Chapter 1

Introduction

Outlook for thesis

The Hedgehog (Hh) signalling pathway is one of the major signalling pathways that governs embryonic and adult development in bilateria (Hooper and Scott 2005). It regulates a wide range of processes including morphogen-mediated patterning, cell proliferation and survival, and cell polarity. Despite the importance of the Hh signalling pathway, many of the direct targets of Hh signalling remain unknown. However, recent work to identify Gli transcription factor binding sites within the enhancer elements of genes has allowed the identification of various target genes directly controlled by Hh signalling. This work included the use of chromatin immunoprecipitation (ChIP) (Vokes et al. 2007), and DNA deletion analyses to narrow down the DNA enhancer sequences required to drive reporter gene expression (Maurya et al. 2011). Identification of these target genes is crucial for the understanding of the action of Hh.

1.1: The Hedgehog signalling pathway

1.1.1: Hedgehog genes and Hedgehog protein function

Hedgehog (*Hh)* genes are conserved from invertebrates to vertebrates, and have been identified in *Drosophila*, lamprey, zebrafish, chick, mice and human (Nusslein-Volhard and Wieschaus 1980; Echelard et al. 1993; Krauss et al. 1993; Riddle et al. 1993; Marigo et al. 1995). Whilst *Drosophila* has just one *Hh* gene, chick, mice and humans have three genes: *Sonic Hedgehog* (*Shh*), *Desert Hedgehog* (*Dhh*) and *Indian Hedgehog* (*Ihh*). Five *Hh* genes are expressed in the zebrafish; two *shh* homologs, *sonic hedgehog* (*shh*) and *tiggywinkle hedgehog* (*twhh*), two *ihh*-like homologs, *echidna hedgehog* (*ehh*) and *ihha*, and also a *dhh* homolog (Ensembl.org) (Krauss et al. 1993; Ekker et al. 1995; Currie and Ingham 1996).

In the early vertebrate embryo, *Shh* is expressed in three main signalling centres: the zone of polarizing activity (ZPA) in the limb bud, the floor plate, and the notochord (Bumcrot et al. 1995; Ingham and McMahon 2001). Shh proteins in these centres act as morphogen signalling molecules capable of acting in a dose-dependent manner at sites distant from their production source (Ingham and Fietz 1995; Ericson et al. 1997; Struhl et al. 1997). In this way, *Shh* is essential for normal embryonic and adult development, playing key roles in the proliferation and survival of cells, and in cell fate decisions and patterning of tissues, including muscles, brain, limb buds, eyes, ears, neural tube, testes and gut (Bitgood et al. 1996; Chiang et al. 1996; Wolff et al. 2003; Hammond et al. 2009; Matise and Wang 2011).

The earliest studies of *Shh* function were carried out in *Drosophila*, where it was discovered that Hedgehog controls the expression of *wingless* (Ingham and Hidalgo 1993). Together, *hedgehog* and *wingless* are required for segmentation patterning of the embryo along the antero-posterior axis (Ingham 1993). In vertebrates, *Shh* is required for the correct formation and specification of somitic compartments (Borycki et al. 1998), and regulates limb bud outgrowth and antero-posterior patterning (Riddle et al. 1993; Ingham and McMahon 2001). *Shh* also controls left/right

asymmetry (Levin et al. 1995). This demonstrates a conserved role for *Shh* in mediating the patterning of tissues. Perhaps the best documented role of Shh as a morphogen is its function within the neural tube, where changes in the concentration of Shh along the dorso-ventral axis generate five distinct domains of ventral neurons (see section 6.2.2). Neurons that form in more ventral regions of the neural tube, closer to the source of Shh in the floor plate and notochord, require higher concentrations of Shh for their induction (Ericson et al. 1997).

Loss of Shh activity in human results in holoprosencephaly, characterised by cyclopia and the loss of ventral cell types in the central nervous system (Ingham and McMahon 2001). Similarly, loss of *Shh* in mouse results in a variety of defects including the loss of ventral cell types in the neural tube, and results in perinatal death (Chiang et al. 1996). Excessive Hh signalling in adult mice and human can lead to a variety of cancers caused by over proliferation of cells. For example, mutation in *Ptc1* can cause Gorlin syndrome, which is characterised by craniofacial and skeletal abnormalities, with an increased risk of developing tumours such as basal cell carcinoma and medulloblastoma (Epstein 2008; Jiang and Hui 2008).

1.1.2: The synthesis and secretion of Hedgehog proteins

In the endoplasmic reticulum, a 45kDa precursor Hedgehog protein is palmitoylated at the N-Terminus by Skinny Hedgehog (Pepinsky et al. 1998; Farzan et al. 2008; Chen et al. 2011), and then auto-cleaved, catalysed by the C-terminus (Lee et al. 1994; Bumcrot et al. 1995). This produces a 19kDa N-terminal fragment, which carries all signalling activities (Lee et al. 1994; Bumcrot et al. 1995; Porter et al. 1996; Porter et al. 1996). Cholesterol modification promotes binding of the N-terminal Hh fragment to the cell membrane (Lee et al. 1994; Bumcrot et al. 1995; Ekker et al. 1995), a process required for its secretion and long-range signalling (Porter et al. 1996).

Hedgehog proteins are released from the cell surface as lipoprotein-associated oligomers, a process requiring Dispatched, a twelve-transmembrane pass protein (Burke et al. 1999; Nakano et al. 2004). Once secreted, Hh is thought to diffuse through the extra-cellular matrix (ECM) to Hh-responding cells. In *Drosophila*, *tout-velu* (*ttv*) and *dally-like* (*dlp*) are genes which encode heparin sulphate proteoglycans (HSPG) and are required for the transport of cholesterol-modified Hh proteins (Bellaiche et al. 1998; Lum et al. 2003). *Shifted*, the *Drosophila* homolog of vertebrate *Wnt Inhibitory factor-1* (*WIF-1*), appears to stabilise Hh in the ECM and to regulate its distribution (Glise et al. 2005). These data suggest that the ECM environment and its modification is crucial for the diffusion and transport of Hh proteins. In zebrafish, *scube2* (encoding an EGF related protein containing a CUB domain thought to be involved in protein-protein interaction and ligand binding) has been proposed to play a similar role in the transport or stabilisation of Hh proteins in the ECM (Grimmond et al. 2000; Woods and Talbot 2005; Johnson et al. 2012), although others suggest it plays a role in Hh protein reception and endocytosis

(Hollway et al. 2006). The role of *Scube2* in mammalian cells is also not fully understood. Some suggest it is required for the release of Shh from Shh-expressing cells (Creanga et al. 2012), whilst others indicate a role for reception of the Shh signal (Tsai et al. 2009).

1.1.3: The reception of Hedgehog proteins

The receptor of Hh is Patched (Ptc), which like Dispatched is a twelve-transmembane pass protein with a sterol-sensing domain (Hooper and Scott 1989; Nakano et al. 1989). *Drosophila* has a single *Ptc* gene, whilst vertebrates have two genes, *Ptc1* and *Ptc2* (Carpenter et al. 1998; Lewis et al. 1999). Three additional cell surface co-receptors, Cdo, Boc and Gas1, have been implicated in Hh binding (Tenzen et al. 2006; Allen et al. 2011). Cdo and Boc have immunoglobulin and fibronectin-like repeats on their extracellular domain, and the fibronectin repeats have been shown to be critical for Hh binding (*Figure 1.1*) (McLellan et al. 2008; Allen et al. 2011). Overall, data suggest that these co-receptors promote Hh signalling and function redundantly (Tenzen et al. 2006; Allen et al. 2011). Ptc however, suppresses *Hh* signalling and *Hh*-target gene transcription, as revealed by the phenotype of *Ptc* mutants in various organisms (Ingham et al. 1991; Dahmane et al. 1997). Mutations in *Ptc* phenocopy *Hh* over-expression (Ingham and Hidalgo 1993). Binding of Hh to Ptc, and the co-receptors Cdo, Boc and Gas1, antagonises Ptc-mediated suppression on the Hh pathway (*Figure 1.1*). The activation of Hh-target gene expression in invertebrates and vertebrates also depends on the presence of the seven-transmembrane pass protein Smoothened (Smo) (van den Heuvel and Ingham 1996; Chen et al. 2001), and in its absence, zebrafish and mouse embryos display a Hh loss-of-function phenotype (van Eeden et al. 1996). In the absence of Hh, Ptc represses Smo by inhibiting the translocation of Smo to the plasma membrane in *Drosophila* (Ingham et al. 2000). Upon Hh binding, Ptc is internalised into vesicles, allowing Smo to translocate to the plasma membrane and to initiate the downstream Hh signalling cascade (Denef et al. 2000; Zhu et al. 2003). Supporting this finding, suppression of Ptc activity blocks the trafficking of Smo to the lysosome (Ingham and McMahon 2001). In vertebrates, Hh signalling components are localised to the primary cilia, which are essential for Hh signal transduction (Huangfu et al. 2003; Corbit et al. 2005; Kim et al. 2010; Tuson et al. 2011). In the absence of Hh ligand, Ptc localisation in the cilia inhibits the accumulation of Smo in cilia (*Figure 1.1*). When Hh binds to Ptc, Ptc leaves the cilia, allowing the translocation of Smo to the cilia tip, via essential intraflagellar transport proteins (*Figure 1.1*) (Huangfu et al. 2003; Huangfu and Anderson 2005). How Ptc restricts the localisation of Smo is not fully understood.

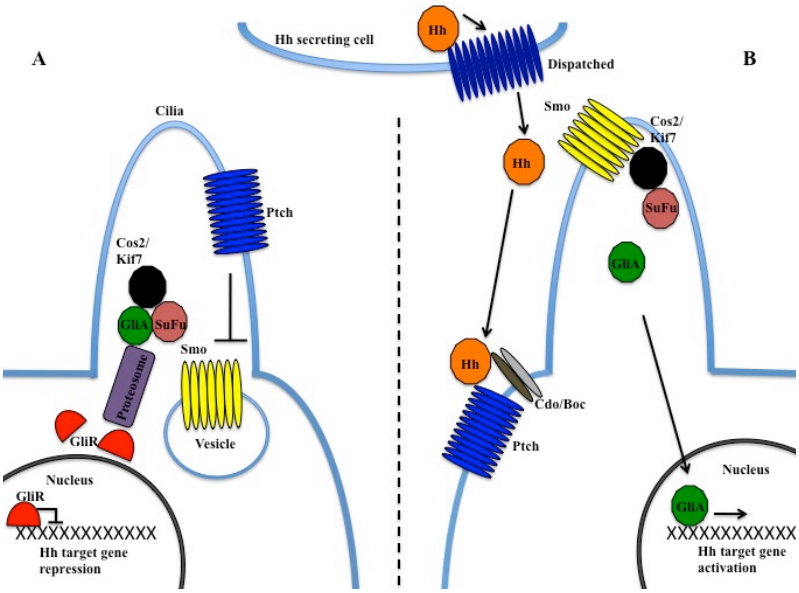


Figure 1.1: *The hedgehog signalling pathway in vertebrates.* A: in the absence of Hh ligand, Patched (Ptc) transmembrane receptor localises to primary cilia and inhibits vesicular trafficking of Smoothened into cilia. As a result, Gli proteins are tethered to microtubules in the cilia in a complex with Kif7 (Cos2), Fused, and Suppressor of Fused (SuFu), are phosphorylated, and cleaved into Gli repressor (GliR). Translocation of Gli R into the nucleus represses Hh target genes. B: Hh ligand is secreted by the transmembrane protein Dispatched, and diffuses through the ECM to a Hh-responsive cell. Hh binds to Patched, and co-receptors Cdo/Boc, and prevents the translocation of Patched into the cilia. This relieves the Patched-mediated repression of Smoothened, allowing the translocation of Smoothened into the cilia. The intracellular domain of Smoothened associates with the microtubule-Kif7 complex, sequestering it in the cilia. This prevents the stabilisation of Gli and subsequent proteolyic cleavage of Gli in the cytoplasm, allowing Gli activator (GliA) to enter the nucleus and activate Hh-target gene transcription.

1.1.4: Hedgehog signal transduction is largely conserved

In *Drosophila*, a single zinc-finger containing transcription factor called Cubitus interruptus (Ci) is responsible for all Hh signalling-mediated transcriptional activity (Forbes et al. 1993; Alexandre et al. 1996). Vertebrates have three Ci homologs, Gli1-3, which have repressor and activator transcriptional functions (Aza-Blanc et al. 2000).

In flies, post-translational modifications of Ci determine whether it acts as an activator or repressor transcription factor. In the absence of Hh, the 155kDa Ci protein is cleaved into a 75kDa N-terminal repressor product, whilst cells transducing the Hh signal maintain Ci in its full-length activator form (Aza-Blanc et al. 1997; Wang and Holmgren 1999).

The mechanism of Ci processing involves phosphorylation by Protein Kinase A (PKA), Glycogen Synthase Kinase 3β, and a CK1 family member (Chen et al. 1999; Price and Kalderon 1999; Jia et al. 2002). Costal2 (Cos2), a kinesin-related protein, also negatively regulates the Hh signalling pathway (Robbins et al. 1997; Sisson et al. 1997; Tay et al. 2005). Cos2 binds both the cytoskeleton and Ci, holding it in the cytoplasm (*Figure 1.1*). Following its phosphorylation, Ci is ubiquitinated by Slimb, allowing its proteolysis by the proteosome (Jiang and Struhl 1998; Theodosiou et al. 1998). Upon Hh signalling, Cos2 directly interacts with the cytoplasmic tail of Smo (Jia et al. 2003). Cos2 recruitment to Smo prevents Ci sequestration in the cytoplasm and its subsequent cleavage, allowing the full-length form of Ci to be trafficked to the nucleus to activate Hh-target gene transcription (*Figure 1.1*).

Suppressor of fused (Su(Fu)) is a cytoplasmic protein, which also negatively regulates Hh signalling by controlling the nuclear-cytoplasmic localisation of Ci (*Figure 1.1*) (Preat 1992; Methot and Basler 2000). Removal of Su(Fu) activity in PKA mutant *Drosophila* further enhances the level of Hh-target gene activation, compared to PKA mutant *Drosophila* with normal Su(Fu) activity. This is caused by the constitutive nuclear import of Ci, suggesting that Su(Fu) normally antagonises this transport (Preat 1992; Methot and Basler 2000). When Hh signalling occurs, a serine/threonine kinase encoded by *Fused* (*Fu*) suppresses the antagonistic action of Su(Fu) (Alves et al. 1998; Therond et al. 1999). Mammalian Su(Fu) appears to have a more significant role than *Drosophila* Su(Fu), as *Drosophila* Su(Fu) mutants do not display a severe phenotype, whereas Su(Fu) null mutant mice die at E9.5 with a Hh gain-of-function phenotype (Svard et al. 2006; Varjosalo et al. 2006).

Together, loss and gain-of-function studies in the zebrafish homologs of Cos2, PKA, Fu and Su(Fu) suggest that Gli processing is regulated by conserved mechanisms in vertebrates and invertebrates (Hammerschmidt et al. 1996; Tay et al. 2005; Svard et al. 2006; Varjosalo et al. 2006). The processing of Gli proteins creates a ratio of Gli activator and Gli repressor proteins, the balance of which is important for precise tissue patterning, such as in the neural tube (Ruiz i Altaba 1998; Jacob and Briscoe 2003).

Gli1 has only activator functions as it cannot be proteolytically cleaved. It is also a target of Hh signalling (Sasaki et al. 1997; Ruiz i Altaba 1998; Dai et al. 1999). Thus, Gli1 cannot be the initial transducer of the Hh signal (Bai et al. 2002), and consistent with this, it is dispensible for mouse development (Park et al. 2000). However in the zebrafish, Gli1 is the main activator of Hh-target genes in the ventral neuroectoderm, and *gli1* mutants (*detour* (*dtr*) zebrafish) display a phenotype that is also observed in *smu* zebrafish, which carry a mutation in Smoothened (Barresi et al. 2000; Varga et al. 2001). This includes abnormal optic chiasm formation, a loss of lateral floor plate cells, and a failure of cranial motor neurons to differentiate (Brand et al. 1996; Karlstrom et al. 1996; Chandrasekhar et al. 1999; Odenthal et al. 2000). However, unlike *smu* zebrafish, slow muscle fibres are present in the somites of *dtr* zebrafish despite the requirement of Hh signalling for slow muscle fibre specification (Brand et al. 1996; Barresi et al. 2000; Ingham and Kim 2005). This suggests that Gli2 can compensate for the lack of Gli1 to specify slow muscle fibres. Mutations in *you-too* (*yot*) encode a C-terminally truncated Gli2 in the zebrafish, which acts as a dominant repressor of Gli mediated transcription (Karlstrom et al. 1999; Karlstrom et al. 2003). In the absence of both Gli1 and Gli2 function in *yot* zebrafish, defective somite patterning and a loss of slow muscle is observed (Karlstrom et al. 2003). Therefore, compared to *dtr* zebrafish, *yot* zebrafish have a greater reduction in Hh-target gene expression and display a greater similarity to *smu* zebrafish (Karlstrom et al. 2003).

In contrast to Gli1 which acts solely as an activator, Gli2 and Gli3 both have activator and repressor functions in the mouse and zebrafish (Borycki et al. 2000; Tyurina et al. 2005; Wang et

al. 2007), although Gli3 predominately acts as a repressor and Gli2 predominantly acts as an activator (Ruiz i Altaba 1998; Litingtung and Chiang 2000; Wang et al. 2000). In zebrafish, Gli3 function overlaps with that of Gli1 to activate Hh-target genes in the ventral neural tube at high levels of Hh signalling, but acts as a repressor together with Gli2 in the dorsal spinal cord, midbrain and hindbrain during late somitogenesis stages (Li et al. 2004).

Taken together, the Hh signalling pathway is conserved and operates to play a key role in the development of invertebrates and vertebrates, at both embryonic and adult stages of development. The vast array of genes regulated by the Gli transcription factors means that Hh is critical to the formation of a wide variety of tissues, and in its absence, severe patterning or proliferation defects are observed.

1.2: Somitogenesis

1.2.1: An overview of zebrafish somitogenesis

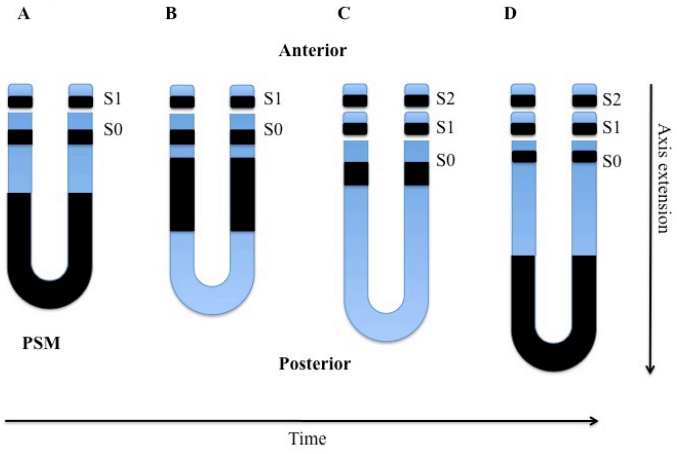
The zebrafish offers an ideal system to study the formation of somites, owing in part to its optical clarity and the external fertilisation of the embryos. Genetic screens for zebrafish mutants with defects in somitogenesis have identified over 50 genes that are essential for normal somite development (van Eeden et al. 1996; Stickney et al. 2000), and so our understanding of how specific gene products regulate cell behaviour during somite development is ever expanding.

The overall process of somitogenesis in zebrafish is similar to that in birds and mammals (Kimmel et al. 1995). Paraxial mesoderm develops from cells which converge towards the dorsal side of the embryo, from the margins of the early gastrula (Kimmel et al. 1990). The first somites form at the end of gastrulation (10.5hpf) when cells in the anterior pre-somitic mesoderm (PSM) undergo a mesenchymal to epithelial transition, forming epithelial balls of cells that surround mesenchymal cells. Bilateral somites are then produced at 30 minute intervals in an anterior to posterior direction, until about 30 somite pairs are formed on either side of the notochord by 24hpf (*Figure 1.2*) (Kimmel et al. 1990).

