Characterising *Cfm2*, a Novel Gene Involved in Vertebrate Development

Samuel James Capon

A thesis submitted for the partial fulfilment of the requirements for the degree of Master of Science

University of Otago
Dunedin
New Zealand

August 2011
**Top left:** a stage 40 *Xenopus laevis* tadpole expressing β-galactosidase on the left hand side of its body only.

**Top right:** a haematoxylin (nuclei) and eosin (cytoplasm) stained transverse section through the heart of a wild type stage 40 *Xenopus laevis* tadpole.

**Bottom:** endochondral ossification occurring in the forelimb (left) and hindlimb (right) of a wild type *Xenopus laevis* froglet stained with alcian blue (cartilage) and alizarin red (bone).
Abstract

The Cfm2 gene is conserved in vertebrates and is linked to various developmental anomalies in humans. This gene is a parologue of another novel gene named Cfm, for caudal forebrain and midbrain. In the mouse, Cfm was found to be expressed in the neuroectoderm, which forms the future caudal forebrain and midbrain, and also in the optic rudiment, first pharyngeal arch, Rathke’s pouch, tongue muscle, lung, inner wall of the alimentary canal, genital tubercle and peripheral nerves (Hirano et al., 2005). Also in the mouse, Cfm2 was found to be expressed in the presomitic mesoderm of the segmenting somite, the optic nerve, otic capsule, peripheral nerves, tegmentum, lung and tongue muscle (Hirano et al., 2005). More recently a yeast-two-hybrid screen found that FAM101A, the human orthologue of the CFM2 protein, binds to filamin A (FLNA) (Gay et al., 2011). Most instances of the otopalatodigital (OPD) syndrome spectrum of disorders are associated with mutations in FLNA, the gene encoding FLNA, however in some instances no such mutations are found (Robertson, 2007). This has led to a hypothesis implicating CFM2, the protein product of the Cfm2 gene, in the OPD syndrome spectrum of disorders through its interactions with FLNA.

In silico analysis of the CFM and CFM2 proteins revealed this family to be very highly conserved in vertebrates with orthologues found in a number of vertebrate Classes. Orthologues were also identified in the lancelet, a primitive chordate with an ancestral relationship to the vertebrates. Phylogenetic analysis of these orthologues found the CFM and CFM2 proteins to form two distinct clades in the resulting tree, representing the CFM and CFM2 sequences respectively. These results strongly support the notion of the CFM and CFM2 sequences having arisen from a duplication event in a common ancestor of the Chordates.

Reverse transcription PCR (RT-PCR) found that the Cfm gene is expressed continuously in Xenopus laevis development, and similarly in the development of the limb in Xenopus. In situ hybridisation found this expression to be located to the lateral plate mesoderm, neural tube floor plate, hypochord and the developing proctodeum at stage 30. Cfm expression was detected in the epidermal layer of the hindlimb during limb development. RT-PCR found that Cfm2 expression is much more restricted during development. In situ
hybridisation found $Cfm2$ expression to be localised in the developing olfactory organ, lateral line precursors, pronephros and the branchial arches at stage 40 in $Xenopus$ development. In the limbs, $\textit{in situ}$ hybridisation revealed $Cfm2$ to be expressed at the joints of the bones of the limb, in a manner consistent with the sequential formation of these structures. No overlapping expression was found between these genes, consistent with the notion of independent evolution in these two genes after the aforementioned duplication event.

Down-regulation of $Cfm2$ using a morpholino oligonucleotide (MO) had diverse effects on $Xenopus$ development. $Cfm2$ MO injected tadpoles developed at a slower rate than their wild type siblings, showed less body pigment and had a number of developmental defects. Among the developmental defects observed were craniofacial and axial malformations consistent with the phenotypes observed in the OPD syndrome spectrum of disorders. However, the most consistent phenotype seen was the irregular development of the ventral fin in $Xenopus$. Exogenous expression of $Cfm$ or $Cfm2$ was also found to have severe effects on $Xenopus$ development, with exogenous expression of $Cfm2$ found to be more severe than $Cfm$.

The expression patterns of $Cfm2$ and the developmental effects seen in $Cfm2$ knockdown in $Xenopus$ appear to confirm an association between $Cfm2$ and the OPD syndrome spectrum of disorders. However, this association is not complete and will require further analysis and understanding of the function of $Cfm2$ in vertebrates. Furthermore, the results from this research have led to a proposed role for $Cfm2$ as an important regulator in neural crest cell migration during $Xenopus\ laevis$ development.
Acknowledgements

First and foremost, I must say a huge thank you to Dr. Caroline Beck. Her knowledge, patience, advice and general thoughtfulness have guided me for the last two years and have made my transition from a clueless young grad student into a budding novice relatively painless and enjoyable.

I would also like to thank the other members of the Beck lab, past and present, it has been fun sharing my experiences with you and having someone to talk to during those long wash steps. To Matthew Downes, Ken Miller and Dr. Martyn Kennedy I owe you all at least one more thank you for your help with the various aspects of histology, photography and phylogenetics.

Last but certainly not least, I would like to thank my friends and family. Especially my parents, without your support, particularly during these last two years, I would never have been able to pursue my interest in science. Although I can never thank you enough, I can promise that there will not be another thesis for you to edit for least a few more years.
Table of Contents

Abstract

Acknowledgements

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Chapter 1  Introduction

1.1  Vertebrate development
1.2  Vertebrate limb development
1.3  Somitogenesis
1.4  Osteogenesis
1.5  Filamin A
1.6  The otopalatodigital spectrum of disorders
1.7  Cfm2
1.8  Aims

Chapter 2  Materials and Methods

2.1  Materials
2.2  In silico analysis
  2.2.1  Orthologues of CFM and CFM2
  2.2.2  Phylogenetic analysis
2.3  Preparing Xenopus laevis embryos
2.4  Raising Xenopus laevis tadpoles
2.5  RNA extraction
2.6  Reverse transcription

Chapter 3  Results

3.1  Experimental validation of Cfm2
3.2  Osteogenesis in Xenopus
3.3  In silico analysis of Cfm2
3.4  Phylogenetic analysis of Cfm2
3.5  Preparing Xenopus laevis embryos
3.6  Raising Xenopus laevis tadpoles
3.7  RNA extraction
3.8  Reverse transcription

Chapter 4  Conclusion

4.1  Summary
4.2  Future directions
4.3  Acknowledgements

List of Figures

List of Tables

List of Abbreviations
2.7 Polymerase chain reaction
   2.7.1 Standard polymerase chain reaction
   2.7.2 Colony screening using polymerase chain reaction
2.8 Whole mount in situ hybridisation
   2.8.1 RNA probe synthesis
   2.8.2 Probe hybridisation
   2.8.3 Probe detection
2.9 Gene cloning
   2.9.1 Site directed mutagenesis of mismatched Xenopus laevis Cfm2
   2.9.2 Sub-cloning sequences into the pCS2+ vector
   2.9.3 Ambystoma mexicanum Cfm2
2.10 Transforming and culturing competent cells
   2.10.1 Transformation
   2.10.2 Cell culturing
2.11 In vitro transcription
2.12 Microinjection
   2.12.1 Morpholino oligonucleotide mediated knockdown
   2.12.2 Cfm2 morpholino/β-galactosidase mRNA co-injections
   2.12.3 Morpholino rescue
   2.12.4 Cfm and Cfm2 overexpression
2.13 Histology
   2.13.1 Wax embedding
   2.13.2 Sectioning and mounting
   2.13.3 Staining
2.14 Fixing and photography
   2.14.1 Fixation
   2.14.2 Bleaching
   2.14.3 Photography
2.15 Statistical analysis
Chapter 3  Homology of *Xenopus laevis* Cfm and Cfm2

3.1  Introduction
3.2  Orthologues of CFM and CFM2
3.3  Sequence conservation
3.4  Phylogenetic relationships of CFM and CFM2
3.5  Summary

Chapter 4  Expression patterns of *Xenopus laevis* Cfm and Cfm2

4.1  Introduction
4.2  RNA extraction
4.3  PCR analysis
4.4  *In situ* hybridisation
   4.4.1  Cfm2 expression in *Xenopus laevis*
   4.4.2  Cfm expression in *Xenopus laevis*
   4.4.3  Cfm2 expression in the axolotl, *Ambystoma mexicanum*
4.5  Summary

Chapter 5  Misexpression of Cfm2 in *Xenopus laevis*

5.1  Introduction
5.2  Cfm2 knockdown
5.3  Developmental markers in morpholino injected tadpoles
5.4  Morpholino knockdown rescue
   5.4.1  Sub-cloning and *in vitro* transcription
   5.4.2  mmCfm2 mRNA rescue
5.5  Overexpression of Cfm and Cfm2
5.6  Summary

Chapter 6  Discussion

6.1  Summary of results
6.2  The CFM/CFM2 family of proteins is conserved in vertebrates
6.3  Cfm2 is expressed in the head, kidney and limb joints during *Xenopus* development
6.4 Cfm2 knockdown has diverse effects on Xenopus development 100
6.5 Developmental markers reveal defects in Cfm2 morpholino injected tadpoles 102
6.6 Exogenous expression of Cfm and Cfm2 results in a lethal phenotype 104
6.7 Conclusions and future direction 105

References  108

Appendices  117
I  NCBI sequence dataset 117
II  Metazome sequence dataset 118
III  Full protein sequence alignment 119
IV  Primer sequences and PCR programs 124
V  5’ phosphorylated primers 127
VI  Plasmid maps 128
VII  List of solutions used 134
VIII  Xenopus laevis stage series 136
**List of Figures**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Figure 1.1</strong></td>
<td>Vertebrate limb development</td>
<td>5</td>
</tr>
<tr>
<td><strong>Figure 1.2</strong></td>
<td>Waves of gene expression in the PSM control boundary formation</td>
<td>8</td>
</tr>
<tr>
<td><strong>Figure 1.3</strong></td>
<td>Endochondral ossification</td>
<td>12</td>
</tr>
<tr>
<td><strong>Figure 1.4</strong></td>
<td>Structural representation of the FLNA dimer</td>
<td>14</td>
</tr>
<tr>
<td><strong>Figure 1.5</strong></td>
<td><em>Xenopus</em> CFM and CFM2 amino acid sequences</td>
<td>17</td>
</tr>
<tr>
<td><strong>Figure 3.1</strong></td>
<td>Multiple sequence alignment subset</td>
<td>49</td>
</tr>
<tr>
<td><strong>Figure 3.2</strong></td>
<td>CFM/CFM2 protein family phylogenetic tree</td>
<td>51</td>
</tr>
<tr>
<td><strong>Figure 4.1</strong></td>
<td>RNA integrity</td>
<td>56</td>
</tr>
<tr>
<td><strong>Figure 4.2</strong></td>
<td>PCR analysis of gene expression in development</td>
<td>57</td>
</tr>
<tr>
<td><strong>Figure 4.3</strong></td>
<td><em>Cfm2</em> expression in the developing tadpole</td>
<td>59</td>
</tr>
<tr>
<td><strong>Figure 4.4</strong></td>
<td><em>Cfm2</em> expression in the developing limb</td>
<td>61</td>
</tr>
<tr>
<td><strong>Figure 4.5</strong></td>
<td><em>Cfm</em> expression in the developing tadpole (I)</td>
<td>64</td>
</tr>
<tr>
<td><strong>Figure 4.6</strong></td>
<td><em>Cfm</em> expression in the developing tadpole (II)</td>
<td>65</td>
</tr>
<tr>
<td><strong>Figure 4.7</strong></td>
<td><em>Cfm</em> expression in the developing limb</td>
<td>66</td>
</tr>
<tr>
<td><strong>Figure 5.1</strong></td>
<td>Morpholino injection survival rates</td>
<td>70</td>
</tr>
<tr>
<td><strong>Figure 5.2</strong></td>
<td>Phenotypic characteristics of microinjected tadpoles</td>
<td>72</td>
</tr>
<tr>
<td><strong>Figure 5.3</strong></td>
<td>Morpholino injection phenotypes</td>
<td>75</td>
</tr>
<tr>
<td><strong>Figure 5.4</strong></td>
<td>Analysis of developmental markers in morpholino injected tadpoles</td>
<td>79</td>
</tr>
<tr>
<td><strong>Figure 5.5</strong></td>
<td>Morphology of morpholino injected tadpoles</td>
<td>80</td>
</tr>
<tr>
<td><strong>Figure 5.6</strong></td>
<td>mRNA production</td>
<td>83</td>
</tr>
<tr>
<td><strong>Figure 5.7</strong></td>
<td>mRNA injection survival rates</td>
<td>85</td>
</tr>
<tr>
<td><strong>Figure 5.8</strong></td>
<td>mRNA injection phenotypes</td>
<td>86</td>
</tr>
</tbody>
</table>
Table 2.1 - SuperScript® III reverse transcription ingredients (I) 24
Table 2.2 - SuperScript® III reverse transcription ingredients (II) 24
Table 2.3 - PCR reagents 25
Table 2.4 - Colony screening PCR reagents 26
Table 2.5 - Restriction enzyme reaction 26
Table 2.6 - *In situ* hybridisation probe transcription reaction 27
Table 2.7 - *Xenopus* proteinase K treatments 28
Table 2.8 - Phusion DNA polymerase reaction 31
Table 2.9 - Klenow fragment reaction 33
Table 2.10 - mRNA synthesis reaction 36
Table 2.11 - Microinjection solutions 38
Table 2.12 - Overexpression experiment injection solutions 40
Table 4.1 - RNA concentration and purity 55
List of Abbreviations

° C degrees celsius
µg microgram(s)
µL microlitre(s)
µM micromole(s)
β-gal β-galactosidase
BCIP 5-bromo-4-chloro-3-indolyl phosphate
BLAST basic local alignment search tool
BMB boehringer mannheim blocking reagent
BMP bone morphogenetic protein
bp base pair(s)
BSA bovine serum albumin
cDNA complementary DNA
CDS coding sequence
Cfm caudal forebrain midbrain gene (also known as FAM101B and RefilinB)
CFM caudal forebrain midbrain protein (also known as FAM101B and RefilinB)
Cfm2 caudal forebrain midbrain 2 gene (also known as FAM101A and RefilinA)
mmCfm2 mismatched Cfm2
CFM2 caudal forebrain midbrain 2 protein (also known as FAM101A and RefilinA)
CHAPS 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
DIG digoxigenin
DNA deoxyribonucleic acid
DNase deoxyribonuclease
dNTP deoxyribonucleotide triphosphate
DTT dithiothreitol
E embryonic day (in the context of mouse development)
E. coli escherichia coli
EDTA ethylene diamine tetraacetic acid
EGTA ethylene glycol tetraacetic acid
EST expressed sequence tag
EtOH ethanol
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLNA</td>
<td>filamin A gene</td>
</tr>
<tr>
<td>FLNA</td>
<td>filamin A protein</td>
</tr>
<tr>
<td>g</td>
<td>gram(s)</td>
</tr>
<tr>
<td>HCG</td>
<td>human chorionic gonadotrophin</td>
</tr>
<tr>
<td>KOH</td>
<td>potassium hydroxide</td>
</tr>
<tr>
<td>L</td>
<td>litre(s)</td>
</tr>
<tr>
<td>LB</td>
<td>luria broth</td>
</tr>
<tr>
<td>M</td>
<td>mole(s)</td>
</tr>
<tr>
<td>MAB</td>
<td>maleic acid buffer</td>
</tr>
<tr>
<td>MABT</td>
<td>MAB + 0.1% tween 20</td>
</tr>
<tr>
<td>MAPK</td>
<td>mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MEMFA</td>
<td>MOPS, EGTA, MgSO4, and formaldehyde</td>
</tr>
<tr>
<td>MeOH</td>
<td>methanol</td>
</tr>
<tr>
<td>mg</td>
<td>milligram(s)</td>
</tr>
<tr>
<td>mL</td>
<td>millilitre(s)</td>
</tr>
<tr>
<td>mM</td>
<td>millimole(s)</td>
</tr>
<tr>
<td>MMR</td>
<td>Marc’s modified ringers</td>
</tr>
<tr>
<td>mmMO</td>
<td>mismatched MO</td>
</tr>
<tr>
<td>MO</td>
<td>morpholino oligonucleotide</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>MQW</td>
<td>milli-Q water</td>
</tr>
<tr>
<td>MS222</td>
<td>ethyl 3-aminobenzoate methanesulfonate salt</td>
</tr>
<tr>
<td>NBT</td>
<td>nitro blue tetrazolium chloride</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram(s)</td>
</tr>
<tr>
<td>nL</td>
<td>nanolitre(s)</td>
</tr>
<tr>
<td>OPD</td>
<td>otopalatodigital</td>
</tr>
<tr>
<td>PBSA</td>
<td>phosphate buffered saline (Dulbecco A)</td>
</tr>
<tr>
<td>PBSAT</td>
<td>PBSA + 0.1 % tween 20</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PFA</td>
<td>paraformaldehyde</td>
</tr>
<tr>
<td>PSM</td>
<td>presomitic mesoderm</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>SSC</td>
<td>standard saline citrate</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris, acetic acid and EDTA</td>
</tr>
<tr>
<td>TEA</td>
<td>triethanolamine</td>
</tr>
<tr>
<td>TGF-β</td>
<td>transforming growth factor-β</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>v/v</td>
<td>volume per volume</td>
</tr>
<tr>
<td>w/v</td>
<td>weight per volume</td>
</tr>
<tr>
<td>Wnt</td>
<td>wingless/integration site-1</td>
</tr>
<tr>
<td>X-gal</td>
<td>bromo-chloro-indolyl-galactopyranoside</td>
</tr>
</tbody>
</table>
Chapter 1 - Introduction

1.1 Vertebrate development

In vertebrates, development can be divided into a number of events beginning with the fertilisation of a single cell and ending with the formation of a complex organism comprised of many cells, each unique in morphology and function and yet each possessing the same genomic copy of DNA. The phenomenal cellular diversity that is achieved during vertebrate development is regulated by the immaculately controlled expression of a number of genes in respect to both time and space.

Vertebrate development begins with the process of fertilisation, the female and male gametes combining to form the fertilised embryo consisting of a single cell that will eventually grow into a fully-fledged organism (Wolpert and Tickle, 2011). Following fertilisation the newly formed embryo will begin to divide at a rapid pace (Wolpert and Tickle, 2011). Through subsequent rounds of cellular division the embryo forms a mass of cells known as the blastula or blastosphere (Wolpert and Tickle, 2011). As the cells of the blastula continue to divide they begin to migrate to defined positions through a process known as gastrulation, to form the gastrula (Wolpert and Tickle, 2011). During gastrulation the tissue of the embryo is segregated into the three germ cell layers; the endoderm, ectoderm and mesoderm (Wolpert and Tickle, 2011). As the cells of an organism grow and divide they became increasingly more specified and fixed in their cellular fate. This occurs as early as gastrulation with the endoderm going on to form the epithelium of the digestive and respiratory tracts as well as the organs of the digestive system, the ectoderm going on to form the epidermis, neural crest and the central nervous system and the mesoderm going on to form the somites, notochord, blood and bone (Wolpert and Tickle, 2011). The embryo then continues to grow and divide, developing the organs through the process of organogenesis and then developing into its final form through the process of morphogenesis (Wolpert and Tickle, 2011).

Vital to development in vertebrates is the construction of some form of a ‘body plan’ or ‘blue print’ that holds the key information for an organism’s symmetry, the number of body segments, limbs and the other information on the fundamental features of a specific
organism’s morphology (Wolpert and Tickle, 2011). This developmental body plan is established by the early patterning of the embryo and the establishment of the dorsal-ventral and the anterior-posterior axes (Wolpert and Tickle, 2011). The body plan is established through the combined efforts of specific signalling centres in the embryo - the Spemann organiser in amphibians and Henson’s node in the chick and mice embryos - and a number of molecular signalling pathways (Wolpert and Tickle, 2011). These include the Wnt, Hedgehog, TGF-β and MAPK signalling pathways that regulate many of the key events of embryogenesis including the establishment of these axes (Wolpert and Tickle, 2011). These pathways play important roles in the processes of proliferation, differentiation, apoptosis and, ultimately, the development of an organism (Wolpert and Tickle, 2011). They are commonly found in vertebrates\(^1\) and even some invertebrates\(^2\) such is their fundamental importance (Wolpert and Tickle, 2011). Often the misregulation of these signalling pathways is found in diseases where developmental control is lost, such as cancer, reinforcing their importance to correct development (Logan and Nusse, 2004). The Wnt and TGF-β pathways are particularly important to the establishment of the body axes in frogs, chick, zebrafish and mice (Petersen and Reddien, 2009). Although other influences such as maternal factors, cell-cell interactions and external influences, such as gravity, can also play their part in the establishment of these axes during embryogenesis in the various vertebrate species (Wolpert and Tickle, 2011). More recent research has also uncovered a novel role for the small non-coding microRNAs in the patterning of the embryo through the complex regulation of specific genes involved in development (Harfe, 2005).

By the end of gastrulation the basic vertebrate body plan has been laid down (Wolpert and Tickle, 2011). The identity of the body segments is then established through regulated gene expression involving the \(Hox\) genes (Wolpert and Tickle, 2011). The \(Hox\) genes are a set of genes which cluster together on the chromosomes (Pearson et al., 2005). These clusters are conserved from nematodes through to humans (Pearson et al., 2005). They are expressed in

\(^1\)Vertebrates are members of the subphylum Vertebrata. These members are Chordates defined by the presence of a spinal column (Wolpert and Tickle, 2011).

\(^2\)Invertebrates comprise all animals which are not classified as Vertebrata. This group represents around 97% of all known animals (Wolpert and Tickle, 2011).
a segmental pattern along the anterior-posterior axis corresponding to their position in the cluster with the most anteriorly expressed genes located at the 5’ end of the cluster, and the most posteriorly expressed genes located at the 3’ end of the cluster (Pearson et al., 2005). The Hox genes are responsible for ensuring the vertebrae develop their correct positional identity (Wolpert and Tickle, 2011) and, similarly, for ensuring the vertebrate limb develops in the correct position and with the correct identity (Tickle and Eichele, 1994).

As development continues, each of the body segments continue to develop towards their terminal identity in an ongoing process that involves the tightly regulated expression of many genes and genetic pathways. Two specific examples of these processes, somitogenesis and limb development, are described in more detail below.

1.2 Vertebrate limb development

As shown in figure 1.1, the vertebrate limb has three axes of symmetry: the anterior-posterior axis, running parallel to the body from thumb to little finger; the dorsal-ventral axis, running from the back to the front (palm) of the limb; and the proximal-distal axis which runs from the base of the limb (shoulder) to the tip (phalanges).

Vertebrate limb development begins with the formation of a limb bud at the site of the future limb (Wolpert and Tickle, 2011). The location of the limb bud is precisely regulated. This is in part due to the expression of Hox genes, which, as mentioned before, are responsible for determining the identity of the body segments of the organism along the anterior-posterior body axis (Tickle and Eichele, 1994). Counter intuitively, it has been found that the limb bud forms as the result of a decrease in proliferation in the cells adjacent to the bud rather than an increase in proliferation at the site of the limb bud (Tickle and Eichele, 1994). Although the Hox genes are thought to establish where a limb bud may form, it is the fibroblast growth factors (FGFs) that induce the formation of the limb bud (Johnson and Tabin, 1997). This has been demonstrated in experiments where beads soaked in FGFs were shown to induce the formation of limb buds when placed at interlimb regions of the lateral plate mesoderm (Cohn et al., 1995).
The early limb bud consists of a mass of mesenchymal cells that are derived from the lateral plate mesoderm, surrounded by a layer of ectodermal epithelial cells (Niswander, 2003). The muscle, nerves and blood vessels of the limb form from mesenchymal cells that migrate into the limb from the somites whereas all other structures of the limb, including the skeletal elements, originate from the mesenchymal cells of the lateral plate mesoderm (Johnson and Tabin, 1997). Once the limb bud has formed it begins to grow distally at a rapid pace through cell proliferation at the tip of the bud where the apical ectodermal ridge (AER) is located (Tickle and Eichele, 1994).

As the limb bud grows mesenchymal condensations form (Wolpert and Tickle, 2011). These condensations are the precursors of the cartilage that will develop into the bones of the limb through endochondral ossification (Marieb and Hoehn, 2006) which will be described in more detail in chapter 1.4. In early limb development mesenchymal condensations of the proximal structures form at the proximal base of the limb (Wolpert and Tickle, 2011) (Fig 1.1 B). As the limb continues to grow distally the more distal structures of the limb begin to form, again through the formation of mesenchymal condensations (Wolpert and Tickle, 2011). Eventually the distal region of the limb takes on a flattened, plate-like appearance as the digits of the limb form (Tickle and Eichele, 1994). In the embryo of the chick, and other mammals, the digits are then separated through programmed cell death by apoptosis in the interdigital regions (Zuzarte-Luis and Hurle, 2005). However, in amphibians digit separation is due to the digits growing more than the interdigital regions of the limb (Cameron and Fallon, 1977).

There are two key signalling regions in the vertebrate limb (Tickle and Eichele, 1994). The first of these is the AER that, as mentioned previously, is located at the distal tip of the limb bud (blue, Fig 1.1 A and B). The other key signalling region is the zone of polarising activity, or ZPA (sometimes referred to as the polarizing region) that is located on the posterior side of the developing limb (red, Fig 1.1 A and B). These two regions are crucial to the correct formation of the vertebrate limb (Tickle and Eichele, 1994).
Figure 1.1. Vertebrate limb development. The developing wing (forelimb) of a chicken embryo from the early limb bud stage (A) with the two signalling regions (apical ectodermal ridge, blue, and the zone of polarising activity, referred to here as the polarizing region, red) and the progress zone (yellow) indicated. As the limb develops the bud grows distally and the progenitors of the limb structures begin to form in a proximal to distal pattern (B, C). Development continues and the structures of the limb are established (C). The limb then continues to grow until the final form of the limb is achieved, not shown. The three axes of symmetry in the vertebrate limb are shown in the top far-right corner as they occur in this figure. Figure modified from Wolpert and Tickle, 2011.

The AER is a thickened layer of ectodermal cells that is required to mediate the outgrowth of the limb and maintain the progress zone, a mass of mesenchymal cells located immediately below the AER, through the secretion of FGFs (Tickle and Eichele, 1994). Surgical removal of the AER from a limb bud results in a truncated limb (Summerbell, 1974). If the AER is removed later in development the proximal structures of the limb will form normally, but the limb will lack distal structures (Summerbell, 1974). If the AER is removed at an earlier stage in limb development the limb will lack both the distal and proximal structures (Summerbell, 1974). The importance of the FGFs to AER signalling has been shown by removing the AER and replacing it with FGF soaked beads (Niswander et al., 1993). In such experiments the limbs continued to develop normally (Niswander et al., 1993).

The ZPA is a region of mesenchymal cells that regulates anterior-posterior patterning through the secretion of the protein sonic hedgehog (SHH) (Niswander, 2003). Experiments where ZPA cells have been grafted onto the anterior of a developing limb result in the mirror-image duplication of the limb digits (Tickle, 1981). A similar affect is seen when SHH protein or SHH producing cells are similarly applied to the anterior of a developing limb (Riddle et al., 1993) demonstrating the importance of SHH to the ZPA and anterior-posterior axis.
The dorsal-ventral axis is established by signals from the limb ectoderm (Niswander, 2003). Experiments where the ectoderm has been rotated 180° around the mesenchyme results in a corresponding reversal of the dorsal-ventral axis (MacCabe et al., 1974). Wnt7a expression in the dorsal limb ectoderm induces the expression of the transcription factor Lmx1b (Chen and Johnson, 1999). This appears to be critical to establishing the dorsal polarity as the absence of Wnt7a leads to the appearance of a bi-ventral limb (Chen and Johnson, 1999). In the ventral ectoderm the expression of the transcription factor engrailed1 (En1) represses Wnt7a expression, the absence of En1 leads to bi-dorsal limb appearance (Chen and Johnson, 1999). En1 expression is induced in the ventral ectoderm by BMP signalling, the loss of BMP signalling also leads to bi-dorsal limb appearance (Chen and Johnson, 1999).

