HI site in construct vector pPNT.

3. The orientation of the fragment in the vector was verified by sequencing (table 3.14)

4. The 1.5kb *Not*I/*Hind*III fragment derived from 20A, was used as template in a PCR reaction using the *Xho*I tagged primers listed in table 3.15. This amplified a 1.4kb fragment.

5. The 1.4kb *Xho*I PCR fragment was cloned into pGEM T-Easy vector and then restriction digested using *Xho*I to excise the fragment from the vector. The excised fragment was then subcloned into the unique *Xho*I site of pPNT already containing the 3kb *Bam*HI fragment.

6. The sequence and orientation of the *Xho*I fragment was verified using primers in table 3.14. The construct was then linearised with restriction enzyme *Not*I.

**Figure 3.69 The steps taken to form the Cds2 construct**

Shown is a schematic illustration of the steps carried out to prepare the Cds2 knockout construct.
After incubating on ice, the cells were suspended in 22ml of fresh media and plated out onto three plates, two plates with 10ml of the cell suspension and a third plate with 2mls diluted out to 10ml final. This third plate was subjected to G418 selection only as a control for transfection and plating efficiency. A mock transfection was also carried out. After two days in culture the plates were changed to media containing 2μM gancyclovir and 200μg/ml G418. Selection was maintained for 8 days by which time there were no surviving colonies in the mock transfected control plates.

A total of 26 ‘healthy’ colonies were picked from plates subjected to positive selection with G418 only. These plates contained a tenth of the cell number plated on the double selection plates. This colony number is therefore equivalent to 260 clones, which is a lower number than expected. Three clones were obtained from double selection, this equated to 1.15% of the total number of transfectants. The number of targeted clones seems acceptable for the percentage of homologous integration events in transfected clones. However, the total number of transfectants initially obtained was a lower than expected. DNA was prepared (see section 2.3.1.7) from the selected colonies.

3.8.4 Screening of potentially targeted ES cells

10μg of each of the potentially targeted DNA samples were restriction digested using enzyme EcoRI and Southern blotted (see section 2.3.1.11). A probe was generated from the 6kb HindIII fragment (derived from clone 20A) that was outside the area used in the construct production. The 750bp PstI/BamHI fragment (figure 3.68) was 32P-radiolabelled (see section 2.3.1.20) and used in a hybridisation reaction (see section 2.3.1.21) with the Southern blot. In wild type DNA the probe fragment should hybridise to a 15kb EcoRI fragment. Figure 3.70 shows the wildtype band produced when wildtype mouse genomic DNA is probed with the 750bp fragment. Use was made of the EcoRI site, which had been incorporated into the construct with the 3kb BamHI fragment when it was inserted into the pPNT vector (figure 3.67). Thus targeted DNA should show two bands on an autoradiograph, one from the wild type gene and another smaller band from the targeted gene. Unfortunately none of the samples showed this double-band profile and it was considered that no homologous recombination had taken place.
Figure 3.10 Control Southern blot for assessment of future targeted ES cells

Shown is a Southern blot autoradiograph of mouse genomic DNA digested with the restriction enzyme EcoRI. The blot was probed with the 750bp fragment derived from intron 9 (figure 3.66). A band of ~17kb was observed. This control shows that future targeted ES cells should produce a band smaller than 17kb when probed with the same fragment due to the EcoRI site that was inserted when the construct was made.
Several further electroporation runs were performed altering some of the variables such as, the quantity of DNA used for electroporation, the number of cell in the cuvette and the concentrations of the antibiotics used. The results from these experiments suggested that the level of G418 selection may be too stringent as the single selection plates should have had dozens of colonies whereas there were only a few.

A dose response curve was prepared comparing another expression vector (PGKneo) to the pPNT vector (empty vectors used). 40μg of linearised plasmid was used and 0.8ml of cells at 1 x 10^7/ml. The DNA was dissolved in sterile 10mM Tris 0.1mM EDTA pH7.5. The rest of the experimental conditions were as described above. Selection was applied on day two after electroporation. The electroporated cells were split onto 5 plates and the following concentrations of active G418 were used: 200, 190, 180 170mg/ml or no selection. The media were changed every day and the colonies assessed after 10 days of selection (figure 3.71). The results from this experiment are shown graphically in figure 3.72.

Based on the data derived from the control experiments described above, the final targeting electroporation was carried out using 50μg of linearised plasmid, either the CDS2 construct or empty pPNT vector, transfected into cells at a concentration of 0.8 ml at 1 x 10^7 cells/ml, with electroporation conditions as previously described. Selection was performed two days after electroporation with either G418 alone at 170μg/ml or dual selection G418 170μg/ml and 2μM Gancyclovir. Photographs of single and dual selection plates, stained in the same way as in figure 3.74, were taken after 10 days of selection (figure 3.73). These show that cells containing the empty vector grow reasonably well, but when Cds2 gene fragments are inserted very few colonies grow. Hence the targeting experiments using the Cds2 construct described above were not pursued.