1.2.2: A molecular clock controls the oscillation of genes in the PSM

A molecular clock has been suggested to function with temporal periodicity to allow somitogenesis to occur at regular intervals, and create uniformly sized somites (Cooke and Zeeman 1976; Stickney et al. 2000; Maroto et al. 2012). Consistent with this theory, components of the Notch (including *Her1*, *C-hairy1* and *Lunatic fringe*), FGF (in mouse) and Wnt pathways have been shown to cycle across the PSM of the mouse and the chick in a posterior to anterior direction (*Figure 1.2*), although Wnt is not believed to be involved in the zebrafish (Dequeant and Pourquie 2008; Murray et al. 2011; Maroto et al. 2012). The Notch pathway is critical for clock gene oscillation and somite formation in the mouse, but is only required for synchronising the oscillations in the zebrafish (Jiang et al. 2000; Ozbudak and Lewis 2008; Ferjentsik et al. 2009).

The mechanisms controlling the periodicity of the oscillations are not fully elucidated, although oscillation of Notch pathway components are generated by negative feedback loops whereby the *Her* genes in the mouse, and *Lunatic fringe* in the chick, are repressed by their own protein products (Bessho et al. 2003; Dale et al. 2003). These proteins also have a rapid turnover, meaning the genes can be re-expressed after protein degradation, contributing to the oscillating expression of the genes (Bessho et al. 2003; Dale et al. 2003).



Adapted from Maroto et al. (2012)

Figure 1.2: *A schematic representation of segmentation and somite formation, as seen from a dorsal view*. The pre-somitic mesoderm (PSM) shows cyclic expression of genes (black) that are likely controlled by a molecular clock. Somite formation occurs where cyclic gene expression meets a permissive environment known as the determination front (S0). *Mesp2* is induced in this region, and leads to epithelialisation and somite boundary formation. Newly formed somite is S1, previously formed is S2, and so on. Blue represents mesodermal tissue (pre-somitic or somitic).

1.2.3: Segmentation of the PSM occurs at the determination front

Posterior to anterior gradients of FGF8 and β-Catenin in the PSM are opposed by an anterior to posterior gradient of Retinoic acid (RA) (Maroto et al. 2012). The point that these two gradients meet is known as the determination front (Dubrulle et al. 2001; Aulehla et al. 2003; Diez del Corral and Storey 2004; Aulehla et al. 2008). It is at this point that the PSM begins segmentation, when a wave of cyclic gene expression passes through the determination front (Cooke and Zeeman 1976; Dubrulle et al. 2001). *Mesp2* is activated in the anterior PSM, in addition to an increased expression of adhesion molecules in this region (Durbin et al. 1998; Takke and Campos-Ortega 1999; Durbin et al. 2000; Sawada et al. 2000; Maroto et al. 2012). *Mesp2* activates Ephrin signalling, resulting in epithelialisation of the anterior PSM, and the creation of a furrow (Durbin et al. 1998). This also involves the formation of a fibronectin-based extracellular matrix (ECM), also initiated by Ephrin signalling (Julich et al. 2009). Following formation of the somite, signals released from surrounding tissues such as the notochord act to specify distinct cell types within the somite (Blagden et al. 1997; Stickney et al. 2000).

1.2.4: Specific cell types divide the somite into distinct compartments

Somites give rise to distinct compartments, which contain precursor cells for the axial skeleton and skeletal muscles. Unlike amniotes, the fish somite is mainly composed of myotome, as the fish does not require a heavy skeleton to support its body weight. The sclerotome, which forms the skeleton, is therefore relatively small in fish. Despite the differences in lineage compartment

size between amniotes and fish, similar molecular mechanisms and signalling pathways are used to specify them. The zebrafish sclerotome consists of a cluster of cells in the ventro-medial domain of the somite that express *twist* and *pax9* (*Figure 1.3*) (Nornes et al. 1996; Morin-Kensicki and Eisen 1997; Stickney et al. 2000). Similar to amniotes, the zebrafish myotome is surrounded by a dermomyotome which becomes distinctive after the formation of the primary myotome (*Figure 1.3*) (Stellabotte and Devoto 2007).Within the myotome, four main classes of muscle cell types can be identified: slow muscle pioneer (MP), medial fast fibre (MFF), fast muscle, and superficial slow muscle fibres (SSF) (*Figure 1.3*) (Ingham and Kim 2005). Unlike amniotes, fast and slow muscle fibres are not inter-mixed, making the zebrafish an ideal model to study the specification of muscle cell types.

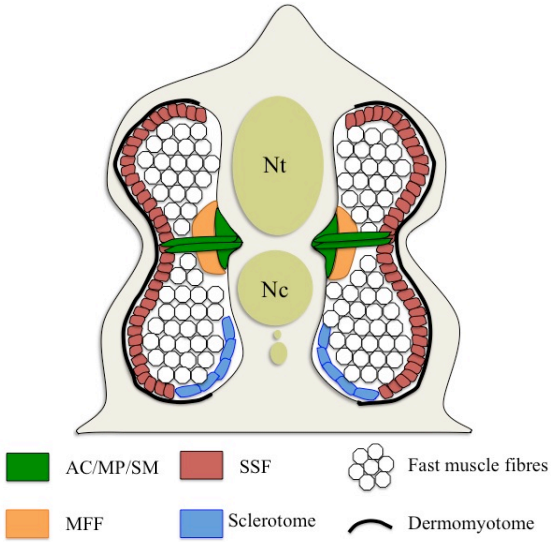


Figure 1.3: *A schematic representation of a transverse section of a 24hpf somite.* Adaxial cells (AC)/muscle pioneer cells (MP) (green) lie adjacent to the notochord in the most medial part of the somite, and develop into slow muscle (SM). Some muscle pioneer cells extend to the lateral surface, resulting in the formation of the horizontal myoseptum (green). Lateral to the adaxial cells/muscle pioneer cells are the medial fast fibres (MFF) (orange), and lateral to this is the fast muscle domain (white), making up the bulk of the somite. On the lateral surface of the somite, superficial slow muscle fibres (SSF) (red) are surrounded by the dermomyotome (shown in black). The sclerotome is located on the ventro-medial portion of the zebrafish somite (blue).

1.3: Muscle cell specification

1.3.1: Myogenic regulatory factors control muscle cell specification

Muscle cells are specified by the myogenic regulatory factors (MRFs), MyoD, Myf5, Mrf4, and Myogenin. These are basic helix-loop-helix (HLH) proteins, which can dimerise with E-class HLH proteins, and interact with E-box consensus sequences (CANNTG) present in the enhancers and promoters of muscle specific genes (Edmondson and Olson 1993). In vivo, MRFs are functionally redundant, although they do play slightly different roles during skeletal myogenesis (Pownall et al. 2002).

In the zebrafish, the first MRF to be expressed is *myoD*, detectable from 70% epiboly, closely followed by *myf5*, at 80% epiboly (Weinberg et al. 1996; Coutelle et al. 2001). Both *myoD* and *myf5* are then expressed in adaxial cells, which are cuboidal shaped cells located adjacent to the notochord (Devoto et al. 1996) (*Figure 1.3*), whilst only *myf5* is expressed in the more lateral cells of the PSM (Coutelle et al. 2001; Pownall et al. 2002). Following somite formation, *myf5*

expression is down-regulated in the lateral paraxial mesoderm cells of the somite, whereas *myoD* expression is initiated in the lateral somite, the prospective fast muscle domain. Lateral myogenesis is therefore driven by *myoD* alone (Hinits et al. 2009). Both MRFs are maintained in the adaxial cells (Weinberg et al. 1996; Coutelle et al. 2001). *Myogenin* expression begins later in the segmentation process, and closely follows the expression pattern of *myoD* in the adaxial cells and the lateral somite (Weinberg et al. 1996; Coutelle et al. 2001). This is in line with data indicating that MRFs can auto-regulate their own expression and cross-regulate the expression of other MRFs (Pownall et al. 2002). *mrf4* is initiated at the 5-somite stage in adaxial cells, and is expressed in both slow and fast differentiating muscle cells in the zebrafish, although it is not expressed in the PSM despite the presence of terminally differentiating adaxial cells (Hinits et al. 2007). *mrf4* does not appear to contribute to early myogenesis in the zebrafish, and *mrf4* null mutants are fertile and viable (Hinits et al. 2009).

1.3.2: Regulation of myogenic regulatory factors

In zebrafish, Shh is necessary for normal expression of *myoD* and *myf5* in adaxial cells, although expression is maintained in the PSM and lateral paraxial mesoderm in zebrafish that are deficient in Shh (Lewis et al. 1999; Coutelle et al. 2001). This suggests that Shh is not required for the initiation of myogenesis, although the loss of MRF expression causes defective slow muscle differentiation (Coutelle et al. 2001). This causes a loss of *mrf4* expression in the slow muscle fibres (Hinits et al. 2007). Medial fast fibres also require *myf5* or *myoD* for myogenesis (Hinits et al. 2009). *myoD* expression in the lateral somite is Shh independent (Hinits et al. 2009), and instead relies on FGF8 signalling for its expression, where it functions to specify the fast muscle lineage (Groves et al. 2005; Hinits et al. 2009).

Overall, different combinations of MRFs are required for myogenesis (Pownall et al. 2002) in the distinct compartments of the zebrafish somite. They are regulated by Hh and FGF8 signalling in the medial and lateral somite, respectively, in addition to auto-regulatory feedback mechanisms (Coutelle et al. 2001; Pownall et al. 2002; Groves et al. 2005; Hinits et al. 2009). The specification of muscle cell types occurs comparatively much earlier in the zebrafish myotome than in amniotes (Ingham and Kim 2005). After muscle cell determination, muscle cells differentiate with specific slow or fast fibre type characteristics.

1.4: An overview of muscle cell differentiation

1.4.1: Fast vs slow muscle fate is governed by *prdm1* expression

Adaxial cells expressing *myf5* and *myoD* develop into slow muscle fibres, generated in response to Hh signalling. Hh signalling activates expression of the zinc finger transcription factor Prdm1, encoded by *u-boot*, in the prospective slow muscle cells (Roy et al. 2001; Baxendale et al. 2004; Liew et al. 2008; von Hofsten et al. 2008). The C-terminal zinc finger domain of Prdm1 mediates

nuclear import, binding to DNA, and the recruitment of co-repressors (Yu et al. 2000). Prdm1 functions to repress fast muscle gene expression (Roy et al. 2001) in addition to indirectly activating slow muscle differentiation, through the repression of the repressive transcription factor Sox6 (von Hofsten et al. 2008). Sox6 is a repressor of slow muscle fibre differentiation (Hagiwara et al. 2007; von Hofsten et al. 2008). Muscle fibres that do not express *prdm1* are fated to become fast muscle fibres (*Figure 1.4*). Therefore, *prdm1* acts as a global switch determining slow or fast muscle fate (Roy et al. 2001). The homeobox gene *prox1* is a marker of differentiating slow muscle, and is implicated in the assembly of slow myofibrils (Roy et al. 2001).

1.4.2: Late myogenesis occurs in the dermomyotome

Similar to amniotes, zebrafish myotomal growth at later stages of development relies on myogenic cells in the dermomyotome (Feng et al. 2006). The dermomyotome lies external to the superficial slow muscle fibres, and expresses *pax3* and *pax7* (Devoto et al. 2006). In the immature somite, these *pax3/7*-expressing progenitor like cells are located in the anterior somite (Hollway et al. 2007). The somite rotates 90° (clockwise for somites on the right of the notochord, anti clockwise for somites on the left of the notochord) which places the *pax3/7*-expressing cells onto the lateral surface, where they remain in a myogenic progenitor like state (Hollway et al. 2007). Hh signalling causes myogenic differentiation of the superficial dermomyotome into fast muscle cells (*Figure 1.4*) (Feng et al. 2006).

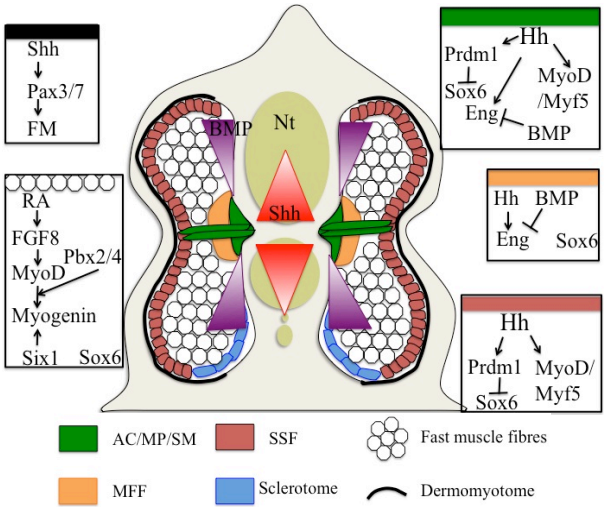


Figure 1.4: *A schematic representation of the signalling pathways that determine specific muscle cell types.* Sonic hedgehog (Shh) concentration is greatest at the notochord and floor plate of the neural tube, its sources of secretion. Bone morphogenetic protein (BMP) concentration is greatest at the most dorsal and ventral parts of the somite, and the Shh and BMP gradients act antagonistically. Muscle pioneer cells (MP), arising from the adaxial cells (AC) require the greatest Shh concentration and are induced next to the notochord, where they express *prdm1* and *eng*. The majority of adaxial cells migrate to the lateral surface of the somite, and also require Shh expression to induce *prdm1*. Following slow muscle cell (SM) migration to the lateral surface, Shh is able to signal to the medial fast fibres (MFF), and enduce *eng*. *eng* expression in MPs and MFFs is restricted by BMP signals. Lateral to the MFFs, the fast muscle domain is Shh independent, and relies on *FGF8* via retinoic acid (RA) signalling. *sox6* expression is also required for the fast muscle differentiation programme. The dermomyotome *pax3* positive progenitor cells are capable of responding to late Shh signals, and contribute to the growing fast muscle domain.

1.5: Specific fibre type differentiation

1.5.1: Adaxial cells form muscle pioneer cells

Slow muscle pioneers arise from adaxial cells (*Figure 1.3*). After somite formation, adaxial cells undergo a complex morphological change in which a pseudo-epithelial sheet composed roughly of a 4 x 5 array of cuboidal adaxial cells elongate whilst located medially. This elongation contributes to the displacement of adaxial cells dorsally and ventrally, creating a 1 x 20 stack of elongated cells, along the dorso-ventral axis (*Figure 1.5*) (Daggett et al. 2007). Whilst differentiating into slow muscle fibres, most of these adaxial cells migrate laterally to the surface of the somite, where they form the superficial slow fibre layer (*Figure 1.6*) (Devoto et al. 1996). Adaxial cells that remain medially located differentiate into the slow muscle pioneer cells. These form the horizontal myoseptum of the somite, separating the ventral and dorsal populations of muscle fibres (*Figure 1.6*) Slow muscle fibres express slow *myosin heavy chain* (*myhc*) and *prox1*, and slow MPs also express Engrailed proteins (Eng1, Eng2) (Blagden et al. 1997; Roy et al. 2001).

Smu, *yot* and *prdm1* (*ubo*) mutant zebrafish all display a similar phenotype. These mutants are characterised by U-shaped somites, caused by the loss of MPs and the horizontal myoseptum (van Eeden et al. 1996). The loss of MPs in *smu* and *yot* embryos is consistent with the idea that Hh induces the slow muscle lineage (Lewis et al. 1999; Barresi et al. 2000). In *ubo* zebrafish, MP cells are lost despite the presence of Hh signalling, and the continued expression of *myoD* and *ptc1* (Roy et al. 2001). This highlights the importance of *prdm1* in the formation of the slow muscle lineage, and places *prdm1* downstream of Hh signalling.

In addition to its role in *prdm1* activation, early Hh signalling is required for *eng* expression in MP cells (Wolff et al. 2003; Maurya et al. 2011). Over-expression of *hh* mRNA or dominant negative PKA mRNA causes an increase in *eng* expressing MP cells, along with an expanded expression of *myoD* and *myogenin*, at the expense of other cell types (Currie and Ingham 1996; Hammerschmidt et al. 1996). In agreement with this, *ptc1/2* double mutant zebrafish show a complete conversion of the somite to the slow muscle lineage (Koudijs et al. 2008). The expansion of the *eng* expressing MP cell population in zebrafish with increased Hh signalling shows a dorso-ventral bias, with the extreme dorsal and ventral somite resisting the expansion compared to regions adajacent to the notochord (Hammerschmidt et al. 1996; Du et al. 1997). These extreme regions are sites of *BMP* and pSmad expression (Du et al. 1997). There is therefore an inverse correlation of *eng* expression and *BMP* expression. Driving nuclear accumulation of pSmad in MPs can suppress *eng* expression, even in the presence of Hh. Hh signalling from the notochord and floorplate is required to restrict the accumulation of pSmad in adaxial cells, allowing MP and MFF formation to occur (Maurya et al. 2011). *Wnt11* is also highly expressed in the adaxial cells, although its role in slow muscle formation is not known (Makita et al. 1998).

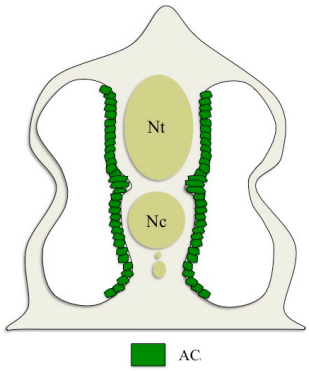


Figure 1.5: *A schematic representation of early adaxial cell (AC) morphogenesis.* After adaxial cell specification, cellular intercalation and migration results in the formation of a stack of adaxial cells along the dorso-ventral axis on the medial surface of the somite.

1.5.2: Fast muscle fibres

1.5.2.1: MFFs are dependent on Hh signalling for their specification

Muscle cells that do not express *prdm1* will form fast muscle fibres. These myoblasts differentiate and express fast *myhc* during the process of slow muscle fibre lateral migration to the superficial surface of the somite (Henry and Amacher 2004), and then fuse together to create around 80 multinucleated fast muscle fibres per somite (Ingham and Kim 2005). A subset of the fast fibres will form MFFs. These are the fibres that are closest to the notochord and MP cells, and like MP cells, express *eng*. BMP and pSmad also restricts the domain of *eng* expressing MFFs (Maurya et al. 2011). MFFs are still present in the *ubo* mutant as judged by expression of *eng* and *myoD* (Groves et al. 2005; Ingham and Kim 2005), and over-expression of *hh* in the *ubo* mutant results in ectopic *eng* expression throughout the somite, without the usual loss of fast muscle fibres through conversion to slow muscle fibres (Roy et al. 2001). This shows that unlike slow muscle fibres, fast muscle fibres are still capable of forming in the *ubo* mutant, and do not require *prdm1* for their differentiation.

Interestingly, unlike the main bulk of lateral fast muscle cells, MFFs show some requirement for Hh signalling, as *smu* and *yot* embryos lack MFF *Eng* expression, although they retain *myoD* expression (Wolff et al. 2003; Groves et al. 2005).

Therefore, the MFFs are a distinct population of fast muscle fibres which do not require *prdm1* for their formation. Similar to the lateral fast muscle domain, MFFs do not require Hh signalling for the expression of *myoD*, however unlike the rest of the fast muscle myotome, they do require Hh signalling for their specification, through induction of *eng*.