Although the establishment of the anterior-posterior and dorsal-ventral axis is well supported, the evidence for the establishment of the proximal-distal axis is less conclusive (Niswander, 2003). As a result two models have been described to explain proximal-distal patterning of the limb. The progress zone model proposes that, as cells leave the AER during limb outgrowth the proximal-distal axis is established with cells that spend longer in the AER becoming the more distal structures (Niswander, 2003). The second model, proposed by Dudley et al. (2002), suggests that opposing FGF and retinoic acid gradients pattern the limb bud at an early stage, establishing the complete proximal-distal axis early in limb development. Both models can be interrupted to explain the formation of the truncated limb as a result of AER removal mentioned earlier. Proponents of the progress zone model suggest that removing the AER at later stages results in the patterning of the distal region of the limb being lost, subsequently the distal structures of the limb do not form (Niswander, 2003). Proponents of the model proposed by Dudley et al. (2002) suggest that, at early stages in limb development, all the proximal and distal progenitors are located in the AER due to the compact nature of the limb bud, whereas at later stages only the distal progenitors are located in the AER. Thus, removing the AER at an early stage results in the loss of the proximal and distal structures whereas at later stages only the distal structures are lost (Niswander, 2003). Neither of these models has been conclusively proven and the mechanism that establishes the dorsal-ventral axis therefore remains a matter of debate.
1.3 Somitogenesis

Somitogenesis is the process by which the somites develop in vertebrates. Somites form as blocks of segmented mesoderm lying each side of the notochord\(^1\); these blocks differentiate into the skeletal muscle, the axial skeleton and parts of the dermis (Pourquié, 2001). The somites begin to form soon after gastrulation through segmentation of the presomitic mesoderm (PSM) along the anterior-posterior axis (Pourquié, 2001). Somitogenesis takes place periodically with each new pair of somites forming at the rostral-most end of the PSM (Pourquié, 2001), this occurs every 90 minutes in the chick embryo but the timing varies in a species specific manner with frogs taking 120 minutes per pair of somites (Dequéant and Pourquié, 2008). The periodic formation of somites continues until the full complement of somites has been formed (Pourquié, 2001). The total number of somites is strictly regulated but also species dependent, in the chick there are 55 somites whereas mice have 65 and snakes upwards of 300 somites (Gomez et al., 2008).

To understand the mechanisms that control segmentation of the PSM, experiments were conducted where the PSM was grafted to various ectopic locations (Kieny et al., 1972). It was found that there was no affect on the segmentation pattern of the PSM suggesting that segmentation is intrinsically defined within the PSM (Kieny et al., 1972). This was further illustrated in experiments where a portion of the rostral region of the PSM was ablated by microsurgery (Kieny et al., 1972). This ablation resulted in a long pause followed by segmentation of the remaining caudal region of the PSM at the precise time expected as if the whole PSM had been left intact (Kieny et al., 1972).

A model was initially proposed by Cooke and Zeeman (1975) to explain the molecular mechanism that controls somitogenesis. The ‘clock and wavefront’ model they described involves an oscillator, or clock, controlling the periodicity of somitogenesis and a wavefront controlling the positional information of the somatic boundaries. Notch signalling has been proposed to account for the clock mechanism of this model (Palmeirim et al., 1997; Jiang et

---

\(^1\) The notochord is a rigid structure found in all Chordates that runs head-to-tail just ventral to the neural tube, the precursor to the central nervous system (Wolpert and Tickle, 2011). Further, the notochord also acts as the foundation for the formation of the spinal column (Wolpert and Tickle, 2011).
al., 2000) while a morphogen gradient formed by *fgf8* and *Wnt3a* expression forms the wavefront (Pourquié, 2011). The formation of the FGF8 gradient is established through the continued production of *fgf8* mRNA from the caudal end of the embryo (Dubrulle and Pourquié, 2004). As the embryo grows the source of *de novo* mRNA (the caudal end) moves further and further away and with time mediated decay a gradient of actively transcribed *fgf8* mRNA is established along the anterior-posterior axis (Dubrulle and Pourquié, 2004). At the same time a second morphogen gradient, consisting of Wnt3a, is produced in a similar manner (Aulehla et al., 2008). Wnt3a is produced at the caudal end of the embryo and as the embryo grows the signal moves further and further away through time mediated decay, a Wnt3a gradient is established parallel to the FGF8 gradient (Aulehla et al., 2008). When the FGF8/Wnt3a gradient dips below a threshold level, differentiation occurs and the boundaries of the somites form (Dequéant and Pourquié, 2008). Through this gradient the positional information of the segmenting somites is defined (Dequéant and Pourquié, 2008).

Oscillations in Notch signalling occurring in the caudal region of the embryo produce waves of gene expression in the PSM that then travel along the anterior-posterior axis of the PSM (Dequéant and Pourquié, 2008) as shown in figure 1.2. These waves are responsible for initiating a cascade of events at the boundary of the segmenting somite (S0), as established by the FGF8/Wnt3a gradient, that leads to the formation of the somitic boundaries (Dequéant and Pourquié, 2008).

**Figure 1.2. Waves of gene expression in the PSM control boundary formation.** Oscillations in gene expression at the caudal end of the embryo create waves of activity (blue) in the PSM that encounter the wavefront of FGF8/Wnt3a (not shown) to form the somite boundaries. Abbreviations: S0, segmenting somite; SI, most recently formed somite; SII, second-most recently formed somite. Figure replicated from Dequéant and Pourquié, 2008.
Wnt3a expression appears to link the clock and wavefront together. As mentioned previously Wnt3a has been proposed to establish the wavefront through co-expression with FGF8 (Dequéant and Pourquié, 2008). The Wnt signalling pathway interacts with the Notch signalling pathway in a manner in which oscillations in the Notch pathway are dependent on Wnt3a signalling (Aulehla et al., 2003). This suggests that Wnt3a controls oscillations in the Notch signalling pathway through oscillations in the canonical Wnt pathway (Aulehla et al., 2003). It is interesting to note that fgf8 expression in the caudal region of the embryo is also controlled by Wnt3a expression, suggesting that Wnt3a may act as the master regulator of somitogenesis (Aulehla and Hermann, 2004).

It has recently been shown that impairment of Notch signalling does not prevent segmentation in zebrafish (Özbudak and Lewis, 2008), suggesting that Notch signalling is not required for somitogenesis. Rather, it seems to be important for the formation of regular and sharp somite boundaries (Özbudak and Lewis, 2008). Contradictory evidence has been found in mice where experiments suggest that Notch signalling is indeed required for segmentation to occur (Ferjentsik et al., 2009). These contradictory results imply that the molecular oscillator may be not as well conserved as previously thought. Oscillations in the caudal region of the PSM have also been described in components of the Wnt, FGF and Notch pathways in vertebrates (Pourquié, 2011). Analysis of these oscillating genes in mice have shown that oscillation in genes related to the Notch and FGF pathways occurs mostly in opposite phase to oscillations reported in genes related to the Wnt pathway (Pourquié, 2011). This suggests that interpathway interactions are important in controlling oscillations (Pourquié, 2011). Yet still, the master regulator of these oscillations has not been identified (Pourquié, 2011).

Once the somites form, their identity is controlled by the action of the Hox genes, as mentioned earlier, and a number of important transcription factors which pattern the somites (Krumlauf, 1994). Mutations in Hox genes can lead to the vertebrae of the axial skeleton acquiring a more anterior or posterior identity (Wellik, 2007). Furthermore, it appears that Hox gene activation is linked to the segmentation clock through components of the Notch and FGF signalling (Dubrulle et al., 2001; Zákány et al., 2001) and there is also evidence that mutations in Wnt3a cause homeotic mutations along the vertebrae in mouse
embryos (Ikeya and Takada, 2001). This information suggests that the Wnt pathway is also linked to the control of Hox gene expression, adding further support to the notion that Wnt3a may be the master regulator of somitogenesis.

1.4 Osteogenesis

Osteogenesis, or ossification, is the process by which the bones form. This can occur via two processes known as intramembranous ossification and endochondral ossification (Marieb and Hoehn, 2006). Both processes begin with the condensation of mesenchymal stem cells at the site of the future bone through the regulated expression of unique cellular adhesion molecules (Kronenberg, 2003). The mesenchymal stem cells arise from three distinct lineages in vertebrates: the neural crest cells give rise to the craniofacial skeleton, the somitic mesoderm gives rise to the bones of the axial skeleton and the lateral plate mesoderm gives rise to the skeleton of the limbs (Olsen et al., 2000). With intramembranous ossification these condensations differentiate into the osteoblasts which lay down a matrix rich in type-1 collagen that forms the bone matrix (Kronenberg, 2003). However, this relatively simple process is the exception and occurs mainly in the formation of the flat bones of the skull (Kronenberg, 2003).

Most condensations proceed through endochondral ossification, shown in figure 1.3, to form the long bones of the vertebrate skeleton (Kronenberg, 2003). The major difference between the two processes is the formation of cartilage as an intermediary. In endochondral ossification the cells of the mesenchymal condensations differentiate to become chondrocytes, sometimes referred to as cartilage cells, which secrete collagen type II and the proteoglycan aggrecan (Mackie et al., 2008). The cells around the perimeter of the cluster form a fibrous membrane known as the perichondrium that is the precursor of the bone collar (Kronenberg, 2003). The chondrocytes at the center of the cartilage mould stop proliferating, but continue to grow and expand through hypertrophic growth (Kronenberg, 2003). During this hypertrophic expansion the chondrocytes begin secreting collagen type X, forming an extracellular matrix consisting of the proteoglycan aggrecan as well as type II and type X collagen (Mackie et al., 2008). Hypertrophic growth coupled with the establishment of the extracellular matrix is responsible for the initial elongation of the bone (Mackie et al., 2008). Hypertrophic chondrocytes then direct the mineralisation of this matrix before
undergoing apoptosis creating large cavities in the bone (Kronenberg, 2003). The transverse struts of the established matrix are then selectively degraded (Mackie et al., 2008). This process of degradation and cell death allows the osteoclasts, osteoblasts and blood vessels to invade the cartilage mould (Marieb and Hoehn, 2006). This establishes the primary center of ossification where the osteoblasts begin replacing the cartilage matrix with the trabeculae of the cancellous, or spongy, bone (Olsen et al., 2000). The osteoclasts assist in the degradation of the cartilage matrix and are also responsible for maintaining calcium level homeostasis during later life through bone resorption (Mackie et al., 2008).

As the primary ossification centre is being established, osteoblasts in the perichondrium begin forming the bone collar around the diaphysis which serves as a precursor to the cortical, or compact, bone (Olsen et al., 2000; Karsentry et al., 2009). The primary centre of ossification immediately begins expanding, slowly encroaching on the cartilage cells and ossifying along the way (Marieb and Hoehn, 2006). As the bone grows the secondary centres of ossification form at the epiphyses of the bone (Marieb and Hoehn, 2006), leaving only a relatively small region of cartilage between each epiphysis and the diaphysis known as the growth plate, or epiphyseal plate (Olsen et al., 2000). The growth plate allows longitudinal bone growth to occur in the growing organism though continued chondrocyte proliferation, hypertrophic growth and calcification (Marieb and Hoehn, 2006).

At the same time the bones widen to strengthen the lengthening bone (Marieb and Hoehn, 2006). This is achieved by osteoblasts located beneath the periosteum, a thin membrane that covers the outer surface of all bones except at the joints of the long bones, which secrete bone matrix on the external surface of the bone while osteoclasts on the inner surface remove excess bone (Marieb and Hoehn, 2006). The rate of bone building usually exceeds the rate of bone break down on the inner surface resulting in thick and strong bones (Marieb and Hoehn, 2006). The bones continue growing with the bones eventually becomes fully ossified during adulthood when the growth plates eventually ossify (Marieb and Hoehn, 2006).
Figure 1.3. Endochondral ossification. The process of endochondral ossification begins with the formation of mesenchymal condensation that differentiate into cartilage (shown in blue), the bone collar then forms around the diaphysis of the future bone (A). Hypertrophic chondrocytes at the primary centre of ossification encourage calcification of the extracellular matrix which then deteriorates as the chondrocytes die (B), allowing the blood vessels, osteoclasts and osteoblasts to invade the inner matrix, osteoblasts begin forming spongy bone (C). As the primary ossification centre enlarges secondary ossification centres are established in the epiphyses and the osteoclasts open up the medullary, or marrow, cavity in the centre of the diaphysis (D). Ossification continues until the only cartilage remaining is located on the epiphyseal surfaces as the articular cartilage, or at the growth plates (E). Figure modified from Marieb and Hoehn 2006.

There are a number of genes and signalling pathways involved in the process of osteogenesis. The FGF and Wnt signalling pathways have been implicated as negative regulators of chondrocyte proliferation (Karsentry et al., 2009) whereas the BMP signalling pathway appears to be a positive regulator of chondrocyte proliferation (Karsentry et al., 2009). Indian hedgehog (Ihh) also appears to be a positive regulator of chondrocyte proliferation along with parathyroid hormone related protein (PTHrP) by preventing chondrocytes becoming hypertrophic (Olsen et al., 2000). Ihh is also an important regulator in the formation of the bone collar and the trabeculae (Olsen et al., 2000). The transcription factor Sox9 is required for chondrogenesis, mutations in Sox9 cause campomelic dysplasia, a rare and severe form of dwarfism in humans (Olsen et al., 2000). Vascular endothelial growth factor (VEGF) is an important growth factor produced by the hypertrophic chondrocytes that encourages angiogenesis, a process vital to the establishment of the ossification centres (Olsen et al., 2000). A number of other important genes and transcription factors are required for the formation and maintenance of the bone such as Runx2, a transcription factor involved in the control of osteoblast differentiation (Karsentry...
et al., 2009) and members of the Dlx, bHLH, leucine zipper and AP-1 families of transcription factors involved in chondrogenesis (Goldring et al., 2006).

1.5 Filamin A
Filamin was first identified over 30 years ago as a non-muscle actin binding protein (Wang et al., 1975). Since then it has been found to be a versatile protein important to many cellular processes (Feng and Walsh, 2004). Three filamin proteins have been discovered in mammals; these have been named filamin A (FLNA), filamin B (FLNB) and filamin C (FLNC) respectively (van der Flier and Sonnenberg, 2001). The three filamin proteins are generally widely expressed during development although FLNC expression is more restricted with its expression limited to the skeletal and cardiac muscles (Feng and Walsh, 2004). Recent evidence has shown that mutations in FLNA and FLNB (the genes encoding FLNA and FLNB respectively) cause a wide range of human developmental abnormalities in the limbs, brain, bone and other organs (Feng and Walsh, 2004).

The structure of a filamin dimer is shown in figure 1.4. Each filamin monomer is composed of an actin-binding domain, located at the amino terminus of the protein, and a long rod-like domain of 24 repeated β-sheets interrupted by two flexible hinge-like loops (van der Flier and Sonnenberg, 2001). The last of the β-sheets, located at the carboxy terminus of the protein, forms the dimerisation domain (van der Flier and Sonnenberg, 2001). Interactions between dimerisation domains gives the dimer a flexible V-shaped structure, as shown in figure 1.4, which is essential to its function (Feng and Walsh, 2004).

Filamins function by binding to actin through the actin binding domain and maintain the orthogonal actin networks\(^1\) that play an important role in cell localisation and migration at the leading edge of motile cells (Stossel et al., 2001). Filamins have also been reported to interact with numerous other proteins through the 24 repeat rod-like domain, suggesting that filamins possess a wide range of functional roles (Stossel et al., 2001). Specifically, interactions between the filamins and membrane receptors indicate that the filamins are

---

\(^1\) Orthogonal actin networks consist of many actin filaments attached perpendicular to one another (Stossel et al., 2001).
important for integrating the cellular actin scaffold with these receptors (Feng and Walsh, 2004). Through these interactions the filamins appear to have important roles in the maintenance of the mechanical stability of the cell membrane; maintaining cell-cell and cell-matrix connections; and in signal transduction (Stossel et al., 2001).

![Figure 1.4. Structural representation of the FLNA dimer.](image)

Each filamin monomer is composed of an actin binding domain and a rod domain consisting of 24 repeated β-sheets, the last of which forms the dimerisation domain which enables the functional FLNA dimer to form. N - amino terminus, C - carboxy terminus. Figure modified from Stossel et al., 2001.

The importance of the filamins as cellular proteins is demonstrated by their conserved nature with filamins being conserved in vertebrates, *Drosophila* and even in prokaryotes (Stossel et al., 2001). The three human filamin proteins also show ~70% homology over the majority of the amino acid sequence, the only exception being the two hinge regions which demonstrate greater divergence as might be expected due to the structural, rather than functional, role they play (van der Flier and Sonnenberg, 2001).

The importance of the filamins to development is demonstrated by their implication in a number of human developmental disorders. Loss-of-function mutations in *FLNA* cause the neural disorder periventricular heterotopia (PVNH) and the majority of cases of the
otopalatodigital syndrome spectrum of disorders (Fox et al., 1998; Robertson et al., 2003). Loss-of-function mutations in \textit{FLNB} are involved in several osteochondrodysplasias including boomerang dysplasia, spondylocarpotarsal syndrome, Larsen syndrome and atelosteogenesis I and III (Lu et al., 2007).

### 1.6 The otopalatodigital syndrome spectrum of disorders

The otopalatodigital (OPD) syndrome spectrum of disorders encompasses four syndromes with an X-linked pattern of inheritance: OPD syndrome type 1 (OPD1), OPD syndrome type 2 (OPD2), frontometaphyseal dysplasia (FMD) and Melnick-Needles syndrome (MNS) (Robertson et al., 2003). These disorders are characterised by skeletal dysplasia and variable anomalies in the brain, craniofacial structures, cardiac, genitourinary and gastrointestinal systems (Robertson, 2006).

Of the four disorders, OPD1 has the mildest phenotype in males and females usually exhibit milder characteristics but can be affected as severely as males (Robertson, 2007). In OPD2 the skeletal dysplasia is more severe than OPD1 and a majority of cases die in the peri-natal period or during infancy (Robertson, 2007). The features of FMD distinguishing it from OPD1 include more severe skeletal and facial abnormalities as well as joint contractures of the phalanges and malformations within the tracheobronchial tree (Robertson, 2007). MNS is almost exclusively observed in females due to embryonic or peri-natal lethality in males but in males that do survive to birth the phenotype is very similar to that of OPD2 (Robertson, 2007). In all four disorders there is some overlap between the associated phenotypes (Robertson, 2007) as might be expected due to their relationship under the OPD spectrum.

Missense mutations in the X-linked gene that encodes FLNA (\textit{FLNA}) gives rise to all four of these syndromes (Robertson et al., 2003). Mutations associated with each of the disorders tend to cluster to specific regions of the FLNA protein with a genotype-phenotype correlation that suggests FLNA is involved in a number of specific yet varied roles in development (Robertson, 2003). Robertson et al. (2006) reported that all individuals with OPD syndrome type 1 and all female individuals with MNS have identifiable mutations in \textit{FLNA}, however, some individuals with OPD syndrome type 2 and FMD apparently lack any such mutations in the \textit{FLNA} gene. It is possible in such cases that these individuals possess
mutations in a gene encoding one of the many proteins that interact with FLNA and that these mutations cause similar phenotypic affects to those found in individuals with mutations in FLNA.

1.7 Cfm2

Cfm2 is a novel gene with no known function. This gene was first identified in mice by Hirano et al. (2005). It was initially identified as a paralogue of another novel gene in mice which was named Cfm, for caudal forebrain midbrain (Hirano et al., 2005). In mice, Cfm expression was initially detected at the 5-somite stage (E 8.0) in a region which corresponds to the future caudal forebrain and midbrain (Hirano et al., 2005). Expression continued in the brain through to E12.5, although expression was reduced to the pineal gland of the forebrain by this stage (Hirano et al., 2005). Interestingly, at E12.5 expression of Cfm was also detected in tongue muscle, lung, alimentary canal, genital tubercle and peripheral nerves suggesting this gene has multiple roles in development (Hirano et al., 2005). Rather surprisingly then, knocking out Cfm had absolutely no detectable affects on development in the mouse (Hirano et al., 2005) and its role in development has since remained unexplored.

Despite the sequence similarities with the paralogous Cfm, Cfm2 expression was not detected in the brain but instead it was found to be uniquely expressed in the presomitic mesoderm at E8.5 (Hirano et al., 2005). This expression was detected in the portion of the PSM that was to be segmented but never in the pre-existing somites (Hirano et al., 2005), suggesting Cfm2 plays a role in somitogenesis. Cfm2 expression was also detected in the hindbrain from E9.0 to E12.5 and the optic nerve, otic capsule, peripheral nerves, tegmentum, lung and tongue muscle at E12.5 (Hirano et al., 2005). However, the function of Cfm2 wasn’t explored any further and its role in development remains a mystery.

Both these genes appear to be well conserved with Hirano et al. (2005) identifying orthologues of both genes in human, chicken, Xenopus and zebrafish. Alignments produced from these genes identified two conserved regions (CR1 and CR2) and a mitogen-activated protein kinase (MAPK) phosphorylation consensus sequence in CR1 as shown in figure 1.5.
Figure 1.5. Xenopus CFM and CFM2 amino acid sequences. An alignment of the Xenopus laevis CFM2 and CFM amino acid sequences showing the two conserved regions, CR1 and CR2, that begin at positions 24 and 90 in the alignment respectively. A second start codon in the CFM2 sequence exists at position 92 of the alignment. The MAPK consensus sequence is underlined in black.

These features were found to be highly conserved in both of these genes and in the identified orthologues of these genes (Hirano et al., 2005) suggesting that these regions play a vital role in the function of CFM and CFM2. MAPK is an integral component of the well-studied and highly conserved MAPK signalling cascade which is involved in the processes of proliferation, differentiation and development (Seger and Krebs, 1995) suggesting that these genes may play an important role in some, or all, of these processes.

More recently Cfm2 has been implicated in the OPD spectrum of disorders. A yeast two-hybrid screen by the research group led by Professor Stephen Robertson revealed that the protein FAM101A, the human orthologue of CFM2, binds to FLNA (Gay et al., 2011). As mentioned earlier mutations in FLNA are associated with the OPD spectrum of disorders, but many patients with these disorders do not possess mutations in FLNA, suggesting that mutations in proteins which interact with FLNA may also be representative of these disorders. These results have led to a hypothesis implicating Cfm2 in the OPD spectrum of disorders.

It has been found that FAM101A is capable of binding to itself suggesting that FAM101A functions as a dimer (Gay et al., 2011). It has also been found that a truncated version of the human FAM101A orthologue lacking CR1 cannot bind to itself but is still capable of interacting with FLNA (Gay et al., 2011). These results suggest that CR2 is responsible for interacting with FLNA whereas CR1, which is deleted in the shortened transcript, is most likely responsible for forming a functional protein dimer.
Consistent with the hypothesis that mutations in \textit{Cfm2} are potentially responsible for phenotypes consistent with the OPD spectrum of disorders, work carried out by Sarah Holman in Dr. Caroline Beck’s lab found that down-regulation of the \textit{Cfm2} gene in \textit{Xenopus laevis} using morpholino oligonucleotides caused anomalies comparable to the syndromes of the OPD spectrum of disorders (Holman, 2007). Further work also confirmed that the \textit{Cfm2} gene is expressed in the somites of the mouse embryo consistent with the results from Hirano et al. (2005), although attempts at analysing the expression patterns in \textit{Xenopus} were unsuccessful (Holman, 2007). \textit{Cfm2} expression was also detected in the developing limbs of both chick and mice and in a number of other tissue, including the heart in chick embryos (Holman, 2007), suggesting that \textit{Cfm2} may have other important functions in organogenesis in the developing embryo.

Three lines of evidence implicate \textit{Cfm2} as a good candidate as a gene which may be responsible, when mutated, for the OPD spectrum of disorders through its interactions with FLNA. Firstly, FAM101A, the human orthologue of CFM2, is known to interact with FLNA (Gay et al., 2011). Secondly, the OPD spectrum of disorders are characterised by skeletal, craniofacial and, in some cases, cardiac abnormalities (Robertson, 2007) consistent with the expression patterns of the mouse orthologue of the \textit{Cfm2} gene as identified by Hirano et al. (2005) and more recent work (Holman, 2007). And lastly, down-regulation of the CFM2 expression in \textit{Xenopus} appears to produce an OPD spectrum related phenotype. Despite this, no mutations in the \textit{FAM101A} gene have been identified in individuals with any of the OPD spectrum disorders suggesting that mutations in this gene are either very rare or lethal. Because of this the importance of \textit{Cfm2} to vertebrate development has since remained largely unexplored.

Very recently a research group in France identified the \textit{Cfm} and \textit{Cfm2} genes (renamed RefilinA and RefilinB, for REgulator of FILamin proteIN) as being regulators of actin filament bundles (Gay et al., 2011). They then identified the RefilinB/FLNA complex as an organiser of the perinuclear actin bundles that form the actin cap, a structure important in controlling the intracellular localisation and the shape of the nucleus during cellular and developmental processes (Gay et al., 2011). They also identified the RefilinB/FLNA complex as an important regulator of the actin cap during changes to the nuclear shape that occur during epithelial to
mesenchymal transition (EMT) (Gay et al., 2011). This report was the first to describe a functional role for Cfm or Cfm2, although their work was largely focussed on Cfm (RefilinB).

1.8 Aims

Cfm2 expression has already, although rather briefly, been examined by RNA extraction in the frog (Xenopus laevis) and by in situ hybridisation in the mouse (M.musculus), chick (G.gallus) and unsuccessfully in the frog. The first aim of my research is to confirm the expression patterns of Cfm2 in Xenopus by reverse transcription PCR and in situ hybridisation, as this information will be useful to forming conclusions as to which functional processes CFM2 is involved in. The function of CFM2 is still unknown, the interaction between CFM2 and FLNA and the expression patterns of Cfm2 in mice suggest possible roles in somitogenesis, skeletal development and organogenesis. My second aim is to determine the affects of reduced Cfm2 expression on these potential roles through MO mediated knockdown. My third aim is then to determine what features of CFM2 are important for its function by carrying out MO rescue experiments with mutated Cfm2 mRNA. To analyse any redundancy in CFM2 function my fourth aim is to determine if there is any overlap between Cfm and Cfm2 expression and whether or not Cfm expression can functionally substitute for a lack of Cfm2 by rescuing embryos injected with Cfm2 MO.
Chapter 2 - Materials and Methods

2.1 Materials

All instruments were provided by the Beck lab, P107, in the Zoology department at the University of Otago unless otherwise stated. Histology solutions and instruments were provided by the histology lab, B108, in the Zoology department at the University of Otago. *Xenopus laevis* stocks were kept in aquariums in the animal suite at the Zoology department at the University of Otago. *Xenopus laevis* were provided and maintained by Dr. Caroline Beck with technical assistance from Amy Armstrong and Kim Garrett.