3.8.5 Discussion point of section 3.8
From the results of all of the above experiments, it appears that the colonies, which were tested from the first round, did not contain the targeted gene. It may be that the clones from the first round of cells were dying and not actively growing at the time of picking from the plate and that the cultures were overgrown with non-G418 resistant cells once the selection was removed. Another interesting result, and will be discussed
Figure 3.71 ES cells transfected with control PGKneo vector
ES cells transfected with PGKneo vector which have been fixed and stained with haematoxylin. The colonies were grown on plates, clockwise from top left, containing 180μg/ml, 170μg/ml, 190μg/ml and 200μg/ml of G418. A fixed size central square from each plate was imaged and the number of objects and their pixel size was assessed using SigmaScan (picture courtesy of Dr K. Wells)
Figure 3.72 PGKneo vs. pPNT colony count/drug dose graph
The number of colonies containing PGKneo vector compared to pPNT vector present on plates containing 170-200μg of G418 antibiotics (courtesy of Dr K. Wells).
Figure 3.73 Dual and single drug selection of Cds2 targeted ES cells
Shown are the fixed, stained plates of dual and single selection of Cds2 construct. Dual selection with gancyclovir and G418 (top left), and single selections with G418 alone of pPNT vector (top right) and Cds2 construct (bottom right) (picture courtesy of Dr K. Wells).
further in chapter 4, was that although the number of colonies on the ‘pPNT vector only’ plate were lower than that of the other PGKneo vector, they were still present in a reasonable amount. However, when the construct containing Cds2 DNA was introduced into the ES cells, the number of colonies on the plates were substantially reduced.
3.9 Patient Screening

Human CDS1 was mapped to chromosome 4q12-22 and CDS2 to 20p13 (Halford et al., 1998, Volta et al., 1999). The region 4q21-22 does not contain any ocular disease loci but two disease loci with ocular phenotypes are localised to the 20p13 region. These are neuron degeneration with brain iron accumulation 1 (NBIA1) which was previously called Hallervorden-Spatz disease, which often shows retinal degeneration and optic atrophy as secondary features, and congenital hereditary endothelial dystrophy of the cornea (CHED2). Recently, mutations in pantothenate kinase were shown to be responsible for NBIA1 (Zhou et al., 2001).

CHED can be inherited as both an autosomal dominant, CHED1, and an autosomal recessive disease, CHED2 (Maumenee, 1960). Both disorders are characterised by diffuse oedema that varies in severity from a mild haze to an opaque cornea and is generally present at birth or shortly after, with little progression. Both CHED1 and CHED2 map to chromosome 20. The gene responsible for CHED1 has been mapped to chromosome 20p11.2 (Toma et al., 1995) and that for CHED2 to 20p13 (Callaghan et al., 1999, Hand et al., 1999). The CHED2 disease locus spans 3.4 Mb of DNA in a relatively gene dense interval (Mohamed et al., 2001).

Using in situ hybridisation, Cds2 was found to be expressed in the cornea (figure 3.46); this data combined with the map location of the gene, suggested that CDS2 is a good candidate gene for CHED2. In order to assess this possibility the genomic organisation of human CDS2 was determined, using methods described in sections 3.1-3.4, and this information was used for mutation screening in a CHED2 family. Table 3.19 provides information about the exon-intron boundaries of human CDS2 generated from PAC clones known to contain the gene (Halford, personal communication). Using the exon-intron boundary information generated from the PAC clones, PCR primers (table 3.20) were designed to amplify the thirteen exons of the CDS2 gene in two affected and one normal member of the CHED2 Pakistani family described by Mohamed et al. (2001). The exons were amplified (section 2.3.1.6), directly sequenced (2.3.1.18), and compared to the known exon sequence. The exon-intron boundaries were also sequenced and analysed. No changes were detected in the exon-intron boundaries or in the coding region of the gene, making it highly unlikely that that CDS2 is responsible for the
<table>
<thead>
<tr>
<th>Exon</th>
<th>Exon size (bp)</th>
<th>5' intron</th>
<th>Exon sequence</th>
<th>3' intron</th>
<th>Intron size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>57*</td>
<td>5'UTR</td>
<td>ATGACAGAGC-----CGAGGACAAG</td>
<td>gtageggcagctcggggtg</td>
<td>46373</td>
</tr>
<tr>
<td>2</td>
<td>137</td>
<td>tctgtatctgcctccattag</td>
<td>GAGTCAGAGT---------TGCTTCAAG</td>
<td>gtaacctgccttaatagtgc</td>
<td>1523</td>
</tr>
<tr>
<td>3</td>
<td>97</td>
<td>tccctgtcctctgaccttag</td>
<td>ATGGAAGAAC--------GATGATAATC</td>
<td>gtaagtgcctatttcacacta</td>
<td>1368</td>
</tr>
<tr>
<td>4</td>
<td>98</td>
<td>gaatgttttgtgtccccagcag</td>
<td>GTGATGTGCG---------CGCTCAGCTG</td>
<td>gtaagctctccgccccaca</td>
<td>2071</td>
</tr>
<tr>
<td>5</td>
<td>140</td>
<td>tgacctgtctttttccattacag</td>
<td>GTACTTTTCTC------TATCTAATAG</td>
<td>gtaagctattacagttcag</td>
<td>3430</td>
</tr>
<tr>
<td>6</td>
<td>59</td>
<td>tctccctccacctcctaactag</td>
<td>GATTCTGCAT--------GTTCTACATG</td>
<td>gtaaggaagactcgtgctg</td>
<td>615</td>
</tr>
<tr>
<td>7</td>
<td>83</td>
<td>tattttttttttttttgcag</td>
<td>TTTGGCTTGG---------GAAATGATCTG</td>
<td>gtaaggaatttcgggatggtg</td>
<td>1714</td>
</tr>
<tr>
<td>8</td>
<td>88</td>
<td>tactttggttcatttctag</td>
<td>GTTCATTTGTC------ACTCATCAAG</td>
<td>gtaaattcgtttaccctag</td>
<td>812</td>
</tr>
<tr>
<td>9</td>
<td>69</td>
<td>aactgtctctctccacagcag</td>
<td>CTGTCCCCGA---------TGGCTTCTG</td>
<td>gtaggtggtgtctctcgct</td>
<td>860</td>
</tr>
<tr>
<td>10</td>
<td>153</td>
<td>tctctgattctctctcctcag</td>
<td>CTGTCTTATG-------CATTGGCTTG</td>
<td>gtagctgtcactcagggg</td>
<td>2227</td>
</tr>
<tr>
<td>11</td>
<td>120</td>
<td>actctattatctctctcctcag</td>
<td>AAAACGGTCC--------TAAAAATCAA</td>
<td>gtagtaaacccttgtactg</td>
<td>548</td>
</tr>
<tr>
<td>12</td>
<td>104</td>
<td>cctctctctctctctaaccag</td>
<td>GACTTTGCCA---------GTTTTATCAG</td>
<td>gtagactctctcactcag</td>
<td>263</td>
</tr>
<tr>
<td>13</td>
<td>133^</td>
<td>ttctccatcttcctgtccag</td>
<td>AGGCCCCTAAC------GGACGAGTAG</td>
<td>3'UTR</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.19 The exon-intron boundaries of human CDS2. (* Length from translational start site, ^ Length to stop codon)
corneal dystrophy seen in this family. This work has subsequently been published Halford et al., 2002 (appendix 1).