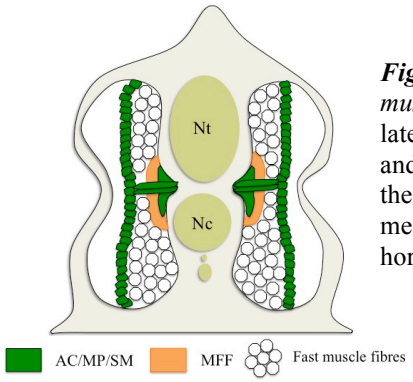


Figure 1.6: *A schematic representation of medial to lateral slow muscle cell migration.* A subset of adaxial cells migrate towards the lateral surface of the somite to form the superficial slow muscle layer, and in doing so, induce a wave of fast muscle differentiation medial to them. Some of the migrating slow muscle cells remain attached to the medial surface of the somite, and so span the entire somite creating the horizontal myoseptum.

1.5.2.2: Lateral somitic fast fibres are Hh independent

The main portion of the somite consists of lateral fast muscle fibres that are Hh-independent. The MP cells create a gradient of Hh by strongly expressing its receptor Ptc in response to Hh signalling, and also express Hhip, another negative regulator of Hh activity (Lewis et al. 1999; Ochi et al. 2006). Ptc and Hhip receptors prevent the lateral diffusion of Hh, and so create a sharp Hh gradient. The resulting drop in Hh concentration is insufficient to allow *prdm1* expression in the MFFs and lateral somitic fast muscle domain (Ochi and Westerfield 2007). Instead, *sox6* is expressed in fast muscle cells, and its over-expression in adaxial cells causes an inhibition of slow muscle fibre gene expression, such as *prox1*. *sox6* is ectopically expressed in the adaxial cells of *ubo* mutants, which explains the loss of slow muscle fibres observed in the *ubo* mutant (Roy et al. 2001).

Unlike in adaxial cells, *myoD* expression in the lateral somitic fast muscle cells, and in MFFs is Hh independent (Groves et al. 2005). Other regulatory networks that function in fast muscle myogenesis include FGF and RA signalling (*Figure 1.4*). *FGF8* expression is required for *myoD,* but not *myf5* expression in the lateral fast somite, and in the absence of *FGF8* in the *acerebellar* (*ace*) mutant, myogenic cells remain in a dermomyotome-like state characterised by *pax3* expression and a lack of fast MyHC expression (Groves et al. 2005). *myogenin* is a direct target of MyoD (Weinberg et al. 1996; Bergstrom et al. 2002), and thus blockage of FGF8 signalling also causes a reduction in *myogenin* expression in the lateral somite. FGF8 signalling is also impaired when Retinoic acid (RA) synthesis is inhibited, revealing that *FGF8* operates downstream of RA in mediating fast muscle cell development (Hamade et al. 2006).

MyoD, and probably Myf5, activate *myogenin* expression and fast muscle differentiation in co-ordination with Pbx2 and Pbx4 homeodomain proteins (Maves et al. 2007). Another homeodomain protein, Six1, is also required for efficient activation of *myogenin* and the fast muscle genes, *myhz1* and *myhz2*, in the zebrafish fast muscle (Bessarab et al. 2008). However, *six1* morpholino-mediated knockdown does not affect *myoD* expression, raising the possibility that *six1* plays a role in the later stages of fast muscle differentiation (Bessarab et al. 2008). Nonetheless, as *myogenin* expression eventually recovers in *six1* morphants, it suggests that other unknown factors are also involved in fast muscle differentiation.

A later role for Hh signalling has also been identified in the formation of mature differentiated fast muscles (Feng et al. 2006). Indeed, Hh acts on the dermomyotome and causes a down-regulation of *pax3* and *pax7* expression (Feng et al. 2006). These dermomyotomal cells then express *myogenin* and undergo myogenic differentiation, contributing to the growth of fast muscles in the zebrafish. In the absence of Hh signalling (in *smu* and *yot* mutant zebrafish) an overabundance of Pax7 positive myogenic progenitor cells is observed, which correlates with a decrease in the formation of late fast muscle fibres (Feng et al. 2006). Conversely, over-

expression of Hh dramatically reduces the number of *pax7* expressing myogenic cells, as demonstrated by injection of *shh* mRNA into the zebrafish embryo (Du et al. 1997).

Overall, the bulk of the somite consists of fast muscle fibres which form independently of Shh. Fast muscle fibre differentiation relies on the combined actions of FGF8, Six1, and Pbx proteins. Myogenic cells originating from the dermomyotome contribute to later growth of the fast muscle myotome, and this is dependent on Hh signalling.

1.5.3: Adaxial cells form the superficial slow fibres

As previously mentioned, following their formation the majority of adaxial cells migrate in a medial to lateral direction throughout the extent of the myotome, and come to lie on the lateral somitic surface (*Figure 1.6*) (Devoto et al. 1996; Blagden et al. 1997). This process depends on a dynamic and differential expression of N and M-cadherins, which are expressed in the adaxial cells and throughout the myotome. A loss of either of these homophilic adhesion molecules causes abnormal migration (Cortes et al. 2003). Migration of slow muscle fibres initiates the differentiation of MFFs and the lateral somitic fast muscle fibres (Henry and Amacher 2004).

1.5.4: Fibre type specification depends on the concentration and timing of Shh exposure

MP cells, MFFs, and SSFs all require Hh signalling for their correct specification and development. However, the concentration and timing of Hh signalling is a critical factor in this process.

MP cells, which form adjacent to the notochord, the source of Shh, require the greatest concentration of Shh for their formation, compared to MFFs and SSFs. Application of 15μM cyclopamine to zebrafish embryos at 5.5hpf causes the loss of MP cells, whilst application of 20μM or above cyclopamine results in the suppression of *Eng* expression in MFFs too. SSFs were only lost when a concentration of 30μM or above cyclopamine was added to the zebrafish, suggesting that MP specification requires a greater Hh concentration than MFFs, which require a greater Hh concentration than SSFs (*Figure 1.7*) (Wolff et al. 2003).

The timing of Hh signalling also plays an important part in the specification of distinct muscle cell types. Application of cyclopamine between the one-cell stage and 7hpf results in the loss of MP cells, whilst MFFs are unaffected. However, after 18hpf, cyclopamine treatment causes the loss of MFF formation, and the MPs remain unaffected (Wolff et al. 2003). The temporal effect of Hh signalling provides an answer as to why the MFF population form fast muscle cell characteristics, despite the high level exposure to Hh. The migration of slow muscle cell precursors away from the midline means that the MFFs will have an increased exposure to Hh, although at this time they are already specified as fast muscle cells, and cannot be converted back to the slow muscle lineage (*Figure 1.8*) (Wolff et al. 2003).

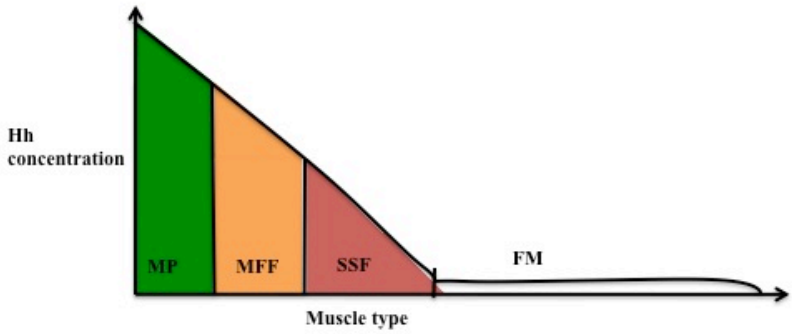


Figure 1.7: *A representation of the hedgehog (hh) concentrations required to induce different muscle cell types.* Muscle pioneer cells (MP) require the greatest concentration of Hh signalling, followed by the medial fast fibres (MFF). Although the superficial slow fibres (SSF) are situated furthest from the source of Hh, they need a higher level of Hh signalling compared to the fast muscle fibres (FM), which only require Hh signalling at a later stage.

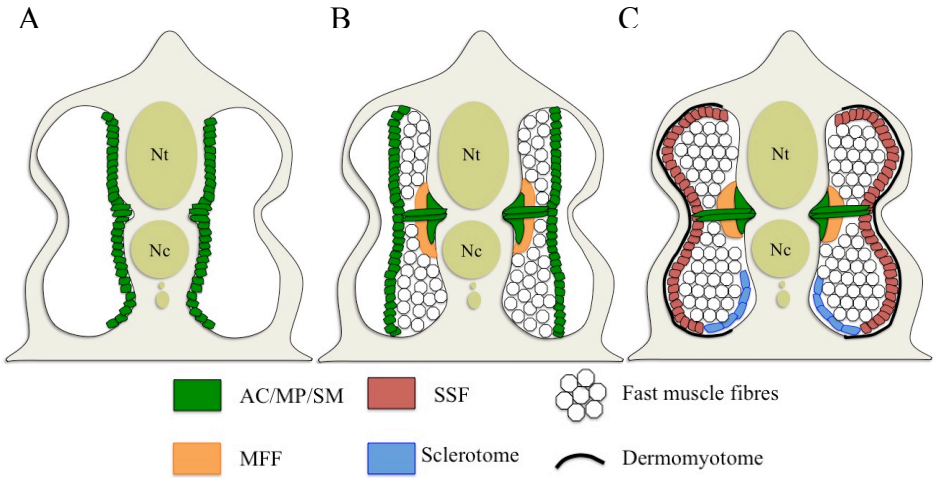


Figure 1.8: *An overview of muscle cell compartmentalisation.* Adaxial cells (AC) and muscle pioneer cells (MP) (green) are specified first (A). The majority of these slow muscle cells migrate to the lateral surface of the somite to form the superficial slow fibre (SSF) layer (C) (red). During the migratory process, the medial fast fibres (MFF) (orange) and the fast muscle (FM) domain (white) are specified and begin differentiation (B). The mature somite is surrounded by a dermomyotome (black), and also contains a sclerotome (blue) that is located in the ventro-medial compartment of the somite (C).

To conclude, muscle specification and morphogenesis is a complex process requiring a range of differing transcription factors and signalling pathways, operating in a spatial, temporal, and concentration dependent manner. Some of the regulators of muscle fate including *hh*, *smad*, *prdm1, sox6* and *FGF8* have been identified, although there are numerous other and essential pathways yet to be defined. Zebrafish ENU-induced mutations and chromatin immunoprecipation analyses using muscle specific genes such as *prdm1,* are likely to lead the way in identifying these important novel genes involved in muscle cell specification and morphogenesis (van Eeden et al. 1996; von Hofsten et al. 2008; Kettleborough et al. 2011).

1.6: Muscle fibre elongation and boundary attachment

During the process of muscle fibre differentiation, muscle fibres elongate in both rostral and caudal directions towards the somitic boundaries. It is to these boundaries that the muscle fibres must attach to generate stable contractile muscle groups.

Fast muscle fibre elongation can be separated into three discrete phases (Snow et al. 2008). After slow muscle migration, the first phase sees fast muscle cells exhibiting random protrusive activity in all directions. In the second phase, the protrusions fill out and thicken, but only in the direction of the extension, towards the anterior and posterior early somite epithelial boundaries. In the third phase, the muscle fibres anchor to the somite boundaries also known as the myotendinous junction (MTJ) (Snow et al. 2008), which contains extra-cellular matrix (ECM) proteins such as Fibronectin, Laminin, and the Laminin receptors (Crawford et al. 2003; Snow and Henry 2009).

In Laminin β1 and Laminin γ1 morphant zebrafish, fast muscle cell elongation is delayed, and cells appear misoriented (Snow et al. 2008). This elongation phenotype could be due to a number of reasons, including the early disorganisation of the fast muscle cells. Laminin could also be required in the ECM surrounding the muscle fibres to generate traction forces that allow the muscle cells to elongate and migrate, by acting as an adhesive substrate. Thirdly, fast muscle elongation could be delayed because slow muscle migration is also delayed in the Laminin morphants (Peterson and Henry 2010). It has been shown that slow muscle fibre migration in response to Hh signalling is necessary and sufficient to induce efficient fast muscle fibre elongation (Henry and Amacher 2004). Therefore, Hh is indirectly required for the recovery of fast muscle elongation in the Laminin γ1 mutant zebrafish, through the specification of slow muscle fibres (Peterson and Henry 2010), and their migration to the lateral somitic surface.

After elongation, some fast muscle fibres do not stop elongating at the MTJ, indicating a failure of muscle fibre capture (*Figure 1.9*). Transplantion of dextran-injected wild-type cells into the Laminin γ1 morphant reveals that the process of fibre capture is a cell autonomous process (Peterson and Henry 2010). Other molecular cues must also play a role in fibre-boundary capture, as 75% of fibres do stop normally at the MTJ in the Laminin γ1 morphant.

Taken together, efficient muscle fibre elongation, migration and differentiation depend on the presence of ECM components. In the absence of ECM components a variety of defects associated with the muscle are observed in the zebrafish and in other vertebrates (Sunada and Campbell 1995; Hall et al. 2007; Jacoby et al. 2009).

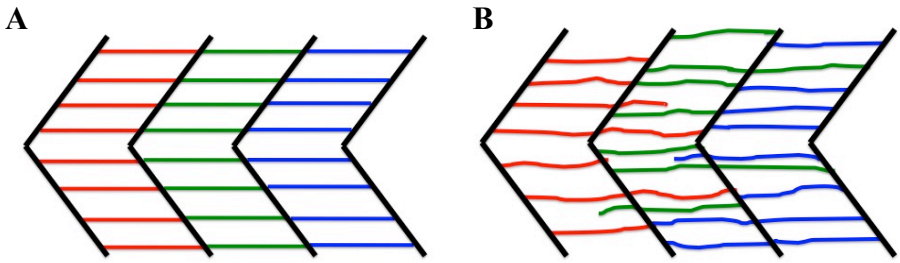


Figure 1.9: *A schematic representation of muscle fibre boundary capture in wild-type (left) or in Laminin β1 or Laminin γ1 mutant zebrafish (right).* In wild-type zebrafish, muscle fibres extend along the length of a single somite, and are captured at the myotendinous junction that creates a distinctive somite. In Laminin β1 or Laminin γ1 mutant zebrafish, there are problems with myotendinous junction formation and as a result, muscle fibres are not correctly captured and bound at the somitic boundaries. The fibres continue to extend into another somite, and can travel across the length of two somites.

1.7: The extra-cellular matrix

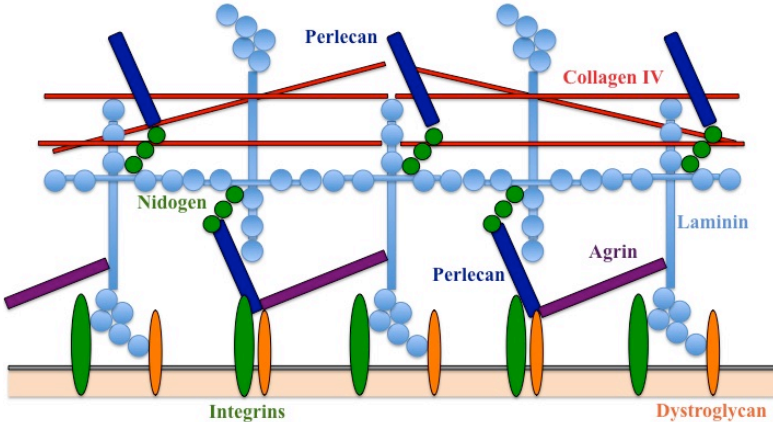
The extra-cellular matrix (ECM) is the non-cellular component that is present in all tissues and organs, comprised of a macromolecular network that is secreted from cells, and adapted to the functional requirements of a specific tissue (Rozario and DeSimone 2010).

The ECM is multifunctional, playing crucial biomechanical and biochemical roles simultaneously. It provides a physical substrate for adjacent cells, allowing cellular migration and the compartmentalisation of different tissues. It can also provide a buffering action to maintain extracellular homeostasis, and regulate gene transcription. In this way, the ECM regulates tissue morphogenesis, differentiation and homeostasis (Frantz et al. 2010). The ECM can directly affect cell behaviour and gene expression through interaction with the cell-surface receptors, Integrins and Dystroglycan. Downstream of these receptors, specific signalling cascades are activated depending upon the specific ECM ligands. Through sequestration of growth factors, including Shh, FGFs, and Myostatin (Horowitz et al. 2002; Blaess et al. 2004), the ECM can also indirectly affect cell fate and behaviour. The ECM is capable of presenting soluble growth factors to cells in a spatially and temporally regulated manner, and is able to create a morphogen gradient of growth factors. This differential presentation of growth factors to cells in a tissue will impact upon cellular differentiation, migration and tissue morphogenesis (Rozario and DeSimone 2010). It is often the cells themselves that secrete factors capable of degrading the ECM such as matrix metalloproteinases (MMPs) and ADAMs (Guerin and Holland 1995; Kherif et al. 1999), resulting in the release of these growth factors. Therefore, the ECM is highly dynamic and modified by the cells that it is in contact with, creating a bi-directional mode of cell-matrix communication (Larsen et al. 2006; Frantz et al. 2010; Rozario and DeSimone 2010). Bi-directional communication that is mediated by the Integrin receptors is known as outside-in, and inside-out signalling (Legate et al. 2009). As a result of bi-directional signalling, the ECM environment

surrounding different tissues varies in its structural and biochemical composition. This is also dependent upon the stage of embryonic and adult development.

The ECM comprises two major types of matrices, interstitial and pericellular. The interstitial matrix consists of Collagens, Elastins, Fibronectins, Tenascins, proteoglycans and glycosaminoglycans (Frantz et al. 2010). To date, 28 types of Collagen have been identified (Myllyharju and Kivirikko 2004; Gordon and Hahn 2010) and this is the main structural element in the interstitial ECM. Most interstitial Collagen is secreted from fibroblasts that inhabit the stroma in the interstitial ECM (De Wever et al. 2008), and they act to provide tensile strength, cell adhesion and migratory substrates for cells (Rozario and DeSimone 2010). Fibronectin also plays a role in the organisation of the interstitial ECM, and it also has an important role in cell attachments (Smith et al. 2007).

Pericellular matrices include Fibronectin matrices and basement membranes (BMs), which form as sheet-like structures in close contact with cells. Fibronectin matrices are formed from Fibronectin-Fibronectin crosslinking, which occurs when the globular Fibronectin molecule is bound to its Integrin receptors (Mao and Schwarzbauer 2005; Singh et al. 2010). BMs are much more highly organised in comparison to the interstitial matrices.



Adapted from Yurchenco et al. (2004).

Figure 1.10: *Schematic representation of the structure of a basement membrane, anchored to a cell through Integrin and Dystroglycan connections.* Interactions between the individual components that form the BM generate a highly organised sheet-like structure that provides physical and biochemical support to the surrounding cells and tissues. Just some of the connections between the numerous components that form the basement membrane are shown.

1.7.1: Basement membranes are multifunctional networks supporting epithelia

Basement membranes (BMs) are conserved sheet-like structures found in both vertebrates and invertebrates, surrounding epithelia, endothelia, muscle cells, Schwann cells, peripheral nerves, the CNS and fat cells (Yurchenco and Patton 2009). They act to separate these cell types from the surrounding connective tissue and interstitial matrix (LeBleu et al. 2007). In addition to their tissue compartmentalisation and mechanical support roles, they are key to several cellular

processes at both embryonic and adult stages of development, including migration, proliferation, survival, differentiation and morphogenesis (Miner et al. 2004; Miner and Yurchenco 2004; Yurchenco and Wadsworth 2004).

The macromolecular structure of BMs, formed from the polymerisation of its main components Collagen IV and Laminins, provides a large substrate on which cells can adhere to. These adhesive interactions are mediated by specific cellular receptors, namely Integrins and Dystroglycan (*Figure 1.10*) (section 1.8.2), to which the ECM binds with varying affinities (Yurchenco and Patton 2009; Gawlik et al. 2010). Differential expression of these receptors in a given cell can cause migration of the cell in a specific direction along the BM. The composition of the BM can also affect the migratory speed of a cell, possibly due to varying strengths of BM-cell interactions (Rozario and DeSimone 2010). Cell locomotion also occurs when a contractile cytoskeleton generates forces on an ECM substrate. The roles BMs play in regulating cell adhesion and migration are crucial to processes such as muscle fibre elongation and neural crest and axonal pathfinding (reviewed in Henry et al. 2005; Rozario and Desimone 2010).