All reagents used were of molecular biology grade. A detailed list of solutions used can be found in Appendix VII. Gel electrophoresis was carried out using 0.7-2% w/v agarose gels with TAE buffer and SYBR safe (Invitrogen). All gels were run with 1Kb Plus DNA ladder (Invitrogen) as a reference. Gels were visualised using ultraviolet light in a UVItec gel doc. Graphs were compiled using the free online software SciDAVis\(^1\).

2.2 In silico analysis

2.2.1 Orthologues of CFM and CFM2

The online search engine BLAST\(^2\) was used to identify potential orthologues of the *Xenopus laevis* CFM and CFM2 sequences. Separate protein BLAST (BLASTp) searches using the *Xenopus laevis* CFM and CFM2 amino acid sequences were conducted. Similarity to the *Xenopus* sequences were evaluated using the E-values calculated in BLAST. The E-value of $5 \times 10^{-9}$ was chosen as the cut-off point as sequences below this value did not appear to represent the CFM/CFM2, or FAM101, protein family. E-values greater than this value were considered to be similar by chance and genetically un-related. To identify further sequences, a separate search was carried out on the Metazome website\(^3\) using the term ‘FAM101’ in a keyword search.

---

1. Downloaded from http://scidavis.sourceforge.net/
3. Found at http://www.metazome.net
2.2.2 Phylogenetic analysis

The protein sequences identified from the online searches were downloaded and aligned using ClustalX\textsuperscript{1}. The resulting sequence alignment was then exported and any obvious errors manually corrected using BioEdit\textsuperscript{2}. The modified alignments were then loaded into the phylogenetic analysis program PAUP*\textsuperscript{3}. Neighbour-joining trees were created in PAUP* using the default distance settings and a root established using mid-point rooting\textsuperscript{4}. To evaluate the level of confidence in the branches and nodes of the tree it was then bootstrapped\textsuperscript{5} using a heuristic search and 1000 sub-replicates. Finally the phylogenetic tree was exported into Mesquite\textsuperscript{6} where the tree was annotated and saved.

2.3 Preparing *Xenopus laevis* embryos

Adult female *Xenopus laevis* were induced to produce eggs by injecting the frogs with HCG the night before the eggs were required. The female frogs were then placed in 1x MMR in temporary tanks, moved to an incubator and left overnight. The following morning an adult male frog was selected and killed by placing the frog in a lethal solution of benzocaine. Once dead his testes were dissected out and placed in 1x MMR, the testes were then kept in this solution at 2°C and used as required. The female frogs were removed from the incubator when ready and eggs collected by gently ‘squeezing’ the frogs, which forces the female to release her eggs. These eggs were then collected in petri dishes with a small amount of 1x MMR to maintain the eggs. 1x MMR is a high-salt solution that mimics the physiological environment and is used to keep the eggs and sperm in an inactivate state until fertilisation is desired.

\textsuperscript{1} Downloaded from http://www.clustal.org/clustal2/
\textsuperscript{2} Downloaded from http://www.mbio.ncsu.edu/bioedit/bioedit.html
\textsuperscript{3} Provided by the Zoology department at the University of Otago
\textsuperscript{4} This method assumes that all operational taxonomic units show the same evolutionary rate on average and establishes the root as the mid-point between the two most divergent operational taxonomic units (Hess and De Moraes Russo, 2007).
\textsuperscript{5} Bootstrapping evaluates the branches of a phylogenetic tree by analysing a number of possible replicate trees and recording how often each particular branch is present (Baldauf, 2003).
\textsuperscript{6} Downloaded from http://mesquiteproject.org/mesquite/mesquite.html
To fertilise the eggs a small piece of the testes was removed and added to a petri dish containing eggs. The testes were then sheared and gently mixed in with the eggs. After approximately one minute the sperm were activated by flooding the dish with MQW to lower the salt concentration and initiate fertilisation. Successful fertilisation was observed by rotation of the eggs so that the darker animal pole faces upwards, which takes 20-30 minutes.

Once this rotation had been observed a 2% cysteine hydrochloride solution was added to the fertilised eggs to ‘de-jelly’ them, as eggs are laid in a gelatinous coating which makes it difficult to manipulate them. The cysteine hydrochloride solution dissolves the coating around the eggs allowing them to be moved freely of one another. Once de-jellying was completed, the eggs were washed three times using 0.1x MMR before being left to cleave in 0.1x MMR in an incubator. After 2-3 hours the embryos were examined and selected for those that had undergone successful division with any abnormal or unfertilised eggs being removed and discarded. Following selection the embryos were returned to the incubator and allowed to develop until they reached the feeding tadpole stage where they were moved to the aquarium and fed on a daily basis. Embryos were regularly examined during development and any dead embryos or tadpoles were removed and discarded. The temperature of the incubator used during this process can be changed to manipulate the speed of development as desired, typically 18-24°C was used for the experiments.

2.4 Raising *Xenopus laevis* tadpoles

Once the developing embryos had reached the feeding tadpole stage they were moved to a fish tank filled with 0.1x MMR. The tadpoles were fed spirulina and a meat supplement on a daily basis by mixing the feed with filtered ‘frog water’ and adding this to the tank. The 0.1x MMR was gradually changed to frog water by removing roughly 10-20% of the tank volume and replacing it with frog water every 2-3 days. After at least one week of water changes the tadpoles were shifted into a temperature controlled aquarium containing frog water. The tadpoles were continually fed as they developed until they reached the desired stage.

\[ \text{1 All *Xenopus laevis* staging was done according to Nieuwkoop and Faber (1967), see appendix VIII for guide.} \]
During this time they were also checked and any dead tadpoles were removed and discarded.

2.5 RNA extraction
Embryos, tadpoles and limbs were collected for RNA extraction as they reached the appropriate stage in development. Embryos/tadpoles younger than stage 40 were collected and placed straight in RNA\textregistered later solution (Invitrogen), samples were then left at 2°C overnight to allow the solution to penetrate the tissue before being moved to -30°C for longer term storage. Tadpoles at stage 40 or older were collected and anaesthetised in MS222, samples were then rinsed in MQW before being transferred to RNA\textregistered later and again left at 2°C overnight and -30°C thereafter as before. Limb samples were collected by lethally anaesthetising tadpoles that had reached the desired limb stage with a high dose of MS222. The limbs were then dissected using microscissors and the limbs placed in RNA\textregistered later and left at 2°C overnight and -30°C thereafter as before.

To extract RNA from these samples the PureLink™ Micro-to-Midi total RNA purification system (Invitrogen) was used. RNA was extracted by removing as much RNA\textregistered later solution as possible from the samples, using a small bore pipette, and then following the manufacturer’s instructions for RNA purification using clean plastic homogenisers to homogenise the tissue. To test the quality of the RNA each sample was analysed using a nanodrop to check the concentration and purity, each sample was also run on a 1% agarose gel to check the integrity of the RNA. Purity is measured by the 260/280 ratio, this measures the protein contamination in the sample. Nucleic acids absorb light most strongly at a wavelength of 260nm giving an indication of the nucleic acid concentration whereas proteins absorb light at a wavelength of 280nm (due to the presence of tyrosine and tryptophan residues) to give an indication of the concentration of proteins. A value of 2.0 indicates pure RNA although values ranging from 1.8-2.2 are acceptable for use.

2.6 Reverse transcription
To reverse transcribe the RNA collected from the various stages of *Xenopus* development SuperScript® III Reverse Transcriptase (Invitrogen) was used. The reverse transcription
reaction was set up in sterile 1.5mL microcentrifuge tubes using the reagents described in table 2.1.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA</td>
<td>1µg total</td>
</tr>
<tr>
<td>oligodT</td>
<td>1µL</td>
</tr>
<tr>
<td>10mM dNTP</td>
<td>1µL</td>
</tr>
<tr>
<td>MQW</td>
<td>to a total volume of 13µL</td>
</tr>
</tbody>
</table>

This initial reaction mix was first heated at 65°C for 5 minutes and then placed on ice for a minimum of 1 minute as the reagents from table 2.2 were added to the reaction.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x First-strand synthesis buffer</td>
<td>4µL</td>
</tr>
<tr>
<td>0.1M DTT</td>
<td>1µL</td>
</tr>
<tr>
<td>RNase OUT</td>
<td>1µL</td>
</tr>
<tr>
<td>SuperScript® III</td>
<td>1µL</td>
</tr>
</tbody>
</table>

The reaction was then mixed gently and heated at 50°C for 1 hour, the reaction was then heat inactivated by heating to 75°C for a further 15 minutes. The resulting cDNA was then stored at -30°C or placed on ice for immediate use.
2.7 Polymerase chain reaction

2.7.1 Standard polymerase chain reaction

Standard PCR reactions were set up using MangoTaq™ DNA polymerase (Bioline) as outlined in table 2.3.

Table 2.3 - PCR reagents

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x Mango buffer</td>
<td>4µL</td>
</tr>
<tr>
<td>50mM MgCl$_2$</td>
<td>0.6µL</td>
</tr>
<tr>
<td>10mM dNTP mix</td>
<td>0.4µL</td>
</tr>
<tr>
<td>10µM forward and reverse primer mix</td>
<td>2µL</td>
</tr>
<tr>
<td>DNA template</td>
<td>1µL</td>
</tr>
<tr>
<td>MangoTaq™ DNA polymerase</td>
<td>0.3µL</td>
</tr>
<tr>
<td>MQW</td>
<td>11.7µL</td>
</tr>
</tbody>
</table>

Reagents were mixed on ice in 0.2mL microcentrifuge tubes using the appropriate DNA template and primers, and PCR carried out using the programs and primers outlined in appendix IV as appropriate.

2.7.2 Colony screening using polymerase chain reaction

Colonies from transformed *E.coli* (see 2.10.1) were screened using PCR to examine which colonies had been transformed with a plasmid containing the desired insert. This was achieved by resuspending each colony in 5µL of MQW in a sterile microcentrifuge tube, and then using these suspensions as the templates for PCR reactions. Reactions were set up as described in table 2.4, the primers used for colony screening and programs used are outlined in appendix IV.
Table 2.4 - Colony screening PCR reagents

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x MANGO buffer</td>
<td>4µL</td>
</tr>
<tr>
<td>50mM MgCl₂</td>
<td>0.6µL</td>
</tr>
<tr>
<td>10mM dNTP mix</td>
<td>0.4µL</td>
</tr>
<tr>
<td>10µM forward and reverse primer mix</td>
<td>2µL</td>
</tr>
<tr>
<td>Colony suspension</td>
<td>2µL</td>
</tr>
<tr>
<td>MangoTaq™ DNA polymerase</td>
<td>0.3µL</td>
</tr>
<tr>
<td>MQW</td>
<td>10.7µL</td>
</tr>
</tbody>
</table>

2.8 Whole mount in situ hybridisation

2.8.1 RNA probe synthesis

Plasmids containing the gene of interest were linearised using restriction enzymes in a reaction set up in a clean, sterile 1.5mL microcentrifuge tube as described in table 2.5. For creation of the anti-sense Cfm and sense (negative control) Cfm2 probes the restriction enzyme NcoI was used with SuRE cut buffer H (Roche). For creation of the sense Cfm and anti-sense Cfm2 probes the restriction enzyme NotI was used with NEBuffer 3 (New England BioLabs).

Table 2.5 - Restriction enzyme reaction

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>2µg total DNA</td>
</tr>
<tr>
<td>Buffer</td>
<td>2.5µL</td>
</tr>
<tr>
<td>Restriction enzyme</td>
<td>1µL</td>
</tr>
<tr>
<td>MQW</td>
<td>to a total volume of 25µL</td>
</tr>
</tbody>
</table>

This reaction was incubated for 2 hours at 37°C and then heat inactivated by heating to 65°C for a further 20 minutes. A 5µL sample was taken from the digest and run on a 1% agarose gel to check the linearisation had been successful along with a sample of the uncut plasmid for reference.

The DNA was then purified using the GENE CLEAN® III kit (MP biomedicals) following the manufacturer’s instructions. The purified linear template was then used in the reaction.
described in table 2.6 to transcribe the DIG-labelled RNA probe using the appropriate polymerase (SP6 or T7) to give an anti-sense RNA product for use as a probe, or a sense RNA product for use as a negative control.

Table 2.6 - In situ hybridisation probe transcription reaction

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear template</td>
<td>~1µg total DNA</td>
</tr>
<tr>
<td>10x transcription buffer</td>
<td>5µL</td>
</tr>
<tr>
<td>10x DIG RNA labelling mix</td>
<td>2.5µL</td>
</tr>
<tr>
<td>RNA polymerase</td>
<td>2µL</td>
</tr>
<tr>
<td>MQW</td>
<td>to a total volume of 50µL</td>
</tr>
</tbody>
</table>

The reaction was incubated for 3 hours at 37°C in an incubator, following this 2.5µL of RNase free DNase I (Roche) was added to the reaction and it was then incubated for a further 20 minutes at 37°C to digest the DNA template. The RNA product was then precipitated by adding 25µL of a 7.5M LiCl, 50mM EDTA solution and leaving it at -30°C overnight.

The following morning the RNA was spun down in a cold centrifuge at maximum speed for 15 minutes. The RNA pellet was washed by removing the supernatant, adding 100µL of cold 70% ethanol and spinning this in the cold centrifuge once more at maximum speed for 5 minutes. The supernatant was then removed and the sample then left at room temperature with the lid off for ~10 minutes to allow any remaining ethanol to evaporate. Finally 50µL of 10mM EDTA was added to resuspend the RNA pellet, this was then mixed using a vortex and heated at 80°C for 5 minutes to dissolve the pellet. The RNA was then analysed by running a sample on a gel to check the synthesis has worked and another sample analysed using a nanodrop to estimate concentration and purity.

The remaining probe was then prepared for in situ hybridisation by heating it at 80°C for 5 minutes to denature the probe, hybridisation buffer pre-warmed to 60°C was then added to the probe to give a final concentration of 1µg/mL and the probe stored at -30°C until required.
2.8.2 Probe hybridisation

Vials containing fixed samples (see 2.14.1) were removed from the -30°C freezer and allowed to warm up to room temperature. For older sample (≥ stage 45) the viscera was removed using microscissors and forceps, for limb stage samples (≥ stage 50) the tail and head were also removed using a scalpel. These samples were then rehydrated using the following series of washes on a nutator:

- 75% EtOH in PBSAT for 10 min
- 50% EtOH in PBSAT for 10 min
- 1x PBSAT for 5 min
- 1x PBSAT for 5 min
- 1x PBSAT for 5 min

The tissue was then permeabilised by treatment with proteinase K. Samples were left in the appropriate concentration of proteinase K at room temperature for the times given in table 2.7.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Proteinase K concentration</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole embryo ≤ Stage 30</td>
<td>10µg/mL</td>
<td>10 minutes</td>
</tr>
<tr>
<td>Whole tadpole stage 31-35</td>
<td>10µg/mL</td>
<td>20 minutes</td>
</tr>
<tr>
<td>Whole tadpole stage 36-40</td>
<td>20µg/mL</td>
<td>20 minutes</td>
</tr>
<tr>
<td>Whole tadpole stage 41-45</td>
<td>20µg/mL</td>
<td>30 minutes</td>
</tr>
<tr>
<td>Whole tadpole stage 46-50</td>
<td>30µg/mL</td>
<td>30 minutes</td>
</tr>
<tr>
<td>Stage 50 limbs</td>
<td>10µg/mL</td>
<td>10 minutes</td>
</tr>
<tr>
<td>Stage 51 limbs</td>
<td>10µg/mL</td>
<td>15 minutes</td>
</tr>
<tr>
<td>Stage 52 limbs</td>
<td>10µg/mL</td>
<td>20 minutes</td>
</tr>
<tr>
<td>Stage 53 limbs</td>
<td>10µg/mL</td>
<td>25 minutes</td>
</tr>
<tr>
<td>Stage 54 limbs</td>
<td>10µg/mL</td>
<td>30 minutes</td>
</tr>
<tr>
<td>Stage 55/56 limbs</td>
<td>20µg/mL</td>
<td>30 minutes</td>
</tr>
</tbody>
</table>

Following permeabilisation the samples were acetylated to reduce background staining by replacing the proteinase K with 2mL of 0.1M TEA, pH 7.8. Samples were left standing at
room temperature for 5 minutes before the TEA was removed and replaced with 2mL of fresh 0.1M TEA, pH7.8, along with 5µL acetic anhydride. The samples were incubated for a further 5 minutes at room temperature with regular swirling to encourage mixing, as acetic anhydride mixes poorly. A further 5µL acetic anhydride was then added to the vials and samples were incubated for 5 minutes at room temperature with regular swirling once more. After acetylation the samples were washed twice with 1x PBSAT for 5 minutes with nutation at room temperature before being refixed in formalin for 20 minutes with nutation. Samples were then washed with 1x PBSAT at room temperature for 5 minutes. This wash was repeated four more times.

Following the final wash, samples were left in 1mL of 1x PBSAT and prehybridised by adding 250µL of hybridisation buffer and leaving the samples undisturbed for 2-3 minutes. The solution was then removed and replaced with 1mL of hybridisation buffer and the samples transferred to a 60°C water bath where they were left for 10 minutes. Following this the solution was removed and replaced with 1mL of fresh hybridisation buffer pre-warmed to 60°C, samples were then moved to an incubator and left for 2 hours at 60°C with gentle shaking. After prehybridisation the solution was removed and replaced with 1mL of pre-warmed probe at a concentration of 1µg/mL and left at 60°C in the incubator overnight with gentle shaking.

The next morning the probe was removed, collected and returned to the -30°C for re-use. The samples were then put through the following series of washes at 60°C, with gentle shaking, to increase the stringency of probe binding:

- 1mL hybridisation buffer for 10 minutes
- 1mL hybridisation buffer for 10 minutes
- 2mL 2× SSC, 0.1% tween 20 for 20 minutes
- 2mL 2× SSC, 0.1% tween 20 for 20 minutes
- 2mL 2× SSC, 0.1% tween 20 for 20 minutes
- 2mL 0.2× SSC, 0.1% tween 20 for 30 minutes
- 2mL 0.2× SSC, 0.1% tween 20 for 30 minutes
2.8.3 Probe detection

To detect the hybridised DIG-labelled probe the samples were first washed three times with 1x MABT for 15 minutes. Following this, blocking was carried out to reduce non-specific antibody-antigen interactions by first preincubating samples in 1x MABT, 2% blocking reagent for 30 minutes at room temperature with nutation. The solution was removed and replaced with 1mL of fresh 1x MABT, 2% blocking reagent, 20% lamb serum and the samples incubated at room temperature for 2 hours with nutation. The solution was then changed to 1mL of 1x MABT, 2% blocking reagent, 20% lamb serum, with a 1:2000 dilution of anti-DIG fab fragments coupled to alkaline phosphatase, and the samples left to incubate overnight at 2°C with nutation.

The following morning the solution was once again removed and the samples washed with nutation at room temperature with 1x MABT for 15 minutes, repeated two more times, and then with 1x MABT for 30 minutes, repeated five more times. After the MABT washes the samples were washed with alkaline phosphatase buffer once for 5 minutes and a second time for 10 minutes with 5mM levamisole added to repress the activity of endogenous alkaline phosphatases. The solution was then replaced with 1mL of NBT-BCIP, again with 5mM levamisole, and left undisturbed at room temperature until colour developed. Once samples were adjudged to have completed staining they were washed in 1x PBSAT for 10 minutes, 100% EtOH for 10 minutes and then in 1x PBSAT for a further 10 minutes before being refixed and stored in formalin.

2.9 Gene cloning

2.9.1 Site directed mutagenesis of mismatched *Xenopus laevis Cfm2*

A plasmid containing a full copy of the mismatched *Cfm2 (mmCfm2)* CDS in the pGEM-T vector was obtained from a previous lab member. This mismatched *Cfm2* sequence contains 5 synonymous mutations around the ATG start codon of the CDS which prevent the MO from binding to the mRNA without affecting the functional amino acid sequence of the CFM2 protein, this allows the *mmCfm2* mRNA to be used for MO rescue experiments.

Mutations were introduced into this mismatched sequence using PCR with specific primers designed for the mutagenesis. 5’ phosphorylated primers were designed and ordered from
Sigma-Aldridge. These primers were then used in a PCR reaction with Phusion® high-fidelity DNA polymerase (New England BioLabs) as outlined in table 2.8 to produce the desired mutated mmCfm2 sequences.

<table>
<thead>
<tr>
<th>Table 2.8 - Phusion DNA polymerase reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reagent</td>
</tr>
<tr>
<td>5x Phusion HF buffer</td>
</tr>
<tr>
<td>10mM dNTP mix</td>
</tr>
<tr>
<td>20µM forward primer</td>
</tr>
<tr>
<td>20µM reverse primer</td>
</tr>
<tr>
<td>DNA template</td>
</tr>
<tr>
<td>Phusion DNA polymerase</td>
</tr>
<tr>
<td>MQW</td>
</tr>
</tbody>
</table>

Two features were targeted by mutating the template sequence; the MAPK phosphorylation site in CR1 and the shortened transcript lacking CR1 completely. To analyse the MAPK phosphorylation site primers were designed that carried substitution mutations that would change the two serine residues in the CFM2 protein to either alanine, which mimics the unphosphorylated form of serine, or aspartate, which mimics the phosphorylated form of serine. To analyse the shortened transcripts primers were designed that would produce a PCR product that lacked nucleotides 1-138 of the CDS. The primer sequences and PCR program used are described in appendix V.

The resulting PCR products were then purified using the High Pure PCR product Purification kit (Roche) and re-ligated together using the Rapid DNA Ligation kit (Roche) to produce the ‘new’ plasmids containing the mutated mmCfm2 sequences. Each ligation product was then used to transform competent E.coli cells (see 2.10.1) and the resulting colonies screened for the correct insert using colony PCR (see 2.7.2). To verify the successfully transformed colonies had the desired mutations the plasmid was cultured (see 2.10.2), isolated and sequenced using the genetics analysis service at the University of Otago.
2.9.2 Sub-cloning sequences into the pCS2+ vector

To produce the mRNA transcripts for the MO rescue experiments the desired sequences needed to be cloned into the pCS2+ vector. This vector has a strong enhancer/promoter and flanking sequences to produce stable mRNA. These sequences also needed to be in the correct orientation to produce a sense RNA molecule that will be translated into a functional protein. To achieve this both the pCS2+ vector and the plasmids containing the Cfm, mmCfm2 and mutated mmCfm2 sequences were digested with two restriction enzymes to leave one ‘sticky’ end and one blunt end that would ensure the sequences could only orientate in one direction in the pCS2+ vector.

To prepare the pCS2+ vector, 2µg of vector was simultaneously digested with the restriction enzymes StuI and XbaI using SuRE cut buffer B (Roche) in a reaction as described previously in table 2.5. StuI produces a blunt end and XbaI a sticky end which allows the inserts to be orientated in a known direction as long as the insert has one matching sticky end and a blunt end.

The pCS2+ digestion reaction was incubated at 37°C for 2 hours and then heat inactivated by heating to 65°C for 15 minutes. The cut DNA was then purified using the GENE CLEAN® III kit following the manufacturer’s instructions. A 2µL sample was run on a gel to check that the digest had worked and the rest of the cut plasmid then dephosphorylated to prevent self religation using shrimp alkaline phosphatase (Roche). This was then incubated at 37°C for 1 hour followed by heat inactivation at 65°C for 15 minutes. The resulting phosphorylated, linear DNA was again purified using the GENE CLEAN® III kit and the DNA concentration measured using a nanodrop and then stored at -30°C until required.

To prepare the Cfm insert the pGEM-T plasmid containing the Cfm CDS was sequentially digested with the Ncol and XbaI restriction enzymes. Digestion with XbaI produces the desired 3’ sticky end, however, Ncol also produces a sticky end which needs to be filled in using Klenow fragment (Roche) to give the desired 5’ blunt end. To achieve this, 3µg of the pGEM-T plasmid containing the Cfm CDS was first digested with Ncol using SuRE cut buffer H (Roche) as previously described in table 2.5
This reaction was incubated at 37°C for 3 hours followed by heat inactivation at 65°C for 15 minutes. The DNA was then purified using the GENECLEAN®III kit and the purified linear plasmid used immediately in another reaction with the Klenow fragment to fill in the overhang as described in table 2.9.

<table>
<thead>
<tr>
<th>Table 2.9 - Klenow fragment reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reagent</td>
</tr>
<tr>
<td>DNA</td>
</tr>
<tr>
<td>10x filling buffer</td>
</tr>
<tr>
<td>10mM dNTP mix</td>
</tr>
<tr>
<td>Klenow fragment</td>
</tr>
<tr>
<td>MQW</td>
</tr>
</tbody>
</table>

This reaction was incubated at 37°C for 15 minutes followed by heat inactivation at 65°C for 10 minutes. The DNA was then purified once again using the GENECLEAN®III kit.

At the same time as the first Cfm digest above was being carried out, 3µg of each of the pGEM-T plasmids containing the desired mmCfm2 and mutated mmCfm2 sequences were digested with the HincII restriction enzyme using NEBuffer 3 (New England BioLabs) as previously described in table 2.5. These reactions were incubated at 37°C for 3 hours followed by heat inactivation at 65°C for 15 minutes. The linear plasmid was then purified using the GENECLEAN®III kit. HincII produces a blunt ends so no filling in was required after these digests.

Once the plasmids containing the Cfm, mmCfm2 and mutated mmCfm2 sequences had been digested, and the sticky ends from the Cfm sequence filled in, the remaining linear template DNA was digested with the XbaI restriction enzyme using SuRE cut buffer B (Roche) as described in table 2.5. These reactions were incubated at 37°C for 3 hours to produce the 5’ sticky ends that would ensure the desired orientation of the inserts was established in the pCS2+ vector. After digestion the reactions were heat inactivated at 65°C for 15 minutes and the DNA then purified for the last time using the GENECLEAN®III kit. A 2µL sample from each digest was run on a gel to check that the digests had produced inserts of the desired
size. The rest of the digested plasmid samples were run on E-Gel® CloneWell™ gels (Invitrogen) using an E-Gel® iBase™ illuminator (Invitrogen) to purify the band of interest corresponding to the insert.

Finally, to produce the pCS2+ plasmid containing the desired sequences the purified inserts were added to the digested pCS2+ vector and ligated using the Rapid DNA Ligation kit (Roche). The ligation product was then used to transform competent *E.coli* cells (*see* 2.10.1), the resulting colonies were screened (*see* 2.7.2) and colonies containing an insert of the correct size then cultured (*see* 2.10.2) and sequenced using the genetics analysis services at the University of Otago to confirm the plasmid contained the correct insert in the desired orientation.

### 2.9.3 *Ambystoma mexicanum* Cfm2

Primers were designed by Dr. Caroline Beck for the amplification of the *Ambystoma mexicanum* (axolotl) *Cfm2* EST. These were used to amplify the axolotl EST from cDNA reverse transcribed from RNA (*see* 2.5 and 2.6) collected from day 5 axolotl embryos using PCR (*see* 2.7.1). The PCR product was run on an agarose gel to check that a correctly sized product was amplified. When this was confirmed the remaining PCR product was cloned into the pCR®II-TOPO® vector using the TOPO® TA Cloning® Kit (Invitrogen) following the manufacturer’s instructions. The resulting plasmid product was used to transform competent top10 *E. coli* cells (*see* 2.10.1), 40µL of 40mg/mL X-gal was also added to the agar plates for this plasmid to allow for blue/white selection. The following morning white and partially blue colonies were selected and screened for the insert (*see* 2.7.2), colonies containing the appropriate sized band were cultured overnight (*see* 2.10.2) and sequenced using the genetics analysis services at the University of Otago to confirm the axolotl EST had been cloned correctly.