<table>
<thead>
<tr>
<th>Exon</th>
<th>Primer sequence</th>
<th>Amplimer (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F 5'-GGCTGCTAAAGGGAACCTGTCAG-3' R 5'-AAGGGTCAGAGAACAAGACGAC-3'</td>
<td>305</td>
</tr>
<tr>
<td>2</td>
<td>F 5'-ATGACGGCAGCTAGACTTCTC-3' R 5'-CCACTGCTAAGCAGACTGATT-3'</td>
<td>383</td>
</tr>
<tr>
<td>3</td>
<td>F 5'-GTGAATTGGAAGGACTGA-3' R 5'-AGTCACAGATGTTCATATGC-3'</td>
<td>369</td>
</tr>
<tr>
<td>4</td>
<td>F 5'-TCATGACCTGTTGGTTTCCA-3' R 5'-CCCTGCTACAGATATTACCT-3'</td>
<td>286</td>
</tr>
<tr>
<td>5</td>
<td>F 5'-ATTCTCACCCTGTTGTC-3' R 5'-ATTCTCACCCTGTTGTC-3'</td>
<td>446</td>
</tr>
<tr>
<td>6</td>
<td>F 5'-CTGCTGGAATAATACTAGCC-3' R 5'-CCCTCATGCTCAGATCAT-3'</td>
<td>305</td>
</tr>
<tr>
<td>7</td>
<td>F 5'-CTGTCTCTATGCATGACT-3' R 5'-CTGTCTCTATGCATGACT-3'</td>
<td>366</td>
</tr>
<tr>
<td>8</td>
<td>F 5'-ACTCCTTAATGGGTACCTCAG-3' R 5'-CTAGATTCTGCTGACTGCC-3'</td>
<td>309</td>
</tr>
<tr>
<td>9</td>
<td>F 5'-GACATGGTTAAGGAATACTG-3' R 5'-AGAAGGCTACCTGCTGATG-3'</td>
<td>433</td>
</tr>
<tr>
<td>10</td>
<td>F 5'-CATGAAACCTCTGACACATA-3' R 5'-AGAAGGCTACCTGCTGATG-3'</td>
<td>429</td>
</tr>
<tr>
<td>11</td>
<td>F 5'-GGCCACAATGAGGATGAAAT-3' R 5'-CTTAATCTTGAGGTGTC-3'</td>
<td>359</td>
</tr>
<tr>
<td>12</td>
<td>F 5'-TGCTCTATATAGCCTCACC-3' R 5'-ATTAACAGATGACAGGAGG-3'</td>
<td>420</td>
</tr>
<tr>
<td>13</td>
<td>F 5'-GCCTCCTGTCATCTGTTAAT-3' R 5'-GTCATTGCTCAAGTCACACC-3'</td>
<td>343</td>
</tr>
</tbody>
</table>

Table 3.20 Primer pairs for CDS2 mutation analysis.
CHAPTER FOUR
DISCUSSION

4.1 RNA MasterBlots™

The results of the MasterBlots™ (see section 3.6.1) showed that \textit{Cds1} and \textit{Cds2} had very different expression patterns. It was apparent that \textit{Cds2} was a ubiquitously expressed gene and was involved in the development of the embryo from as early as 7.5\textit{d.p.c} whereas \textit{Cds1} had a more restricted pattern of expression and was not expressed in the embryonic samples. However previous reports indicate that \textit{Cds1} is expressed in the embryo although only at 12d.p.c and only in the thymic rudiment (Volta, \textit{et al.}, 1999). Volta \textit{et al.} (1999) also reported the embryonic expression of \textit{Cds2}. They report that expression appears from 12.5d.p.c onwards, yet the blot shown in figure 3.43c clearly shows expression from a much earlier time point. Only by studying timed embryo sections more closely will these finding be confirmed.