The BM microenvironment can also regulate the activation state of stem cells and satellite cells. For example, dermomyotomal cell interaction with the dermomyotomal BM is important for their maintenance in an undifferentiated and proliferative state, in the developing mouse somite. Positional cues from the dermomyotomal BM also participate in the polarisation of cells and can promote symmetric cell divisions (Bajanca et al. 2006). Satellite cells, which reside just beneath the skeletal muscle BM, can be activated to proliferate or differentiate, depending on the mechanical forces applied to the BM, caused by muscular contraction or extension (Girgenrath et al. 2005; Grossi et al. 2007; Boonen and Post 2008).

Mutations affecting components of the BM are therefore associated with arrested development at different embryonic stages, and a variety of post natal diseases affecting muscles, nerves, the brain and eyes, skin, vasculature, and the kidneys (Yurchenco 2011). Despite the essential known structural and signalling roles that BMs play in tissue morphogenesis, the mechanisms controlling their formation remain poorly understood. Analysis of the components that constitute the BM and their functions will help our understanding of how BMs form, and the roles that they play during development.

1.7.2: Basement membranes are composed of glycoproteins

BMs consist mainly of the glycoproteins Laminin and Nidogen (Entactin), Collagen type IV, and heparan sulphate proteoglycans (HSPG) such as Perlecan and Agrin (Yurchenco et al. 1992).

1.7.2.1: Nidogen acts as a bridging molecule in the basement membrane

Nidogens are widely expressed and highly conserved sulphated glycoproteins (Ho et al. 2008). Vertebrates have two Nidogen genes, whilst invertebrates have only one (Ackley et al. 2003;

Yurchenco and Patton 2009). Nidogen consists of epidermal growth factor-like modules, low-density lipoprotein receptor YWTD modules, and Nidogen like (NIDO) modules (Mann et al. 1989; Nagayoshi et al. 1989; Kohfeldt et al. 1998). The globular NIDO domains have an unknown function (Ho et al. 2008), whilst the six YWTD modules fold to produce a concave surface, to which a loop of Laminin γ1 LE III 4 can bind with high affinity (Takagi et al. 2003) (see section 1.8.1.3). Nidogen also binds to Collagen type IV, thereby mediating a link between the Laminin and Collagen networks (Baumgartner et al. 1996; Stetefeld et al. 1996). Nidogen also acts as a bridging molecule between Collagen type IV and Perlecan via an EGF module and β-barrel domain (Hopf et al. 2001; Ries et al. 2001), thereby playing an important role in BM stability during development. Consequently, loss of both Nidogen-1 and 2 causes perinatal death caused by mild changes in BMs and organ structure (Bader et al. 2005; Ho et al. 2008).

1.7.2.2: Perlecan acts as a bridging molecule in the basement membrane

Perlecan is an ubiquitous proteoglycan that has five domains, and has the ability to dimerize and oligomerize (Whitelock et al. 2008). Its modular protein core is capable of interacting with numerous ECM components and growth factors, including Laminin α1, Collagen type IV, Nidogen-1/2, Fibronectin, Fibulin-2, Heparins and Heparan Sulphates, FGF2, Shh, BMP-2, vascular endothelial growth factors (VEGF), Wnt, and platelet derived growth factor (PDGF) (Iozzo 2005; Whitelock et al. 2008). The C-terminal domain of Perlecan interacts also with the ECM receptors β-Integrin and α-Dystroglycan (Whitelock et al. 2008). Perlecan can therefore function to regulate the pericellular concentration of growth factors and morphogens, thus may have a profound influence on cellular behaviour. Consistent with this and with the ubiquitous expression pattern of Perlecan, Perlecan null mice demonstrate a complex phenotype not restricted to one tissue or organ, and die at E11.5 (Arikawa-Hirasawa et al. 1999; Costell et al. 1999; Whitelock et al. 2008) highlighting the importance of Perlecan and BMs in tissue morphogenesis.

1.7.2.3: Collagen type IV networks are assembled into basement membranes

Collagen type IV is restricted to and found in all BMs, and is composed of six individual chains, which assemble into three distinct isomers via triple helical interactions; α1,α1,α2; α3,α4,α5 and α5,α5,α6 (Boutaud et al. 2000; Hudson et al. 2003). The N-terminal domain of each Collagen trimer is rich in cysteine and lysine residues, allowing for the cross-linking of four Collagen trimers through disulphide bonds and lysine-hydroxylysine links, observed using electron microscopy and X-ray crystallography (Yurchenco and Furthmayr 1984; Glanville et al. 1985; Siebold et al. 1987). Its C-terminal globular end is capable of self-polymerisation, allowing the generation of a large Collagen type IV network (Weber et al. 1984). Lateral interactions between

the triple helical coils also promote a super coil formation caused by supramolecular twisting of the network, providing further stability to Collagen type IV and the BMs they are associated with (Yurchenco and Ruben 1987; Yurchenco and Patton 2009).

The α1,α1,α2 isoform is detected earliest in the mouse embryo and is expressed ubiquitously (Dziadek and Timpl 1985). Later in development, it is partially replaced by the other Collagen networks in specific tissues, altering the local structure and function of the ECM (Ninomiya et al. 1995; Kruegel and Miosge 2010).

Many cell types can directly bind to Collagen type IV, such as platelets, hepatocytes, keratinocytes, and endothelial, mesangial, pancreatic and tumour cells, as shown in cell attachment and spreading assays (Aumailley and Timpl 1986; Khoshnoodi et al. 2008). These connections are mediated by Integrin and non-Integrin receptors (Khoshnoodi et al. 2008). Loss of Collagen type IV in the mouse results in late embryonic lethality at E10.5 - E11.5 due to impaired BM stability (Guo and Kramer 1989; Guo et al. 1991; Poschl et al. 2004). Unsurprisingly, mutations in Collagen type IV chains can lead to a variety of diseases (Yurchenco and Patton 2009). In human, mutation in the Collagen IV α5 gene results in Alport's syndrome, a disease characterised by renal failure (Hudson et al. 2003).

1.8: Laminins are the main component of basement membranes

1.8.1: The structure of Laminin

Laminins are heterotrimeric glycoproteins, conserved amongst vertebrates and invertebrates (Timpl et al. 1979). Their origin is believed to be from a single gene, and there are now 4 Laminin subunits in invertebrates and 12 known subunits in vertebrates (Miner and Yurchenco 2004). *Drosophila* have two α chains (α1/2 and α3/5), one β chain and one γ chain, which form two Laminin molecules (Urbano et al. 2009). Vertebrates have five α chains, four β chains, and three γ subunits, capable of forming at least 16 Laminin heterotrimers, composed of one subunit of each chain type (*Table 1.3*) (Miner and Yurchenco 2004; Sztal et al. 2011). Additional trimers can also be formed as a result of mRNA splicing of the Laminin α3 and γ2 chains (Airenne et al. 1996; Ferrigno et al. 1997; Lee et al. 2001). Laminin trimers (α,β,γ) such as Laminin-111 (α1, β1, γ1), form a cross-shaped or a T-shaped three dimensional structure (Engel et al. 1981; Beck et al. 1990). This structure is common to most of the Laminin trimers, except Laminins-332, -311, -321, -411, -421 and -423, which contain one or more of the N-terminally truncated chains α3a, α4, β3, and γ2 (*Figure 1.11*) (Tunggal et al. 2000; Hamill et al. 2009).

Individual Laminin chains are composed of domains I to VI (in a C-terminal to N-terminal direction) and include N-terminal globular and C-terminal coiled-coil rod like domains (*Figure 1.12*) (Engel et al. 1981; Beck et al. 1990). The α chain is the largest of the three chain types, with a molecular mass of approximately 400kDa, whilst the β1 and γ1 chains have a molecular mass of

approximately 200kDa. This is because the C-terminal long arm of the α chain contains several globular LG domains (known collectively as the G domain), which mediate the interaction of the Laminins with cell membrane receptors (Talts et al. 1998; Andac et al. 1999).

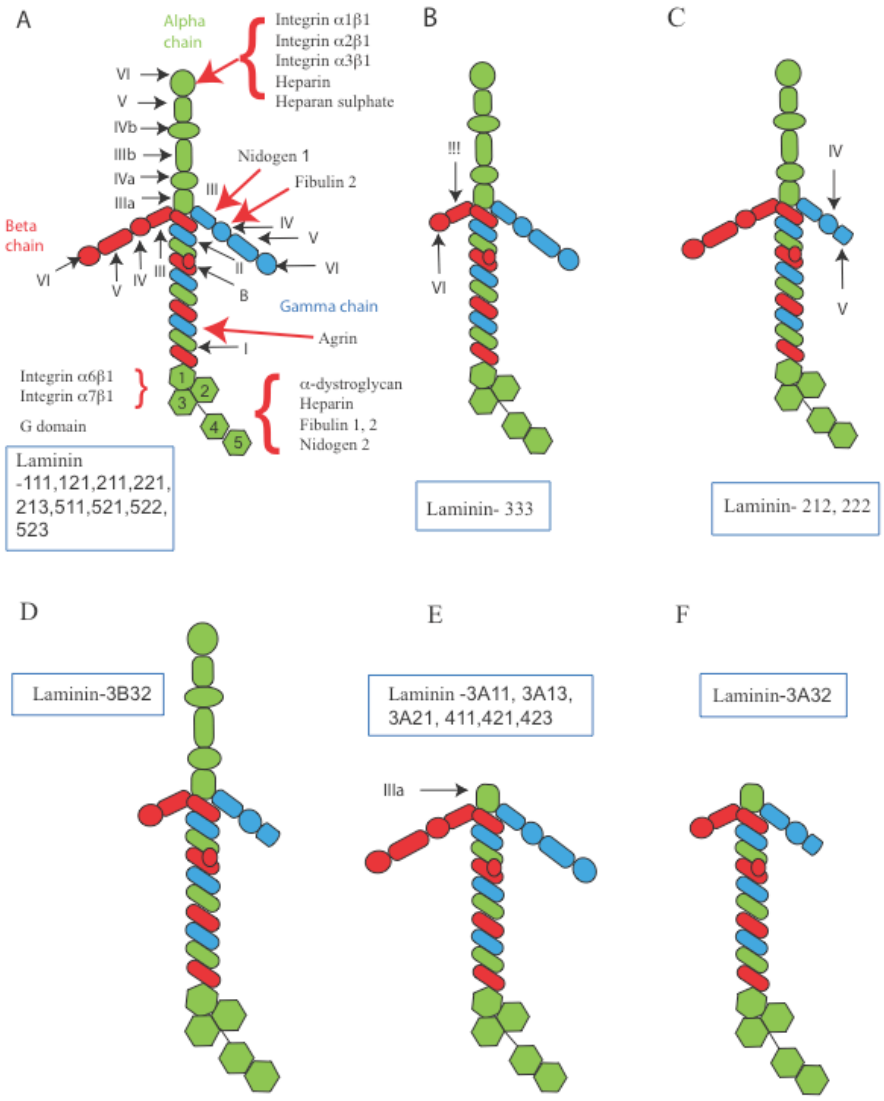


Figure 1.11: *A schematic representation of the different Laminin heterotrimers.* (A) shows the classic Laminin shape that allows for self polymerisation, enabling basement membrane formation. Laminin α3A, and α4 contain a truncated N terminus (E and F), as do Laminin β3 and γ2 (B, C, D and F). Black arrows indicate specific Laminin domains; red arrows/brackets indicate domains involved in the interaction with other ECM proteins and cell surface receptors.

Adapated from Miner and Yurchenco (2004).

1.8.1.1: The G domain is found at the C-terminal of Laminin α chains

The LG (Laminin Globular) domain is composed of 5 subdomains termed LG1 to LG5 (Sasaki et al. 1988). Of these, the LG3 domain diverges the most from the common LG domain template (Carafoli et al. 2009). The β-sandwich protein structure of these domains can be modified by N-glycosylation (Timpl et al. 2000; Carafoli et al. 2009), which affects the interaction of Laminin with other proteins. The primary function of these domains is to interact with cell-surface receptors, mainly the Integrins and Dystroglycan. Different LG domains in a given Laminin chain bind to the same, or to different receptors on the cell surface, and they do so with varying affinities (see section 1.8.2.1). For example, Laminin α1 LG4, but not Laminin α1 LG5, strongly

binds to Heparin (Talts et al. 1999), whilst Laminin α1 LG1-3 and LG4 binds to Integrins and Dystroglycan, respectively. The LG domains also contain a number of cleavage sites, and cleavage can result in modified interaction and adhesion between a cell and the ECM, altering cellular migration (Ryan et al. 1996; Talts et al. 1998; Talts and Timpl 1999; Timpl et al. 2000; Smirnov et al. 2002).

1.8.1.2: Domains I and II consist of a coiled coil

Domains I and II are about 77nm in length (Engel et al. 1981) and are formed from α-helical coiled coil repeats (Hunter et al. 1990), which contain a series of heptad repeats (Beck et al. 1990; Conway and Parry 1990). The length of these domains enables long-range direct interaction between the G domain and the Integrins or Dystroglycan. Several cell adhesion recognition sites are found within these coiled coil domains, such as a motor neuron-specific attachment site (Engvall and Wewer 1996), and an Agrin binding site (Denzer et al. 1998; Meinen et al. 2007).

1.8.1.3: Domain III and V are situated on the Laminin N-terminal arms

Domains III and V (LEa), located in the short arm of Laminins (*Figure 1.12*), are Laminin-type epidermal growth factor-like (LE) domains. These domains act as spacer elements, and so similar to domains I and II, they help to increase the distance that Laminin can interact with other molecules, enabling the possibility of a greater number of interactions. LE module 4 in domain III of Laminin γ1 is also known to associate with Nidogen-1 (Poschl et al. 1996). Laminins α1, α2, α3b and α5 chains have domain IIIa and IIIb (also known as domains LEc and LEb, respectively) (*Figure 1.11*), whilst the β and γ chains have only domain III (LEb) (Beck et al. 1990; Tunggal et al. 2000; Tzu and Marinkovich 2008).

1.8.1.4: Domain IV binds to other ECM components

The globular domains interspersed with the LE domains are known as domain IV. Domains IVa and IVb in full-length α chains are homologous to each other within both the α1 and α2 chains. Domains IVa and IVb differ to each other in the α3b and α5 chains, whilst α3a and α4 do not contain any domain IV (*Figure 1.11*) (Tunggal et al. 2000; Tzu and Marinkovich 2008). Domain IV is referred to as the LF domain in β1 and β2 chains, and as L4 modules in the γ chains (Tunggal et al. 2000; Aumailley et al. 2005). These domains contain binding sites for other ECM components such as Fibulin-2, which binds L4 in the γ2 chain (Allamand et al. 1997; Utani et al. 1997).

1.8.1.5: Domain VI binds to Integrins

Domain VI consists of the N-terminal most globular domain, known as the LN domain. It is the most conserved part of Laminin, and is also capable of binding Integrins α1β1, α2β1, and α3β1 (Hall et al. 1990; Colognato et al. 1997).

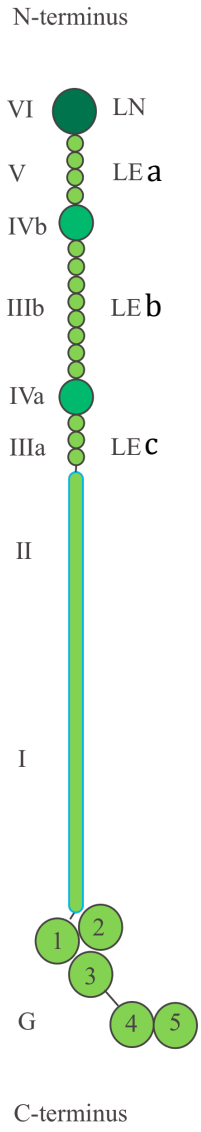


Figure 1.12: *A schematic representation of the Laminin α1 chain* (adapted from (Tunggal et al. 2000). Domain numbers are indicated on the left of the image. Abbreviations: Laminin-type epidermal growth factor-like domains (LE), Laminin module (LN).

1.8.1.6: Laminin α1 is a conserved protein

Lama1 encoding Laminin α1 protein, is a conserved gene amongst vertebrates (*Table 1.1*) (Miner et al. 1997). Human, mouse and zebrafish *Lama1* all contain 63 exons (Zinkevich et al. 2006), whilst chicken *Lama1* contains 64 exons. Human, mouse and zebrafish Laminin α1 proteins are composed of 3075, 3083 and 2680 amino acids, respectively (http://www.ensembl.org).

	Zebrafish Laminin α1			**Mouse Laminin α1**		
	Human	Chicken	Mouse	Human	Chicken	*Drosophila* α1/2 subunit
Overall homology %	49	50	49	76	64	29

Table 1.1: *The amino acid sequence homology (%) of zebrafish Laminin α1 compared to human, chicken and mouse, and of mouse Laminin α1 compared to human, chicken, and the Drosophila Laminin α1/2 subunit* (Garrison et al. 1991; Henchcliffe et al. 1993).

The structure of the C-terminal end of Laminin α1, investigated by electron microscopy (Tashiro et al. 1989), shows a conserved globular domain, comprised of 5 Laminin globular (LG) subdomains. Even in *Drosophila*, the individual globular subdomains of Laminin α1/2 are related to their analogous vertebrate counter parts (Garrison et al. 1991), indicating the importance of these individual subdomains. The *Drosophila* G domain shares 26% similarity at the amino acid level to both mouse and human Laminin α1 (Henchcliffe et al. 1993).

Like all other Laminin chains, Laminin α1 is composed of domains I and II. (Beck et al. 1990; Tunggal et al. 2000). A conserved IKVAV sequence important for cellular adhesion, neurite outgrowth and angiogenesis, is located in the C-terminal end of domains I and II in the Laminin α1 chains (Tashiro et al. 1989; Grant et al. 1992; Nomizu et al. 1995). It has been shown to associate with two 110kDa proteins; the cell membrane-associated Laminin-binding protein (LBP110), and nucleolin. LBP110 is a member of the β-amyloid precursors protein (APP) family, and binds to the IKVAV sequence independently of nucleolin (Kibbey et al. 1993; Kibbey et al. 1995; Kasai et al. 2007). This motif is largely conserved in *Drosophila* too, where an IKVGV sequence exists in domains I and II. Unlike Laminin α3a, α4 and the β and γ chains, Laminin α1 contains domains IIIa (LEc) and IIIb (LEb) (Beck et al. 1990; Tunggal et al. 2000). Domains IVa, IVb and VI are also situated within the Laminin α1 chain (*Figure 1.12*) (Colognato and Yurchenco 2000). It is domains I and II in each Laminin subunit which appear to be the most important domain for the process of Laminin trimer oligomerisation.

1.8.1.7: Laminin subunits are assembled into a heterotrimer

Individual Laminin subunits are glycosylated with high mannose oligosaccharide side chains at the N-terminal end, within the rough endoplasmic reticulum (Morita et al. 1985). This glycosylation acts to stabilise the Laminin subunits, and offer protection from degradation. Varying glycosylation levels between the different Laminin subunits (Champliaud et al. 2000) therefore regulates the amount of each subunit produced, and as a result, the Laminin trimer formed. Glycosylated Laminin subunits are assembled into trimers within the Golgi apparatus (Yurchenco et al. 1997). The first stable association within the trimer occurs when domain I of β and γ chains are disulphide bonded through a pair of cysteines at the N and C-terminal long arms (Paulsson et al. 1985). The α chain then binds with this dimeric complex through disulphide bonds at the N-terminal long arm (Tokida et al. 1990; Yurchenco et al. 1997; Ekblom et al. 2003; Miner and Yurchenco 2004). The C-terminus of domain I in γ chains may have a role in the specific Laminin isoform assembled (Macdonald et al.), as substituting a γ1 chain for either a γ2 or γ3 chain alters the specificity for the α chains.