### 2.10 Transforming and culturing competent cells.

#### 2.10.1 Transformation

Microcentrifuge tubes containing competent Top10 *E.coli* cells were collected from storage in the -70°C freezer and thawed on ice. Approximately 50ng of plasmid DNA was then added to 50µL of competent cells in a microcentrifuge tube and mixed gently by hand. The cells
were then incubated on ice for 30 minutes, followed by a 30 second heat-shock in a 42°C water bath to induce the uptake of the plasmid DNA.

250µL of LB medium was then added to each tube and these then incubated at 37°C for 1 hour with shaking at 200rpm to allow cells that had been transformed to begin transcribing the plasmid DNA. Following incubation 50µL from each tube was plated on LB agar plates with a selective antibiotic present at 50ng/µL. The plates were allowed to dry and then incubated overnight at 37°C. The next morning any colonies from successful transformation were resuspended and screened for the correct insert (see 2.7.2). Successfully transformed colonies were then cultured as described below.

**2.10.2 Cell culturing**

Colonies for culturing were grown up overnight at 37°C with shaking at 200rpm, in 5mL LB medium with selective antibiotic at a concentration of 50ng/µL. The following morning the cultures were retrieved and the plasmid DNA isolated using the Wizard® Plus SV Miniprep DNA Purification System (Promega) following the manufacturer’s instructions. The resulting samples were analysed using a nanodrop to determine concentration and purity and then sequenced using the genetics analysis services at the University of Otago to ensure the correct sequence had been obtained.

**2.11 In vitro transcription**

To synthesise mRNA for the overexpression and MO rescue experiments the mMMessage mMachine® SP6 kit (Ambion) was used. To use this kit the desired CDS was first cloned into the pCS2+ vector (see 2.9.2) as this vector has flanking UTRs to stabilise the RNA. The kit also produces capped RNA that mimics most eukaryotic mRNAs due to the presence of a 7-methyl guanosine cap structure at the 5’ end of the synthesised RNA. When injected into a living cell the RNA produced by this kit is stable, actively transcribed and will be passed to the daughter cells as these cells continue to divide during development.

To begin this process 3µg of the pCS2+ plasmid containing the desired CDS was linearised by a restriction enzyme digest as previously described in table 2.5. For the Cfm templates this was done using the NotI restriction enzyme and NEBuffer 3 (New England BioLabs) and KpnI...
restriction enzyme and React buffer 4 (Invitrogen) for the \textit{mmCfm2} and mutated \textit{mmCfm2} templates.

These reactions were incubated at 37°C for 3 hours and a 2µL sample was then taken and run on a 0.7% agarose gel to check that the plasmid had been successfully linearised. The rest of the reaction was then cleaned up using the GENECLEAN® III kit following the manufacturer’s instructions. The 3’ overhang of the \textit{mmCfm2} and mutated \textit{mmCfm2} templates, created by cutting with KpnI, were then filled in with Klenow enzyme (Roche) as previously described in table 2.9. This was necessary to avoid the novel transcription products that occur from SP6 polymerases when combined with a template that has 3’ overhang as has previously been reported (Schenborn and Mierendorf, 1985).

This reaction was incubated at 37°C for 15 minutes followed by heat inactivation at 65°C for 10 minutes. The DNA was then purified once again using the GENECLEAN®III kit and the resulting purified linear DNA from the \textit{Cfm}, \textit{mmCfm2} and mutated \textit{mmCfm2} plasmid templates then used to set up the mRNA synthesis reactions in a clean, sterile, 1.5mL microcentrifuge tube as described in table 2.10.

\textbf{Table 2.10 - mRNA synthesis reaction}

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>1µg total DNA</td>
</tr>
<tr>
<td>10x Rxn buffer</td>
<td>2µL</td>
</tr>
<tr>
<td>2x NTP CAP mix</td>
<td>10µL</td>
</tr>
<tr>
<td>Enzyme mix</td>
<td>2µL</td>
</tr>
<tr>
<td>MQW</td>
<td>to a total volume of 20µL</td>
</tr>
</tbody>
</table>

This reaction was then incubated at 37°C for 3 hours to synthesise the mRNA. Following this 1µL of TURBO DNase (Ambion) was added to the reaction tube and the reaction incubated for a further 15 minutes at 37°C to digest the DNA template. The RNA was then precipitated by adding 30µL of MQW and 30µL of a 7.5M LiCl, 50mM EDTA solution, mixing this and then leaving the tube at -30°C overnight.
The following morning the RNA was spun down in a cold centrifuge at maximum speed for 15 minutes. The RNA pellet was then washed by removing the supernatant, adding 200µL of cold 70% ethanol and spinning the sample in the cold centrifuge again at maximum speed for 8 minutes, this wash was repeated once more. The RNA pellet was then air-dried by removing as much supernatant as possible and leaving the reaction tube at room temperature with the lid off for 10 minutes. Finally the RNA pellet was dissolved in 30µL of nuclease free water. A sample was then analysed on the nanodrop to check concentration and purity. The remaining mRNA was then aliquoted and stored at -70°C until required.

2.12 Microinjection

2.12.1 Morpholino oligonucleotide mediated knockdown

For the morpholino knockdown and rescue experiments microinjection was required, for this a Drummond nanoject II microinjector was used. The anti-sense Cfm2 MO and control mismatched MO (mmMO) were obtained from a previous lab member who had originally ordered them from Genetools LLC. The Cfm2 MO, 5’GCAAATGCAGGTGACCTACCATGTC3’, was designed over the major ATG start codon (underlined), the mmMO, 5’GCaATcCAGGTcACCTACgATcTC3’, contains 5 mismatched bases (indicated by lower case) which prevent the mmMO from binding to and subsequently inhibiting translation of the Cfm2 mRNA.

*Xenopus* embryos were obtained as described in 2.3. Embryos were collected in multiple petri dishes which were then left to develop at different temperatures to stagger development. Needles for the microinjections were created by pulling glass capillaries using a needle puller prior to use and these stored in clean petri dishes. On the day of the injections the microinjector was prepared by firstly using fine forceps to break off the tip of one of the needles, the needle was then filled with paraffin liquid (Pharmacy health™) using a syringe and then fitted onto the microinjector. The oil was then ejected and the needle loaded with the desired solution by placing a drop of the solution onto a clean petri dish and loading this into the needle.

Once embryos reached the two-cell stage they were placed in 6% ficoll 1x MMR in an agar lined petri dish with a narrow well cut into the agar. Embryos were lined up along the well
and both cells of each of the embryos then injected, the solution and volume injected is described in table 2.11.

Table 2.11 - Microinjection solutions.

<table>
<thead>
<tr>
<th>Solution</th>
<th>Concentration</th>
<th>Volume injected per cell</th>
<th>Final concentration per embryo</th>
</tr>
</thead>
<tbody>
<tr>
<td>MQW</td>
<td>-</td>
<td>13.8nL</td>
<td>-</td>
</tr>
<tr>
<td>Cfm2 MO</td>
<td>1.087ng/nL</td>
<td>9.2nL</td>
<td>20ng</td>
</tr>
<tr>
<td>Cfm2 MO</td>
<td>1.087ng/nL</td>
<td>13.8nL</td>
<td>30ng</td>
</tr>
<tr>
<td>mmMO</td>
<td>1.087ng/nL</td>
<td>9.2nL</td>
<td>20ng</td>
</tr>
<tr>
<td>mmMO</td>
<td>1.087ng/nL</td>
<td>13.8nL</td>
<td>30ng</td>
</tr>
</tbody>
</table>

After being injected the embryos were moved to a new agar lined petri dish filled with fresh 6% ficoll 1x MMR, at the end of the day the embryos were moved to 3% ficoll 0.1x MMR in agar lined plates and the embryos left to develop overnight. The following day the solution was changed to 0.1x MMR for the rest of development, the embryos were observed on a daily basis with any dead embryos being removed and the numbers recorded until they reached stage 40 as described by Nieuwkoop and Faber (1967). At stage 40 embryos were fixed overnight in 4% PFA and then stored in 100% alcohol at -30°C for further analysis.

MOs were designed for mammalian systems where the temperature is 37°C (Moulton and Yan, 2008). They are still functional at cooler temperatures although they are more suited to warmer temperatures (Moulton and Yan, 2008). In initial experiments, the injected *Xenopus* embryos, including controls, didn’t develop well when kept at 24°C so they were subsequently allowed to develop at 18°C for the first 36 hours before being shifted to 24°C. This allows the embryos to handle the fragile early stages of development better. Further, as no Cfm2 expression was detected until stage 35 the activity of the MO at the lower temperature during these early stages should not affect the phenotype.

2.12.2 Cfm2 morpholino/β-galactosidase mRNA co-injections

Further injections were carried out where the same Cfm2 MO was co-injected with β-gal mRNA. The β-gal mRNA will be actively translated during development, this can then be visualised to ensure that the injected solution is being evenly distributed throughout the
embryo. As a nuclear β-\textit{gal} mRNA version was used the somites were also able to easily be visualised for defects after staining for β-\textit{gal} expression.

Embryos were injected as previously described (see 2.12.1) although only one cell was injected at the two-cell stage, 1ng of β-\textit{gal} mRNA and 15ng \textit{Cfm2} MO were injected per embryo. As the first cellular division separates the future left and right sides of the tadpole this meant only one half of the embryo would be affected by the MO injection. When the injected tadpoles reached stage 40 they were again fixed, however MEMFA was used instead of PFA as the fixative for these samples and the tadpoles were fixed for only 2 hours. The tadpoles were then stained using X-gal, which is cleaved by β-gal to produce a substrate which is subsequently oxidised to produce an insoluble blue substrate. This staining allowed me to determine which side of the tadpole had been affected by the MO.

To begin the staining the tadpoles were first washed with PBSAT supplemented with 2mM MgCl\textsubscript{2} for 10-15 minutes, this wash was repeated once. Next, 1mL pre-warmed X-gal staining buffer was added to vials containing the tadpoles and the vials moved to a 37°C incubator for 30 minutes in the dark. Once the embryos had turned a bright blue on one side they were refixed in MEMFA for 1 hour at room temperature, followed by a 10 minute wash with 100% MeOH and then stored in fresh 100% MeOH at -30°C for further analysis.

2.12.3 Morpholino rescue

To test whether the phenotype observed by \textit{Cfm2} MO injection was caused by knockdown of \textit{Cfm2}, and not non-specific affects of the MO injections, further microinjection experiments were designed where \textit{mmCfm2} mRNA was co-injected with the \textit{Cfm2} MO in an attempt to rescue the MO phenotype. The \textit{mmCfm2} mRNA contains 5 synonymous mutations introduced around the start codon of the \textit{Cfm2} transcript where the \textit{Cfm2} MO usually binds. These mutations should prevent the MO binding to the injected \textit{Cfm2} mRNA which will then be actively translated in the injected tadpoles as they develop.

Further rescue experiments were designed to test the functional regions of the \textit{Cfm2} gene. For these experiments \textit{Cfm} mRNA, or, \textit{mmCfm2} mRNA containing either (1) a large deletion cutting out CR1, (2) point mutations causing two serine-alanine amino acid changes at the
detected MAPK phosphorylation site, or (3) point mutations causing two serine-aspartate amino acid changes at the detected MAPK phosphorylation site, were co-injected with the Cfm2 MO. These injections were done using the same protocol as before for carrying out the injections and raising the injected tadpoles (see 2.12.1).

### 2.12.4 Cfm and Cfm2 overexpression

To further examine the effects of Cfm2 expression on the developing tadpoles the final set of microinjection experiments were designed to test the affects of Cfm and Cfm2 overexpression. These injections, and the raising of the injected tadpoles, were carried out as previously described (see 2.12.1), however only Cfm or mmCfm2 mRNA were injected as outlined in table 2.12.

#### Table 2.12 - Overexpression experiment injection solutions

<table>
<thead>
<tr>
<th>Solution</th>
<th>mRNA concentration</th>
<th>Volume injected per cell</th>
<th>mRNA concentration per embryo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cfm mRNA</td>
<td>0.03623ng/nL</td>
<td>13.8nL</td>
<td>1ng</td>
</tr>
<tr>
<td>Cfm mRNA</td>
<td>0.1812ng/nL</td>
<td>13.8nL</td>
<td>5ng</td>
</tr>
<tr>
<td>Cfm mRNA</td>
<td>0.3623ng/nL</td>
<td>13.8nL</td>
<td>10ng</td>
</tr>
<tr>
<td>mmCfm2 mRNA</td>
<td>0.03623ng/nL</td>
<td>13.8nL</td>
<td>1ng</td>
</tr>
<tr>
<td>mmCfm2 mRNA</td>
<td>0.1812ng/nL</td>
<td>13.8nL</td>
<td>5ng</td>
</tr>
<tr>
<td>mmCfm2 mRNA</td>
<td>0.3623ng/nL</td>
<td>13.8nL</td>
<td>10ng</td>
</tr>
</tbody>
</table>
2.13 Histology

2.13.1 Wax embedding

Samples to be sectioned were embedded in paraffin wax by first washing the sample in 1x PBSA for 10 minutes, repeated once. The samples were then dehydrated in EtOH using the following series of washes:

- 50% EtOH for 30 minutes
- 70% EtOH for 30 minutes
- 80% EtOH for 30 minutes
- 90% EtOH for 30 minutes
- 95% EtOH for 30 minutes
- 100% EtOH for 30 minutes
- 100% EtOH for 30 minutes

Samples were then transferred to Xylene and left in a 60°C oven for a further 30 minutes. Following this, samples were moved into melted paraffin wax and left under vacuum in the 60°C oven for 20 minutes. The wax was then replaced with fresh molten wax and the samples were returned to the 60°C degree oven under vacuum for a further 20 minutes. Finally the samples were mounted in wax by placing the sample in a metal mould with fresh molten wax, a labelled cassette was then added and the mould left overnight to allow the wax to set.

2.13.2 Sectioning and mounting

The wax mounted samples were sectioned using a Leica RM 2125 RT microtome, sections were cut 7 - 12μM thick as desired. The wax ribbons, containing these sections, were then placed in cold water for 30 seconds, a labelled frosted glass microscope slide was then used to ‘catch’ the ribbons which were then positioned on the slide and immersed in a 45°C water bath briefly to allow the wax to stretch and smooth out. The slides were then left to dry on a 40°C hot block overnight before proceeding with staining.

2.13.3 Staining

Sections from in situ hybridisations were left unstained as staining these slides, even when stained with eosin alone, was found to mask the staining produced by in situ hybridisation.
Therefore, *in situ* hybridisation sections were washed six consecutive times in xylene, 1-2 minutes per wash, to de-wax the samples and then mounted for long term storage using DPX and a coverslip.

Where the morphological characteristics of the tissue were of more interest, sections were stained with haematoxylin (nuclei) and eosin (cytoplasm). This was done by immersing the slides in the following series of solutions:

- Xylene for 2 minutes
- Xylene for 2 minutes
- Xylene for 2 minutes
- 100% EtOH for 1 minute
- 100% EtOH for 1 minute
- 70% EtOH for 1 minute
- Distilled water for 1 minute
- Gills haematoxylin for 30 seconds
- Tap water for 1 minute
- Schott’s water for 1 minute
- Tap water for 1 minute
- Eosin in 70% EtOH for 30 seconds
- 70% EtOH for 1 minute
- 100% EtOH for 1 minute
- 100% EtOH for 1 minute
- 100% EtOH for 1 minute
- Xylene for 1 minute
- Xylene for 1 minute
- Xylene for 1 minute

The stained sections were then mounted with a coverslip and DPX for long term storage.

### 2.14 Fixing and photography

#### 2.14.1 Fixation

Samples were fixed in glass vials using 4% PFA unless otherwise stated. Younger samples (≤stage 40) were placed straight into PFA, older samples (≥stage 41) were anaesthetised first.
using MS222 and then rinsed with MQW before being placed in PFA. Vials were then placed on a nutator for continuous mixing and swirling at room temperature for at least 1 hour, older stage tadpoles were left in fixative overnight. Following fixation the samples were washed with 100% EtOH for 5 minutes at room temperature on a nutator. The EtOH was then removed and replaced with fresh 100% EtOH and the samples stored at -30°C until required.

2.14.2 Bleaching
To bleach the older pigmented tadpoles for photography, the fixed tadpoles were first washed with 1x PBSA for 10 minutes, repeated three more times. The tadpoles were then bleached by placing the tadpoles in a glass vial with 5% v/v H₂O₂ in 1x PBSA and leaving these on a nutator at room temperature in direct light until the pigment disappeared. Following this the tadpoles were washed once more in 1x PBSA for 10 minutes and then refixed by replacing the solution with formalin for long term storage.

2.14.3 Photography
For photography of embryos and tadpoles, samples were transferred to an agar lined petri dish and photographed using a Leica DF632 camera on a Leica M2FLIII microscope with Leica image manager IM50 software. For clearing, samples were photographed in methyl salicylate in glass petri dishes using the same equipment.

Sectioned samples were photographed using an Olympus DP25 digital camera on an Olympus BX51 compound microscope with Olympus’ DP2-BSW software.

2.15 Statistical analysis
Phenotypes for the MO experiments were measured by the presence or absence of a number of characteristics. To evaluate the difference between these characteristics, measured as a proportion of the total number of injected tadpoles, z-scores were calculated by an online calculator¹ provided by Vassar College which uses the formula

\[ z = \frac{p_1 - p_2}{s_{p_1 - p_2}} \]

1 http://faculty.vassar.edu/lowry/propdiff_ind.html
where $p_1 - p_2$ is the difference between the two compared proportions and $s_{p_1 - p_2}$ is the estimated standard error of the difference between the proportions. The estimated standard error is calculated using the formula $s_{p_1 - p_2} = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$ where $p$ is a weighted average of $p_1$ and $p_2$, calculated using the formula $p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$, and $n_1$ and $n_2$ represent the total number of tadpoles sampled for the two populations. The subsequent z-scores were then used to produce a two-tailed probability using a $z$ table, the resulting probabilities were used to evaluate the null hypothesis that the two proportions are equal (i.e. $p_1 = p_2$). Where the probability was less than the significance level of 0.05 the null hypothesis was rejected and the difference between the two proportions considered statistically significant. To ensure validity, this test was only used for proportions where both $n(p)$ and $n(1 - p)$ were equal to or greater than 5.

To estimate the 95% confidence interval for each proportion a separate online calculator provided by Vassar College was used. This calculator estimates the standard error using the formula $\sigma_p = \sqrt{\frac{p(1-p)}{N}}$, where $p$ represents the proportion and $N$ represents the total number of embryos/tadpoles sampled. This value was then used to calculate the standard error by using the equation $p \pm z \sigma_p$ where $z$ represents the $z$-score for the confidence interval, in this case a 95% confidence interval was chosen, therefore $z = 1.96$. As proportions were used here a correction for continuity was applied to improve the confidence interval, this correction was calculated by dividing the value 0.5 by the total number of embryos/tadpoles sampled ($N$) in each situation. The correction was then applied by subtracting the calculated value from the lower limit of the confidence interval and adding the correction to the upper limit to give the final 95% confidence interval.

---

1 http://faculty.vassar.edu/lowry/prop1.html
Chapter 3 - Homology of *Xenopus laevis* Cfm and Cfm2

3.1 Introduction
The *Cfm* and *Cfm2* genes have not been extensively examined. At present very little is known about the homology of these genes or their phylogenetic relationships. After *Cfm* was initially indentified in mice the paralogous *Cfm2* gene was found (Hirano et al., 2005). Orthologues of both genes were subsequently found in human, chick, *Xenopus* and zebrafish (Hirano et al., 2005). Here I used the CFM and CFM2 protein sequences from *Xenopus laevis* to identify homologous sequences using the online search tools BLAST and Metazome. Protein sequences were used as these sequences are more conserved than DNA sequences due to the ambiguity of the genetic code. To assess the conservation of the identified sequences a multiple sequence alignment was produced and a phylogenetic tree constructed from this alignment. From the resulting phylogenetic tree it is possible to determine the relationship of the identified sequences in an evolutionary context revealing information about the origins of these sequences.

3.2 Orthologues of CFM and CFM2
A protein BLAST (BLASTp) search using the *Xenopus laevis* CFM and CFM2 amino acid sequences identified thirty-two sequences (see Appendix I) that were similar to the *Xenopus* sequences. Similarity was evaluated using E-values calculated by the BLAST search engine. The arbitrary value of $5 \times 10^{-9}$ was chosen as the cut-off point with E-values greater than this considered to be similar by chance and genetically unrelated. The sequences identified were from *Xenopus laevis*, *Xenopus tropicalis*, human, chick, mouse, rat, zebrafish, chimpanzee, rhesus macaque, opossum, dog, cattle, horse, giant panda, wild boar, zebra finch, platypus, salmon and the green-spotted puffer. Interestingly, all nineteen of these species belong to the Chordate subphylum Vertebrata, more commonly known as the vertebrates.

The identified sequences were then classified as either CFM or CFM2 orthologues using the names given to the sequences in the NCBI database where possible. In cases where the descriptive title was unable to discriminate between paralogues the NCBI database provided a list of orthologues of each sequence which could be used to identify the sequence as either CFM or CFM2. In rare cases there were sequences which were still unable to be classified as
either an orthologue of CFM or CFM2, however some of these sequences were then able to be classified as such based on the original sequence used in the BLASTp search which identified them as a potential orthologue. Of the thirty-four sequences in this dataset, thirty could be classified as either CFM or CFM2 using this identification process leaving only four that could not be distinguished. The thirty sequences that could be classified as CFM or CFM2 represented fifteen species, each of which possessed CFM and CFM2 orthologues, the four other sequences were from four separate species - platypus, wild boar, salmon and the green-spotted pufferfish - with each of these only possessing a single orthologue.

To identify any further sequences a separate search was carried out using the Metazome project search engine found online. The Metazome project consists of twenty-four Metazoan proteomes organised by the Joint Genome Institute and the Center for Integrative Genomics ‘to facilitate comparative genomic studies amongst the metazoans’. A keyword search on the Metazome website using ‘FAM101’ as the search term revealed twenty-seven sequences (see Appendix II) belonging to the CFM/CFM2, or FAM101, protein family. These twenty-seven sequences belonged to thirteen different species. Five of these species were not identified as having CFM or CFM2 orthologues in the BLAST dataset - neither below or above the cut-off point - representing a total of nine additional sequences. Two of these newly identified sequences were from a species of lancelet, *Branchiostoma floridae*, and another of the sequences belonged to the owl limpet, *Lottia gigantea* - a sea snail. The six other ‘new’ sequences belonged to three species that are all members of the Vertebrata subphylum. Each of these three species - fugu fish, three-spined stickleback and medaka fish - possessed two homologous sequences. None of the nine newly identified sequences in the Metazome dataset were able to be classified as either CFM or CFM2 orthologues based on the available information, however combining the datasets and constructing a multiple sequence alignment allowed these sequences to be classified as such (see 3.3).

Combining the NCBI and Metazome datasets gave a total of forty-three sequences that belonged to the CFM/CFM2, or FAM101, family of proteins. All forty-three sequences identified were eukaryotic with forty of these sequences belonging to vertebrates. The forty-three sequences represented twenty-four species; nineteen of these species possessed two separate sequences, the other five species each possessed a single sequence. A total of
thirteen sequences were not able to be classified as either CFM or CFM2 orthologues with the information immediately available. Four vertebrate Classes are represented in the combined dataset; the Actinopterygii (a Class of bony fish represented in the combined dataset by zebrafish, the three-spined stickleback, medaka fish, fugu fish, salmon and the green-spotted puffer); the Amphibia (comprised of the amphibians and represented here by the two *Xenopus* species, *X. laevis* and *X. tropicalis*); the Aves (comprised of birds and represented here by the zebra finch and the chick); and lastly the Mammalia (comprised of mammals and represented in the combined dataset by human, chimpanzee, rhesus macaque, mouse, rat, dog, giant panda, horse, cattle, wild boar, opossum and platypus).

### 3.3 Sequence conservation

The sequences from the BLAST and Metazome datasets were downloaded and a multiple sequence alignment of all forty-three protein sequences was produced using ClustalX. The resulting alignment was then modified using BioEdit to correct any obvious errors. Figure 3.1 shows the resulting alignment of a subset of these sequences while the full alignment is attached in Appendix III. The alignment shows a visible, although not strikingly obvious, split between the vertebrate CFM and CFM2 sequences allowing the vertebrate sequences that were previously unable to be distinguished as CFM or CFM2 orthologues to now be classified as such. The unclassified orthologues from the BLAST dataset were identified as orthologues of CFM, salmon and platypus, or orthologues of CFM2, wild boar and green-spotted puffer. The pairs of sequences from the fugu fish, three-spined stickleback and medaka fish in the Metazome dataset were also able to be separately classified as CFM or CFM2 orthologues. These classifications were further confirmed by the clustering of these sequences in the phylogenetic tree (see 3.4). The three invertebrate sequences, from the lancelet and the owl limpet, are much less conserved, as might be expected, and are not readily distinguished as either CFM or CFM2 orthologues in the alignment.

The two conserved regions, CR1 and CR2, identified by Hirano et al. (2005) appear widely conserved in the majority of the sequences identified. These regions are shown shaded in the subset alignment in figure 3.1. However, the alignment also showed the presence of a number of shortened sequences that apparently lack CR1 as exemplified by the wild boar and chimpanzee CFM2 sequences shown in figure 3.1. There are six shortened vertebrate
sequences - wild boar, cattle CFM2, rhesus macaque CFM2, chimpanzee CFM2, giant panda CFM2 and horse CFM2 - that align with the CFM2 sequence cluster in the full alignment. All of these sequences are missing CR1, but they still align well despite the apparent lack of a ‘full’ sequence. All of these shortened CFM2 sequences possess a conserved methionine residue that acts as a secondary start site in vitro. This secondary start site is present in all bar three (mouse, rat and opossum) of the twenty vertebrate CFM2 orthologues and is even found in one of the two lancelet sequences although, as stated earlier, the lancelet sequences align poorly making it difficult to distinguish any true characteristics of these sequences. Eight of the shortened vertebrate sequences align with the CFM sequence cluster, four of these belong to species that also have shortened CFM2 sequences (cattle, chimpanzee, giant panda and horse). No obvious pattern is observed for these shortened CFM sequences in relation to their position in the alignment. It is also of note that none of the twenty CFM sequences possess the secondary start site indicating it is a feature unique to CFM2 sequences.

The dog CFM2 sequence possesses a long amino acid sequence that, through the first 183 residues, does not align well to any other sequences in the alignment. However, residue 184 is a methionine residue that aligns to the secondary start site present in the CFM2 sequences, the rest of the sequence aligns very well to the other vertebrate orthologues over CR2 and the N-terminal end of the protein. Two further vertebrate sequences, green-spotted pufferfish and chick CFM, also possess relatively long C-terminus ends that, in both cases, precede a methionine residue that aligns well to the major start site. With both of these sequences the rest of the sequence aligns well to the other vertebrate orthologues over both CR1 and CR2.