The \textit{E.coli} DNA negative control was briefly discussed in section 3.6.1.2, further to the discussion in the previous section, an alignment of the \textit{E.coli} \textit{cds} and mouse \textit{Cds1} and \textit{Cds2} amino acid sequences were performed (figure 4.1a and 4.1b). The figures show that there is a region of high homology between residues 216 and 222 (using the \textit{E.coli} sequence as reference) equating to an identity of potentially 20 nucleotides. Since oligonucleotide sequences as short as 20 base pairs can be used as primers or probes, this is possibly a large enough region of homology to allow hybridisation to occur. Since the \textit{E.coli} genome does not have introns, this region of homology is continuous and the cDNA probes generated from the mouse \textit{Cds} clones may provide sufficient homology to result in hybridisation. This is reinforced by the alignments shown in figure 3.32, where the yeast samples have a greater degree of homology to \textit{Cds} than \textit{E.coli}, but due to the presence of introns in the yeast genome, hybridisation to the control yeast sample does not occur. In addition to the evidence provided by the negative controls, the data gathered from the MasterBlots™ for \textit{Cds1} and \textit{Cds2} are validated since the genes display very different patterns of expression, hence implying that cross-hybridisation did not occur. The same clones, 801-h9 and 1247-c2, which were used to make the cDNA probes used on the MasterBlot™, were used to make the riboprobes which were used for the \textit{in situ} hybridisation studies shown in figures 3.44
Figure 4.1a  *E.coli* cds and mouse Cds1 amino acid alignment

Shown is the alignment between the two cds amino acid sequences, the overall homology is ~20% with the *E.coli* sequence spanning a much shorter distance. The shaded area shows regions of similarity and the boxed and shaded area shows regions of identity. The area highlighted in red is a stretch of seven amino acids, which are identical. This region of the sequence may be the reason why the *E.coli* DNA control on the MasterBlot™ many have been positive.
**Figure 4.1b E.coli cds and mouse Cds2 amino acid alignment**

Shown is the alignment between the two cds amino acid sequences, the overall homology is 20% with the E.coli sequence spanning a much shorter distance. The shaded area shows regions of similarity and the boxed and shaded area shows regions of identity. The area highlighted in red is a stretch of seven amino acids, which are identical. This region of the sequence may be the reason why the E.coli DNA control on the MasterBlot™ many have been positive.
to 3.55 (discussed in section 4.3). *In situ* hybridisation on tissue sections using riboprobes is generally regarded as being more sensitive than techniques such as the Masterblot™ since RNA probes detect gene transcripts when using *in situ* hybridisation compared to cDNA probing RNA samples on the MasterBlot™. No cross-hybridisation was observed in the *in situ* hybridisation studies. Therefore, the pattern of expression derived from the MasterBlot™ should be regarded as an accurate interpretation of the expression pattern of murine *Cds1* and *Cds2*.

### 4.3 Northern Blots

The Northern blots shown in figures 3.42 and 3.43 provide evidence for more than one transcript for each gene. This is most likely due to the presence of different size splice variants that exist for each gene. There is evidence that the human *CDS* gene has splice variants (Volta *et al.*, 1999; S. Halford, personal communication). Hence the true transcript size is most likely to be ~4kb in length, with the smaller bands representing the splice variant sizes where certain exons have been excluded. The data presented in sections 3.6.1 and 3.6.2 differs from data reported by Lykidis *et al.* (1997), who report low levels of expression of human *CDS1* in the brain, lung, placenta, small intestine, ovary and testis using Northern blotting. In these tissues they identified two transcripts of 5.6kb and 3.9kb. Lykidis *et al.* (1997) used human tissue however mouse and human *Cds1* transcripts are of a similar length as regards to the coding and 3' UTR sequences (described in 3.1). Therefore, unless the 5' UTR is much greater in human than in mouse, these findings confer differences between the two sets of results. Another finding reported by Lykidis *et al.* (1997) that differs from results reported in this thesis is that when foetal tissue was used to form the Northern blot, transcripts were identified in lung, kidney and brain samples. These results are surprising since clearly there was no expression in foetal samples in figure 3.41a when probed with *Cds1* and the only report to suggest that *Cds1* is expressed in the embryo identifies a specific time point and with expression only present in the thymus. Volta *et al.* (1999) reported findings from Northern blots probed with human *CDS2*, in which three transcript lengths, 9.5, 4.4 and ~3kb, were identified and found to be expressed at different levels. Again, these results differ from the finding reported in this thesis. In order to verify these previously reported findings, various improvements to the Northern blot filters prepared
in this study should be performed, this will be discussed in chapter five which will detail all future work.

4.4 In situ hybridisation studies

The results reported in section 3.6.3, were interesting since both Cds genes evolved from a single ancestral Drosophila gene (figure 3.33), they display very different expression patterns, including very different locations within the ocular tissues. Cds2, shares more sequence homology (57%) than Cds1 (56%) to Drosophila eye-cds but does not have an expression pattern to suggest that it plays a direct role in phototransduction. This variant, may have a more general signalling role. The data gathered from the corneal sections shows that Cds2 is expressed in this tissue. As described in the introductory chapter the chromosomal map location of human CDS2, 20p13, includes various ocular diseases including autosomal recessive congenital hereditary endothelial dystrophy (CHED2). CHED2 patients experience diffuse bilateral clouding of the cornea caused by an abnormality in the endothelium with many patients requiring corneal transplantation to restore vision (Hand et al., 1999). Since the expression data indicated a role for Cds2 in the cornea, screening of CHED2 patients could eliminate Cds2 as the causative gene. Results previously reported by Volta et al. (1999) claimed that Cds2 was not expressed in the adult retina. This finding is very much the converse of data reported in this thesis (figure 3.48).