The heptad repeats within domains I and II of each Laminin subunit type act to stabilise the α-helical coiled coil repeats of the newly formed trimer (Hunter et al. 1990). The heptad repeats (termed a-g) carry different electrical charges; a and d are hydrophobic, e and g are charged, and b, c and f are hydrophilic residues. The residues assume a conformation that is energetically favourable, whilst strengthening the coiled coil (Beck, Hunter et al. 1990; Conway and Parry 1990). There are also glycosylation sites on positions b, c and f, although they appear not to be necessary for chain assembly (Wu et al. 1988).

Both cell culture and in vivo experiments in the mouse and *Drosophila* have shown that α chains can be secreted from cells as a single subunit (Kumagai et al. 1997; Yurchenco et al. 1997), and that the α chain is needed to drive the secretion of the Laminin trimer from the cell (Yurchenco et al. 1997). The α chain is therefore the rate-limiting step in the formation of the trimer. Despite the importance of the Laminin α subunit in driving the secretion of the Laminin trimer from the cell, the mechanism of how it does this is unknown. As the Laminin α chain is the only chain to have a G domain, it is plausible that this domain is involved in the secretion process. The Laminin α G domain could directly or indirectly interact with proteins that are involved in the secretory process. Alternatively, the Laminin α chain could initiate a signalling cascade that results in secretion of the Laminin heterotrimer. The G domain is also responsible for Laminin signalling, mostly through Integrins and Dystroglycans.

1.8.2: Laminin signals through Integrin and Dystroglycan receptors

Integrins and Dystroglycan are the main Laminin receptors (Mercurio 1995). However, Laminins can also bind to Heparin and Heparan sulphates such as Syndecans, and to lipid sulphatides (Roberts et al. 1985; Hall et al. 1990; Colognato et al. 1997; Talts et al. 1999). These interactions can be mediated through Laminin domain VI (*Figure 1.11*) (Colognato et al. 1997). Laminins bind to Integrins and Dystroglycan via the Laminin globular domain at the C-terminus of the protein, although some Integrins can also interact with domain VI at the N-terminus of Laminin, such as Integrins α1β1, α2β1, and α3β1 (Hall et al. 1990; Colognato et al. 1997).

1.8.2.1: Integrin receptors bind Laminin within varying specificity

Integrins are composed of non-covalently associated α and β subunits (Hynes 2002), of which there are 18 α subunits and 8 β subunits capable of forming 24 different heterodimers in vertebrates (*Table 1.2*) (Takada et al. 2007). The α subunit determines ligand specificity, and Integrin heterodimers can bind to multiple ligands (Barczyk et al. 2010). A specific Laminin, or Collagen type IV molecule, can bind to multiple types of Integrin combinations. For example, Laminin-111 and -511 bind to Integrins α1β1, α2β1, α3β1, α6β1, α7β1, α9β1 and α6β4 (Colognato and Yurchenco 2000; van der Flier and Sonnenberg 2001). Of these, Integrins α3β1, α6β1, α7β1 and α6β4 bind primarily to the LG1-3 subdomains on the Laminin α subunits (Nishiuchi et al. 2006). The coiled coil formed from the three Laminin chains is also required for Integrin binding (Deutzmann et al. 1990; Navdaev et al. 2008). In particular, a glutamic acid in the third position from the C-terminal end of the γ chain is important for Integrin binding (Ido et al. 2007).

The LG domains of different Laminin α subunits bind Integrins with varying affinities (Andac et al. 1999; Talts et al. 1999; von der Mark et al. 2007). For example, Laminin α2 in skeletal muscle

primarily binds to Integrin α7β1 and weakly to Integrin α6β1, whilst Laminin α1 binds strongly to Integrin α6β1 and less strongly to Integrin α7β1 (Talts et al. 1999; Gawlik et al. 2010). Both of these Laminin interactions are through LG domains 1-3 only.

Integrin	Expression	Knock-out phenotype
α1	Mouse myotome from E10.5	Viable, fertile
α3	Chick adult myofibres	Kidney, lung, skin defects. Die neonatally
α4	Mouse dermomyotome from E9.5, Myotubes from E12.5-14.5	Heart and placenta defects. Die at E11-14
α5	Mouse epithelial somites, differentiated myocytes in myotome, MPCs migrating to limbs, trunk muscle at E12.5. Zebrafish medial somite.	Defective somite formation and muscular dystrophy in mice, die at E10-11. Zebrafish defective somite boundary maintenance
α6	Mouse dermomyotome and myotome, E13.5–E14.5 trunk and limb muscles. Chick myotome, limb MPCs and myoblasts, primary myotubes	Loss of hemidesmosomes, skin blisters. Die neonatally
α7	Mouse E10.5 myotome, E13.5 intercostal and pectoral muscles, myotendinous junctions	50% viable and fertile, with progressive muscular dystrophy
α9	Mouse E12.5 diaphragm and tongue, E14.5 myotubes	Respiratory defects. Die neonatally
β1	Mouse myogenic cells and all muscle fibres. Zebrafish muscle	Die at the pre-implantation stage
β2	Zebrafish muscle	
β3	Mouse myotendinous junction at E14.5, when β1 is absent. Zebrafish muscle	Haemorrhaging & placental defects. Viable and fertile.

Table 1.2: *Integrin expression in skeletal muscle development, and Integrin subunit inactivation phenotypes*. Adapted from Thorsteinsdóttir et al. (2011).

1.8.2.2: Dystroglycan binds to different Laminin LG domains with varying affinities

Like Integrins, Dystroglycan is composed of two subunits, α and β, although both are encoded by a single gene, *Dag-1* (Ibraghimov-Beskrovnaya et al. 1992). Laminins bind to the extracellular α-Dystroglycan subunit, which is non-covalently linked to the transmembrane β-Dystroglycan subunit (Barresi and Campbell 2006). The LG domains of different Laminin α subunits bind to α-Dystroglycan with varying affinities. Laminin α2 LG1-3 binds to α-Dystroglycan with an affinity three times higher than that of Laminin α2 LG4-5 (Hohenester et al. 1999). Laminin α1 LG4 is responsible for the binding of the LG4-5 domain to α-Dystroglycan, and this interaction is 4 times lower than that of Laminin α2 LG4-5 (Andac et al. 1999). Laminin α1 LG1-3 does not bind to α-

Dystroglycan (Talts et al. 1999). Laminin α1 LG4 and Laminin α2 LG4-5 also bind to other glycoconjugated receptors including Heparan sulphates and Galactosylsulphatides (Harrison et al. 2007). Dystroglycan is also capable of binding the LG domains contained in other components of the ECM, such as Perlecan and Agrin (Gesemann et al. 1996; Friedrich et al. 1999).

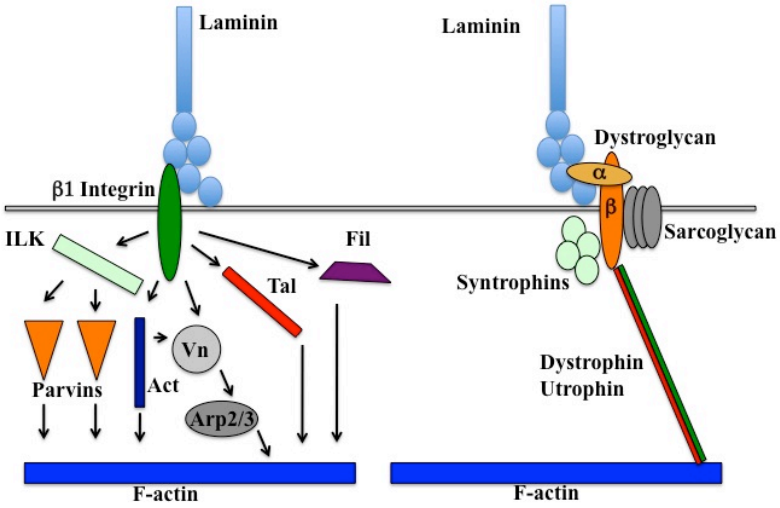
1.8.2.3: Laminin signalling can alter cell shape or cell fate

Integrins and Dystroglycans serve to connect the ECM to the cellular cytoskeleton, and are therefore crucial for correct cytoskeletal organisation and cellular adhesion (Berrier and Yamada 2007). A combination of different Integrins can act together in the activation of specific members of the Rho GTPase family, which play a key role in the organisation of the actin cytoskeleton (Danen and Sonnenberg 2003; Huveneers and Danen 2009). Other important factors in the process include the Integrin-associated adaptor proteins, including Integrin linked kinase (ILK), Focal adhesion kinase (FAK), α and β-Parvins, Talin, α-Actinin and Filamin, and Dystroglycan-associated proteins including Dystrophin and Utrophin (*Figure 1.13*) (Yurchenco and Patton 2009). The cytoplasmic tail of β-Dystroglycan is also associated with many components of similar pathways that are activated by the Integrins, such as Mitogen activated protein kinases (MAPK), and Growth factor receptor-bound protein 2 (Grb2), suggesting similar signalling roles to the Integrins (Cavaldesi et al. 1999; Winder 2001; Barresi and Campbell 2006).

Laminin interaction with Integrins and Dystroglycan not only influence the structural stability and organisation of tissues, but there is also evidence that most ECM adherent cells depend on Integrin-mediated signalling for their survival. Indeed, Integrin-mediated signalling activates the Phosphatidylinositide 3-kinase (PI3K) pathway leading to suppression of tumour protein 53 (p53) and Caspase activity (Boudreau et al. 1995; Zhang et al. 1995).

Mouse embryoid body studies have also revealed that Integrins and Dystroglycans are essential for epiblast survival and differentiation (Li et al. 2002). Similarly, Integrin-mediated cell adhesion is required for the terminal differentiation of chicken myoblasts (Menko and Boettiger 1987).

Integrins also stimulate cyclin D1 expression and suppress cyclin-dependent kinase inhibitors. This leads to the S-phase entry within a cell cycle, and proliferation of a cell (Zhu et al. 1996; Roovers et al. 1999). The Laminin and ECM environment, in combination with a tissue specific set of Laminin receptors, is therefore crucial in determining cell fate and behaviour through the specific activation of intracellular signalling pathways, at both embryonic and post natal stages of development (Danen and Sonnenberg 2003; Givant-Horwitz et al. 2005). Understanding what regulates the Laminin-ECM environment and its receptors will help to develop our knowledge on why a cell behaves or functions in a specific manner, and allow the opportunity to exploit this system and alter a cell's fate or behaviour.



Adapted from Yurchenco and Patton (2009).

Figure 1.13: *Schematic representation of some of the signalling pathways under the control of β1-Integrin that regulate F-actin cytoskeleton dynamics, and the connections between β-Dystroglycan and F-actin.* Most BM components bind to β1-Integrin, which can in turn bind to the cytoskeletal intermediates Integrin-linked kinase (ILK), α- and β-Parvin, α-Actinin (Act), Talin (Tal), Vinculin (Vn), Arp2/3, and Filamin (Fil). Laminin, Agrin and Perlecan can bind to α-Dystroglycan, which binds to the transmembrane protein β-Dystroglycan. β-Dystroglycan binds to F-actin through Dystrophin and Utrophin, and is also in a complex with Sarcoglycans and Syntrophins in muscle.

1.8.3: Laminin polymerises through interacting LN domains

Laminin polymerisation into a stable network is an essential prerequisite for BM assembly (Miner 2008). This process is restricted to Laminin molecules that contain three full-length chains and three N-terminal LN domains (domain VI) (Cohen et al. 1997; Colognato and Yurchenco 2000; Tzu and Marinkovich 2008). This is known as the three-arm interaction model (Yurchenco et al. 1985; Yurchenco and Cheng 1993). The LE domains (domains III and V) may also play an important part in polymerisation, by acting as a spacer element between the LN domains, allowing for long distance interactions between molecules within the ECM (Engel 1989). Data suggest a possible role for these LE domains in Laminin stabilisation (McKee et al. 2009). This is supported by evidence that LE domains IV and VIII of the γ1 chain can bind to Nidogen-1 (Poschl et al. 1996; Stetefeld et al. 1996). Homologous self-interactions of the Laminin chains are very weak or non-existent (Yurchenco and Cheng 1993), although others have reported some homotypic α–α interactions (Odenthal et al. 2004). Polymerisation of the Laminin network is affected by a temperature-dependent oligomerisation step, and a calcium-dependent polymerisation step (Schittny and Yurchenco 1990). In vitro, Laminin polymerisation is triggered by an acidic surface. In vivo, an acidic microenvironment is present, provided by the sialic acid and sulphate groups of the ECM components surrounding cells (Freire and Coelho-Sampaio 2000).

The BM and Laminin network formed within a tissue is therefore dependent upon a whole host of factors, including the expression and availability of individual Laminin subunits, the expression of receptors and other ECM molecules capable of binding to and stabilising Laminin, and a permissive microenvironment that allows the polymerisation of Laminin and Collagen type IV networks. These are all tightly regulated mechanisms, and are poorly understood. For example, it is unknown why Laminin α5 subunit is able to partially compensate for the loss of Laminin α1 in Reichert's basement membrane, but not in the embryonic basement membrane (Miner et al. 2004). Both are full-length Laminin chains, and contain C-terminal LG subdomains 1-5. However, differences in the binding affinities to different Laminin receptors are known to exist between α subunits, with respect to their LG subdomains, and their N-terminal LN domains. A temporal differential expression of Laminin receptors in a tissue during development therefore encourages the transition of one Laminin polymer type to another, a process that is key to normal development and tissue morphogenesis. Understanding what regulates the expression of Laminin genes is therefore essential to understand embryonic and post-natal ECM structure, BM composition, and tissue and organ morphogenesis during normal and diseased conditions.

1.8.4: Laminin is necessary for basement membrane assembly

The first forming basement membranes in the mouse embryo are assembled between the visceral endoderm and the developing epiblast (embryonic BM), and also underneath the parietal membrane (Reichert's BM) (Leivo et al. 1980). Differentiation of epiblast cells and proamniotic cavitation require the presence of these BMs (Murray and Edgar 2001). BMs are also associated with follicle cells in the ovaries of *Oryzias latipes* (Medaka), and so play a role in teleost oocyte development (Tesoriero 1977).

In vitro and in vivo studies indicate that Laminins are essential for BM assembly, and that Laminin polymerisation into a network is a prerequisite for BM assembly (Yurchenco and Wadsworth 2004). In Laminin β1 and γ1 knockout mice, there is a failure to form the embryonic and Reichert's BMs, due to the absence of Laminin-111 and -511 (Smyth et al. 1999; Miner et al. 2004). This results in embryonic lethality at day E5.5, through an arrest of blastocyst differentiation (Smyth et al. 1999; Miner et al. 2004). Mutations in Laminin α1 also result in embryonic lethality, although at a slightly later stage of development (Miner et al. 2004). This is because Laminin α5 is able to partially compensate for the loss of Laminin α1 in the embryonic BM, although there is an absence of Reichert's BM (Miner et al. 2004).

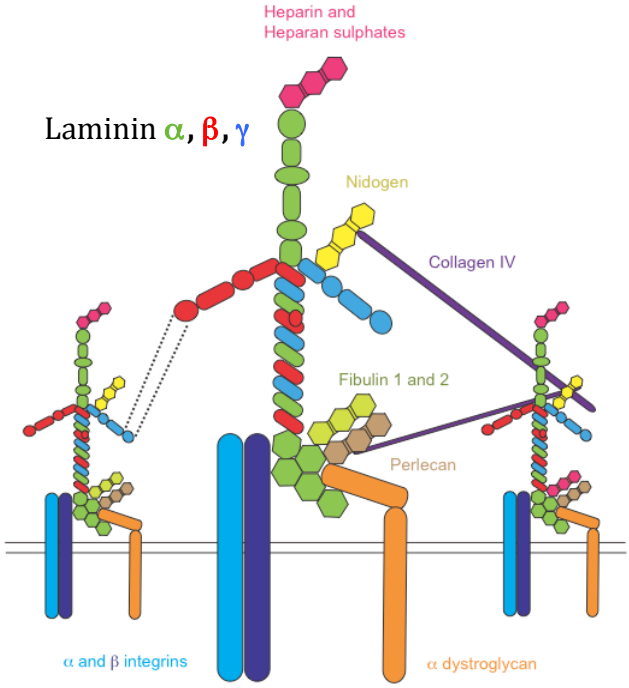
Mouse embryos that are mutant for Integrin β1 (*Itgb1*) and Dystroglycan (*Dag-1*) also fail to form the embryonic BM and Reichert's BM (Williamson et al. 1997; Aumailley et al. 2000), confirming that Laminin binding to Integrin and Dystroglycan is critical for BM formation. However, Li et al. (2002) have shown that Integrin and Dystroglycan are not required for the

initiation of BM assembly in embryoid bodies when Laminin is added. In support of this, embryoid bodies formed using Laminin γ1-null mouse ES cells do not assemble the embryonic BM (Murray and Edgar 2000; Li et al. 2002). The addition of Laminin-111 to the culture was sufficient to restore the embryonic BM (Li et al. 2002), confirming a direct requirement for Laminin in early embryonic BM formation. Blocking the Laminin LG4-5 subdomains also prevents the formation of the embryonic BM (Li et al. 2002), highlighting the importance of Laminin interaction with its binding partners for the assembly of BMs.

Nidogen or Perlecan do not appear essential for the initial formation of BMs,

(Mayer et al. 1998; Murshed et al. 2000; Dong et al. 2002; Schymeinsky et al. 2002; Yurchenco et al. 2004) as despite their widespread distribution within BMs, targeted mutations of Nidogen-1 or 2 (Kohfeldt et al. 1998; Salmivirta et al. 2002) or Perlecan (Arikawa-Hirasawa et al. 1999; Costell et al. 1999) in mice do not cause early lethality prior to implantation, as loss of Laminin γ1 causes (Smyth et al. 1999). However, loss of Nidogen-1 or 2 or Perlecan causes widespread disorganisation of tissue and organ structures including the vascular system, respiratory system, and skeletal and cartilage tissues, due to defects in BM stability which result in death by E11.5 (Arikawa-Hirasawa et al. 1999; Costell et al. 1999; Bader et al. 2005). Likewise, studies in *C.elegans* and mice have revealed that Collagen type IV is not required for initial BM assembly, but is needed later in development to enable BMs to withstand mechanical stress (Guo and Kramer 1989; Guo et al. 1991; Poschl et al. 2004).

Taken together, data from mouse mutant embryos and embryoid bodies suggest that Nidogen, Perlecan and Collagen type IV are not required until the later stages of development with regard to BM formation. This suggests that Laminins, secreted by primitive endodermal cells, are both necessary and sufficient to form the nascent BM, to which the remaining components are then incorporated (Yurchenco et al. 2004). Data now support a model in which Integrins and Dystroglycans anchor and cluster Laminin to the cell surface (*Figure 1.14*), facilitating its polymerisation and the formation of a nascent Laminin network (Henry et al. 2001; Yurchenco et al. 2004; Weir et al. 2006). The Laminin heterotrimers bind to each other via the N-terminal globular LN (VI) domains (*Figure 1.14*) (Yurchenco and Cheng 1993), and anchorage is further enhanced by the binding of the α chain LN (VI) domain to sulphated glycolipids and Integrin α1β1 and Integrin α3β1 (Yurchenco et al. 2004; Yurchenco and Patton 2009). Nidogens, Perlecan, and Agrin are then incorporated into the forming BM, and provide stabilisation between the independently assembled networks of Laminin and Collagen type IV (*Figure 1.14*) (Yurchenco et al. 2004). These additional BM components also bind to β1 Integrins, and also enable the tethering of tissue-specific growth factors (Friedrich et al. 1999; Yurchenco and Patton 2009).