Similarly there are two vertebrate sequences, Xenopus laevis CFM (shown in Fig 3.1) and three-spined stickleback CFM2, which possess a sequence that is noticeably longer at the N-terminus ends than any of the other vertebrate sequences and also appear less conserved over the N-terminal regions of the sequence when compared to the other vertebrate sequences.
Figure 3.1 - Multiple sequence alignment subset. Protein sequence alignment showing a subset of the sequences identified in the combined dataset. Sequences are as labelled. The major start site occurs at position 45 in the alignment; the secondary start site found in the CFM2 sequences is underlined by a blue line at position 140. CR1 is indicated by the shaded region from position 71-82 and CR2 by the shaded region from position 139-240. The MAPK consensus sequence in CR1 is underlined in black. Amino acid residues are coloured according to the default settings in BioEdit.
3.4 Phylogenetic relationships of CFM and CFM2

The sequence alignment containing all forty-three protein sequences was then transferred to the computer program PAUP* and a phylogenetic tree constructed. A neighbour-joining tree was created using the default distance settings and a root established using mid-point rooting. To evaluate the accuracy of the tree, each of the branches was bootstrapped using a heuristic search and 1000 sub-replicates. The resulting phylogenetic tree is shown in figure 3.2. Bootstrap values below 50% are not indicated on the tree due to the support for these nodes being considered weak as they occur in less than half the replicates analysed (Baldauf, 2003), whereas values above 70% are considered “likely to indicate reliable groupings” (Baldauf, 2003).

The tree shown in figure 3.2 comprises two distinct clades representing the CFM (shown in red) and CFM2 (shown in blue) sequences respectively. The tree also confirms the identity of the nine sequences that were unable to be classified as CFM or CFM2 orthologues from the information initially available. The species relationships with respect to their taxonomic ranks are demonstrated in the tree; *X.laevis* and *X.tropicalis* CFM and CFM2 sequences cluster together in each of the CFM and CFM2 clades as do the mouse and rat CFM and CFM2 sequences. The six species from the Actinopterygii cluster together in both CFM and CFM2 clades although the salmon and green-spotted pufferfish only have one orthologue each. Similarly the Mammalia CFM and CFM2 sequences, with the exception of the opossum, cluster together in both CFM and CFM2 clades although the position of the dog CFM and CFM2 sequences switches in the two clusters. In the CFM clade the dog CFM sequence is indicated as being most closely related to mouse and rat CFM sequences, however, in the CFM2 clade the dog CFM2 sequence indicated as being most closely related to the giant panda CFM2 sequence. The Avians also cluster together in the tree although again some ambiguity is present with the chick and zebra finch CFM sequences separated by a significant difference relative to the corresponding CFM2 sequences.
Figure 3.2 - CFM/CFM2 protein family phylogenetic tree. Sequences are as labelled, those shown in red belong to the CFM clade and those in blue belong to the CFM2 clade. Sequences marked with * were unclassified orthologues identified in the BLAST dataset. Sequences marked with ** were unclassified orthologues identified in the Metazome dataset. M1 and M2 labels were added to Metazome sequence names where paralogues were present to distinguish the sequences. Bootstrap values above 50% are shown for nodes on the corresponding branch. The scale represents 0.1 amino acid changes.
In the majority of cases the bootstrap values in figure 3.2 indicate support for the nodes of the tree. A high level of confidence is shown in the relationship between the CFM and CFM2 sequences with a bootstrap value of 100% indicating definitive support for the presence of the two distinct clusters. There is similar support in both clusters of the tree separating the Actinopterygii from the other three Classes (the Amphibia, Aves and Mammalia) represented in this tree. However, less support is shown for the divisions within the Actinopterygii and also within the Mammalia. There is little support for the node separating the Avian CFM sequences from the Mammalian CFM sequences and there is no support for the node separating the Mammalian, Avian and Amphibian CFM2 sequences.

The three invertebrate sequences are separated by some distance from the vertebrate sequences making it difficult to distinguish these sequences as members of the CFM/CFM2, or FAM101, protein family. The owl limpet sequence appears to be very distantly related, if at all, to the vertebrate sequences. Similarly the lancelet -M2 sequence appears to have little relationship to the vertebrate sequences. However, the lancelet -M1 sequence is more closely related to the vertebrate sequences although there is still a large distance between the lancelet -M1 and the vertebrate sequences indicating that any possible relationship is likely to be quite distant. Again the bootstrap values for the nodes separating these invertebrate sequences support the occurrence of these nodes.

3.5 Summary

Using the online search engines BLAST and Metazome, a total of 43 protein sequences were identified as belonging to the CFM/CFM2 family of proteins. Of these, 40 sequences belonged to vertebrates whereas only three invertebrate sequences were identified. Notably, no orthologues were found in any of the well studied invertebrate species, including Drosophila melanogaster, suggesting that this protein family has evolved in vertebrates. Of the three invertebrate sequences identified, two belong to the lancelet, the closest known invertebrate relative of the vertebrates (Putnam et al., 2008), which is further suggestive of this protein family having evolved in vertebrates.

Phylogenetic analysis of the 43 sequences revealed two distinct clades corresponding to the CFM and CFM2 sequences, which is indicative of a duplication event in a common ancestor.
The three invertebrate sequences clustered separately indicating an, at best, distant relationship to the vertebrate sequences. The bootstrap values calculated indicate definitive support for the associations found in this tree. Similarly, the species relationships with respect to their taxonomic ranks found in the tree are, in the most part, in agreement with the sequence relationships found here.
Chapter 4 - Expression patterns of *Xenopus laevis* Cfm and Cfm2

4.1 Introduction

Previous attempts at examining the expression patterns of *Cfm2* in *X. laevis* have been unsuccessful (Holman, 2007). The only published report on the expression patterns of *Cfm2* was carried out by Hirano et al. (2005). This paper used *in situ* hybridisation to detect *Cfm2* mRNA in mice. Expression was seen in the presomitic mesoderm of the segmenting somite, the optic nerve, otic capsule, peripheral nerves, tegmentum, lung and tongue muscle (Hirano et al., 2005). Here, I initially used RNA collected from a number of stages in *Xenopus* development to determine when *Cfm2* is expressed by reverse transcribing the RNA into cDNA and probing this for the *Cfm2* sequence using PCR. Samples were then collected from corroborating stages in development and these then analysed by *in situ* hybridisation using an anti-sense RNA probe to determine where *Cfm2* is expressed in the developing tadpole.

Similarities between the CFM and CFM2 sequences indicate that these two genes may share a similar function. To examine this possibility *Cfm* expression was also analysed using the same methods to analyse any potential overlap in expression that may relate to a common function. *Cfm* has also been examined in mice where it was found to be expressed in the neuroectoderm, which forms the future caudal forebrain and midbrain, and also in the optic rudiment, first pharyngeal arch, Rathke’s pouch, tongue muscle, lung, inner wall of the alimentary canal, genital tubercle and peripheral nerves (Hirano et al., 2005), again, this is the only published report on the expression patterns of the *Cfm* gene. Determining the expression patterns of these genes in *Xenopus* and other model organisms, is essential to understanding the developmental processes they are involved in and elucidating the functional roles they play and the genetic pathways they contribute to.

4.2 RNA extraction

A number of samples were collected from stages 1-15 (pooled sample), 15, 20, 25, 30, 35, 40, 45 and 50. Limb samples were also collected from stage 50 hindlimbs, 51 hindlimbs, 52 hind- and forelimbs, 53 hind- and forelimbs, 54 hind- and forelimbs and 55 hind- and forelimbs. All samples were stored in RNA*later*® solution and the RNA extracted at a later date using the PureLink™ Micro-to-Midi total RNA purification system. The resulting RNA
The absorbance ratios for the samples collected here are, in most cases, close to the ideal 260/280 ratio of 2.0, indicating a high degree of purity was achieved in these samples. To create the cDNA 1µg of RNA was needed, each of the RNA samples collected provided enough total RNA for this purpose although some, including a number of the limb samples, had much lower concentrations than desired due to smaller amounts of starting material. To assess the integrity of the RNA in each of these samples a small amount was also run on a 1% agarose gel as shown in figure 4.1. The gels show two bands corresponding to the
abundant 18S and 28S rRNA in each of the samples analysed indicating that the RNA is intact and suitable for further use. Some of the bands appear fainter than others due to the varied concentrations of RNA in each sample. The 18S bands in the RNA samples collected from the whole tadpole samples are, in general, very faint and difficult to see here without significantly overexposing the gel.

Figure 4.1 - RNA integrity. RNA integrity from the samples collected during embryonic development as indicated (A). RNA integrity from the samples collected during limb development as indicated (B). In both gels 1Kb Plus DNA ladder was run in the far left and far right columns as a reference. St refers to the stage as described by Nieuwkoop and Faber (1967) (see appendix VIII). In both gels the bands seen in each sample represent 18S rRNA (~650bp) and the larger 28S rRNA (~850bp).

4.3 PCR analysis

cDNA was created from the RNA samples collected using SuperScript® III Reverse Transcriptase (Invitrogen) and an oligo(dT) oligonucleotide primer. The oligo(dT) oligonucleotide primer binds to the poly-A tail of mature mRNA molecules allowing all mature mRNAs expressed at any given time in a cell to be reverse transcribed into representative cDNA. Two reactions were simultaneously set up to act as negative controls. One of these had no RNA included in the reaction to check for contamination in the reagents, and the other had no reverse transcriptase included in the reaction to check for any genomic DNA contamination in the RNA samples. PCR reactions were set up with Cfm2, Cfm or β-actin primers using the resulting cDNA as a template. Primer sequences and PCR programs are shown in appendix IV. The products of these reactions were run on an agarose gel as shown in figure 4.2.
Figure 4.2 - PCR analysis of gene expression in development. PCR analysis of $\beta$-actin expression (A, B), Cfm expression (C, D) and Cfm2 expression (E). Each lane is as labelled. For all gels 1Kb Plus DNA ladder was run in the far left and far right columns as a size reference. Expression at a number of stages in embryonic development as labelled (A, B, C). Expression at a number of stages in limb development as labelled (B, D). St refers to the stage as described by Nieuwkoop and Faber (1967) (see appendix VIII). The negative (-ve) controls in A and B also apply to D and E respectively.

To ensure that the cDNA reflects gene expression accurately $\beta$-actin expression was analysed as a positive control. The results for $\beta$-actin expression are shown in panels A and B of figure 4.2. A single band of 196bp was detected in all samples except for the negative controls, as expected.

The Cfm PCR profiles are shown in panels C and D of figure 4.2. Cfm expression was detected by the presence of a single band at 782bp as shown. These results show that Cfm is constitutively expressed in the embryo during development up until at least stage 50 and similarly in the limb from stage 50-55 of the hindlimb and 52-55 of the forelimb.

The Cfm2 PCR profile for the samples collected from whole embryos is shown in panel E of figure 4.2. Cfm2 expression was detected by the presence of a single band at 775bp as shown. These results show that Cfm2 is expressed from stage 35 to at least stage 45 during development. No samples from stages 1-15 and 50 were analysed for Cfm2 expression and none of the samples collected during limb development were able to be analysed. Similarly, no working negative control from samples without reverse transcriptase present in the reaction was obtained. This was due to a number of issues encountered specifically with the Cfm2 PCR. The clean, although incomplete, profile shown in figure 4.2 was obtained using primers designed by a previous lab member. Attempts at reproducing this Cfm2 profile using the same primers and PCR program resulted in large smears in the gel due to primer
degradation. New primers were designed and ordered to solve this issue, however, despite producing clean bands the new primers also produced bands in the negative controls indicating contamination had occurred. After changing the reagents of the PCR and using new cDNA samples the contamination issue was not resolved. Despite the inability to reproduce this Cfm2 PCR profile it was subsequently validated by in situ hybridisation as discussed in the following section.

4.4 In situ hybridisation of Cfm and Cfm2

Whole mount in situ hybridisation was carried out on developing tadpoles using anti-sense DIG-labelled RNA probes prepared, as needed, from plasmid constructs containing the gene of interest. Embryos were prepared and raised as described earlier, when they reached the stage of interest - where Cfm and/or Cfm2 expression had been detected by PCR analysis - they were fixed in 4% PFA for a minimum of 1 hour. Cfm expression was analysed at a range of stages from stage 20 up to stage 40 in the developing tadpole and from stage 50 up to late stage 54 in limb development. Cfm2 expression was analysed at a range of stages from stage 20 up to stage 40 in the developing tadpole and also from stage 50 up to stage 57 in limb development.

4.4.1 Cfm2 expression in Xenopus laevis

The results from the Cfm2 in situ hybridisation experiments in the developing tadpole are shown in figure 4.3. At stage 40 staining was observed in the head and the pronephros, at this stage staining was clearly localised to specific regions in the head of the tadpoles allowing this to be described with a high degree of confidence.
Figure 4.3 - *Cfm2* expression in the developing tadpole. *Cfm2* expression at stage 40 (A, B, E, F, G), B is a magnified image of A; E-G are 12µm sections of a stage 40 tadpole showing *Cfm2* expression in the approximate positions indicated in A. Staining is indicated in the olfactory organ (green arrowheads, A, B), the lateral line precursors (black arrowheads, A, B, G), pronephros (red arrowheads, A, B), ventral branchial arch region (blue arrowheads, A, B, F, G) and head mesenchyme (orange arrowheads, A, B, E, F). Negative control using a sense RNA probe at stage 40 (C, D), D is a magnified image of C. Nonspecific staining is indicated in the otic vessel (white arrowheads, C and D). Scale bars represent 1mm in A-D, 100µm in E and F and 200µm in G. Lateral view is shown in each panel with anterior left (A-D) and dorsal top (A-G).

Figure 4.3 shows this staining at stage 40, when expression is observed in the pronephros (red arrowheads, Fig 4.3 A and B) and in the head in four distinct regions. The staining seen
in the head corresponds to the olfactory organ (green arrowheads, Fig 4.3 A and B), the lateral line precursors\(^1\) (black arrowheads, Fig 4.3 A, B and G), the ventral region of the branchial arches (blue arrowheads, Fig 4.3 A, B, F and G) and the head mesenchyme (orange arrowheads, Fig 4.3 A, B, E and F). Some staining is present in the negative controls in the otic capsule (white arrowheads, Fig 4.3 C and D) due to probe trapping. This issue is common in *Xenopus* from these stages and unpreventable. There is also some non-specific staining present at the outer edges of the tadpoles (Fig 4.3 A, B, C and D). The staining in the olfactory organ and the pronephros has not yet been confirmed in sections although convincing evidence of this expression is found in the whole mount samples.

The results from the *Cfm2* in situ hybridisation experiments in the developing limb are shown in figure 4.4. Staining was initially detected in the hindlimbs at stage 52 as a narrow stripe of expression at the knee joint and as a separate, stronger, band at the ankle joint (white arrowheads, Fig 4.4 A). A small protrusion from the ankle staining is also visible at the base of digit IV as it begins to form. At stage 53 staining in the hindlimb was restricted to two small regions of digits III and IV (white arrowheads, Fig 4.4 B) that correspond to the approximate position of the phalangeal joints of the forming digits. At stage 54 staining was still present at the phalangeal joints of digits III and IV and a similar band of staining was also visible at the phalangeal joint of digit V (white arrowheads, Fig 4.4 C). In the late stage 54/early stage 55 hindlimb staining was present in all five of the digits of the hindlimb at the sites of the phalangeal joints (white arrowheads, Fig 4.4 G), in digit III a second, similar, band of staining was also visible. No expression was detected in the limbs collected from later stages in development up to stage 57, not shown, although it is possible that these samples were under-fixed for in situ hybridisation of these older stages. Although a stage specific match for the late stage 54/early stage 55 hindlimb (Fig 4.4 G) was not found in any of the samples analysed, no staining was seen in any of the negative controls examined.

\(^1\) The lateral line is a sensory organ found in aquatic species that is believed to detect pressure and vibrations (Cernuda-Cernuda and García-Fernández, 2006).
In the forelimb staining was initially detected at stage 53 as three distinct stripes at the shoulder, elbow and wrist joints (white arrowheads, Fig 4.4 H). By stage 54 the staining had become restricted to the wrist joint (white arrowheads, Fig 4.4 I) similar to what was observed in the hindlimb at stage 52. Again, no staining was observed in any of the negative controls examined. Because the forelimb develops under a flap of skin in *Xenopus* and its development lags behind the hindlimb expression was not examined as extensively in the forelimb as it is both less convenient and more time consuming to analyse.

**Figure 4.4 - Cfm2 expression in the developing limb.** *Cfm2* expression in the hindlimb at stage 52 (A), 53 (B), 54 (C) and 54/55 (G) and in the forelimb at stage 53 (H) and 54 (I). White arrowheads indicate staining in A-C and G-I. Negative controls from an albino frog, lacking pigment, using a sense RNA probe are shown in D, E, F, J and K at matched stages. Scale bar represents 1mm in all panels. Digits are indicated as I, II, III, IV or V as they form in the hindlimb. Limbs were dissected for photography, lateral view is shown with proximal top and anterior left.
4.4.2 Cfm expression in *Xenopus laevis*

The results from the *Cfm in situ* hybridisation experiments in the developing tadpole are shown in figure 4.5. Expression was first detected both dorsally and ventrally in the posterior region of the embryo at stage 21 (white arrowheads, Fig 4.5 A), a small region of staining at the anterior end of the embryo was also apparent (black arrowhead, Fig 4.5 A) although this staining was inconsistent in samples examined from the same experiment. At stage 23 any anterior staining disappeared but the posterior staining became much stronger and was again located both dorsally and ventrally (white arrowheads, Fig 4.5 B).

At stage 25 some staining was observed at the anterior end of the embryos once more (black arrowhead, Fig 4.5 C) although this staining was again inconsistent in samples from the same experiment. The posterior staining became much more refined at this stage with two visibly distinct areas of expression in the dorsal and ventral regions of the tadpole (white arrowheads, Fig 4.5 C) although some overlapping staining is apparent between the two regions. By stage 26 the anterior staining was again lost and the posterior staining along the dorsal surface had expanded significantly to cover much of the tadpole (red arrowheads Fig 4.5 G). The ventral staining remains at this stage but appears as a comparatively smaller stripe along the posterior-ventral surface of the endoderm (green arrowhead, Fig 4.5 G).

The staining pattern remained relatively unchanged in stage 27 embryos although a thin strip of staining (blue arrowheads, Fig 4.3 H) was now visibly distinct in a region just dorsal to the previously identified staining in this region (red arrowheads, Fig 4.3 H). At stage 29/30 a similar staining pattern remained (Fig 4.5 I). As the tail bud began to form the staining along the dorsal surface straightened and eventually became fully distinct from the staining in the ventral region (Fig 4.5 I).

At stages 31 (Fig 4.5 M) and 34 (Fig 4.5 N) the staining patterns remained unchanged although the staining observed in stage 34 was noticeably fainter. However, *in situ* hybridisation is not quantitative and this may reflect individual differences between the samples analysed. By stage 40 the broad staining patterns observed earlier had disappeared. A small region of staining was observed at the dorsal end of the proctodeum (white arrowhead, Fig 4.5 O) although this was inconsistently observed in samples examined
from the same experiment once more. Two very fine strips of staining were observed at the posterior end of the tadpoles (blue arrowheads, Fig 4.5 O) and some staining was present in the anterior region (black arrowheads, Fig 4.5 O) although this was difficult to reconcile with the non-specific staining seen in the same region of the negative controls at stage 40 (Fig 4.5 R).

Samples were then sectioned to determine what tissue the staining seen in stage 30 and 40 tadpoles located to. The resulting sections are shown in figure 4.6. Sections from the stage 30 tadpoles revealed the dorsally located thin strip of staining was expressed in the ventral region of the neural tube known as the floor plate (blue arrowheads, Fig 4.6 A - C). Staining was also apparent in the hypochord\(^1\) (black arrowheads, Fig 4.6 A- C) although this would have been masked in the whole mount photographs due to the broad stripe of expression present. This broad stripe of expression corresponds to the dorsal lateral plate mesoderm (red arrowheads, Fig 4.6 A - C). In more rostral sections (Fig 4.6 A and B) this staining appears much more dorsally located although this is likely still lateral plate mesoderm. The staining seen at the posterior-ventral surface of the stage 30 tadpoles appears to be localised in the developing proctodeum although this is yet to be confirmed in sections. Sections from stage 40 tadpoles (Fig 4.6 D) show the two thin strips of expression to be localised to the neural tube floor plate (blue arrowhead) and hypochord (black arrowhead) as was seen at stage 30, although this staining is much more difficult to see in the stage 40 sections.

\(^{1}\) The hypochord is a transient structure found in fish, lampreys and amphibians (Cleaver et al., 2000). It has an unknown function but it has been proposed to play a role in the positioning of the dorsal aorta (Cleaver et al., 2000).
Figure 4.5 - *Cfm* expression is the developing tadpole (I). *Cfm* staining at stages 21 (A), 23 (B), 25 (C), 26 (G), 27 (H), 29/30 (I), 31 (M), 34 (N) and 40 (M). See text for an in depth description of the staining observed. D, E, F, J, K, L, P, Q and R are stage-matched controls using a sense RNA probe. Scale bars represent 1mm. Lateral view is shown in each panel with anterior left and dorsal top.
Figure 4.6 - *Cfm* expression is the developing tadpole (II). Transverse sections showing *Cfm* expression at stage 30 (A-C) and stage 40 (D). Staining is indicated in the floor plate (blue arrowheads), the hypochord (black arrowheads) and lateral plate mesoderm (red arrowheads). The approximate positions of sections in A-D are indicated in E and F. The structures of a mid-trunk stage 30 transverse section are shown in G (modified from Cleaver et al., 2000). Scale bars represent 100µm in A, B, and C and 200µm in D, dorsal is top in each panel.

The results from the *Cfm* in situ hybridisation experiments in the developing limb are shown in figure 4.7. Staining was detected as early as stage 50 as a band of expression around the edge of the hindlimb bud (white arrowheads, Fig 4.7 A), this edge effect staining is indicative of staining in the epidermal layer of the limb bud. As the staining is seen right the way around the limb bud at this stage, it is likely to indicate *Cfm* expression over the entire epidermal surface of the limb. Staining continues at stage 51 but is now more spread out, with the edges at the more proximal regions of the limb staining stronger than in the distal region (white arrowheads, Fig 4.7 B). At stage 52 the more distal regions of the limb continue to show faint staining although strong staining continues in the proximal regions of the limb (white arrowheads, Fig 4.7 C). By stage 53 staining is not apparent in the distal region of the limb where the flattened plate-like morphology forms as the foot begins to take shape (Fig 4.7 G). However, staining still persists in the more proximal regions of the epidermal layer (white arrowheads, Fig 4.7 G) although it is obscured by the pigmentation of the limb along the posterior edge of the limb. By stage 54 the same staining pattern
continues (white arrowheads, Fig 4.7 H) with no expression seen in the distal region of the limb as the digits begin to form.

**Figure 4.7 - Cfm expression in the developing limb.** A, B, C, G and H show Cfm expression in the hindlimb at stage 50 (A), 51 (B), 52 (C), 53 (G) and 54 (H). White arrowheads indicate staining. Negative controls using a sense RNA probe on matched limb stages from albino frogs lacking pigment are shown in D, E, F, J and K. Scale bar represents 1mm in all panels. Lateral view is shown with proximal left and anterior top in all panels. The location of the developing digits is indicated at stage 54 (H) for reference.

### 4.4.3 Cfm2 expression in the axolotl, *Ambystoma mexicanum*

*In situ* hybridisation was also attempted in the axolotl, *Ambystoma mexicanum*, using an anti-sense probe produced from the Cfm2 EST. This was successfully cloned from cDNA prepared from RNA collected from day 5 axolotl embryos. However no *in situ* hybridisation results were obtained. This problem was not limited to Cfm2 in the axolotl as many
attempts at performing \textit{in situ} hybridisation with various probes were made without success. As our lab has only recently started working with axolotls we have unfortunately not yet found an effective method for carrying out \textit{in situ} hybridisation in the axolotl.

\subsection*{4.5 Summary}

\textit{Cfm2} expression was detected from stage 35 to at least stage 45 in the whole tadpole by amplifying cDNA constructed from RNA collected at a number of stages in development. This expression profile was confirmed by \textit{in situ} hybridisation at stage 35 and stage 40. At stage 40, \textit{Cfm2} expression was detected in the olfactory organ, lateral line precursors, the pronephros, the ventral region of the branchial arches and the head mesenchyme. Expression was also detected in the developing limbs by \textit{in situ} hybridisation. In the hindlimb \textit{Cfm2} expression was initially found at the knee and ankle joints, expression was also found in the phalangeal joints in a sequential manner consistent with the sequential formation of the digits. \textit{Cfm2} expression appeared to be transient in the joints of the limbs whereas expression in the head and pronephros appeared more stable.

\textit{Cfm} expression was detected in a pooled sample from stage 1-15 through to at least stage 50 in the cDNA collected from the developing tadpole indicating that it is constitutively expressed. Similar results were seen in the cDNA collected from the limbs during development with \textit{Cfm} expression detected in all the collected samples from stage 50 through to 55 for the hindlimb and from stage 52 through to stage 55 for the forelimb. \textit{In situ} hybridisation revealed that \textit{Cfm} is expressed in the hypochord, the floor plate, proctodeum and the lateral plate mesoderm for long periods during development. Expression was also observed in the anterior region in a more temporal manner during certain stages of development. In the limb, \textit{in situ} hybridisation revealed that \textit{Cfm} is expressed in the epidermal layer of the early limb bud, and similar expression was observed in the more proximal regions of the limb during later stage of limb development.

No overlapping expression was detected between the \textit{Cfm} and \textit{Cfm2} genes at any point during development of the tadpole or in limb development although there is some evidence that \textit{Cfm} is expressed in the head of stage 40 tadpoles at the same time \textit{Cfm2} is expressed in the head. The expression of \textit{Cfm} in the head was hard to accurately define due to a large
amount of background staining present in the negative controls although such possible overlapping expression is intriguing.
Chapter 5 - Misexpression of \textit{Cfm2} in \textit{Xenopus laevis}

5.1 Introduction

One of the advantages of working with the \textit{Xenopus laevis} model organism is that they are an oviparous animal\textsuperscript{1} making them excellent for developmental studies due to the ease of which their embryos can be manipulated and also because it is possible to observe the whole process of development without intrusion. These characteristics make \textit{Xenopus} embryos amenable to microinjection with morpholino oligonucleotides (MOs\textsuperscript{2}). Using MOs is an efficient, relatively easy and much quicker method to control gene expression than creating transgenic lines, although this method results in gene ‘knockdown’ rather than a complete ‘knockout’.

Previously morpholino oligonucleotide (MO) mediated knockdown of the \textit{Cfm2} gene has been shown to have severe affects on \textit{Xenopus} development (Holman, 2007). Here, I continued these MO knockdown experiments in \textit{Xenopus} by microinjecting \textit{Cfm2} MO into a number of fertilised wild type embryos and observing the effects on development in an attempt to determine the developmental processes affected by \textit{Cfm2} misexpression. I then attempted MO rescue experiments by co-injecting MO and mismatched \textit{Cfm2} mRNA to show that the MO phenotype was specifically caused by a lack of \textit{Cfm2} expression. The effects of exogenous \textit{Cfm2} or \textit{Cfm} mRNA were also examined by injecting \textit{Xenopus} embryos with \textit{Cfm2} or \textit{Cfm} mRNA to further examine the developmental processes affected by \textit{Cfm2} and also \textit{Cfm} misexpression.

5.2 \textit{Cfm2} knockdown

Embryos were collected and fertilised as described earlier. At the two-cell stage in development these embryos were injected with either (1) \textit{Cfm2} MO, (2) a control mismatched (mm)MO, or, (3) MQW as a second negative control. As a third negative control

\textsuperscript{1} Oviparous animals are animals that lay eggs with little or no development in the mother.

\textsuperscript{2} MOs are small synthetic molecules that bind to mRNA targets to prevent cellular proteins binding subsequently inhibiting essential processes such as translation and splicing depending on the target sequence (Moulton and Yan, 2008).
a small number of embryos were collected and treated along with the other embryos although they were not injected at all. A comparison of the survival rate of these treatment groups, measured as the number of embryos that survived to stage 40 as a proportion of the number of embryos in each treatment group, is shown in figure 5.1.