Another finding by Volta et al. (1999), which disagrees with results reported here, is that Cds1 is expressed in the inner segments of the photoreceptor layer of the retina, whereas as reported here, there is also expression in the ganglion cell layer (figure 3.45). The location of the Cds1 molecule in the inner segments of the photoreceptors suggests that this gene may play a direct role in phototransduction. The inner segments contain the cytoplasmic machinery for the synthesis of proteins, the production of energy-rich nucleoside triphosphate and the assembly of disc membranes. Other molecules known to play an important role in phototransduction, such as rhodopsin, show a similar spatial and temporal expression pattern in the inner segments (Ishiguro et al., 1987; Bowes et al., 1988). The RT-PCR and in situ studies using the rd mice proved interesting as it was shown that Cds1 is expressed only in the rod
photoreceptors. However, the function of \textit{Cds1} is likely only to be defined by the production of a knockout mouse model.

Chapter five will describe some of the future work that is required to elucidate the type of photoreceptor cells expressing \textit{Cds1}, what the function of both the \textit{Cds1} and \textit{Cds2} genes are and whether they can be excluded from causing specific ocular diseases.

\textbf{4.5 Epitope tagging and subcellular localisation of \textit{Cds1} and \textit{Cds2}}

The set of experiments performed in section 3.6.5 were done in order to elucidate the molecular weight of the Cds proteins after translation and to localise them subcellularly. As described in section 3.6.5 the Western blots prepared to elucidate the size of the protein produced a number of non-specific bands, which were at first thought to be lack of specificity on behalf of the commercial antibody against c-myc. The cells were then transfected and immunostained prior to examination by fluorescence microscopy. The cells had an appearance that would suggest that an apoptotic event had occurred. This suggested that the antibody may have been working but the cells had not produced any myc-tagged protein hence the non-specific bands.

Apoptosis may have occurred in the transfected cell as a result of an accumulation of mis- or un-folded proteins in the endoplasmic reticulum (ER). The ER is a vast structure, which post-translationally modifies proteins and folds them into their tertiary structure before directing them to their final locations. The unfolded protein response (UPR) constitutes stress signalling from the ER lumen to the nucleus in response to accumulation of mis-folded or unfolded polypeptides (reviewed in Kaufman, 1999; Pahl, 1999; Mori, 2000). By pathways not fully characterised in mammals, the UPR may induce transcription of additional chaperone molecules (to assist folding), attenuation of translation (to prevent generation of more mutant protein), and mechanisms for protein degradation. Additional death-inducing signals are generated by the ER overload response (EOR). It is therefore, conceivable that the tags on the Cds proteins are interfering with the correct folding and overloading the ER hence inducing a stress signal culminating in programmed cell death (Plempner and Wolf, 1999).
In chapter five various future experiments will be outlined in order to re-address some of the points discussed in this section, such as ways around using tagged proteins, or using different expression systems and also ways to verify if the cells described are really apoptotic.

4.6 Mutant fly rescue system

The object of this set of experiments was to identify which if either of the mouse Cds genes conferred a function similar enough to that of the Drosophila eye-cds gene so as to rescue the mutant phenotype in the fly. This experiment has been successfully performed using a murine homologue of the Drosophila gene rdgB (Chang et al., 1997), which is a gene required for phototransduction in the fly. Sequence information from both mouse genes suggest that Cds2 is more similar to Drosophila eye-cds but Cds1 has a retinal expression pattern which suggests it is more likely to be directly involved in phototransduction. Constructs were prepared for both Cds1 and Cds2 (see section 3.7 for details). The results show, using RNA extracted from transgenic flies, that both the Cds genes and the green fluorescent protein (GFP) tag fused to the gene are present and being transcribed. However, no protein was detected by fluorescence, embryo immunostaining or by Western blots prepared from transgenic fly heads. This could be because the proteins are being translated but being rapidly degraded but more likely the information required to form viable proteins is not complete. The experiment carried out in this thesis was performed in the standard way, but recently as more genes are being analysed in this manner it has been found that many fail to form viable proteins. These genes are found only to form viable proteins when intronic sequence is present in the original construct (R. Padinjat, personal communication). The likely explanations for these findings is that enhancers of the genes are present within introns and are required for proper protein expression or that the genes require all genomic information to produce all splice variants of the gene, without intronic information this is not possible. To date not enough is known about splicing mechanisms and the full role of the introns to speculate further in these findings. The construct used in this thesis was produced using cDNA before these observations were made. Results from the Northern blots and information about the human genes implies that splice variants exist for murine Cds1 and Cds2. The experiment was not completed in section 3.7, since no
GFP expression was detected in the transgenic flies there was no evidence to suggest completing the crosses would rescue the *cds* mutant fly phenotype.

In the light of recent findings from analysis of other genes being examined in a similar way, chapter five details some future work that could change the outcome of this type of experiment and elucidate whether the murine *Cds* genes convey a similar function to those of *Drosophila cds*.

### 4.7 Cds2 knockout mouse

Data gathered in this thesis after the production of the *Cds2* knockout construct, shows that the method used may not be the most optimal for the production of the *Cds2* mouse, and *Cds1* appears to be a much better option for the method used. The MasterBlot™ results clearly show that *Cds2* is expressed in all tissues and is present in very early embryonic samples. This indicates that removal of *Cds2* may result in embryonic lethality. The embryonic tissues that would be affected are not known and future *in situ* hybridisation studies could be performed to determine the exact time points and tissues expressing *Cds2*. A better approach for the production of the *Cds2* mouse is described later in this section.