Adapted from Yurchenco and Patton (2009).

Figure 1.14: *A schematic model of Laminin, and its interacting partners in the basement membrane*. Laminin trimers and Collagen type IV make up the bulk of the basement membrane, along with other bridging molecules including Nidogen, Perlecan, Fibulin and other glycoproteins such as Heparan sulphates. The Collagen network is connected to the Laminin network through Nidogen, which binds to domain III of the Laminin γ chain. Laminins bind to cellular membrane-associated Integrins and Dystroglycans via the LG domains of the Laminin α subunit. The LG domains can also bind to Perlecan, Nidogen, Heparin, and Fibulin -1 and -2.

1.8.5: The distribution and function of Laminins

Laminins are usually restricted to basement membranes separating epithelia and endothelia from surrounding connective tissues, although it is often difficult to be sure of the exact Laminin trimer present (*Table 1.4*). This is because in situ hybridisation and immunohistological experiments reveal only the location of individual Laminin subunits (*Table 1.3*), and not the location of the Laminin heterotrimeric molecules (Miner et al. 1997). The temporal and spatial expression of the Laminin subunits is also very dynamic, making it harder to predict exactly where a Laminin heterotrimer will be present at a given time point (Tzu and Marinkovich 2008). There have also been problems regarding the specificity of different Laminin chain antibodies (Scheele et al. 2006).

In general Laminin α1, α3 and α5 subunits are found in BMs supporting epithelial structures, whilst α2 and α4 are often found in endothelial or mesenchymal tissues (*Table 1.3*) (reviewed in Tunggal et al. (2000). In comparison, the β and γ Laminin subunits show a less restricted expression pattern during vertebrate embryonic development, and are mostly expressed ubiquitously (*Table 1.3*) (Tunggal et al. 2000).

Laminin subunit	Laminins	Non-muscle expression	Null phenotype
α1	Lm-111, 121	M: embryonic BM, Reichert's BM, myotomal BM, kidney, eye, brain, NT, vasculure. ZF: NC, OV, brain, eye, NT, Pf	M: Loss of Reichert's BM and pre-implantation death. ZF: abnormal muscle development, defective NC differentiation, brain and & defects
α2	Lm-211, 221, 213, 212/222	M, ZF & H: brain, eye, peripheral nerves, capillaries. ZF OV, Pf	M & H: post natal muscular dystrophies. ZF: loss of muscle fibre attachment from somitic boundaries
α3	Lm-332, 311, 321, 3B32, 333	M & H: skin, bladder, lungs, oesophagus. ZF: skin, eye, brain, NT, Pf	M & H: epidermolysis bullosa
α4	Lm-411, 421, 423	M: vasculature, mesenchymal cells, lungs, adult mouse NMJ. ZF: NC BM, vasculature, OV, brain, Pf	M & ZF: haemorrhaging
α5	Lm-511, 521, 522, 523	M: embryonic BM, Reichert's BM, myotomal BM, neural tube, kidney, lungs, brain, nerves, adult mouse NMJ, ectoderm BM, vasculature. ZF: eye, brain, NT, Pf	M: death at E14-17, placental vasculature defects, exencephaly, defective kidney and limb development. ZF: delayed NC differentiation
β1	Lm-111, 211, 213, 212, 313, 411, 511	M: most tissues. ZF: eye, brain, NT, NC, OV, Pf	M: pre-implantation death. ZF: defective NC differentiation & shortened body axis, abnormal muscle development and MTJ fibre capture, eye defects
β2	Lm-121, 221, 222, 321, 421, 423, 521 522, 523	M & H: NMJ, glomerulus, eye. ZF: eye, OV, brain, NT, Pf	Defects in the NMJ, glomerulus and eye
β3	Lm-332, 3B32, 333	M & H: skin, epithelia. ZF: not identified	M & H: epidermolysis bullosa
β4	Unknown	M: not detected. ZF: eye, brain, NT, OV, Pf	
γ1	Lm-111, 121, 211, 221, 311, 321, 411, 421, 511 521	M: all BMs. ZF: eye, brain, NT, NC, OV, Pf	M: pre-implantation death. ZF: defective NC & shortened body axis. abnormal muscle development and MTJ fibre capture, eye defects
γ2	Lm-212/222, 332, 3B32, 522	M: skin, epithelia. H: lung, kidney, epithelia. ZF: not identified	M & H: epidermolysis bullosa
γ3	Lm-213, 333, 423, 523	M: brain, kidneys, skin, testes nerves. ZF: eye, OV, brain, NT, Pf	

Table 1.3: *Expression of Laminin subunits (excluding muscle expression) during development of mouse (M), zebrafish (ZF) or human (H), and the phenotypes caused by a null mutation in the Laminin subunit*. Laminin trimers formed from the specific Laminin subunit are indicated. Abbreviations: basement membrane (BM), notochord (NC), otic vesicle (OV), neural tube (NT), pectoral fin (Pf), myo-tendinous junction (MTJ).

1.8.5.1: Laminin α1 and α5 are expressed in the embryonic BM

Laminin α1 and Laminin α5 are the earliest Laminins to be expressed during mouse embryogenesis, prior to gastrulation (Ekblom et al. 2003; Miner et al. 2004). They are detected in the embryonic BM between the epiblast and visceral endoderm, as Laminin-111 and -511 trimers (Leivo et al. 1980). In *Lama5* (Laminin α5) null embryos, the embryonic BM still forms, suggesting that Laminin α1 can compensate for the absence of Laminin α5 (Miner et al. 2004). However, *Lama5* mutations cause defects in digit separation, vascularisation, lung lobe separation, and intestinal smooth muscle differentiation, in addition to abnormalities in neural tube closure and kidney formation described below. These defects correlate with the specific expression pattern of *Lama5* mRNA in these tissues (Miner et al. 1998). As a result, *Lama5*$^{-/-}$ mice die at E17.

1.8.5.2: Laminins have a role in CNS development

Laminins have roles in brain and central nervous system development and differentiation, in *Drosophila*, *Xenopus*, *C.elegans*, mice and humans (Garcia-Alonso et al. 1996; Lallier 1996; Chun et al. 2003), correlating with their expression in the brain vasculature, the choroid plexus, and the pial BM surrounding the central nervous system (Miner and Yurchenco 2004). Mutations in *Lama2*, encoding Laminin α2, are known to cause altered neuromuscular junction morphology and a reduced myelination in the brain and central nervous system (Sunada et al. 1995; Mercuri et al. 1996), which can lead to epilepsy in human (Miyagoe-Suzuki et al. 2000). This is likely a result of reduced interaction between Laminin-211 and Integrin α6β1, as this interaction in cultured oligodendrocytes has been shown to enhance myelination (Buttery and ffrench-Constant 1999). The most severe defects in brain development are observed in mice with mutations in *Lama5* (Miner et al. 1998). These mice exhibit exencephaly caused by a failure of the anterior neural tube to close.

1.8.5.3: Laminins are required for normal eye development

Correct formation and patterning of the eye is dependent upon Laminin expression in BMs such as the retinal BM (Halfter et al. 2005). Mutation in Laminin subunits α1, β1, β2 or γ1 all lead to ocular developmental phenotypes (Libby et al. 1999; Zenker et al. 2004; Gross et al. 2005; Semina et al. 2006). Mutation in human Laminin β2 results in a variety of ocular defects including microcoria, lenticonus, glaucoma, cataracts and microphthalmia (Zenker et al. 2004). Studies in mouse and human have also revealed that the Laminin β2 chain is essential for proper synapse formation, in the outer plexiform layer of the eye (Libby et al. 1999). Mutations in zebrafish Laminin β1 (*grumpy*) or γ1 (*sleepy*) also result in multiple eye abnormalities including

retinal blowout, disorganisation of optic nerves, lens hypoplasia and corneal defects (Neuhauss et al. 1999; Gross et al. 2005).

It is believed that the retinal BM is required to prevent ganglion cell apoptosis by anchoring radial cells to the retinal surface, bringing them into close contact with the ganglion cells. These radial cells are critical for synapse formation and ganglion cell survival, but the radial cells retract when the retinal BM is disrupted, through Laminin γ1 mutation (Halfter et al. 2005).

All three Laminin chains and Nidogen-1 and Collagen type IV are synthesised in the lens and ciliary body of the embryonic mouse eye, and not by the ganglion cells or retinal glia cells (Sarthy and Fu 1990; Dong and Chung 1991). Laminins are also expressed throughout the vasculature of the eye (Libby et al. 2000).

1.8.5.4: Laminins have a role in the maturation of the kidneys

A dynamic expression pattern of Laminins occurs during the differentiation of the kidney. Laminin α5, in the trimers -511 and -521 is detected in the glomerular BM, where it replaces Laminin-111 as development proceeds (Miner et al. 1998). A mutation in *Lama5* results in the breakdown of the glomerular BM, leading to disorganisation of glomerular cells and a failure of glomerular vascularisation (Miner and Li 2000). Laminin β2 expression replaces Laminin β1 (Noakes et al. 1995) in the glomerular mesangium of the maturing kidney. These BMs are essential for the proper filtration processes occurring in the glomerulus (Miner and Li 2000), perhaps by providing a stable structure to withstand the high hydrostatic pressures (Timpl 1996). Full-length Laminin α3B is also detected in the kidney, in addition to the lungs and brain of the mouse (Champliaud et al. 1996; Miner et al. 1997). An intense but transient expression of Laminin α4 is also associated with kidney tubule differentiation from nephrogenic mesenchyme (Iivanainen et al. 1997).

1.8.5.5: Laminins are crucial for vasculature BM stability

The endothelial BM of the vasculature in mice and zebrafish is largely composed of Laminin α4. It is strongly expressed in the mouse aorta and contributes to Laminin-411 and -421 (Frieser et al. 1997), and in the zebrafish dorsal aorta (Pollard et al. 2006). Mice that lack Laminin α4 have weakened capillary basement membranes leading to rupturing and widespread haemorrhaging during the embryonic and neonatal periods, highlighting the important structural role that Laminins in the BMs play during these embryonic stages (Thyboll et al. 2002). Laminin α5 is thought to compensate for the lack of Laminin α4 at later stages (Sorokin et al. 1997).

1.8.5.6: Laminins are essential for correct skin development

Unlike Laminin-111, Laminin-511 expression is often retained to adulthood (Ekblom et al. 1998; Sztal et al. 2011), and is strongly expressed in the ectodermal BM. Mice with a lack of Laminin α5 exhibit developmental defects in this BM (Miner et al. 1998), and have fewer hair follicles, possibly as a result of disrupted developmental signalling, as suggested by decreased expression levels of Shh and Gli1 (Li et al. 2003). Basement membranes of basal keratinocytes and hair follicles also contain Laminin α2 (Schuler and Sorokin 1995). Laminin β3 is found in the basement membrane of the skin, in the trimer -332 as part of a hemidesmosomal complex (Pulkkinen et al. 1994; Utani et al. 1997) but an ortholog in the zebrafish has not been identified (Sztal et al. 2011). A mutation in any of the subunits of Laminin-332 leads to a skin blistering condition in humans, known as junctional epidermolysis bullosa (Pulkkinen and Uitto 1999). Laminin α3A is also detected in the skin, as well as in the bladder, lungs and oesophagus as Laminin-3A32 and -3A11 (Miner et al. 1997).

1.8.5.7: Laminin β1 and γ1 subunits are required in most BMs

The Laminin β and γ subunits are the most widely expressed subunits. Laminin β1 is more widely spread compared to the other β subunits, and Laminin γ1 is found in all BMs within the mouse (Smyth et al. 1999). Unlike in the mouse, early zebrafish development does not require Laminin β1 or γ1 subunits and so the later effects of the loss of these two subunits can be observed (Parsons et al. 2002). Both subunits are maternally expressed in zebrafish embryos, and are more prominent along the midline by the tailbud stage. By the fifteen-somite stage, expression is abundant in the chordamesoderm and somites, and highly expressed in the caudal fin fold, gut, notochord and hypochord. Laminin γ1 is also expressed in the horizontal myoseptum of the somite (Parsons et al. 2002). Both Laminin β1 and γ1 mutations exhibit near identical phenotypes with a loss of BM surrounding the notochord, resulting in the failure of the notochord to differentiate. The failure to form a properly vacuolated notochord results in a shortened body axis. *Grumpy* and *sleepy* mutants also fail to form intersomitic blood vessels. There is no blood flow between the somites at later stages as a result (Parsons et al. 2002).

Overall, the existence of multiple Laminin subunits with specific and dynamic temporal and spatial regulation provides a mechanism by which BM composition is modulated in different organs during development. This is essential for normal development and tissue morphogenesis. Laminins are critical to the functioning of almost all tissues in the embryo and adult, as exemplified by Laminin mutations which affect the anterior CNS, kidney, eye, skin, and the vasculature. Laminins are also extremely important for BMs associated with skeletal muscles, at both embryonic and adults stages of development.

Laminin isoform	Main sites of expression in mouse	Receptors
-111	Developing epithelia, embryonic BM, Reichert's BM, myotomal BM, somites, neural tube, PSM, adult epithelium, kidney, liver, testis, ovary, brain blood vessels	Integrins α1β1, α2β1, α6β1, α6β4, α7β1, α9β1, α-Dystroglycan
-121	Myotendinous junction	
-211	Embryonic and adult muscle, heart, eye, peripheral nerves, Schwann cells, testes	Integrins α1β1, α2β1, α3β1, α6β1, α6β4, α7β1, α-Dystroglycan
-221	Muscle, heart, peripheral nerve, Schwann cells, neuromuscular junction, glomerulus	Integrins α1β1, α2β1, α3β1, α6β1, α6β4, α7β1, α-Dystroglycan
-213	Placenta, testes, ciliated epithelia	
-212/222	Peripheral nerve	
-3A11	Epidermis, amnion	
-3A21	Epidermis	
-3B23	Central nervous system, retina	
-3A32	Epidermis, placenta, mammary gland	Integrins α3β1, α6β4, α6β1
-3A33b	Testes	
-3B32	Skin, uterus, lung	
-411	Skeletal muscle, endothelium, smooth muscle, peripheral nerve, adipose tissue	Integrin α6β1
-421	Endothelium, smooth muscle, neuromuscular junction, glomerulus	
-423	Central nervous system, retina	
-511	Developing epithelia, embryonic BM, Reichert's BM, myotomal BM, somites, neural tube, mature epithelium, mature endothelium, smooth muscle, skeletal muscle	Integrin α3β1, α6β1
-521	Mature epithelium, mature endothelium, smooth muscle, neuromuscular junction, glomerulus	Integrin α3β1, α6β1
-522	Bone marrow	
-523	Central nervous system, retina	

Table 1.4**:** *The expression of Laminin isoforms during mouse embryonic and adult development, and the known associated Laminin receptors.*

Laminin subunit	Expression in skeletal muscle	Reference
α1	M & ZF: embryonic	(Yurchenco et al. 2004; Sztal et al. 2011)
α2	M & ZF: embryonic and adult	(Schuler and Sorokin 1995; Sztal et al. 2011)
α3	M & ZF: Not detected	
α4	M & ZF: embryonic. M: adult NMJ	(Miner et al. 1997; Sztal et al. 2011)
α5	M: embryonic. Adult NMJ. ZF: not detected	(Miner et al. 1997; Yurchenco et al. 2004)
β1	M & ZF: embryonic. M: adult	(Schuler and Sorokin 1995; Yurchenco et al. 2004; Sztal et al. 2011)
β2	M & ZF: embryonic. M: adult NMJ. ZF: adult	(Yurchenco et al. 2004; Sztal et al. 2011)
β3	M & ZF: not detected	
β4	M: not detected. ZF: embryonic	(Sztal et al. 2011)
γ1	M & ZF: embryonic and adult	(Schuler and Sorokin 1995; Yurchenco et al. 2004; Sztal et al. 2011)
γ2	M & ZF: not detected	
γ3	M & ZF: embryonic. M: adult	(Gersdorff et al. 2005; Sztal et al. 2011)

Table 1.5**:** *Expression of Laminin subunits within the embryonic and adult skeletal muscle, of the mouse (M) and zebrafish (ZF).*

1.8.5.8: A dynamic Laminin expression pattern is observed in muscle

A major site of Laminin expression is muscle, where Laminins are critical for myogenic progenitor cell differentiation, migration, elongation and myotome patterning during embryonic development, and for structural integrity in the adult muscle. Different Laminin heterotrimers play important functions in skeletal and smooth muscle during differing developmental stages, and as a result, the expression pattern of Laminin subunits is very dynamic in these tissues. During embryonic stages, Laminin α1, α2 and α4 are detected in the skeletal muscle of the mouse and zebrafish (Leivo and Engvall 1988; Liu and Mayne 1996; Anderson et al. 2009; Sztal et al. 2011). Laminin α2 and α4 are also expressed in the cardiac muscle of the mouse (Leivo and Engvall 1988; Lefebvre et al. 1999). Laminin α5 is expressed in skeletal muscle of the mouse, but not zebrafish embryos (Sorokin et al. 1997; Anderson et al. 2009; Sztal et al. 2011). Laminin α4 and Laminin α5 are also expressed in the smooth muscle, and in endothelium of embryonic mice (Sorokin et al. 1997; Lefebvre et al. 1999).

In E11.5 mouse intercostal muscles, Laminin α1, α2, α5, β1 and γ1 chains are the first Laminin chains to be expressed (Patton et al. 1997). At this time, Laminin α3, α4, β2, β3, β4, γ2 and γ3 chains are undetectable (Patton et al. 1997). The expression of Laminin α4 and β2 increases rapidly after E11.5 in the skeletal muscle and is expressed throughout by E15, whereas Laminin α1 is down-regulated (Patton et al. 1997). In addition to their expression in the skeletal muscle BMs, Laminin α2, α4, α5, β1 and γ1 chains are present at the neuro-muscular junctions by E15,

although Laminin α5 is expressed only weakly (Patton et al. 1997). Laminin β2 also appears at these synaptic sites (Hunter et al. 1989) in Laminin trimers -221, -421 and -521 (Miner 2008). After birth, Laminin α1, α4 and α5 are no longer detectable in the skeletal muscle BMs, although expression of Laminin α5 is up-regulated at the synaptic sites, whilst Laminin α4 and β2 continue to be specifically expressed at the synaptic sites (Patton et al. 1997; Sorokin et al. 1997). In contrast, Laminin β1 is lost from this region whereas Laminin α2 and γ1 remain highly expressed throughout both the skeletal muscle and the synaptic site regions throughout development.

1.8.5.8.1: Laminins in zebrafish muscle

In agreement with mouse data, Laminin α1 and α4 are strongly expressed in the muscles of the trunk in the zebrafish up to 24hpf and 72hpf, respectively, at which point they become strongly down-regulated (Sztal et al. 2011). Expression of both subunits is observed in the pectoral fins up to 72hpf (Sztal et al. 2011). In both mouse and zebrafish skeletal muscle BMs, Laminin α2 is expressed in the embryo and the adult, in the trimers Laminin-211 and -221 (Sasaki et al. 2002; Sztal et al. 2011). These are the main Laminin isoforms expressed in the muscle fibre BMs. Their importance for muscle integrity (Sunada et al. 1994) is illustrated by the fact that mutations in *Lama2* cause muscular dystrophies in mice (Xu et al. 1994; Huh et al. 2005), and congenital muscular dystrophy (CMD) in humans (Miner and Yurchenco 2004). *Candyfloss* mutant zebrafish, which carry a mutation in the globular Dystroglycan binding domain of Laminin α2 (Hall et al. 2007), also develop muscular dystrophy, providing an insight into the processes controlled by Laminins involved in the disease. However, *Itgα7* null mice also display the same dystrophic phenotype as *Lama2* null mice (Mayer et al. 1997), indicating that it is not only defective Dystroglycan binding that leads to muscular dystrophy. Early myogenesis and muscle innervation is normal in *candyfloss* zebrafish, suggesting that proliferation, fusion of myoblasts, and motor stimulation of muscle are not impaired and thus are likely to not be the cause of muscular dystrophy in CMD patients. However, mechanically induced fibre detachment from the myotendinous junction is observed resulting in fibre retraction and apoptosis (*Figure 1.15*) (Hall et al. 2007).