![Survival rate](image)

**Figure 5.1 - Morpholino injection survival rates.** Bar graph showing the number of embryos in each treatment group that survived until stage 40 in development as a proportion of the total number of embryos from each treatment group. Error bars show the 95% confidence interval. No injection control, n (total number) = 172; 13.8nL MQW injection, n = 101; 30ng mmMO injection, n = 193; 30ng MO, n = 114; 20ng MO, n = 116. * indicates significant differences (p-value < 0.05), ** indicates highly significant differences (p-value < 0.01).

Statistical comparison of the survival rates for the three control groups - no injection, 13.8nL MQW injection and 30ng mmMO - reveals a statistically significant difference between the no injection control and the embryos injected with MQW (p-value = 0.0396), and also between the embryos injected with MQW and those injected with 30ng mmMO (p-value < 0.001), but not between the no injection controls and those injected with 30ng mmMO (p-value = 0.1892). Such a high survival rate for embryos injected with the mmMO was unexpected but reflects unavoidable differences between the quality of embryos from each frog. The survival rates for the embryos injected with 20 or 30ng of Cfm2 MO were 87.9% and 69.3% respectively. A statistically significant difference is found between these two figures (p-value < 0.001) suggesting that a higher dose of MO is more lethal to the embryos.
as might be expected. A comparison of the survival rates between the 30ng MO injected embryos (69.3%) and the 30ng mmMO injected embryos (94.3%) also show a statistically significant difference (p-value <0.001) with a decrease seen in the survival rate of the MO injected tadpoles indicating that the negative control had worked as desired. The 20ng MO treatment group was not compared with the 30ng mmMO treatment group due to the dosage differences. The Cfm2 MO injected treatment groups were not compared to the MQW water injected treatment group of the uninjected control treatment group as these were only used to establish the mmMO negative control was effective.

The injected embryos were allowed to develop until they reached stage 40. It was notable that MO injected tadpoles developed slower than the controls. As all tadpoles were fixed at the same time, to keep variables to a minimum, the MO injected tadpoles did not reach the same stage in development as the controls when fixed despite being the same age. The resulting tadpoles were scored for each of the following characteristics: the abnormal appearance of the ventral tail fin; the lack of, or a drastically shortened, tail; the occurrence and severity of axial bending in the body of the tadpole; the presence of a ‘kink’ in the tail; abnormal craniofacial features; abnormal, or missing, eyes; and a ‘bloated’ body, which indicates an edema. Examples of each of these features are shown in figure 5.2. Eye deformities ranged from tadpoles missing an eye or eyes, those with a small eye or eyes and a complete lack of pigment in one or both eyes. It was also noticeable that the MO injected tadpoles had much less body pigmentation than mmMO injected or uninjected tadpoles.
Figure 5.2 - Phenotypic characteristics of microinjected tadpoles. WT stage 40 tadpole with basic features indicated (A). Abbreviations; Cg, cement gland; Df, dorsal fin; Ey, eye; Pd, proctodeum; Vf, ventral fin; Vs, viscera. Positions of white asterisks indicate approximate position of rostral- and caudal-most somites. Injected tadpoles (B-H) with abnormal characteristics indicated, white arrowheads; abnormal ventral fin and kinked tail (B); no tail (C); mildly bent body (D); bent body (E); severely bent body (F); bloated body and craniofacial deformities (G); eye deformity (H). Each panel shows the lateral view with anterior left and dorsal top - with the exceptions of F, which shows the dorsal view, and I, where anterior is right.

Figure 5.3 shows the occurrence of each characteristic as a proportion of the total number of tadpoles from each of the treatment groups. For the proportions of wild type tadpoles measured in the controls (Fig 5.3 A) the difference between the MQW injected embryos and the no injection controls was found not to be statistically significant (p-value = 0.1164), neither was the difference observed between the MQW injected embryos and the 30ng mmMO injected embryos (p-value = 0.9474). However, the difference seen between the no injection controls and the 30ng mmMO injected embryos is statistically significant (p-value = 0.0478).

Of the nine abnormal phenotypic characteristics measured for the control groups only two characteristics, bloated body (Fig 5.3 H) and eye deformities (Fig 5.3 J), were able to be compared statistically. The rest of the characteristics were not able to be compared as requirements for validity in the statistical test (i.e. \( n(p) \geq 5 \) and \( n(1-p) \geq 5 \)) were not met. In each of the two comparable characteristics one of the control treatment groups is not able to be compared. Again this is due to the data not meeting the validity requirement. For tadpoles with eye deformities only the difference between the 30ng MO injected embryos and the 30ng mmMO injected embryos can be tested. The difference here is not statistically significant (p-value = 0.7849), and for tadpoles with a bloated body only the
difference between the 30ng mmMO injected embryos and the no injection controls can be tested. The difference here is statistically different (p-value <0.001) with a small increase seen in the 30ng mmMO treatment group. Although p-values cannot be calculated for the vast majority of these characteristics, there is an obvious increase in nearly all the characteristics measured in the 30ng mmMO injected embryos relative to the other negative controls. This is particularly obvious with the proportions measured for a shortened/lost tail, severely bent tadpoles, bloated tadpoles and those with craniofacial deformities indicating that even the mmMO induces developmental defects.

Comparison of the 30ng mmMO injected embryos with the 30ng MO injected embryos revealed a statistically significant increase in the number of tadpoles injected with the Cfm2 MO for: an abnormal ventral fin (Fig 5.3 B, p-value <0.001), a bent body (Fig 5.3 E, p-value <0.001), a kinked tail (Fig 5.3 G, p-value <0.001), and eye deformities (Fig 5.3 J, p-value <0.001). No significant difference was found for tadpoles with: a shortened or no tail (Fig 5.3 C, p-value = 0.3340), severely bent bodies (Fig 5.3 F, p-value = 0.1375), a bloated body (Fig 5.3 H, p-value = 0.0963) or craniofacial deformities (Fig 5.3 I, p-value = 0.5775). Once again the requirements for validity of the statistical testing were not met by the values for wild type (Fig 5.3 A) or mildly bent tadpoles (Fig 5.3 D) although a very obvious difference is observed for wild type tadpoles with nearly 80% of the 30ng mmMO injected tadpoles being wild type and none of the 30ng MO injected tadpoles classified as wild type.

Comparison of the 30ng MO and 20ng MO injected embryos revealed a statistically significant increase in the 30ng MO treatment group for tadpoles with a bent body (Fig 5.3 E, p-value <0.001) or a bloated body (Fig 5.3 H, p-value = 0.0286). Although increases were also observed for tadpoles with a severely bent body (Fig 5.3 F), facial deformities (Fig 5.3 I) and eye deformities (Fig 5.3 J), none of these differences were found to be statistically significant (p-values = 0.7726, 0.0612 and 0.5439 respectively). A statistically significant decrease is seen in 30ng MO treatment group relative to 20ng MO treatment group for tadpoles with an abnormal ventral fin (Fig 5.3 B, p-value = 0.0116) or a kinked tail (Fig 5.3 G, p-value <0.001). The requirements for validity of the statistical testing were once more not met for tadpoles with shortened or no tail (Fig 5.3 C) and tadpoles with a mildly bent body (Fig 5.3 D) although a large increase is seen in the 30ng MO treatment group (17.5%) relative
to the 20ng MO treatment group (3.4%) for tadpoles with a shortened or no tail, and a large
decrease is seen in the 30ng MO treatment group (2.5%) relative to the 20ng MO treatment
group (16.4%) for tadpoles with a mildly bent body.

Attempts were made to grow some of the *Cfm2* MO injected tadpoles up to the limb stages
to examine the effects of *Cfm2* knockdown on limb development but the tadpoles, even
those injected with lower doses of the MO, had very limited mobility and died before
reaching the limb stages.
Figure 5.3 - Morpholino injection phenotypes. Bar graphs showing the occurrence of a number of characteristics, measured independently, in each treatment group as a proportion of the total number of embryos from each treatment group. The legend in A applies to all graphs. Error bars show the 95% confidence interval. No injection control, n (total number) = 172; 13.8 nL MQW injection, n = 101; 30 ng mmMO injection, n = 193; 30 ng MO, n = 114; 20 ng MO, n = 116.
5.3 Developmental markers in morpholino injected tadpoles

To test the genetic affects of \textit{Cfm2} knockdown the expression of two well-characterised developmental genes - \textit{Shh} and \textit{Sox3} - were examined by \textit{in situ} hybridisation (figure 5.4). To ensure no somitic defects were present a nuclear $\beta$-galactosidase mRNA transcript was co-injected with the MO in a number of embryos at a concentration of 1ng per embryo. This was later visualised by staining the tadpoles post-fixation, as this staining is located to the nuclei which outlines the somites to reveal any obvious somitic defects (figure 5.4). \textit{Shh} encodes an important ligand of the hedgehog signalling pathway. It is known to be an important regulator of organogenesis, including the development of the vertebrate limb and brain, and it is also known to be involved in the patterning of the central nervous system. In \textit{Xenopus} \textit{Shh} is expressed in the branchial arches, olfactory organ, notochord and floor plate in later stage tadpoles (Ekker et al., 1995). \textit{Sox3} is another important developmental gene, it encodes a transcription factor expressed in the branchial arches, lens, forebrain, lateral line precursors, olfactory organ and the posterior neural tube during \textit{Xenopus} development at later stages (Penzel et al., 1997). These two genes were chosen due to their overlapping expression with \textit{Cfm2} as described earlier (see 4.4.1).

\textit{Shh} and \textit{Sox3} \textit{in situ} hybridisations were initially planned to be carried out in the embryos co-injected with \textit{Cfm2} MO and $\beta$-\textit{gal} mRNA. As only one cell was injected at the two-cell stage for these experiments only one side of the developing tadpole is affected, this would allow the MO and the wild type expression pattern to be examined in one tadpole. However, the intensity of the $\beta$-\textit{gal} staining, required for determining which is the injected side, masked the \textit{in situ} hybridisation staining. As a result the \textit{Shh} and \textit{Sox3} \textit{in situ} hybridisations had to be carried out on the embryos injected in both cells at the two-cell stage with \textit{Cfm2} MO alone. Future experiments using a lower concentration of $\beta$-\textit{gal} mRNA may make the \textit{Cfm2} MO, $\beta$-\textit{gal} mRNA co-injection approach more informative.

Figure 5.4 shows the results of analysis of these markers in MO injected tadpoles. In a wild type, uninjected control, tadpole \textit{Shh} expression is detected in the branchial arches,
olfactory organ, floor plate and notochord\(^1\) (Fig 5.4 A) as was expected. *Shh* expression in the MO injected tadpoles remained relatively unchanged in all the samples examined regardless of the phenotypic defects seen in these tadpoles (Fig 5.4 D, G, J, J', M, M'). Slight changes in expression were observed however, this is most likely due to individual differences and also developmental stage differences as the MO injected tadpoles developed at a slower rate than the uninjected controls. A noticeable difference in gene expression is seen in the tadpole shown in J and J', however the abnormal *Shh* expression is limited to one side of the tadpole only (Fig 5.4 J). Why this difference has occurred is unclear but the affected side also lacks a developed eye suggesting that development on that side of the tadpole has been limited in general. As similar eye deformities were noticed in mmMO injected tadpoles it appears that this defect is caused by injecting the embryos rather than the nature of the solution injected itself.

A wild type, uninjected control, tadpole showing *Sox3* expression is shown in figure 5.4 B. Expression was detected in the branchial arches, olfactory organ, forebrain, midbrain, hindbrain, lateral line precursors and posterior neural tube as was expected. The *Sox3* expression patterns in the MO injected tadpoles (Fig 5.4 E, H, K, N) were similar to the expression patterns observed in the uninjected control tadpoles. Expression, particularly in the posterior neural tube, was noticeably weaker in the MO injected tadpoles (Fig 5.4 E, H, N). This is most obvious in figure 5.4 E where expression in the neural tube is barely visible, the same tadpole shows no signs of expression in the branchial arches whereas this expression is present, although reduced, in the other tadpoles examined. No noticeable differences in the expression of *Sox3* were observed in the olfactory organ and in the fore-, mid- and hindbrain of MO injected tadpoles. Expression was detected in the lens of the MO injected tadpoles (white arrowheads, Fig 5.4 E, H, K, N), however this expression is also observed in younger wild type tadpoles and again reflects the differences in the stage of the tadpoles rather than a specific difference in gene expression caused by the MO.

\(^1\) The floor plate and notochord sit on top of each other, thus it is impossible to distinguish expression between these two structures in whole mount samples.
\(\beta\)-\textit{gal} expression is shown in a mmMO injected tadpole in figure 5.4 C, the first three somites are indicated. As a nuclear copy of \(\beta\)-\textit{gal} was used the somites are clearly indicated although \(\beta\)-\textit{gal} expression is not just limited to the somites. MO injected tadpole expressing \(\beta\)-\textit{gal} are shown in figure 5.4 F, I, L and O. A number of the different phenotypes observed are represented here, noticeably none of these phenotypes show any general defects in the formation of the somites. The rostral somites of a bloated tadpole (Fig 5.4 I) are not visible, however this appears to be due to defects in the surrounding tissue inhibiting visualisation of the somites rather than the formation of the somites as the caudal somites appear unaffected. Similarly, the ‘kinked tail’ tadpole (Fig 5.4 L) shows a somitic defect at the site of the kink (white arrowhead), as this problem is limited to a small region it appears to indicate a specific defect rather than a general defect in somitogenesis.
Figure 5.4 - Analysis of developmental markers in morpholino injected tadpoles. *In situ* hybridisation of Cfm2 MO injected tadpoles using a *Shh* probe (D, G, J, J’, M, M’) or a *Sox3* probe (E, H, K, N); and tadpoles co-injected with β-gal mRNA and Cfm2 MO stained for β-gal expression (F, I, L, O). WT tadpoles in A and B show normal *Shh* or *Sox3* expression with tissue labelled, a control β-gal mRNA, mmMO co-injected wild type tadpole is shown in C with the first three somites indicated. J, J’ show the left and right sides of a tadpole expressing *Shh* on one side only. M, M’ show two different views of a severely affected tadpole still expressing *Shh* normally. Scale bars represent 1mm in all panels. Lateral view is shown in each panel with dorsal top, anterior is left in A-E, G, H, J, K-N and right in F, I, J’,O. Abbreviations used in A, B and C: Ba, branchial arches; Fp, floor plate; Fb, forebrain; Hb, hindbrain; Ll, lateral line precursors; Mb, midbrain; Nc, notochord; Nt, neural tube; Ol, olfactory organ; Sm, somite.
To analyse any morphological abnormalities in the Cfm2 MO injected tadpoles transverse sections were taken and stained using haemotoxylin and eosin (figure 5.5). These sections were used to examine any morphological abnormalities in the structures of the face (C, D), the heart (E, F), the pronephros (G, H), the mid-trunk region (I, J) and the tail region (K, L).

Figure 5.5 - Morphology of morpholino injected tadpoles. Wild type (A) and MO (B) injected tadpoles with the approximate positions of the corresponding sections indicated. Structures are as indicated for WT (C, E, G, I, K) and MO injected sections (D, F, H, J, L) as appropriate. Abbreviations: Df, dorsal fin; Gt, gut; Ht, heart; Hb, hindbrain; Ln, lens; Mb, midbrain; Nt, neural tube; Nc, notochord; Ph, pharynx; Pn, pronephros; Sm, somite; Vf, ventral fin. Scale bars represent 200µm in C-K and 100 µm in L, dorsal is top in each panel.
No major morphological differences were seen in the sectioned MO injected tadpole in any of the structures where *Cfm2* expression was previously detected by *in situ* hybridisation (Fig 5.5 D and F). Visible differences between wild type and MO injected tadpoles are observed in the ventral fin of the MO injected tadpole which developed irregularly (Fig 5.5 L). The MO injected tadpole is also lacking a hypochord (Fig 5.5 H and J) and both the dorsal and ventral fins appear to be lacking blood vessels (Fig 5.5 L). Only the abnormal ventral fin phenotype was examined in figure 5.5 due to time constraints. Analysis of the other phenotypes observed, particularly those with apparent abnormal facial features, may reveal more morphological differences, particularly relating to the craniofacial stuctures and the pronepros, in the MO injected tadpoles.

5.4 Morpholino knockdown rescue

5.4.1 Sub-cloning and *in vitro* transcription

*mmCfm2* and *Cfm* sequences were obtained in the pGEM-T vector from previous lab members. To use these sequences for *in vitro* transcription they had to be sub-cloned into the pCS2+ vector. This was achieved by digesting the pGEM-T vector containing the desired sequence, purifying the digested product and religating this into the pCS2+ vector. To ensure the desired sequence had been cloned into the desired vector plasmids were sequenced and analysed.

To create the mutated *mmCfm2* constructs, to be used to test the functional regions of the *Cfm2* gene in the MO rescue experiments, the PCR products from the Phusion DNA polymerase reaction described in chapter 2.9.1 were run on a gel to confirm the amplification of the desired template. Each sample showed the presence of a band corresponding to the expected size. PCR products were purified and then religated to form the ‘new’ plasmid construct. These ligation products were then used to transform competent *E.coli* cells, and the resulting colonies were then screened using PCR for those which possessed the desired plasmid.

Colonies containing the desired sized bands were picked and cultured overnight in LB medium. The following morning the plasmids were isolated from these cultures using the Wizard® Plus SV Miniprep DNA Purification System (Promega) and stored at -30°C for future
use. To ensure that the desired sequences were obtained, including the mutations in the mutated \textit{mmCfm2} sequences, these plasmids were sent for sequencing. Each of the mutated \textit{mmCfm2} sequences were found to contain the desired mutations, however they also contained a point mutation at nucleotide 347 (A→T) in CR2 of the coding sequence. This mutation is non-synonymous, resulting in a change at amino acid 116 from a lysine residue to an isoleucine residue. This mutation was also found in the \textit{mmCfm2} CDS sequence in the plasmid template used for these mutations, however this mutation was not present in the cloned \textit{Cfm2} sequence indicating that the \textit{mmCfm2} sequence obtained from a previous lab member was the source of the mutation. Although the affected amino acid is in the CR2 region of the CFM2 protein, I decided to continue with these sequences to save time, money and effort. It is possible that the mutation does not affect the protein’s function, meaning that starting the process again may not improve any results. However, it must also be acknowledged that a change from a lysine to an isoleucine residue also represents a change from a positively charged polar amino acid to a hydrophobic residue, which could quite possibly affect the stability and functionality of the protein product.

These mutated \textit{mmCfm2} sequences were then cloned into the pCS2+ vector, along with the \textit{mmCfm2} and \textit{Cfm} sequences. Plasmids containing an appropriately sized insert were sequenced once more to ensure that the desired sequence was obtained in the pCS2+ vector.

Analysis of the sequences in the pCS2+ vector showed that the desired \textit{mmCfm2} sequences, minus the point mutation mentioned above, were successfully cloned into the pCS2+ vector. However, analysis of the \textit{Cfm} sequence showed that there were two mutations in this sequence also. A point mutation at nucleotide 247 (C→T) was detected, this mutation results in an amino acid change from a leucine residue to a phenylalanine residue in CR2. The second mutation detected in the \textit{Cfm} sequence was an insertion of an adenine at nucleotide 620. This mutation causes an amino acid change from a phenylalanine residue to a tyrosine residue. As this mutation is an insertion it causes a frameshift, however the resulting change to the amino acid sequence improves the homology of the C-terminal end of the CFM amino acid sequence, particularly with respect to the CFM sequence from the closely related species \textit{Xenopus tropicalis}. Both these mutations were detected in the original \textit{Cfm} clone sequence obtained from a previous lab member suggesting that these
disparities may reflect mistakes in the Cfm sequence stored in the NCBI database rather than be true mutations.

Despite the concerns that these missense mutations could potentially reduce, or even destroy, protein function these Cfm and mmCfm2 sequences were then used to produce mRNA for use in the overexpression and MO rescue experiments. The Cfm mRNA had a concentration of 952.7ng/µL and a 260/280 absorbance ratio of 2.16, the mmCfm2 mRNA had a concentration of 886.2ng/µL and a 260/280 absorbance ratio of 1.92. Both these concentrations are high, and the 260/280 ratios are adequate for further use.

The gels shown in figure 5.6 showed the expected bands for each of the mRNA samples although there was a fainter, larger band visible in the mmCfm2 mRNA sample most likely due to transcription from a template that had an incompletely filled 3’ end (Schenborn and Mierendorf, 1985). The mRNA was aliquoted and stored at -70°C, and aliquots were defrosted for use as needed.

5.4.2 mmCfm2 mRNA rescue

A number of unsuccessful attempts were made at rescuing the MO phenotype by co-injecting Cfm2 MO and mmCfm2 mRNA. It was noticeable that the majority of the injected embryos died during the second night of incubation, just after embryos had been taken out of 3% ficoll, 0.1x MMR and placed in 0.1x MMR indicating that some contaminant may be causing the mass death observed. Despite sterilising the microinjection gear and autoclaving the solutions used during the rearing and injecting of the embryos, no improvements were
seen in the survival rate. Another attempt was made where the antibiotic, gentamycin sulphate, was added to the 0.1x MMR when the embryos were changed out of the ficoll, unfortunately this was also unable to increase survival. After repeated failure, subsequent rescue attempts included a number of embryos which were injected with the MO only, these embryos also died at the same time as those injected with MO and mRNA - even when gentamycin sulphate was added as before - suggesting that the MO may be the source of the problem. Unfortunately, due to time constraints this problem was unable to be resolved. Returning to the MO stock solution should resolve the issue if the problem is due to the MO as appears likely.

5.5 Overexpression of Cfm and Cfm2

The mRNA used for the overexpression experiments was the same as that used for the MO rescue experiments. After producing the mRNA, embryos were again collected and subsequently injected at the two-cell stage with Cfm or mmCfm2 mRNA. Three different concentrations were used for the injections to give 1ng, 5ng or 10ng of mRNA per embryo, injected embryos were raised to stage 40 as before and then measured for the appearance of the same characteristics described in 5.2. The resulting proportions of each characteristic were graphed and are shown in figures 5.7 and 5.8.

Figure 5.7 shows the survival rate for each of the treatment groups. As before, the survival rate was measured as the number of embryos that survived to stage 40 as a proportion of the number of embryos injected. There is an unexpected increase in the survival rate of embryos injected with 5ng Cfm mRNA compared to that seen in 1ng Cfm mRNA although the survival rate of the embryos injected with 10ng Cfm mRNA is the lowest as expected. Comparing the survival rates for the Cfm mRNA injections revealed a significant difference between all of the groups (1ng-5ng, 5ng-10ng and 1ng-10ng) with p-values of 0.0299, <0.001 and 0.0105 respectively.

As the concentration is increased a decrease is seen in the embryos injected with mmCfm2 mRNA. Comparing the survival rates of these injections revealed significant differences between the 1ng-10ng and the 5ng-10ng groups, both p-values are <0.001, but not the 1ng-5ng groups, p-value of 0.6796.
Comparing the survival rates of the different mRNA molecules when injected at the same concentration (i.e. comparing 1ng Cfm mRNA to 1ng mmCfm2 mRNA) reveals no statistically significant difference for the embryos injected with 1ng of mRNA (p-value = 0.4765). However, comparing the values for 5ng or 10ng of mRNA reveals a significant difference with a p-value of 0.0428 for the 5ng mRNA injected embryos, and a p-value of <0.001 for the embryos injected with 10ng mRNA. In both of these cases the survival rate is reduced in the embryos injected with mmCfm2 mRNA indicating that overexpression of Cfm2 is more lethal than overexpression of Cfm.

The same phenotypic characteristics used to classify the MO knockdown tadpoles were used here to classify the tadpoles injected with mRNA. No other ‘new’ characteristics were observed in these tadpoles, the only exception being a single tadpole injected with 10ng mmCfm2 mRNA that survived to stage 40 but was ‘unclassifiable’ due to the severity of the phenotype. The results for all other tadpoles are shown in figure 5.8.
Figure 5.8 - mRNA injection phenotypes. Bar graphs showing the occurrence of a number of characteristics, measured independently, in each treatment group as a proportion of the total number of embryos from each treatment group. The legend in B applies to all graphs. Error bars show the 95% confidence interval. 1ng Cfm mRNA, n (total number) = 27; 5ng Cfm mRNA, n = 54; 10ng Cfm mRNA, n = 57; 1ng mmCfm2 mRNA, n = 58; 5ng mmCfm2 mRNA, n = 44; 10ng mmCfm2 mRNA, n = 110.
Comparing the different concentrations of \textit{Cfm} mRNA injected shows no difference at all in the proportions of wild type tadpoles from embryos injected with 1ng or 5ng \textit{Cfm} mRNA (Fig 5.8 A). However there is a statistically significant decrease in the proportion of wild type tadpoles injected with 10ng \textit{Cfm} mRNA when compared to those injected with 5ng \textit{Cfm} mRNA (p-value <0.001) and also when compared to those injected with 1ng \textit{Cfm} mRNA (p-value <0.001). The majority of the rest of the characteristics measured do not appear frequently enough in the \textit{Cfm} mRNA injected treatment groups to be able to be compared. The few characteristics that do show up in at least two of the \textit{Cfm} mRNA treatment groups are not statistically comparable as the sample sizes and the proportions are both low meaning the tests would not be statistically valid. The one exception is the percentage of tadpoles with eye deformities (Fig 5.8 J) in the 5ng and 10ng \textit{Cfm} mRNA injections, however the difference seen here is not statistically significant, p-value = 0.1770.

Comparing the different concentrations of \textit{mmCfm2} mRNA shows that there is a decrease in the proportion of wild type tadpoles as the concentration of \textit{mmCfm2} mRNA increases (Fig 5.8 A) as expected. However the difference seen between the 1ng and 5ng \textit{mmCfm2} mRNA injected embryos is not significant (p-value = 0.1250) and no valid statistical comparison is able to be made for the 10ng \textit{mmCfm2} mRNA injected embryos due to the very low proportion of tadpoles that were wild type, although the difference is very large and does appear to be significant. Unfortunately with such a low number of embryos that survived to stage 40 when injected with 10ng \textit{mmCfm2} mRNA, none of the data collected from these embryos can be validly used to test for statistical significance. There were only five characteristics where both the 1ng and 5ng \textit{mmCfm2} mRNA treatment groups were simultaneously represented and other than the wild type characteristic, which as mentioned above was not significant, none of them meet the criteria to be validly tested.

It is possible to compare the figures from the \textit{Cfm} injected embryos with those from the embryos injected with the same concentration of \textit{mmCfm2} mRNA. However, the same problem of finding values where the statistical test used is valid remains. The only figures that can be validly compared are for the proportion of wild type tadpoles between 1ng \textit{Cfm} or \textit{mmCfm2} mRNA and 5ng \textit{Cfm} or \textit{mmCfm2} mRNA, and neither of the differences observed are statistically significant with p-values of 0.4734 and 0.4883 respectively.
5.6 Summary
MO mediated knockdown of the *Cfm2* gene had a number of effects on development; tadpoles developed slower, were less pigmented and various phenotypic defects were observed. The most significant of these defects was the development of an abnormal ventral fin, bent bodies, kinked tails and eye deformities. Although other defects were observed, none of these were statistically significant. Notably the number of wild type tadpoles was reduced to zero when embryos were injected with 30ng of the *Cfm2* MO indicating that although no obvious phenotype was observed, development was severely impaired. This affect appeared dose dependent with the proportion of wild type tadpoles increasing when the embryos were injected with a lesser dose, 20ng of *Cfm2* MO, although the proportion itself remained stubbornly low. The survival rate also appears to be affected in a dose dependent manner with increasing *Cfm2* MO concentrations resulting in a lowered survival rate as one would expect if development is being severely impaired.