An interesting finding to arise from this study is that the vector which was used (pPNT) has a lower transfection efficiency than some others which are available, and in future studies involving the production of knockout mice, pPNT should not be used. Table 4.1 lists some of the other vectors available with the PGKneo being the one of future choice. Although the vector is not as strong as some listed in table 4.1, as figure 3.73 illustrates, colonies are produced when empty vector is transfected into ES cells. However, when the vector contained some *Cds2* sequence hardly any colonies grew. This suggests that the truncated sequence may be producing a poison sequence, which kills the cells. If this gene is to be pursued as a candidate for a knockout study, different regions of the gene will have to be examined to identify a more suitable site for the introduction of the neomycin cassette. In future studies, detailed in chapter 5, use should made of p-lox sites when making the knockout constructs.
The Cre/lox system is highly efficient for inducing genomic changes at a specific time in specific tissues. The Cre (cyclization recombination) recombinase is a 38kDa bacteriophage P1 protein that catalyses recombination between two 34bp loxP (locus of X-over of P1) recognition sites without the need for any cofactors. Depending on the relative orientation and location of the loxP sites, Cre-mediated recombination results in DNA inversion, excision/integration, and translocation, thereby conferring flexibility (figure 4.2).

<table>
<thead>
<tr>
<th>Vector Name</th>
<th>G418 resistant colonies per 10^7 cells</th>
<th>Relative efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMCineopA</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>RV4.0</td>
<td>1632</td>
<td>125</td>
</tr>
<tr>
<td>MCineobpA</td>
<td>464</td>
<td>36</td>
</tr>
<tr>
<td>TKneobpA</td>
<td>124</td>
<td>10</td>
</tr>
<tr>
<td>Pol2sneobpA</td>
<td>324</td>
<td>25</td>
</tr>
<tr>
<td>Pol2neobpA</td>
<td>788</td>
<td>61</td>
</tr>
<tr>
<td>pPNT</td>
<td>260</td>
<td>3</td>
</tr>
<tr>
<td>PGKneobpA</td>
<td>940</td>
<td>72</td>
</tr>
</tbody>
</table>

**Table 4.1** Various knockout construct vectors with the average number of resistant colonies and their relative efficiency (taken from Gene Targeting – A Practical Approach.).

The Cre/lox-directed engineering uses a construct containing loxP sites which introduced the sites into a specific chromosomal location by homologous recombination in ES cells and then deliver Cre to recombine them as required. By using this method, any embryonic lethality can be avoided so the function of the gene in the adult can be examined (Araki et al., 1995; Rajewsky et al., 1996; Sauer et al., 1998; Metzger and Feil, 1999; Metzger and Chambon, 2001; Ryding et al., 2001). The system will be of most use for disruption of Cds2, where expression of the transgene will not be initiated until adulthood. It is possible to use specific Cre, whereby activity could be concentrated in a specific tissue (table 4.2), by introducing Cre driven by a tissue specific promoter, such as expression in the retina via a rhodopsin driven Cre. The specific Cre recombinases can be introduced at the transfected cell stage for immediate gene disruption effect, or can be introduced by breeding the inactive transgene containing mice with specific Cre-expressing mice, to produce progeny lacking the specific gene in a specific tissue. This procedure gives insight into the
Figure 4.2 The Cre/lox site-specific recombination system.

Cre promotes site-specific recombination between two 34bp loxP recognition sequences (triangles). (a) The loxP sequence consists of two 13bp inverted repeats (horizontal arrows) flanking an 8bp asymmetric core region that confers an overall directionality. The two sites cleaved by Cre are indicated by vertical arrows. (b) Recombination between two loxP sites inserted into the same DNA molecule (intramolecular recombination) in opposite orientation leads to inversion of the intervening DNA segment, whereas (c) recombination between directly repeated loxP sites results in excision of the flanked DNA region (circular product) leaving one loxP site behind. When the loxP sites are located on separate DNA molecules (intermolecular recombination), (c) DNA integration or (d) translocation can be achieved.
functional role of the gene under study without introducing the effect of embryonic lethality.