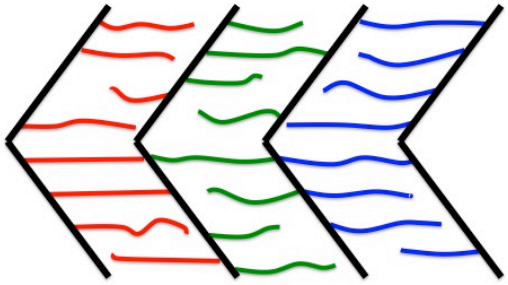


Figure 1.15: *A schematic representation of the somites of the candyfloss zebrafish.* Lateral view of the somite shows muscle fibres have retracted from the MTJ (black chevrons) and will eventually apoptose.

Laminin β1 contributes to the highly expressed Laminin heterotrimers -111, -211, -213 and -411 in embryonic mouse and fish muscle, and to Laminin -511 in embryonic mouse muscle BMs. Laminin β1 is also expressed in Laminin heterotrimers -211, -213, -411 and -511 in the BM of adult mouse muscle (*Tables 1.4* and *1.5*). Laminin β2 is expressed in the skeletal muscle BM of mice during only the embryonic stages, in the Laminin trimers -121, -221, -421, -521. Conversely, zebrafish do not express Laminin β1 at adulthood, but do express Laminin β2 throughout zebrafish development in trunk and pectoral fin muscle. Zebrafish with mutations in Laminin β2 (*softy*) reveal the important role that Laminin β2 plays in the structural integrity of muscles. *Softy* zebrafish also display dystrophic muscles, with a loss of muscle fibre adhesion at the myotendinous junctions (MTJ) (Jacoby et al. 2009). Interestingly, unlike other dystrophic mutants, the myotome of *softy* is able to partially recover due to the formation of vertical myoseptum-like structures that form within the middle of the somite (*Figure 1.16*), and are dependent on the availability of Laminin α2. They can provide anchorage and maybe even survival factors to retracted fibres, and some fibres can elongate across this myoseptum-like structure and correctly attach to the MTJ. This recovery process could have important implications for the therapy of CMD and other BM disorders (Jacoby et al. 2009).

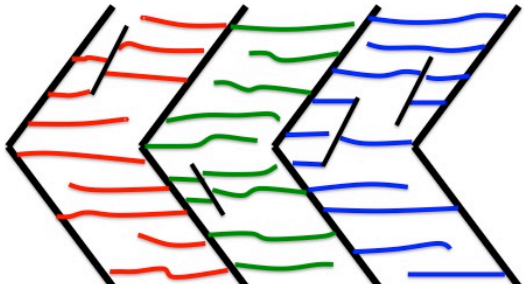


Figure 1.16: *A schematic representation of the somites of the softy zebrafish.* Lateral view of the somite shows muscle fibres have retracted from the MTJ (black chevrons) and will eventually apoptose. Some muscle fibres are able to re-attach to vertical myoseptum-like structures (black lines), which form within the somite of the *softy* mutant.

Laminin γ1 is strongly expressed in the muscle of mice and zebrafish (*Table 1.5*) (Smyth et al. 1999; Sztal et al. 2011). Not only does it contribute to a variety of different Laminin trimers (*Tables 1.3* and *1.4*), but it plays an essential role in the patterning of the zebrafish myotome (Dolez et al. 2011). Laminin γ1 has been shown to modulate the pattern of phosphorylated Smad1/5/8 (pSmad) expression via heparan sulphate proteoglycans, and to restrict pSmad distribution from the midline of the somite. This is critical for normal Engrailed expression in the muscle pioneer cells and medial fast fibres next to the notochord (Dolez et al. 2011). In the absence of Laminin γ1 in *sleepy* zebrafish, there is a loss of Engrailed expression in the central domain of the somite due to an expanded expression domain of pSmad. The expression pattern of Laminin γ1 therefore affects the specific muscle fibre type that is produced within the zebrafish somite. Laminin γ1 can also indirectly affect zebrafish myogenesis through regulation of muscle cell behaviour. In *sleepy* zebrafish, slow muscle migration and fast muscle fibre elongation and

differentiation are delayed (Peterson and Henry 2010; Dolez et al. 2011). Shh signalling is indirectly required for the recovery of fast muscle fibre elongation, most likely through the specification of slow muscle fibres (Henry and Amacher 2004), showing an interaction between Shh signalling and Laminin during early muscle development. Mutation in Laminin γ1 also causes alterations to normal muscle fibre orientation, and muscle fibres may extend past the MTJs following the recovery of muscle fibre elongation (Snow et al. 2008). These muscle defects contrast with the later dystrophic fibre detachment phenotypes observed in the Laminin α2 mutant (*candyfloss*) zebrafish. This highlights the discrete roles that different Laminin chains play during muscle development (Snow et al. 2008).

Taken together, the expression pattern of Laminins in the muscle and NMJs of mice and zebrafish is very dynamic. Laminin α1 and α2 are highly expressed during the earliest stages of muscle development in both species, whilst Laminin α2 dominates in the adult muscle. Unlike Laminin α2, the lack of Laminin α1 expression in the adult muscle means it is unable to maintain skeletal muscle integrity. However, Laminin α1 does have important functions in other tissues of the adult, discussed in more detail below.

1.8.5.9: The distribution and function of the Laminin α1 subunit

Lama1 is first expressed in the mouse primitive endoderm and parietal endoderm, for the synthesis and incorporation of Laminin α1 into the embryonic BM and Reichert's BM (Miner 2008). By E9.5 in the mouse, *Lama1* transcripts are detected in the anterior CNS, synthesised by meningeal cells. *Lama1* is also observed in the neural tube, the somites and the pre-somitic mesoderm, the forming kidney and the glomerulus (Miner et al. 2004; Anderson et al. 2009). Within somites, *Lama1* is synthesised in the sclerotome (Anderson et al. 2009). *Lama1* is strongly expressed in the forming eye at the site of retinal ganglion cell development, in the liver, and later in the adult reproductive organs (Ekblom et al. 2003; Miner et al. 2004). The zebrafish embryo also expresses *lama1* within the eye, anterior CNS, neural tube, somites and the pre-somitic mesoderm. The chicken also shows a similar *Lama1* expression pattern, with early expression detected in the neural plate, and later *Lama1* expression in the anterior CNS, dermomyotome of the somites, and the pro-nephric tubules. However, there is a lack of *Lama1* expression in the sclerotome in HH stage 10-11 embryos, although there may be transient expression of the *Lama1* gene here (Zagris et al. 2000).

Laminin-111, in addition to Laminin-511, is highly expressed around the somites in the dermomyotomal BM of the mouse (Bajanca et al. 2006; Anderson et al. 2009). This BM has an important role in maintaining the epithelial and proliferative nature of the dermomyotome (Ben-Yair and Kalcheim 2005; Bajanca et al. 2006; Cinnamon et al. 2006).

Within the somite, Laminin-111 and Laminin-511 are detected in the mouse myotomal BM separating the sclerotome and the myotome (Miner et al. 2004). The myotomal BM forms when epaxial myogenic progenitor cells (MPCs) from the dorsal medial lip of the dermomyotome enter the myotome, and upregulate the Laminin receptors Integrin α6β1 and Dystroglycan (Bajanca et al. 2004; Bajanca et al. 2006; Anderson et al. 2009). Integrin α6β1 expression in the MPCs promotes the deposition of a Laminin BM by binding to Laminin α1, providing a substrate for MPC and neural crest cell migration (Tosney, Dehnbostel et al. 1994). The myotomal BM is also required for the correct patterning of the myotome, and for muscle fibre elongation (Bajanca et al. 2006; Anderson et al. 2009; Thorsteinsdottir et al. 2011). Within elongating muscle fibres, Laminin α1 is mainly restricted to the ends of the myotubes, where it is speculated that Laminin α1 plays a role in the fusion of myoblasts with the myotubes (Patton et al. 1997).

Loss of Laminin α1 leads to the loss of the myotomal BM (Anderson et al. 2009). A defect in myotomal BM deposition causes myotomal cells to spread medially, and to ectopic ventral and dorsal positions (Tajbakhsh et al. 1996; Bajanca et al. 2006; Anderson et al. 2009).

Within the mouse CNS, Laminin α1, in the heterotrimer Laminin-111, is associated with the pial or meningeal BM that covers the outer surface of the CNS. Laminin α1 is also detected in the ependymal layer of the spinal cord (Miner et al. 2004). Interaction of Laminin-111 with Integrin α6β1 is required for neural tube closure in *Xenopus* (Lallier et al. 1996) and for normal cerebral cortex development in mice (Georges-Labouesse et al. 1998). However, these results are based on studies in which there is a lack of Integrin α6 expression, and so other signalling pathways and interactions will also be affected, such as the interaction of Integrin α6 with Laminin-511.

The interaction of Laminin-111 with Integrin α6β1 is also required for the process of rapid neurite outgrowth from developing retinal ganglion cells in the mouse eye (Cohen et al. 1986). This growth response caused by interaction with Laminin-111 does not occur during later stages of maturation, correlating with the down-regulation of Integrin α6β1 (de Curtis and Reichardt 1993). Expression of Laminin α1 is also reported in the murine lens and sclera (Falk et al. 1999; Libby et al. 2000). The embryonic lethality of *Lama1* null mice embryos prevented analysis of the role of Laminin α1 in the eye and the anterior CNS. However, conditional *Sox2-Cre*; *Lama1*$^{flox/flox}$ knock out mice can be used to explore the function of *Lama1* at later stages of development (Edwards et al. 2010). *Sox2* is expressed throughout the inner cell mass, epiblast and extraembryonic ectoderm (Hayashi et al. 2002). By E6.5, *Sox2-Cre*; *Lama1*$^{flox/flox}$ display complete Cre-mediated recombination and the loss of *Lama1* specifically within cells derived from the epiblast (Hayashi et al. 2002). These mice are viable and fertile, but display a range of eye and cerebellar defects (Edwards et al. 2010; Ichikawa-Tomikawa et al. 2012). In the eye, development of the retinal vasculature is disrupted, and the retinal BM does not form. Loss of this BM likely leads to the disorganisation and cell loss that occurs in the retinal ganglion cell layer by

4 weeks (Edwards et al. 2010). Laminin α1 is also essential for the development of granule cell precursors in the cerebellum, as loss of Laminin α1 causes a decrease in the amount of proliferation and migration in these cells (Edwards et al. 2010; Ichikawa-Tomikawa et al. 2012). Laminin α1 is also required for the formation of Bergmann glial processes, and for dendritic formation on Purkinje cells (Ichikawa-Tomikawa et al. 2012).

Taken together, the anterior CNS and neural tube, eye, somites, and the kidney are all tissues that express *Lama1* and require Laminin α1 for their correct morphogenesis, development and functioning. The requirement of Laminin α1 for early mammalian development prevents the characterisation of potential ocular and axonal defects in human carrying mutation in *Lama1*. However, zebrafish do not require Laminin α1 to complete gastrulation, and so the zebrafish is a very useful tool to study the effects caused by a loss of this protein.

1.8.5.9.1: Severe morphological defects are associated with a loss of Laminin α1 function in the zebrafish

Four mutant alleles for *lama1* have been characterised in the zebrafish. These mutants are known as *bashful*, and their allele names are bal^{m190}, bal^{a69}, bal^{ar1} and bal^{uw1}(Stemple et al. 1996; Vihtelic et al. 2001; Paulus and Halloran 2006; Pollard et al. 2006; Semina et al. 2006). All *bashful* larvae die by 12 days post-fertilisation (Semina et al. 2006). The bal^{m190} allele is a non-sense mutation causing a G to T transition at amino acid 890, and is fully penetrant and thought to be the strongest allele (Stemple et al. 1996; Pollard et al. 2006). The bal^{a69} allele is a T-to-A point mutation which introduces a serine residue instead of a cysteine, at position 56 of the Laminin α1 protein (Semina et al. 2006). This cysteine is likely to be involved in disulphide bridge formation and Laminin trimer assembly (Kumagai et al. 1997). The bal^{ar1} mutant (Vihtelic et al. 2001) causes a truncation at amino acid 99 of Laminin α1, whilst the bal^{uw1} allele (Paulus and Halloran 2006) has a splicing mutation, which leads to the insertion of 100 bases, resulting in a protein truncation at amino acid 1424 (Semina et al. 2006). All *bashful* zebrafish exhibit a similar phenotype to that of the *grumpy* (Laminin β1) and *sleepy* (Laminin γ1) zebrafish (Pollard et al. 2006; Semina et al. 2006). Specifically, loss of Laminin α1, β1 or γ1 impairs the production of Laminin-111 and the formation of the notochordal BM, causing failure of notochord cells to differentiate and their death by apoptosis (Parsons et al. 2002; Pollard et al. 2006). This contributes to the shortened body axis observed in the *bashful*, *grumpy* and *sleepy* zebrafish (Parsons et al. 2002; Pollard et al. 2006). However, in contrast to *grumpy* and *sleepy* zebrafish in which the entire notochord fails to differentiate, only the anterior notochord fails to differentiate in *bashful* zebrafish. This indicates that the *bashful* mutant is a weaker mutant that *grumpy* or *sleepy* mutant zebrafish. Likewise, intersegmental blood vessels fail to form in *grumpy* and *sleepy* zebrafish, but this process is unaffected in the *bashful* mutant (Pollard et al. 2006).

Thus, as already observed in embryoid bodies formed from mouse Laminin mutant ES cells or in early mouse development, loss of Laminin α1 does not cause a phenotype as severe as loss of Laminin β1 of γ1. This is likely due to compensatory function from other Laminin α chains. An illustration of this is the fact that *lama4* morpholino-mediated knock down in *bashful* zebrafish results in defects in notochord differentiation and intersegmental vessel formation as severe as those observed in *grumpy* and *sleepy* zebrafish (Pollard et al. 2006). Thus, Laminin α1 and α4 subunits are required for the correct migration of endothelial cells between somites and can compensate for each other (Pollard et al. 2006).

Bashful zebrafish display a complex ocular phenotype, also observed in mice carrying a mutation in *lama1*, with abnormalities in the development of the lens, cornea, retina, and vasculature (Semina et al. 2006; Zinkevich et al. 2006). *Bashful* zebrafish eyes can be distinguished from wild-type eyes at 48hpf by the presence of an irregularly shaped pupil and a degenerating lens (Semina et al. 2006). Down-regulation of focal adhesions, causing defective cellular migration, in the lens epithelium and cornea in *bashful* zebrafish is likely to be at the origin of this phenotype (Semina et al. 2006). Phosphorylated focal adhesion kinases are normally clustered together following Integrin activation by Laminin binding (Wozniak et al. 2004). Laminin α1 absence also results in multiple axonal projection defects in the eye and a disorganisation of photoreceptor cells (Semina et al. 2006).

In addition to the retinal ganglion cell axon projection defects in the eye, *bashful* zebrafish exhibit a range of axonal pathfinding defects throughout the CNS (Paulus and Halloran 2006). Anterior commissures that connect telencephalic neurons fail to cross the brain in up to 80% of *bashful* zebrafish, despite the normal expression pattern of other axon guidance molecules such as Netrin1a, Ephrin, Slit, Robo 1-3 and Semaphorin (Paulus and Halloran 2006). Defasciculation of axons in the brain also occurs, and many branchiomotor neurons fail to migrate to their correct positions. Other Laminins such as Laminin-411 may be able to compensate and guide some of the normally projecting axons (Paulus and Halloran 2006). In the spinal cord, there is excessive branching of motor neuron axons, although axons in the peripheral nervous system are mostly normal (Paulus and Halloran 2006). Several in vitro studies have revealed the role of Laminin-111 as a permissive growth substrate, and as a molecule capable of providing directional information to growing axons (Patton et al. 1997; Kuhn et al. 1998). It is likely therefore that the axonal defects observed in the *bashful* zebrafish are due to a loss of Laminin α1 containing-permissive substrate which would allow the projection and guidance of axons, or a loss of adhesive substrate which would normally function to hold axonal tracts together.

A detailed analysis of muscle development has not been performed in *bashful* zebrafish, although several markers of early somite patterning including *gli1*, *gli2*, *myoD* and slow myosin heavy chain appear normal in this mutant at 24hpf (Paulus and Halloran 2006). However, a slight defect

in muscle fibre orientation is reported in *bashful* zebrafish (Paulus and Halloran 2006), suggesting that *bashful* zebrafish display a similar, but weaker, muscular phenotype to that observed in *grumpy* and *sleepy* zebrafish (Snow et al. 2008; Peterson and Henry 2010). *Grumpy* and *sleepy* zebrafish display disorganisation of muscle fibres and a delay in fast muscle cell elongation. This is followed by the failure of elongating muscle fibres to stop at the MTJs, resulting in fibre extension into the adjacent somite (*Figure 1.9*) (Snow et al. 2008; Peterson and Henry 2010). This may occur in *bashful* zebrafish to some degree, although it has not been reported. Recently, a significant amount of muscle fibre detachment from the MTJ has been observed in *lama1* morphant zebrafish embryos (Sztal et al. 2012). This phenotype is similar to that observed in *candyfloss* (*lama2*$^{-/-}$) zebrafish, suggesting that these two Laminin subunits have a similar function. Indeed, loss of both *lama1* and *lama2* in the zebrafish causes an increased severity of muscle fibre detachment (Sztal et al. 2012).

Thus, the expression pattern of *lama1* in the zebrafish is conserved, with the main sites of expression being in the anterior CNS and neural tube, the eye and the somites. Loss of Laminin α1 subunit causes defects in each of these organs during development often caused by the lack of BM formation. The conserved expression of *lama1* in the zebrafish indicates that conserved mechanisms operate to control the expression of this gene.

1.8.6: Several signalling pathways may be associated with the regulation of *Lama1*

The *lama1* locus is conserved between zebrafish, mouse, human and the chicken, showing synteny with the genes *lrrc30*, *ptprm* and *ARHGAP28* (*Figure 1.17*). In each species, these two adjacent genes are transcribed in the opposite direction to *lama1*. The close proximity of *Lrrc30* to the transcriptional start site of *lama1*, particularly in the zebrafish, could mean that the regulatory elements controlling *lama1* expression interfere with the regulation of *lrrc30*, or vice-versa.

Figure 1.17: *A comparison of the lama1 loci in the zebrafish, mouse, human and chicken.* Screen shots taken from Ensembl.org show a conserved synteny at the *lama1* locus in each of these species. Zebrafish *lama1* is located on chromosome 24 at position 43,450,377 - 43,542,737, and *Lrrc30* is located just 10,096 bases upstream of the *lama1* start site.

Before searching for DNA sequences which may have the potential to bind transcription factors and activate *lama1* transcription, it is useful to understand which signalling pathways are known to alter the expression levels of *lama1*. The downstream mediators of these signalling pathways could have the ability to bind to *lama1* enhancer sequences.