Similar phenotypic affects were seen in tadpoles when injected with *Cfm2* mRNA. The most significantly affected statistic was the survival rate which again proved to be dose dependent with larger amounts of mRNA resulting in more lethality. Despite the presence of a number of phenotypic affects similar to the *Cfm2* MO knockdown tadpoles, none of them were statistically relevant although there was a significant decrease in the number of wild type tadpoles as the amount of mRNA injected was increased.

*Shh* and *Sox3* expression were examined in the *Cfm2* MO knockdown tadpoles and found to be expressed relatively normally. A small number of inconsistent abnormalities in gene expression were found but none of these appeared specifically related to the *Cfm2* MO. Similar results were observed in embryos co-injected with *Cfm2* MO and β-*gal* mRNA with no significant correlation observed between embryos injected with MO and somite formation as indicated by β-*gal* staining. Further, no obvious differences were found in the morphology of the regions corresponding to *Cfm2* expression in sectioned MO injected tadpoles.

Despite the production of a number of templates for *in vitro* transcription, including *Cfm* and mutated *mmCfm2*, the MO phenotype was unable to be rescued by *mmCfm2* mRNA despite
a number of attempts. Although unsuccessful the problem seemed to be due to the MO itself rather than the mRNA, suggesting that with a bit more time this issue should be able to be resolved. Showing that the Cfm2 MO phenotype can be rescued with mmCfm2 mRNA will then allow the Cfm and mutated mmCfm2 mRNA transcripts to be tested for ‘rescue ability’. This is an important step in showing that the MO affects are caused specifically by Cfm2 knockdown and for testing some of the conserved features of the CFM/CFM2 proteins for their importance to functionality of the protein product.
Chapter 6 - Discussion

6.1 Summary of results

This research set out to establish a role for Cfm2 in the development of the amphibian model organism Xenopus laevis. Despite being described in little detail the information available suggests that this gene plays an important role in the development of a number of vertebrate species, including humans.

Using the protein sequences of CFM2, and its parologue CFM, a total of 43 sequences were identified as members of the CFM/CFM2 protein family. These sequences represented a number of species and four major vertebrate classes. Notably only two invertebrate species were represented by three of these sequences. An alignment was produced and a phylogenetic tree then constructed. The tree showed that the vertebrate sequences form two distinct clades representing the CFM and CFM2 protein sequences. The tree also showed that the three invertebrate sequences had little, if any, relationship to the vertebrate sequences identified.

PCR was used to probe cDNA created from RNA extracted at a number of stages during Xenopus development. Cfm2 was found to be expressed from stage 35 to at least stage 45. Whole mount in situ hybridisation of samples collected at corresponding stages found this expression to be located to the olfactory organ, lateral line precursors, pronephros, ventral regions of the branchial arches and the head mesenchyme. In situ hybridisation performed on the developing limbs also found Cfm2 to be transiently expressed at the joints of the limb as the digits form. Analysis of the Cfm parologue by the same methods found it to be expressed constitutively during the development of whole tadpoles up to stage 50 and in the limbs from stage 50 to stage 55. Cfm expression was detected in the proctodeum, lateral plate mesoderm, neural tube floor plate and hypochord by in situ hybridisation. In situ hybridisation experiments performed on developing limbs found Cfm to be expressed in the epidermal layer all over the limb during the initial limb bud stages. As the limb grew distally Cfm expression become more restricted to the proximal epidermal layer.
Producing MO mediated Cfmd knockdown tadpoles revealed a number of phenotypic defects. Although facial deformities were observed in these MO injected tadpoles, consistent with the expression patterns of Cfmd, the occurrence of this phenotype was not found to be statistically significant. Surprisingly, the most consistent, and statistically significant, effect of Cfmd MO treated tadpoles was the presence of an abnormal ventral fin. Overexpression of Cfmd resulted in similar phenotypes, although these were observed at a low frequency. The survival rate of tadpoles overexpressing Cfmd was significantly lowered and dose dependent. Similar results were obtained with overexpression of Cfmd although these tadpoles survived much better when higher doses of mRNA were administered.

6.2 The CFM/CFM2 family of proteins is conserved in vertebrates

The CFM/CFM2 protein family appears to be well conserved in vertebrates with forty of the forty-three sequences identified belonging to the Vertebrata Class. Furthermore, two of the three invertebrate sequences belong to the Florida lancelet, Branchiostoma floridae. The lancelet is a Chordate that has a very primitive skeleton with a number of similarities to the vertebrates (Putnam et al., 2008). Although the lancelet is not considered a ‘true’ vertebrate, it is a primitive Chordate and is often used in the study of the evolutionary origins of vertebrates (Putnam et al., 2008). The other invertebrate sequence identified in the sea snail, Lottia gigantea, did not align well in the multiple sequence alignment. In the phylogenetic tree subsequently produced, the snail sequence was separated from the vertebrate sequences by a large distance indicating it has little relationship to the vertebrate CFM/CFM2 sequences. These results indicate that the CFM/CFM2 family of proteins is directly linked to the evolution of vertebrates.

It is possible that the snail sequence identified is not related to the CFM/CFM2 family of proteins. This is evidenced by the fact that when a multiple sequence alignment was attempted using the corresponding nucleotide sequences, an error was encountered by ClustalX in that it was unable to fit the snail sequence in the alignment. Although the nucleotide sequences are less conserved than the protein sequences, due to the ambiguity found in the genetic code, it is still reasonable to expect that the sequence can be fitted into an alignment. It is also worth mention that the Metazome website, which was used to identify this snail sequence as an orthologue of the CFM/CFM2 sequences, does not provide
a measure of support for the associations found on the search page. Therefore, the
evidence supporting the association between this sequence and the CFM/CFM2 protein
family may be scarce at best, allowing for speculation as to the true relationship between
this sequence and the conserved vertebrate CFM/CFM2 protein sequences.

As the study of genes and their cognate proteins continues to proceed at a rapid pace, new
sequences are constantly being added to the databases. Another search using the same
methods was conducted on the 23rd of August 2011. Although the Metazome project did
not identify any ‘new’ sequences, searches using the same *Xenopus* sequences with the
BLASTp search engine revealed a number of new orthologues as described below, with
accession numbers indicated in brackets. CFM and CFM2 orthologues were identified in: the
Carolina anole lizard, *Anolis carolinensis* (XP_003227934 and XP_003222784), the wild
turkey, *Meleagris gallopavo* (XP_003211856 and XP_003210941), the common marmoset,
*Callithrix jacchus* (XP_002764006 and XP_002753192), the northern white-cheeked gibbon,
*Nomascus leucogenys* (XP_003280381 and XP_003276297), and the Sumatran orangutan,
*Pongo abelii* (XP_002824002 and XP_002826836). The lancelet-M2 sequence, previously
only identified in the Metazome search, also showed up in the most recent BLASTp search
(XP_002606374) and interestingly a sequence from the invertebrate acorn worm,
*Saccoglossus kowalevskii*, was identified (XP_002731229). The majority of these newly
identified sequences came from primates, further suggestive of this protein family being
vertebrate specific. The identification of potential orthologues in the Carolina anole lizard is
also revealing as this species has been chosen as a model organism for the reptilians. The
Reptilia represent a fifth major Class of the Vertebrata subphylum found to possess
CFM/CFM2 orthologues. Adding the worm sequence to the alignment showed that this
sequence aligns poorly with the other, previously identified, vertebrate sequences as might
be expected. However, how the other newly identified sequences align to the previously
identified sequences was not examined due to time constraints. It was noted, however, that
the CFM2 orthologue identified in the orangutan was identical to one of the human CFM2
sequences also identified in this search (NP_859060) indicating that, at least, this sequence
is highly conserved.
In the most recent search, further sequences were identified in species previously found to possess CFM and/or CFM2 orthologues. A CFM sequence was identified in the pig (XP_003131866) to go with the CFM2 sequence already identified and an additional third sequence was identified in: *Xenopus laevis* (NP_001091173), the dog (XP_852573), the rat (EDM05248), mice (BAB29164), cattle (XP_002695757) and the opossum (XP_001364772). Further sequences were identified in the giant panda (XP_002918105 and XP_002913186) and in humans (EAW90664, BAC11049, BAC86601 and NP_859060). A brief examination indicated that the majority of these sequences differed most notably in length rather than sequence. Some of the sequences identified were dramatically shorter than the sequences previously identified although there were also sequences identified that were longer than the original sequences identified. Although I did not examine any of these sequences in detail it is questionable how many of these sequences are ‘true’ orthologues and how many are just randomly occurring variants of the true sequences. As the Cfm and Cfm2 genes are characterised in more detail I expect a number of these shortened sequences apparently lacking CR1 to be discarded due to the highly conserved nature of CR1 in the majority of the sequences identified so far. However, it is also possible that a number of isoforms of these genes exist and these sequences may merely be representing this possibility.

The forty vertebrate sequences originally identified clustered into two distinct clades in the phylogenetic tree. These two clades represent the CFM and the CFM2 sequences. The two distinct clades that form indicate that these two sequences have arisen from a duplication event in a common ancestor and have subsequently evolved independently of one another. As mentioned before, two sequences were found in the lancelet which is considered to be an ancestor of modern day vertebrates (Putnam et al., 2008). The formation of these two distinct clades then raises the possibility that the two sequences identified in the lancelet could represent the ancestral CFM and CFM2 sequences. Interestingly, a conserved second methionine residue is found in all bar three of the twenty CFM2 sequences identified. One of the lancelet sequences, lanclet-M1, possesses a second methionine residue that aligns to the aforementioned methionine residue. As this feature is conserved in so many of the vertebrate CFM2 sequences and uniquely in one of the two lancelet sequences it is plausible that this lancelet sequence is ancestral to the vertebrate CFM2 sequences identified with the other lancelet sequence, lanclet-M2, being ancestral to the vertebrate CFM sequences.
This adds further support to the hypothesis that these sequences have evolved from the sequences found in the lancelet.

However, there is also the possibility that the CFM/CFM2 protein family arose from a duplication event in an ancestral species of the lancelet, and the lancelet has simply inherited these sequences in a passive fashion. The information interpreted from the tree only indicates that these sequences have arisen from a gene duplication event at some point occurring close to the evolutionary split of vertebrates from other taxa. Interestingly the reported expression of the \textit{Cfm2} gene in the segmenting somites of mice raised the possibility that this gene was important in the acquisition of the defining characteristic of vertebrates, namely the formation of the vertebrae via somitogenesis. Unfortunately this expression pattern does not appear to be conserved in \textit{Xenopus laevis} suggesting that the \textit{Cfm2} gene has either acquired this function in the mouse and related vertebrates, or this expression is not essential in somitogenesis and has subsequently been lost in some vertebrate species. Although these sequences are still likely to have evolved in vertebrates, these two hypotheses do not support the notion of an essential role for the CFM/CFM2 protein family in vertebrate evolution.

Hirano et al. (2005) described two conserved regions (CR1 and CR2) that were found to be conserved in the CFM and CFM2 sequences identified. A number of sequences in the protein alignment were apparently missing a large portion of the CFM or CFM2 sequence. However, in all of these sequences it was notable that none of them lacked CR2. Six of the vertebrate CFM2 sequences identified here - from the wild boar, cattle, rhesus macaque, chimpanzee, giant panda and the horse - exemplify this. Each of these sequences lacks CR1 but still aligns extremely well over CR2. Furthermore they each possess the second methionine residue, mentioned earlier, that is a unique feature of the CFM2 sequences. Similarly the sequence identified as a CFM2 orthologue in the dog apparently lacked CR1 although the sequence itself extended well beyond the N-terminus of any of the other sequences in the alignment. It is worth noting here that as the CFM/CFM2 family is still relatively unstudied, the majority of the sequences identified in the BLAST dataset, including the dog CFM2 orthologue, were listed as ‘predicted’ or ‘hypothetical’ sequences in the NCBI database. Therefore it is extremely likely that a number of the shortened CFM2 sequences are simply incomplete,
and that a number of the other sequences are inaccurate. Determining exactly which of these ‘hypothetical’ and ‘predicted’ sequences are indeed true members of the CFM/CFM2 family will rely on more information from the studies being conducted on the genomes and proteomes of these species. Interestingly, the latest search performed on the 23rd of August using the *Xenopus laevis* CFM2 sequence returned two identical 135 amino acid sequences from humans that also lacked CR1 (accession numbers BAC86601 and NP_859060). These sequences had the exact same sequence as the C-terminal region of the full human CFM2 sequence, including CR2. This indicates that the shortened sequences identified in other species during these searches are also likely to possess longer, full, sequences that simply haven’t been described as yet. These findings also raise the possibility that the shortened sequences described could represent tissue specific transcripts. This may explain some of the differences observed in the expression patterns of *Xenopus laevis* compared with mice and the chick as the *in situ* hybridisation probe I have used was from the full transcript which would not bind the shortened transcripts, if they are indeed ‘true’.

Similar to the long N-terminal end identified in the dog CFM2 orthologue, the green-spotted pufferfish orthologue, chick CFM and horse CFM orthologues each possess noticeably long N-terminal sequences. There are also a small number of sequences that appear to have an extended C-terminal sequence. One of the sequences with an extended C-terminal sequence is the *Xenopus laevis* CFM orthologue. During my experiments it was found that the *Cfm* sequence I was using contained two differences from the published source sequence in the NCBI database. One of these was a non-synonymous point mutation. The other mutation detected was an insertion occurring at nucleotide 620 in the sequence. This insertion resulted in a frameshift which caused a change in the amino acid sequence with a phenylalanine residue being changed to a tyrosine residue and a stop codon introduced immediately following this residue. Interestingly, this brings the C-terminal sequence into near perfect agreement with the CFM orthologue found in the closely related species *Xenopus tropicalis*. This suggests that the *Xenopus laevis* sequence found in the NCBI database is inaccurate. The most recent search, performed on the 23rd of August 2011 with the *Xenopus laevis* CFM sequence found another new protein sequence had been entered into the NCBI protein database that had this insertion included in its sequence (accession number NP_001091173), bringing the *Xenopus laevis* C-terminal sequence into alignment
with the *Xenopus tropicalis* sequence and strongly suggesting that we have found the correct *Cfm* sequence and that the original *Xenopus laevis Cfm* sequence identified contains errors.

The occurrence of sequences missing CR1 was not limited to the CFM2 sequences with another seven of the CFM sequences, from the chimpanzee, cattle, horse, giant panda, platypus, opossum and zebra finch, apparently lacking CR1. Interestingly four of these sequences belong to species which also possessed a sequence lacking CR1 in their corresponding CFM2 orthologues. This suggests that these species lack a fully annotated proteome. As these sequences were nearly all found to belong to mammals, the only exception being the zebra finch CFM sequence, it must also be acknowledged that these sequences may in fact represent a true evolutionary shift. This would indicate that CR1 is dispensable for the function of these sequences, at least in mammals where the proteins may have evolved a different function. It will be interesting to see how these theories unfold as more information becomes available with the study of these genes, their protein counterparts and the genomes and proteomes of the species involved.

### 6.3 *Cfm2* is expressed in the head, kidney and limb joints during *Xenopus* development

Amplification of the cDNA found *Cfm2* to be expressed from stage 35 to stage 45 in *Xenopus* development. Subsequent analysis by *in situ* hybridisation at stage 40 found that *Cfm2* is expressed in the head and in the pronephros. Expression at stage 40 was also able to be mapped more specifically, with expression found to be located in the olfactory organ, the lateral line precursors, the ventral region of the branchial arches and the head mesenchyme.

According to Nieuwkoop and Faber (1967) at stage 40 the mouth of the tadpole is broken through and blood begins to circulate in the gills, various parts of the skull and skeleton are also distinguishable at this stage as mesenchymal condensations. In the head, development proceeds with each of the branchial arches now separated and the branchial skeleton laid down. The branchial arches are the equivalent of the mammalian pharyngeal arches and contribute to the craniofacial skeleton, as well as the musculature of the face and other derivatives. The development of the lateral line begins relatively late in development, at stage 33/34, and proceeds at a rapid pace with the precursors beginning to differentiate into the individual sense organs by stage 40. Also at stage 40, the pronephros is functional but
continues to develop and the olfactory organ and Jacobson’s organ segregate from what was the olfactory placode.

The initial interest in the study of CFM2 came from the implication of this gene in the OPD syndrome spectrum of disorders through its interactions with filamin A (Gay et al., 2011). The specific expression patterns of Cfm2 in Xenopus can be related to the phenotypic characteristics of the OPD spectrum of disorders (Robertson et al., 2006). Genitourinary defects are observed in the OPD spectrum, which correlates with the expression of Cfm2 observed in the developing kidney of Xenopus. Similarly, the expression seen in the branchial arches of Xenopus can be related to the OPD spectrum through the craniofacial defects seen in the OPD syndrome spectrum of disorders. The expression of Cfm2 in the lateral line and the olfactory organ are harder to relate to the OPD spectrum. As the lateral line is a sensory organ only found in aquatic organisms (Cernuda-Cernuda and García-Fernández, 2006) this has no direct relationship to the OPD syndrome spectrum of disorders. Although no defects are seen in the olfactory organ in the OPD spectrum, cleft palate is observed (Robertson, 2006). Cleft palate is a disorder where the roof of the mouth does not form properly, resulting in an open connection of the nasal and oral cavities (Dudas et al., 2007). As Cfm2 expressions was detected in Xenopus as the olfactory organ splits to form Jacobson’s organ, there is an interesting relationship here which warrants further investigation as Jacobson’s organ is found on the roof of the mouth (Døving and Trotier, 1998). It is a chemoreceptor organ that is used to detect pheromones and has been found in many animals, including a number of mammals (Døving and Trotier, 1998). However there is conflicting evidence as to whether this organ exists in humans (Trotier et al., 2000; Kjaer and Fischer Hansen., 1996). Despite these conflicting reports, there is a possible link here between the expression of Cfm2 in the olfactory organ of Xenopus and human disorders, such as cleft palate, which may relate to Jacobson’s organ that is intriguing.

Further evidence of a link between Cfm2 and the OPD spectrum of disorders was found with the analysis of Cfm2 expression by in situ hybridisation in the developing limb. Here I have shown that Cfm2 is transiently expressed at the joints of the limb. This expression was consistent with the sequential formation of the structures of the limb with respect to both the proximal to distal establishment of the limb structures and the sequential formation of
the digits (Nieuwkoop and Faber, 1967). Cfm2 expression was detected at the approximate position of the knee joint connecting the femur to the tibia and fibula at stage 52. At this stage there was also a stripe of expression at the approximate position of the ankle joint connecting the tibia and fibula to the tarsus. At stage 52 the femur is in an early procartilaginous state and the tibia and fibula are still mesenchymal condensations (Nieuwkoop and Faber, 1967). This suggests that the expression of Cfm2 at this stage is somehow patterning the knee and ankle joints, before the bones themselves have formed. Similar expression patterns were observed in the digits with Cfm2 expression detected at the phalangeal joints around stage 53 to late stage 54. This timing again corresponds to the mesenchymal/procartilaginous state of the metatarsals (Nieuwkoop and Faber, 1967). The phalanges are still cartilaginous at stage 56 (Nieuwkoop and Faber, 1967) suggesting that they will similarly be present as mesenchymal condensations or procartilaginous at stage 54. This expression can be linked to the OPD spectrum of disorders where various bone disorders in the limbs have been reported (Robertson et al., 2006). This expression can also, perhaps even more tightly, be linked to mutations in FLNB. Mutations in FLNB are responsible for a number of human skeletal disorders which are characterised by, amongst other things, the disruption of joint formation (Krakow et al., 2004; Lu et al., 2007). Interestingly CFM2 is also capable of binding to FLNB (Gay et al., 2011) suggesting that CFM2 has multiple functions mediated by different interactions. Furthermore, expression of the Cfm2 in the developing mouse limb shows a near identical expression pattern in the phalanges (Dr. Esther Pearl, unpublished) and a similar pattern is observed in the limbs of chick embryos (Holman, 2007) suggesting that this expression is well conserved in tetrapods.

The detection of Cfm2 at these stages is interesting, though the functional role this gene plays in development is unclear. Gay et al. (2011) recently reported CFM and Cfm2 as interacting with FLNA to organise the perinuclear actin cap. They found that Cfm2 (referred to as RefilinA) was upregulated at the commitment of neural precursor cells into glial progenitor cells. Gay et al. (2011) proposed a hypothesis whereby the FLNA-Refilin (CFM/CFM2) interaction was important to perinuclear actin cap dynamics. Interruption of these interactions could lead to defective developmental processes where nuclear rearrangement or re-organisation is important such as during epithelial-to-mesenchymal transition (EMT) or radial neuronal migration (Gay et al., 2011). Interestingly, another
known human disorder, periventricular nodular heterotopia (PVNH), is characterised by mutations in *FLNA* (Fox et al., 1998). In PVNH the neurons fail to migrate appropriately during development resulting in severe developmental defects (Fox et al., 1998). This information taken together is suggestive of another possible link to a human disorder for the *Cfm2* gene. EMT is also important in a number of processes including gastrulation, somitogenesis and development of the precursors of the urogenital system (Thiery et al., 2009). Disruptions to EMT have further been linked to a number of developmental defects as well as cancer progression, particularly with respect to metastasis (Thiery et al., 2009).

Although the expression patterns of *Cfm2* can be linked to the OPD spectrum through correlation with phenotypic characteristics seen in OPD spectrum, the expression of *Cfm2* in *Xenopus* fails to describe many characteristics of these genetic disorders. This may indicate differences in the expression patterns of *Cfm2* between *Xenopus* and humans which may be expected due to the taxonomic distance between these two species. There are numerous differences between the expression patterns of *Cfm2* reported in mice (Hirano et al., 2005) and those found here in *Xenopus*. The obvious examples already mentioned are the lack of expression in the somites and brain of *Xenopus*. It is possible that *Cfm2* is expressed in the somites of *Xenopus* and this expression has simply been missed due to the expression level falling below the detectable threshold of *in situ* hybridisation and PCR. This is unlikely as PCR is a sensitive technique that should pick up low levels of expression that may be missed by *in situ* hybridisation. However, this possibility must be acknowledged as I encountered issues with the PCR analysis which prevented me from obtaining the replicates necessary to completely rule this out. The observed differences in expression pattern most likely reflect a true difference in the expression, and likely the functional roles, of *Cfm2* in *Xenopus* relative to mice, and extending further to humans. However, further work will be required to fully characterise conflicting expression patterns in these and other species.

The expression of *Cfm* showed no overlap with the expression of *Cfm2* in *Xenopus*. *Cfm* was initially identified as a gene expressed in the brain of mice during development (Hirano et al., 2005) although no evidence of that has been found here. *Cfm* expression was found to be expressed continuously in the developmental, and limb specific stages, analysed here. This expression was found to be located in the neural tube floor plate, hypochord and lateral
plate mesoderm at stage 30. At stage 40, Cfm expression was restricted to the floor plate and hypochord. In the limb, Cfm expression was detected in the epidermal layer during the initial limb bud stages. This expression persisted in the limb but became more proximally restricted as the limb developed.

The lack of overlap suggests that these genes have evolved different roles in development although the conserved regions of these genes suggest they may still share common functional mechanisms. Gay et al. (2011) reported that both CFM and CFM2 function by binding to FLNA and regulating the actin cap dynamics although they noted that the two genes were upregulated in different processes. It would be interesting to see if exogenous Cfm expression is capable of rescuing the Cfm2 MO phenotype. Unfortunately, this was one of the experiments sacrificed due to time constraints and the issues encountered with the MO rescue experiments. Previously it has been shown that knocking out Cfm expression in mice does not impair development (Hirano et al., 2005). Similar results were found by Gay et al. (2011) who reported that although a delay was seen in nuclear remodelling in human cells where Cfm expression was abolished, cell viability was not affected. Although the effects of a lack of Cfm expression have been studied no information has been published on the consequences of a lack of Cfm2 expression.

6.4 Cfm2 knockdown has diverse effects on Xenopus development

Cfm2 MO mediated knockdown was achieved in Xenopus laevis embryos using a MO designed to bind to Cfm2 mRNA over the major start site. MO knockdown resulted in a number of severe effects on the developing tadpoles. Cfm2 MO injected tadpoles had a lack of pigment, developed slower and were phenotypically abnormal. Melanophores are derived from the neural crest cells (Le Douarin and Dupin, 2003). Neural crest cells also migrate to form numerous structures of the face (Le Douarín and Dupin, 2003). As Cfm2 was found to be expressed in the branchial arches, which contribute to the craniofacial structures, and the MO injected tadpoles lacked pigment this appears to link the neural crest cells to the function of Cfm2. Interestingly, delamination of the neural crest cells, an essential part of EMT, is required for cell migration and Gay et al. (2011) reported an upregulation of Cfm during EMT. This provides a further link between the neural crest cells and Cfm2, although it should be noted that the upregulation during EMT was reported in
Cfm expression, not Cfm2, which was reported to be upregulated during the commitment of multipotent neural precursor cells to glial progenitor cells.

Although the Cfm2 MO phenotype lacked a specific phenotype there were obvious developmental defects in the resulting tadpoles as no wild type tadpoles developed in the 30ng Cfm2 MO treatment group. The phenotypic effects did not always appear dose dependent which could indicate that these effects are not specifically caused by the MO. Further evidence for this was found in the mmMO control group which also showed a number of these phenotypic defects. The results of these experiments potentially reflect a broad developmental role for Cfm2. However, the expression profile of Cfm2 points to more specific roles in development. The phenotypic defects observed in these MO injected tadpoles did show some correlation with the phenotypes observed in the OPD spectrum, although the phenotypic effects did not appear as specific as had been hoped. Facial deformities in Xenopus tadpoles seemed the most likely phenotypic effect to reflect phenotypic characteristics of the OPD spectrum. Although this was observed in the MO injected tadpoles this was not found to be statistically significant. The most significant effect of the MO injected tadpoles was the presence of an abnormal ventral fin. Although this appeared unrelated to Cfm2 expression in Xenopus at first, it has been reported that the core mesenchyme of the dorsal and ventral fins of Xenopus tadpoles originate from the neural crest cells, although the mesenchyme of the ventral fin has dual origins (Tucker and Slack, 2004). The neural crest cells also act as the inducer of dorsal fin formation but it has been shown that these cells are not the inducer of ventral fin formation (Tucker and Slack, 2004). Although this information appears to add further support to a link between Cfm2 and the neural crest cells, no obvious defects were observed in the formation of the dorsal fin in the MO injected tadpoles although it did appear to lack blood vessels, a derivative of the mesenchyme, in the small number of samples examined by histology. This may reflect the differences in the formation of the dorsal and ventral fins, but could also indicate that no relationship exists between Cfm2 and the neural crest cells. As mentioned earlier, Gay et al. (2011) reported that there was a delay in nuclear remodelling in cells lacking Cfm expression. Such a delay may explain the defects specifically observed in the ventral fin as the neural crest cells are located to the dorsal trunk region. As these cells must migrate further to the ventral fin, a small delay in development could potentially be magnified in the ventral fin as
opposed to the dorsal fin because of the distance these cells must navigate. Taking this information into consideration it is tempting to propose a hypothesis whereby Cfm2 expression is important for neural cell migration in Xenopus. Analysing neural crest cell markers in Cfm2 MO injected tadpoles by in situ hybridisation will help refine and test this hypothesis.