4.7 Conclusions

In this thesis the isolation and characterisation of both homologues of the *Drosophila eye-cds* gene have been reported. The *Drosophila eye-cds* gene plays a key role in phototransduction and when mutated results in light induced retinal degeneration. In addition to the characterisation of these genes, an expression study was performed and it was found that murine *Cds1* is expressed only in a small number of adult tissues in comparison to *Cds2* which was expressed ubiquitously in both embryonic and adult tissues. More specifically in the eye, *Cds1* is confined to the inner segments of the retina whereas *Cds2* is expressed in the cornea, cilary body and outer nuclear layer of the retina. In an attempt to rescue the *cds* mutant fly phenotype, it was found that the cDNA construct approach was not always successful and evidence suggests that genomic DNA may be required to rescue the phenotype. In the production of a knockout model of *Cds2* it was found that the region of DNA used for the construct was effectively killing the ES cells. Hence, new strategies would have to be adopted if the functional roles of the murine *Cds* genes are to be elucidated.
<table>
<thead>
<tr>
<th>Strain</th>
<th>Promoter</th>
<th>Main tissue of recombination</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tissue-Specific Cre mice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mo-Cre</td>
<td>αA-crystallin</td>
<td>Eye Lens</td>
</tr>
<tr>
<td>CaMKIIα-Cre</td>
<td>Calcium/calmodulin-dependent protein kinase IIα</td>
<td>Forebrain (CA1 pyramidal cells)</td>
</tr>
<tr>
<td>PO-Cre</td>
<td>PO gene</td>
<td>Schwann cells</td>
</tr>
<tr>
<td>POMC-Cre</td>
<td>Pro-opiomelanocortin</td>
<td>Pituitary gland (intermediate lobe)</td>
</tr>
<tr>
<td>IRBP-Cre</td>
<td>Interphotoreceptor retinoid binding protein</td>
<td>Retina (photoreceptor cells)</td>
</tr>
<tr>
<td>Wnt1-Cre</td>
<td>Wnt-1</td>
<td>Nervous system</td>
</tr>
<tr>
<td>En2-Cre</td>
<td>Engrailed-2</td>
<td>Nervous system</td>
</tr>
<tr>
<td>Lck-Cre</td>
<td>Proximal lck</td>
<td>T cells</td>
</tr>
<tr>
<td>CD19-Cre</td>
<td>CD19 (knock in)</td>
<td>B cells</td>
</tr>
<tr>
<td>αMHC-Cre</td>
<td>α myosin heavy chain</td>
<td>Heart (ventricular myocytes)</td>
</tr>
<tr>
<td>MLC2v-Cre</td>
<td>Myosin light chain 2v (knock in)</td>
<td>Heart (ventricular myocytes)</td>
</tr>
<tr>
<td>MCK-Cre</td>
<td>Muscle creatine kinase</td>
<td>Skeletal muscle, heart</td>
</tr>
<tr>
<td>RIP-Cre</td>
<td>Insulin</td>
<td>Pancreas (β cells)</td>
</tr>
<tr>
<td>Alb-Cre</td>
<td>Albumin</td>
<td>Liver</td>
</tr>
<tr>
<td>WAP-Cre</td>
<td>Whey acidic protein</td>
<td>Mammary gland, brain</td>
</tr>
<tr>
<td>BLG-Cre</td>
<td>β-lactoglobulin</td>
<td>Mammary gland</td>
</tr>
<tr>
<td>AP2-Cre</td>
<td>Adipose protein 2</td>
<td>Adipose tissue</td>
</tr>
<tr>
<td>K5-Cre</td>
<td>Keratin 5</td>
<td>Skin (basal keratinocytes)</td>
</tr>
<tr>
<td><strong>Inducible Cre mice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mx-Cre</td>
<td>Mx1 (interferon-inducible)</td>
<td>Widespread</td>
</tr>
<tr>
<td>CMV-tTA/tetO-Cre</td>
<td>TetO/cytomegalovirus minimal 9tetracycline</td>
<td>Muscle, skin</td>
</tr>
<tr>
<td><strong>Tamoxifen-activated Cre Mice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMV-CreERβ</td>
<td>Cytomegalovirus</td>
<td>Widespread</td>
</tr>
<tr>
<td>Wnt1-CreERβ</td>
<td>Wnt-1</td>
<td>Nervous system</td>
</tr>
<tr>
<td>Eμ/SV40-CreERβ</td>
<td>Ig heavy chain enhancer/SV40 early minimal promoter</td>
<td>B cells</td>
</tr>
<tr>
<td><strong>RU86-activated Cre mice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CaMKIIα-CrePR1</td>
<td>Calcium/calmodulin-dependent protein kinase IIα</td>
<td>Brain (cortex, hippocampus)</td>
</tr>
<tr>
<td>Thy1-CrePR1</td>
<td>thy-1</td>
<td>Brain</td>
</tr>
</tbody>
</table>

**Table 4.2** Examples of tissue-specific and inducible Cre transgenic mice (Metzger and Feil, 1999).
CHAPTER FIVE
FUTURE WORK

There are a number of questions raised by the work in this thesis and future experiments on this project should be set out to answer them. Through the work carried out and presented here it is apparent that the genes are far more complex than first anticipated. If the exciting goal of producing animals lacking the *Cds1* or *Cds2* gene is to be met, then some further studies on these genes are required.

5.1 Gene structure studies

The potential presence of splice variants in both genes has affected the outcome of several of the experiments. If the *Drosophila* rescue experiments (more detail is given in section 5.5), are to proceed to identify the mouse *Cds* gene that is most similar in function to the *Drosophila* eye-*cds*, then any splice variants are required to be confirmed and the data considered when making a new rescue construct. The splice variants could be identified using a panel of different tissue specific cDNAs as template in a PCR reaction with primers designed to the start and stop codons.

In order to form viable knockout constructs (more detail given in section 5.6), the intronic sequence from each gene is required to be examined for enhancers, promoters, or any other sequence motif which may affect the outcome of the construct’s viability. As reported earlier the gene fragments chosen previously to form the construct have been shown to kill the ES cells and the reason for this needs to be identified prior to the formation of a new construct.

To date the sizes of the full-length transcripts of mouse *Cds1* and *Cds2* remains unknown. In order to identify the transcript lengths, 5' and 3' RACE (rapid amplification of cDNA ends) should be performed to identify the size of the 3' and 5' UTRs. The 5' RACE experiment uses a primer which is designed to allow amplification from dC-tailed cDNA. Using this primer and two gene specific primers a nested PCR can be performed to amplify the 5' end. The same principles apply for 3' RACE only using an oligo dT primer which will bind to the poly A+ tail of the gene. In both cases mouse brain cDNA should be used as template. The information gathered about the
splice variants and the transcript lengths, should go some way to confirm the Northern blot findings in section 3.6.2.

5.2 Northern blot studies
The Northern blots produced in section 3.6.2 implied splice variants existed for each of the Cds genes. This information was valuable in deducing some reasons to explain the outcome of the fly rescue experiments. To finalise the data a new blot should be prepared, using poly A⁺ RNA with greater purity than the previous attempt to verify the data reported. Ideally a commercial blot should be used to ensure a constant quantity and quality of RNA used, unfortunately as yet there are no blots available that contain RNA from eye. The probes used on the blots should also be reconsidered in light of splice variant information that might be found.