1.8.6.1: *Sonic hedgehog* is required for *Lama1* expression in the sclerotome and neural tube of mice

In vivo studies have shown that Shh signalling is required for *Lama1* expression and the formation of the myotomal BM in the mouse (Anderson et al. 2009). Indeed, expression of *Lama1* in the sclerotome and in the neural tube is lost in E9.5 $Shh^{-/-}$ embryos (Anderson et al. 2009). In *Shh:Gli3* double mutant mice embryos, a progressive restoration of myotomal BM is observed, and this correlates with a delayed revovery of *Lama1* expression (Anderson et al. 2009). This suggests that the expression of *Lama1* is more sensitive to the absence of Gli-activator (Gli2) rather than the presence of Gli-repressor. Currently, there is no evidence that Shh directly regulates *Lama1* expression, although an indirect role for Shh in the regulation of myotomal BM assembly has already been reported. Indeed, Shh (and Wnt) signalling induces *Myf5* expression in the MPCs of the epaxial lip of the mouse dermomyotome (Borycki et al. 1999). Myf5 is also required for Integrin $\alpha6\beta1$ expression in epaxial MPCs, and Integrin $\alpha6\beta1$ is essential for Laminin deposition into a myotomal BM (Bajanca et al. 2006). Consistent with this, mutation in *Myf5* also leads to a disruption in myotomal BM formation (Tajbakhsh et al. 1996).
Taken together, *Lama1* expression is required for the assembly of the myotomal BM membrane in E9.5 mice. Shh signalling regulates the expression of *Lama1* in the sclerotome and neural tube, although it is currently unknown whether this regulation is direct or indirect.

1.8.6.2: *Dmrt2* may have a role in Laminin $\alpha1$ synthesis

Other mutations in the mouse that cause myotomal BM disruption have given an insight into potential genes that may regulate *Lama1* expression. Both *Paraxis* (Wilson-Rawls et al. 1999) and *Dmrt2* (Seo et al. 2006) mutant embryos show a similar myotomal BM and myotome phenotype to that observed in $Shh^{-/-}$ embryos (Anderson et al. 2009). In $Paraxis^{-/-}$ embryos, *Lama1* mRNA is still expressed, but organisation of the protein is disrupted (Wilson-Rawls et al. 1999). This causes a disorganisation of myoblasts in the medial myotome, whilst myoblasts do not appear in the lateral myotome until after E10.5. In constrast to the altered Laminin protein localisation observed in $Paraxis^{-/-}$ embryos, $Dmrt2^{-/-}$ embryos show a loss of Laminin $\alpha1$ expression in the dermomyotome at E10.5 and E11.5 (Seo et al. 2006). $Dmrt2^{-/-}$ is associated with defective somite morphogenesis and death soon after birth, caused by abnormal rib and sternal development leading to the inability to breathe (Seo et al. 2006). These mutant studies reveal the

importance of Laminin-containing BMs for correct myogenesis and myocyte migration. However, it remains possible that altered Laminin expression could be the result of altered somite morphogenesis. Nevertheless, results suggest that *Dmrt2*, but not *Paraxis*, may have a role in Laminin α1 production.

1.8.6.3: Integrin β1 is required for Laminin α1 expression in embryoid bodies

The regulation of ECM components and *Lama1* expression has been explored in various embryoid body and cell line studies (Aumailley et al. 2000; Li et al. 2002). In wild-type embryoid bodies, Laminin-111 accumulates extra-cellularly, but this does not occur in embryoid bodies derived from Integrin β1-null embryonic stem cells (Aumailley et al. 2000). In these mutant embryoid bodies, the expression of Laminin α1 is down-regulated, and this causes the failure of the secretion of the Laminin β1 and γ1 subunits, despite their normal synthesis (Aumailley et al. 2000). This results in failure of early BM formation (see BM assembly section 1.8.4). Laminin α5 is known to be sufficient to allow embryonic BM formation in the absence of Laminin α1, yet the absence of early BM formation in Integrin β1-null embryoid bodies indicates a loss of Laminin α5, in addition to loss of Laminin α1 (Li et al. 2002; Miner et al. 2004). Overall, data suggest that Integrins operate a feedback regulation on ECM component synthesis and BM assembly (Aumailley et al. 2000; Ekblom et al. 2003).

1.8.6.4: Retinoic acid up-regulates *Lama1* expression in differentiating F9 cells

The role of Retinoic acid (RA) in regulating Laminin-111 expression was first explored in F9 embryonal carcinoma cells (Wang et al. 1985). RA-treated F9 cells differentiate into primitive endoderm and parietal endoderm-like cells (Wang et al. 1985). RA treatment revealed a simultaneous increase in *Lama1*, *Lamb1* and *Lamc1* mRNA expression (Kleinman et al. 1987), suggesting a co-ordinated control of the expression of these Laminin chains by RA in F9 cells.

To identify putative genes controlling *Lama1* expression in parietal endoderm differentiation, F9 embryonal carcinoma cells were treated with RA and dibutyryl cyclic AMP to trigger their differentiation into parietal endoderm-like cells (Futaki et al. 2004). Treatment resulted in the up-regulation of *Gata-4* and *Gata-6*, and the Sry-box family proteins, *Sox7* and *Sox17*, amongst various other transcription factors (Futaki et al. 2004). Using siRNA expression vectors to target these genes, *Gata-4*, *Gata-6* and *Sox7* were found to be critical for parietal endoderm differentiation and *Lama1* expression. Results indicate that *Sox7* is essential for *Gata-4* and *Gata-6* induction, and that *Gata-4* and *Gata-6* have a redundant function in the expression of *Lama1* during parietal differentiation of F9 cells (Futaki et al. 2004). Enhancer sites capable of binding *Sox7* and *Sox17* have been shown to activate the expression of *Lama1* in F9 cells, indicating that

Sox7 may regulate *Lama1* expression by direct and indirect mechanisms (see section 1.8.7) (Niimi et al. 2003).

1.8.6.5: FGF signalling is required for *Lama1* expression in embryoid bodies

FGF signalling has also been implicated in the regulation of Laminin-111, based on studies using embryonic stem (ES) cell-derived embryoid bodies carrying a dominant negative mutation in FGFR2 (Li et al. 2001). These mutants fail to produce a BM between the endoderm and ectoderm. This is likely due to the specific absence of Laminin α1, β1 and Collagen type IV at both the mRNA and protein level (Li et al. 2001). A downstream effector of FGF signalling, Akt/PKB, has been reported to be defective in the FGFR2 embryoid bodies described above (Chen et al. 2000). Combined with data that reveals up-regulation of Laminin-111 and Collagen type IV in a constitutively active Akt/PKB C2 myoblast cell line (Li et al. 2001), it has been suggested that FGF signalling, through Akt/PKB, controls the expression of Laminin α1 (Li et al. 2001). However, Akt is downstream of many signalling pathways, including the Shh pathway where it regulates PKA-mediated Gli inactivation (Riobo et al. 2006). Akt is also activated by the Dystrophin Glycoprotein complex (DGC) (Zhou et al. 2012). Thus, FGF signalling is required for the synthesis of BM components and for the differentiation of the ectoderm (Li et al. 2001), and loss of all FGF signalling in embryoid bodies leads to down-regulation of *Lama1* transcription. However, it is unknown whether FGF is directly responsible for *Lama1* transcription through Akt signalling. Constitutively active Akt/PKB could in fact be up-regulating Gli-mediated gene transcription, although it remains to be established whether Shh directly regulates *Lama1* expression through Gli (section 1.8.6.1).

1.8.6.6: The effect of TGF-β signalling on Laminin-111 synthesis is cell type dependent

TGF-β signalling is known to regulate various genes encoding ECM components involved in cell adhesion, migration and proliferation in normal and pathological conditions, such as *Lamc2* in human colon carcinoma cells (Roberts et al. 1992; Olsen et al. 2003). However, the effect of TGF-β treatment on Laminin gene expression is not always consistent and appears to be context-dependent. For example, TGF-β1 stimulates the expression of Laminin-111 subunits in rat liver sinusoidal endothelial cells (Neubauer et al. 1999), whilst no effect is observed in human skin fibroblast cells (Kahari et al. 1991; Lankat-Buttgereit et al. 1991). In fact, TGF-β1 treatment decreases the expression of Laminin-111 genes in human pulp fibroblastic cell cultures (Shiba et al. 1998). TGF-β therefore has a role in the regulation of multiple ECM components including the subunits of Laminin-111, although whether *Lama1* expression is up-regulated or down-regulated depends on the cell type receiving TGF-β signalling.

Overall, in vivo mouse studies and embryoid cell studies have revealed that the expression of *Lama1* is controlled by a wide range of signalling pathways including Shh, RA, FGFs and TGF-β, which may or may not converge together to regulate the output of an individual signalling pathway. Alternatively, the downstream transcription factors of these signalling pathways may act synergistically or in combination to regulate the expression of the *Lama1* gene. With the exception of *Sox7* and *Sox17*, the data described above does not suggest direct regulation of *Lama1*, and Laminin-111 expression. For example, alterations in Laminin-111 expression levels could be due to changes in tissue morphogenesis and differentiation, which may cause an increase or decrease in the number of cells synthesising Laminin-111. Treatment of embryoid bodies or cell lines with specific factors and proteins may also have an indirect effect by acting on other genes or ECM components, which then affect the expression of Laminin genes via feedback mechanisms, as alluded to previously (Aumailley et al. 2000; Ekblom et al. 2003). To identify the signalling pathways and signalling molecules that directly regulate *Lama1* expression, the promoter and enhancers responsible for the control of *Lama1* gene expression need to be identified. Some of the transcription factors that bind to the *Lama1* gene are likely to be downstream of the signalling pathways discussed above.

1.8.7: Conserved mechanisms control *Lama1* transcription

Although most of the mechanisms controlling *Lama1* transcription remain obscure, a few regulatory sequences within the mouse *Lama1* locus have now been identified (*Figure 1.18*) (Niimi et al. 2003; Niimi et al. 2004; Piccinni et al. 2004). Interestingly, there appears to be common transcriptional mechanisms controlling several Laminin genes. Like *Lamc1*, the *Lama1* promoter in both human and mouse does not contain a TATA box (Niimi et al. 2003; Niimi et al. 2004; Piccinni et al. 2004). A minimal promoter fragment that drives *Lama1* expression in F9 cells is located within -103 to -178bp upstream of the *Lama1* transcriptional start site (Niimi et al. 2003). In the mouse Caco2-TC7 intestinal epithelial cell line, an Sp1 binding GC rich box at position -73bp is also crucial for basal promoter activity for *Lama1* expression (Piccinni et al. 2004). This promoter region contains a Krüppel-like element (-97bp), to which KLF4 and KLF5 can bind as shown by EMSA experiments (Piccinni et al. 2004). Sp1/3 and KLF4/5 are expressed endogenously in the Caco2-TC7 cell line, and by using luciferase assays following over-expression techniques, it was shown that the Sp1/3 transcription factors and KLF5 activate *Lama1* expression, whilst KLF4 inhibits *Lama1* expression (Piccinni et al. 2004). This inhibition is likely due to KLF4 competing for both KLF5 and Sp1/3 binding sites. However, in a *Drosophila* S2 cell line, which does not express endogenous Sp1 factors (Du et al. 1998), Sp1/3 and KLF4 and KLF5 increase *Lama1* expression in a dose dependent manner, when expressed exogenously (Piccinni et al. 2004). Within human JAR choriocarcinoma cells, Sp1/3 and KLF4 and KLF6 bind to a minimal promoter fragment that is -206 to +31 of the *Lama1* transcriptional

start site, and these factors are required for *Lama1* expression (Niimi et al. 2006). Data suggests that Sp1/3 and KLF transcription factors regulate *Lama1* promoter activity, but their effects are different depending on the cellular context.

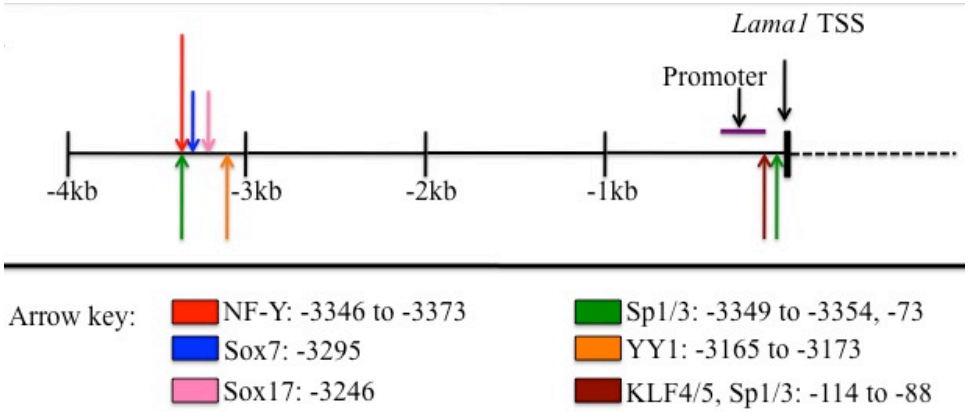


Figure 1.18: *A schematic representation of the mouse Lama1 gene, and the identified promoter and enhancer sites that control Lama1 expression in Caco2-TC7 cells, and differentiating F9 cells.* A minimal promoter fragment -103 to -178bp upstream of the transcription start site (TSS) (horizontal purple line) drives *Lama1* expression in F9 cells, whilst KLF4/5 and Sp1 transcription factors (-114 to -73bp) bind to the *Lama1* promoter in Caco2-TC7 cells. The enhancers responsible for *Lama1* expression during parietal endoderm differentiation in F9 cells (-3516 to -3082) bind to Sox7, Sox17, NF-Y, YY1 and Sp1/3 transcription factors.

Further upstream from the *Lama1* promoter sites, a 435-base parietal endoderm-specific enhancer situated at -3516 to -3082bp and capable of binding the Sp1/3 transcription factors has been identified (Briggs et al. 1986; Lietard et al. 1997; O'Neill et al. 1997; Niimi et al. 2003). As Sp1/3 transcription factors are ubiquitously expressed, their binding to the *Lama1* enhancer is not sufficient to drive *Lama1* expression in the parietal endoderm alone. Two Sox-binding sites have been identified within the *Lama1* enhancer at positions -3295 and -3246, respectively. They act synergistically to allow trans-activation by Sox7 and Sox17 respectively, in undifferentiated F9 cells (Niimi et al. 2003; Niimi et al. 2004). Sox mediated transcription in F9 cells is dependent on the presence of intact Sp1/Sp3 sites, as demonstrated by mutation analyses and luciferase recordings (Niimi et al. 2004). Consistent with these findings, microarray analyses have revealed an up-regulation of *Sox7* and *Sox17* during F9 cell differentiation into parietal endoderm-like cells (Futaki et al. 2003). Knock-down of *Sox7* in F9 cells causes suppression of Laminin-α1 synthesis and BM formation, causing an inhibition of parietal endoderm differentiation (Futaki et al. 2004). Perhaps unsurprisingly, mice carrying mutations in *Sox17* lack a definitive endoderm in the embryonic gut, although no early developmental abnormalities are observed in the extra-embryonic endoderm (Kanai-Azuma et al. 2002). This may be due to compensation from *Sox7*. Indeed, results suggest that Sox7 is a stronger activator of *Lama1* expression in F9 cells than

Sox17 (Niimi et al. 2004). The expression of Laminin α1 or the formation of BMs in *Sox17* null mice has not been analysed.

NF-Y and YY1 transcription factors are also associated with the transcriptional activation of *Lama1*, and have been found to bind to the 435-base parietal endoderm specific enhancer characterised in F9 cells (Niimi et al. 2003). Mutation of the NF-Y binding site causes a significant decrease in *Lama1* activation, whilst YY1 does not appear to play a major role (Niimi et al. 2003). Interestingly, NF-Y is a CCAAT-binding factor, and the CCAAT enhancer-binding protein β (C/EBPβ) has recently been associated with *Lamc1* transcription (Ramathal et al. 2011). Results indicate therefore that *Lama1* transcription in F9 cells is regulated by a combination of ubiquitously expressed transcription factors (Sp1/3 and NF-Y) and the parietal endoderm specific factors Sox7 and Sox17.

Overall, conserved mechanisms operate to regulate different Laminin genes, and in particular, Sp1/3 proteins bind to the promoters and enhancers of *Lama1* and *Lamc1* genes, and KLF proteins and C/EBPβ bind to their promoters (Lietard et al. 1997; Higaki et al. 2002). It is likely that these proteins could also be associated with the transcription of *lama1* in the zebrafish.

1.8.8: *Lama1* as a therapeutic for congenital muscular dystrophy

Laminin-211 is the main Laminin isoform in adult skeletal muscles, and mice with mutations in *Lama2* are models for human congenital muscular dystrophy (section 1.8.5.8.1). Recent studies have shown a marked improvement of the dystrophic phenotype when exogenous Laminin-111 is provided to several models of congenital muscular dystrophy, including *Lama2* and *Integrin α7* knock-out mice (Gawlik et al. 2006; Gawlik et al. 2010), but not in a model of Duchenne muscular dystrophy, the *mdx* mouse (Gawlik et al. 2011). The improvement is likely caused by the addition of Laminin α1, as Laminin β1 and γ1 are already present in adult mice muscle. This indicates that *Lama1* replacement therapy could be an important therapeutic tool to alleviate symptoms in this group of incurable diseases (Gawlik et al. 2010). Thus, understanding the mechanisms regulating *Lama1* expression may have invaluable implications for future therapies involving the re-expression of this embryonic form. For example, determination of the enhancer region(s) that control *Lama1* expression within the skeletal muscle should provide insight into the transcription factors and signalling pathways that activate *Lama1* in muscles. Manipulation of identified pathways could cause the re-expression of *Lama1* in the diseased muscle, and could partially rescue the degenerating muscle phenotype.

1.9: Conclusions and aims of the thesis

The Shh signalling pathway is one of the key signalling pathways responsible for tissue patterning and morphogenesis in both the embryo and adult, playing a role in morphogen-mediated patterning, cell proliferation, cell survival, and cell polarity. It does so through the regulation of a wide range of Hh target genes, many of which are unknown.

It has been demonstrated that Shh signalling controls the expression of *Lama1*, encoding Laminin α1, in the somites and neural tube of the mouse embryo. However, it is unknown if Shh signalling directly regulates the expression of *Lama1* in the mouse embryo.

Laminin α1 is required for the formation of the myotomal basement membrane in the mouse, and in its absence in the *Shh* null mutant, myogenic progenitor cells migrate to aberrant positions within the somite and fail to differentiate (Anderson et al. 2009). A conditional loss of *Lama1* in the mouse is also associated with defective eye and cerebellum formation (Edwards et al. 2010; Ichikawa-Tomikawa et al. 2012), highlighting the importance of this Laminin subunit. Interestingly, it has recently been shown that exogenous application of Laminin α1 can alleviate the dystrophic muscle phenotype caused by loss of Laminin α2 in mouse models of muscular dystrophy (Gawlik et al. 2006; Gawlik et al. 2010). A role for Laminin α1 has now been demonstrated for maintaining muscle fibre attachments to the myotendinous junctions in the zebrafish (Sztal et al. 2012), indicating that Laminin α1 may have a conserved role in mediating muscle fibre attachments and in preventing a dystrophic phenotype.

In this thesis, I aim to investigate the expression pattern of *lama1* and whether Hh signalling regulates *lama1* expression during zebrafish development.

To assess whether *lama1* is directly or indirectly controlled by Hh signalling, I also aim to identify some of the enhancer elements controlling *lama1* expression in the zebrafish. This work is expected to provide an in-depth understanding of the regulatory mechanisms governing *lama1* expression, which will aid the development of therapeutic strategies to alleviate the symptoms of muscular dystrophy and other phenotypes associated with a loss of *lama1* expression

It has also been demonstrated that Shh signalling controls the expression of *C125*, a novel gene, in the mouse embryo. *C125* was identified in a subtractive hybridisation screen to identify targets of Gli2, and a progressive loss of *Gli* alleles is associated with a progressive loss of *C125* expression. It is hypothesised that *C125* plays a role in cell fate specification and motor neuron formation.

In this thesis, I aim to examine the expression pattern of *c125* and whether Hh signalling regulates *c125* expression. I also aim to investigate the function of *c125* in the zebrafish embryo, testing the hypothesis that *c125* plays a role in cell fate specification.

These studies could provide a greater understanding of how Shh signalling determines cell fate specification, cell migration, and cell differentiation, through the identification of Shh target genes. These studies will also reveal whether these genes are regulated by Shh signalling in a conserved manner.