6.5 Developmental markers reveal defects in Cfm2 morpholino injected tadpoles

To examine specific effects caused by Cfm2 knockdown in Xenopus, Shh and Sox3 expression were examined by in situ hybridisation in the MO injected tadpoles and haematoxylin and eosin staining was performed on transverse sections taken from MO injected tadpoles. Shh and Sox3 were chosen due to their well-characterised expression patterns which overlap with those observed for the Cfm2 gene. MO injected tadpoles and uninjected wild type controls were sectioned and the resulting transverse section stained using haematoxylin and eosin to analyse any morphological differences between the samples. Sections of the head, heart, pronephros, mid-trunk region and tail were compared.

In situ hybridisation using an anti-sense Shh probe performed on the MO injected tadpoles showed few differences to those carried out in wild type uninjected tadpoles. The only observable difference was seen in one of the MO injected tadpoles which showed a complete lack of Shh expression on one side of the tadpole only. Expression appeared normal on the opposing side suggesting that this was a non-specific effect of the MO injection that is not caused by the MO itself, which would have affected the whole tadpole. This appears to be related to a gastrulation defect where the blastopore does not close completely and the notochord only forms on one side of the embryo which would also explain the lack of expression observed on one side of the tadpole.

In situ hybridisation using an anti-sense Sox3 probe performed on the MO injected tadpoles showed some differences to those carried out in wild type uninjected tadpoles. The MO injected tadpoles all seemed to show reduced expression in the branchial arches relative to the wild type tadpoles. There was also inconsistent expression in the lateral line precursors
of the MO injected tadpoles which may reflect stage differences. Expression was observed in the lens of the MO injected tadpoles. Expression is observed in the lens placode\(^1\) of wild type tadpoles (Schlosser and Ahrens, 2004), however this expression is not detected at stage 40 in wild type tadpoles which indicates significant stage differences exist between the wild type controls and the MO injected tadpoles. The branchial arches should be well developed, even at the slightly younger stages observed in the MO injected tadpoles, and strong Sox3 has previously been reported in stage 33/34 *Xenopus* tadpoles (Schlosser and Ahrens, 2004). Therefore, this consistently reduced staining appears to indicate that Sox3 expression is reduced in the branchial arches of the MO injected tadpoles. This is consistent with the expression of *Cfm2* in the ventral region of the branchial arches and appears to indicate that there is a defect in the formation of these structures in the MO injected tadpoles. However, *Shh* expression did not appear abnormal in the branchial arches which indicates that *Shh* acts upstream of, or in a different pathway to, *Cfm2*. Transcription factors are the enforcers of signalling pathways, thus it is not surprising to find that *Shh* expression, which (as it encodes a ligand of the Hedgehog signalling pathway) is a part of the initial stages of signalling, is unaffected whereas Sox3, which encodes a transcription factor, does appear to be affected. These results may alternatively indicate that Sox3 and *Shh* act in different pathways with *Cfm2* expression linked to the same pathway as Sox3, but not *Shh*.

Sections of the head, heart, pronephros, mid-trunk region and tail were compared in MO injected and uninjected wild type controls. No noticeable differences were found in the head or pronephros. This is particularly surprising given the strong expression of *Cfm2* detected in these structures. The heart was also examined as *Cfm2* expression has previously been detected in the hearts of chick embryos (Holman, 2007) however no differences were observed here either. The tail sections showed that the ventral fin defect results from a ‘curled’ fin rather than a completely undeveloped fin, suggesting that the defect seen is in ventral fin development rather than induction. This is consistent with the proposed hypothesis that the ventral fin defects are specifically caused by migration of the

\(^{1}\) A placode is a thickening of the epithelial layer at the site of a future organ or structure (Wolpert and Tickle, 2011).
neural crest cells whereas the formation of the ventral fin is not affected. The MO injected tadpole also lacked the hypochord and both the dorsal and ventral fins appeared to be missing blood vessels. The lack of the hypochord was surprising and difficult to reconcile with the expression patterns of Cfm2. However, the lack of blood vessels in the fins can also be interpreted as a resulting defect in the migrating neural crest cells which contribute to the mesenchyme of the dorsal and ventral fins as mesenchymal cells give rise to the circulatory system. It is also worth noting that the hypochord is thought to be required for the formation of the dorsal aorta (Cleaver et al., 2000), adding another possible link to the circulatory system in the Cfm2 MO injected tadpoles. As only the abnormal ventral fin phenotype was analysed using these histological methods, and even this in very small numbers, it is difficult to draw widespread conclusions as to the morphological effects of Cfm2 knockdown. Analysis of more MO injected samples will provide more insight into the structures affected and, in the case of the ventral fin, how they are affected (i.e. are all the ventral fin defects observed a curled fin defect?). Similarly, sections in a different plane may help identify the various abnormalities that appear present in these tadpoles. The analysis of markers of the circulatory system by in situ hybridisation or antibody staining may also be revealing of defects in light of these findings.

6.6 Exogenous expression of Cfm and Cfm2 results in a lethal phenotype

To examine the effects of upregulation of Cfm2, mmCfm2 mRNA was injected into wild type Xenopus embryos. Logic predicts that overexpression of Cfm2 should reverse the effects seen in the MO knockdown tadpoles. However, similar phenotypes were observed in the tadpoles overexpressing Cfm2 to those seen in the tadpoles from the MO knockdown experiments. These phenotypes appeared infrequently in the tadpoles overexpressing Cfm2, the most notable effect was seen in the survival rates of the tadpoles. As the dose of mRNA was increased from 1ng to 5ng and then to 10ng a decrease was seen in the survival rate. The survival rate of the 1ng and 5ng mmCfm2 mRNA injected tadpoles did not differ greatly although it was reduced in the 5ng mmCfm2 mRNA injected tadpoles. However, a large decrease was seen in the 10ng mmCfm2 mRNA injected tadpoles. This indicates that the Xenopus embryos were able to handle 5ng mmCfm2 mRNA, but 10ng mmCfm2 mRNA resulted in severe disruption of development. There was a notable increase in the number of exogastrulated embryos when injected with higher concentrations of the mmCfm2 mRNA.
indicating that these embryos were unable to gastrulate properly. This may again reflect migratory defects in these embryos caused by the misexpression of \textit{Cfm2}. The proportion of wild type tadpoles also decreased with an increasing dose of mRNA. However, tadpoles that did not survive were counted as having a mutant phenotype. Thus, a direct comparison of the proportions of wild type tadpoles for the different concentrations of mRNA may be misleading as a decrease in the survival rate will be reflected in a decrease in the proportion of wild type tadpoles. Alternatively, not including those tadpoles which die in the proportion of wild type tadpoles may also be misleading as it may indicate a disproportionately high number of wild type tadpoles in some cases.

When embryos were injected with the same concentrations of \textit{Cfm} mRNA, the effects were less notable. Again, some of the phenotypes observed in the MO injected tadpoles were infrequently represented in the tadpoles expressing exogenous \textit{Cfm} mRNA. Comparison of the survival rates of these tadpoles showed the survival rate to unexpectedly increase from 1ng to 5ng \textit{Cfm} mRNA. This increase was also found to be significant. However, the difference can be explained, to some degree, by differences in the quality of each embryo and the small sample size, particularly for the 1ng \textit{Cfm} mRNA group. If the sample size was increased I would expect this difference to be negligible. The tadpoles injected with 10ng \textit{Cfm} mRNA showed another sharp decrease in the survival rate as was expected. This decrease was not as large as that seen in the 10ng \textit{mmCfm2} mRNA injected tadpoles indicating that overexpression of \textit{Cfm} mRNA is more tolerable than overexpression of \textit{mmCfm2} mRNA. This result is not surprising as it has previously been shown that knocking out \textit{Cfm} in mice has no apparent effects on development (Hirano et al., 2005) which indicates this gene has a limited effect on development.

6.7 Conclusions and future direction

CFM and CFM2 are well conserved in a diverse array of vertebrate species. These two sequences are likely to have arisen from a duplication event in a common ancestor of the Chordates and have subsequently evolved separately. The new sequences identified in the most recent search provide further support for the CFM/CFM2 family of proteins being specific to the vertebrates and closely related Chordates. How the newly identified invertebrate sequence from the acorn worm fits into the phylogeny of this protein family will
be interesting. However, the initial observations from a protein alignment produced with this sequence suggest it bears little relationship to the vertebrate sequences.

The expression patterns of Cfm2 and the phenotypes observed in the MO knockdown experiments show some correlation to the OPD syndrome spectrum of disorders in humans. However, this correlation is by no means complete. The results from the MO knockdown and overexpression of Cfm2 experiments indicate that this gene is important to broader development with a potential role in neural crest migration. No mutations in the human orthologue of Cfm2 have been identified in patients with these disorders (S. Robertson, personal communication). This suggests that the Cfm2 gene is either rarely mutated, completely lethal when mutated, or does not produce an OPD spectrum specific phenotype when mutated.

No overlapping expression was observed between the Cfm and Cfm2 genes. This indicates that these two genes play different roles in the development of Xenopus laevis. These results are not surprising considering the distinct separation observed in the phylogenetic tree which suggests they have evolved independently. These genes may still have similar functions, as indicated by the work published by Gay et al. (2011) which found that both human CFM and CFM2 protein orthologues function through binding to FLNA and regulating actin cap dynamics, although the roles of these genes in development do not appear to overlap.

Although the work carried out in this study has not established a functional role for Cfm2 in Xenopus laevis development, it has revealed a number of intriguing details. It will be interesting to examine neural crest cell and EMT markers such as xslug, xtwist and foxd3 in the Cfm2 MO injected tadpoles to determine if these specific processes are affected by Cfm2 knockdown. The analysis of circulatory system markers aplnr and cdh5 by in situ hybridisation would also help characterise the defects in Cfm2 MO injected tadpoles.

Future work will require the MO rescue experiments to be successfully conducted. Although this should be straightforward these experiments can be time consuming. It should also be noted that the apparent mutation found in the mmcfm2 mRNA template could affect the
functional protein product and the template sequence may therefore need to be recloned to get these experiments working. The MO rescue experiments will be useful in examining the functional aspects of the CFM2 protein, particularly the N-terminal region of this protein containing CR1. It would also be interesting to create transgenic lines in *Xenopus* and/or mice to examine full *Cfm2* knockouts. As MOs only achieve gene knockdown these transgenic lines may be more revealing of specifically phenotypic defects caused by a lack of *Cfm2* expression. These transgenic lines would also be of great interest to limb development as the MOs were unable to help examine this. The creation of inducible knock-outs may be required to examine limb development if *Cfm2* knockouts proved lethal at an early stage of development.

*In situ* hybridisation in other model organisms will also help resolve the differences observed between mice and *Xenopus*. *In situ* hybridisation probes are already available for the axolotl and chicken making these organisms the most likely candidates for further *in situ* hybridisation studies. Previously attempts at *Cfm2 in situ* hybridisation in the chicken revealed expression in the heart and limbs of the chick embryo (Holman, 2007) although only a limited range of older stage embryos were analysed. Attempts at *in situ* hybridisation in the axolotl will require the development of a working protocol. Protocols are available for the axolotl, however adapting and optimising these protocol can also be a time consuming process. *In situ* hybridisation using a probe designed over the 3’ region of the *Cfm2* gene may help describe further expression patterns of *Cfm2* as more and more of the shortened transcripts are being identified in the various vertebrate species raising the possibility of expression of tissue specific isoforms. Transgenic lines expressing GFP, or another reporter gene, under the control of the *Cfm2* promoter in mice or *Xenopus* may also be of interest to examine the expression profiles of this gene in more detail.

Although a number of questions regarding *Cfm2* remain unanswered, the experiments carried out in this study are essential to understanding the role of *Cfm2* in the development of *Xenopus laevis* and other vertebrates and will help to elucidate the function of *Cfm2* in vertebrate development in the future.
References


Appendices

Appendix I - NCBI sequence dataset


<table>
<thead>
<tr>
<th>Species</th>
<th>Common name</th>
<th>CFM</th>
<th>CFM2</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Xenopus laevis</em></td>
<td>African clawed frog</td>
<td>NP_001087339</td>
<td>NP_001086164</td>
</tr>
<tr>
<td><em>Xenopus tropicalis</em></td>
<td>Western clawed frog</td>
<td>NP_001008059</td>
<td>NP_001011008</td>
</tr>
<tr>
<td><em>Homo sapien</em></td>
<td>Human</td>
<td>AAH31341</td>
<td>EAW98447</td>
</tr>
<tr>
<td><em>Gallus gallus</em></td>
<td>Chick</td>
<td>XP_415867</td>
<td>XP_001235028</td>
</tr>
<tr>
<td><em>Mus musculus</em></td>
<td>Mouse</td>
<td>NP_083934</td>
<td>AAH53444</td>
</tr>
<tr>
<td><em>Rattus norvegicus</em></td>
<td>Rat</td>
<td>NP_001007612</td>
<td>NP_001103017</td>
</tr>
<tr>
<td><em>Danio rerio</em></td>
<td>Zebrafish</td>
<td>XP_002664671</td>
<td>AA13867</td>
</tr>
<tr>
<td><em>Pan troglodytes</em></td>
<td>Chimpanzee</td>
<td>XP_511242</td>
<td>XP_001134722</td>
</tr>
<tr>
<td><em>Macaca mulatta</em></td>
<td>Rhesus macaque</td>
<td>XP_001083814</td>
<td>NP_001181140</td>
</tr>
<tr>
<td><em>Monodelphis domestica</em></td>
<td>Opossum</td>
<td>XP_001377624</td>
<td>XP_001375730</td>
</tr>
<tr>
<td><em>Canis familiaris</em></td>
<td>Dog</td>
<td>XP_853944</td>
<td>XP_853093</td>
</tr>
<tr>
<td><em>Bos taurus</em></td>
<td>Cattle</td>
<td>XP_875886</td>
<td>NP_001192546</td>
</tr>
<tr>
<td><em>Equus caballus</em></td>
<td>Horse</td>
<td>XP_001502268</td>
<td>XP_001916186</td>
</tr>
<tr>
<td><em>Sus scrofa</em></td>
<td>Wild boar</td>
<td>-</td>
<td>XP_001926696</td>
</tr>
<tr>
<td><em>Ailuropoda melanoleuca</em></td>
<td>Giant panda</td>
<td>EFB17956</td>
<td>EFB19400</td>
</tr>
<tr>
<td><em>Taeniopygia guttata</em></td>
<td>Zebra finch</td>
<td>XP_002196282</td>
<td>XP_002189180</td>
</tr>
<tr>
<td><em>Ornithorhynchus anatinus</em></td>
<td>Platypus</td>
<td>XP_001518827</td>
<td>-</td>
</tr>
<tr>
<td><em>Salmo salar</em></td>
<td>Salmon</td>
<td>NP_001167364</td>
<td>-</td>
</tr>
<tr>
<td><em>Tetraodon nigroviridis</em></td>
<td>Green-spotted puffer, or pufferfish</td>
<td>-</td>
<td>CAF91070</td>
</tr>
</tbody>
</table>
Appendix II - Metazome sequence dataset

**Appendix table II - Potential CFM/CFM2 orthologues identified by the Metazome search.** The website that sequences were downloaded from is shown in the far right column. To distinguish between sequences from a single species an -M1 or -M2 label was added to each paralogous sequence arbitrarily. None of these sequences were identified as CFM or CFM2 orthologues prior to analysis.

<table>
<thead>
<tr>
<th>Species</th>
<th>Common name</th>
<th>-M1</th>
<th>-M2</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Fugu rubripes</em></td>
<td>Fugu fish, or pufferfish</td>
<td>553850</td>
<td>594123</td>
<td>gdi.joe.gov</td>
</tr>
<tr>
<td><em>Gasterosteus</em></td>
<td>Three-spined stickleback</td>
<td>ENSGACP00000011076</td>
<td>ENSGACP00000016342</td>
<td>ensembl.org</td>
</tr>
<tr>
<td><em>Oryzias latipes</em></td>
<td>Medaka fish</td>
<td>ENSORLP00000010435</td>
<td>ENSORLP00000017576</td>
<td>ensembl.org</td>
</tr>
<tr>
<td><em>Branchiostoma</em></td>
<td>Lancelet</td>
<td>124309</td>
<td>118517</td>
<td>gdi.joe.gov</td>
</tr>
<tr>
<td><em>Lottia gigantea</em></td>
<td>Owl limpet</td>
<td>LgGsHFWreduced.8019</td>
<td>-</td>
<td>metazome.net</td>
</tr>
</tbody>
</table>
Appendix III - Full protein sequence alignment
Appendix figure I - Full multiple sequence protein alignment. * indicates sequences that were identified in the BLAST dataset but were unable to be distinguished as either CFM or CFM2. ** indicates sequences that were identified in the Metazome dataset but were unable to be distinguished as either CFM or CFM2. M1 and M2 labels were added to Metazome sequence names to distinguish between paralogues. CR1 occurs from position 124 to position 140 in the alignment. CR2 occurs from position 218 to position 300 in the alignment.

<table>
<thead>
<tr>
<th></th>
<th>310</th>
<th>320</th>
<th>330</th>
<th>340</th>
<th>350</th>
<th>360</th>
<th>370</th>
<th>380</th>
<th>390</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild boar</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cattle CFM1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhesus CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chimpanzee CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giant panda CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horse CFM2</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dog CFM2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rat CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X. laevis CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X. tropicalis CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chick CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zebra finch CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opossum CFM2</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green spotted puffer*</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fugu fish-M1**</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medaka fish-M2**</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zebrafish CFM2</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three-spined stickleback</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three-spined stickleback Medaka fish-M1**</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salmon*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhesus CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rat CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dog CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chick CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X. laevis CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X. tropicalis CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fugu fish-M2**</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chimpanzee CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cattle CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horse CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panda CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platypus*</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opossum CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zebra finch CFM</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lancelet-M1**</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lancelet-M2**</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oul Enfrag**</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* indicates sequences that were identified in the BLAST dataset but were unable to be distinguished as either CFM or CFM2. ** indicates sequences that were identified in the Metazome dataset but were unable to be distinguished as either CFM or CFM2. M1 and M2 labels were added to Metazome sequence names to distinguish between paralogues. CR1 occurs from position 124 to position 140 in the alignment. CR2 occurs from position 218 to position 300 in the alignment.
Appendix IV – Primer sequences and PCR programs

*Xenopus laevis Cfm:*

Forward primer: 5’GGTACCATGGTGTCAGACTGAAC3’

Reverse primer: 5’CGTCTAGAGTCCATTTTCAGTGAAG3’

Appendix table III - PCR program used for *Xenopus laevis Cfm*

<table>
<thead>
<tr>
<th>Function</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heated lid</td>
<td>110°C</td>
<td>-</td>
</tr>
<tr>
<td>Initial denature</td>
<td>96°C</td>
<td>4 minutes</td>
</tr>
<tr>
<td>Denaturing</td>
<td>96°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Annealing</td>
<td>57°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>60 seconds</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>

*Xenopus laevis Cfm2:*

Forward primer: 5’GTTCTGCCAGACATGGTAGGTGC3’

Reverse primer: 5’GGTCCCAATGTTCTCTTGT3’

Appendix table IV - PCR program used for *Xenopus laevis Cfm2*

<table>
<thead>
<tr>
<th>Function</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heated lid</td>
<td>110°C</td>
<td>-</td>
</tr>
<tr>
<td>Initial denature</td>
<td>96°C</td>
<td>4 minutes</td>
</tr>
<tr>
<td>Denaturing</td>
<td>96°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Annealing</td>
<td>57°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>60 seconds</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>
**Xenopus laevis β-actin:**

Forward primer: $5^{\prime}$GCCGCATAGAAAGGAGAGACAC$3^{\prime}$

Reverse primer: $5^{\prime}$TTTTGTCCCATTTCCAACCAT$3^{\prime}$

**Appendix table V - PCR program used for Xenopus laevis β-actin**

<table>
<thead>
<tr>
<th>Function</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heated lid</td>
<td>110°C</td>
<td>-</td>
</tr>
<tr>
<td>Initial denature</td>
<td>96°C</td>
<td>4 minutes</td>
</tr>
<tr>
<td>Denaturing</td>
<td>96°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Annealing</td>
<td>53°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>

**Ambystoma mexicanum Cfm2:**

Forward primer: $5^{\prime}$ACTGGCCTCCTCACACTGT$3^{\prime}$

Reverse primer: $5^{\prime}$CTCCTCCTGAGTTCCCTGTGC$3^{\prime}$

**Appendix table VI - PCR program used for Ambystoma mexicanum Cfm2**

<table>
<thead>
<tr>
<th>Function</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heated lid</td>
<td>110°C</td>
<td>-</td>
</tr>
<tr>
<td>Initial denature</td>
<td>96°C</td>
<td>4 minutes</td>
</tr>
<tr>
<td>Denaturing</td>
<td>96°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Annealing</td>
<td>60°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>60 seconds</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>
M13 primers:

Forward primer: 5'CCCAGTCACGACGTTGAATAAC3'

Reverse primer: 5'AGCGGATAACAATTCACACAGG3'

Appendix table VII - PCR program used for M13 primers

<table>
<thead>
<tr>
<th>Function</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heated lid</td>
<td>110°C</td>
<td>-</td>
</tr>
<tr>
<td>Initial denature</td>
<td>96°C</td>
<td>5 minutes</td>
</tr>
<tr>
<td>Denaturing</td>
<td>96°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Annealing</td>
<td>50°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>60 seconds</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>
Appendix V – 5’ phosphorylated primers

For the CR1 point mutations the 5’ phosphorylated primers used to create the serine - alanine mutations were:

$^5'TAGAAAGGAGGAtcGGGGtcCGGTGGAAGA^3'$ (substitutions shown in lower case)
and $^5'CTCAATCTCACCTGGAATTGCTGAAGG^3'$

For the CR1 point mutations the 5’ phosphorylated primers used to create the serine - alanine mutations were:

$^5'TAGAAAGGAGGAgcGGGagcCGGTGGAAGA^3'$ (substitutions shown in lower case)
and $^5'CTCAATCTCACCTGGAATTGCTGAAGG^3'$

For the deletion the 5’ phosphorylated primers used were:

$^5'GTATGGAATCTAQTTCGCGCC^3'$
and $^5'ATGCACCGTTGTATTTGGAGAAAGC^3'$

Appendix table VIII - PCR program used for each 5’ phosphorylated primer set

<table>
<thead>
<tr>
<th>Function</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heated lid</td>
<td>110°C</td>
<td>-</td>
</tr>
<tr>
<td>Initial denature</td>
<td>98°C</td>
<td>2 minutes</td>
</tr>
<tr>
<td>Denaturing</td>
<td>98°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Annealing</td>
<td>60°C</td>
<td>30 seconds</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>2 minutes</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>
Appendix figure II - *Xenopus laevis* complete *Cfm* coding sequence cloned into the pGEM-T vector. Courtesy of Dr Esther Pearl.
Appendix figure III - *Xenopus laevis* complete *Cfm2* coding sequence cloned into the pGEM-T vector. Courtesy of Dr Esther Pearl.
Appendix figure IV - *Xenopus laevis* mismatched (mm)Cfm2 sequence cloned into the pGEM-T vector. Five substitutions, underlined, have been introduced into the sequence around the start codon, shown in grey - CCA\_CATG\_GTCATCTACATTAC. Courtesy of Tania Cumming.
Appendix figure V - *Xenopus laevis* mismatched (mm)Cfm2 sequence cloned into the pCS2* vector. Five substitutions, underlined, have been introduced into the sequence around the start codon, shown in grey - CCATACATGGTTGGTCATCTACATTTAC.

The mmCfm2 S-A and S-D sequences have mutations in the CR1 region which change the amino acid sequence at the MAPK consensus phosphorylation site from PPSPSPP to PPAPAPP and PPDPDPP respectively.
Appendix figure VI - *Xenopus laevis* mismatched (mm)Cfm2 nt1-138 deletion sequence cloned into the pCS2+ vector.
Appendix figure VII - *Ambystoma mexicanum Cfm2* expressed sequence tag (EST) cloned into the pCRII-TOPO vector.
Appendix VII – List of solutions used

10% blocking reagent
10% w/v Boehringer Manheim blocking reagent in MAB, heated to dissolve, autoclaved at 121°C for 15 minutes, stored at -30°C.

10% Formalin
4% v/v formaldehyde in PBSA.

10x MAB
1M maleic acid; 1.5M NaCl, adjusted to pH 7.5 with NaOH, autoclaved at 121°C for 15 minutes.

10x MMR
1M NaCl; 20mM KCl; 10mM MgSO₄; 20mM CaCl₂; 50mM HEPES; 1mM EDTA pH 8.0; in MQW, pH adjusted to 7.8 and autoclaved at 121°C for 15 minutes.

10x PBSA
50x phosphate buffered saline (Dulbecco A) tablets in 500mL MQW, autoclaved at 121°C for 15 minutes, used as 1x PBSA v/v.

20x SSC
3M NaCl; 0.3M NaCitrate in 500mL MQW, pH adjusted to 7.0.

Alkaline phosphatise buffer
100mM Tris Cl pH 9.5; 50mM MgCl₂; 100mM NaCl; 0.1% v/v tween20 in MQW.

Cysteine hydrochloride
2% or 5% w/v cysteine hydrochloride in MQW and adjusted to pH 7.8.

Ficoll
3%, or 6%, w/v ficoll in 50mL 0.1x MMR, or 1x MMR.
LB agar
14.8g LB agar dissolved in 400mL MQW and autoclaved at 121°C for 15 minutes, plated with ampicillin at a concentration of 50ng/µLand allowed to set, plates stored at 2°C.

LB medium
12.5g LB dissolved in 500mL MQW and autoclaved at 121°C for 15 minutes, stored at 2°C.

MABT
1x MABT v/v with 0.1% (v/v) tween 20.

MEMFA
100mM MOPS pH 7.4, 2mM EGTA, 1mM MgSO₄, 3.7% v/v Formaldehyde in MQW.

MS222
2.5% w/v MS222 dissolved in MQW, used 1:50 for anaesthetising.

NBT-BCIP
1 tablet per 10mL MQW.

Paraformaldehyde
4% w/v paraformaldehyde in PBSA, adjusted to pH 7.4.

PBSAT
1x PBSA v/v with 0.1% (v/v) tween20.

X-gal staining buffer
20µM potassium hexacyanoferrate II, 20µM potassium hexacyanoferrate III, 2mM MgCl₂, 1mg/mL X-gal in MQW.
Appendix VIII - *Xenopus laevis* stage series

Appendix figure VIII - *Xenopus laevis* stages 1 - 5, figures from Niuekwoop and Faber (1967).
Appendix figure IX - *Xenopus laevis* stages 6 - 9, figures from Niuekwoop and Faber (1967).
Appendix figure X - *Xenopus laevis* stages 10 - 15, figures from Niuekwoop and Faber (1967).
Appendix figure XI - *Xenopus laevis* stages 16 - 23, figures from Niuekwoop and Faber (1967).
Appendix figure XII - *Xenopus laevis* stages 23 - 28, figures from Niuekwoop and Faber (1967).
Appendix figure XIII - *Xenopus laevis* stages 29/30 - 42, figures from Niuekwoop and Faber (1967).
Appendix figure XIV - *Xenopus laevis* stages 42 - 46, figures from Niuekwoop and Faber (1967).
Appendix figure XV - *Xenopus laevis* stages 47 - 54, figures from Niuekwoop and Faber (1967).
Appendix figure XVI - *Xenopus laevis* stages 55 - 61, figures from Niuekwoop and Faber (1967).