5.3 In situ Hybridisation studies
Possibly the most exciting data to be reported in this thesis was the expression patterns found in the adult tissue for Cds1 and Cds2. It would be interesting and useful to pursue the in situ hybridisation studies in timed embryos. The data that would be gathered from the embryos section would give a good idea of the tempo-spatial expression of Cds2 in the development of the mouse. The developmental data would be useful when manipulating the expression of a knockout construct in a future Cds2 mouse model, to influence the point at which the disrupted gene is expressed to avoid certain critical time points or allow expression only in certain adult tissues. In addition to using the same probes to carry out the experiments listed above, any splice variants which were identified could be used to form riboprobes to screen both the embryos and the adults to isolate the location of expression of particular transcripts. Human sections could also be used to examine the expression of the Cds genes to proceed further with the in situ hybridisation studies. The data gathered from these studies would show a much more in depth time course for Cds gene expression and may help deduce a possible functional role.

5.4 Protein studies
In this thesis attempts to find the subcellular location of the Cds2 protein in mammalian cells resulted in cells which appeared to be apoptotic. This study would prove
interesting if the subcellular location was deduced as this may imply a functional role for the protein. The first experiment that is required to be performed is the establishment of the cause of cell death. The problem could be solved by TdT-mediated dUTP digoxigenin nick end labelling (TUNEL) staining of the transfected cells which would identify truly apoptotic cells. If they proved to be apoptotic, the most likely cause is endoplasmic reticulum overload response due to mis-folded protein that results in a stress signal to the nucleus and culminates in apoptosis. This problem however, may be limited to mammalian cells, so yeast and bacterial expression systems should be transfected with the myc-tagged construct. If the outcome was as before then it would most likely be assumed that the myc-tag is interfering sufficiently with the protein folding to culminate in apoptosis. By the production of native antibodies raised against Cds1 and Cds2, the problems with the tag could be eliminated and the subcellular location of the protein determined and hence would identify the subcellular structures that had interactions with the Cds proteins.

5.5 cdr mutant Drosophila rescue studies
In chapter 3.7 an attempt to rescue the mutant cdr fly phenotype was performed. It would still be very interesting to find out which of the mouse Cds genes had a sufficiently similar function that they could rescue the mutant fly phenotype, and hence imply a functional role for the gene in the mouse. Taking into consideration all the data that has arisen since the first rescue construct was formed, future construct should be prepared not from cDNA, but instead from genomic DNA. In order to perform this, either the existing PAC clones (detailed in section 3.2 and 3.4) will have to be characterised and the relevant regions cloned, or new genomic phage clones will have to be identified. This again produces another question that needs to be addressed. How large is intron 1 of Cds2? At the moment the answer is possibly up to 43kb. Probably only when the sequence of the mouse PAC clone 439-n24 becomes available online will the size be deduced. With an intron of this size, the sequence data derived will be very interesting. As has been shown recently, the introns of one gene may contain whole different genes which either run in a different direction or in a different frame from the original (Halford et al., 2001). It will be interesting to analyse Cds2 intron 1 for the presence of other genes, enhancers or promoters that may have an effect on the expression of the mouse gene in other species.
5.6 Knockout mouse models.
Some very interesting data arose from this project which showed that many genes are far more complex than often originally anticipated. A very exciting aspect of the project was the production of the knockout mouse model of Cds2. This study showed that the DNA used in these studies should be analysed closely as was seen in this project, the DNA used cause the ES cells to die. Data from this project showed that in order to form a viable construct various new considerations are required. The future work to more closely examine the intron structure will hopefully elucidate the reason why the first construct resulted in ES cell death. However, should the construct be remade with the MasterBlot™ data in mind, the knockout model should be made taking advantage of the Cre/lox system (described in chapter 4). This system will allow the activation of the transgene in a more controlled manner to hopefully avoid embryonic lethality due to the likelihood that Cds2 is a developmentally regulated gene. There may be more advantages to using Cds1 to form a knockout construct, the main advantage being the problems of embryonic lethality expected from the Cds2 construct would be avoided since the MasterBlot™ showed no sign of embryonic expression when probed with Cds1. The data gathered from PAC clone 572-a24, should identify the most appropriate regions of Cds1 for the formation of the knockout construct. The pPNT vector should not be used in future studies, but instead pGKNEO (table 4.1) should be used as this vector appears to be more efficient.

The final issue to be addressed from work performed in this thesis, is the screening of patients with diseases which have ocular phenotypes and a chromosomal map position close to that of CDS1 or CDS2 (4q21 or 20p13). Currently there are no diseases of that nature described that map to a position close to 4q21, but this should be made an ongoing study. Even although CDS2 has been eliminated as the causative gene for CHED2, more possible candidate genes with ocular phenotypes should be screen when disease loci arise. Panels of RP patients with no known causative gene should also be screened.
REFERENCES


gene with wide tissue expression and identification of embedded and flanking genes on chromosome 1q43. *Genomics*, 72, 203-208.


287


288


Vollrath, D., Feng, W., Duncan, J. L., Yasumura, D., D'Cruz, P. M., Chappelow, A., Matthes, M. T., Kay, M. A., LaVail, M. M. (2001). Correction of the retinal dystrophy phenotype of the RCS rat by viral gene transfer of MerTK. Proceedings of the National Academy for Science USA, 98, 12584-12589.


APPENDIX 1