Functional Analysis of the Cytoskeletal Protein Talin2 using Gene Disruption

Thesis submitted for the degree of
Master of Philosophy
at the University of Leicester

By

Yasmine El Jaï
Department of Biochemistry
University of Leicester

September 2006
A la mémoire de Muema,
à qui j'aspire toujours à ressembler.

A la mémoire de tonton Jaouad,
qui m'a inspiré à vouloir mieux comprendre le monde du vivant.
Table of Contents and Figures

Table of Contents i
List of Figures v
List of Tables vi
Abstract viii
Acknowledgments ix
Abbreviations x

Chapter 1: Introduction 1
1.1. Interactions between cells and their environment 1
  1.1.1. Cells interact with adjacent cells and the ECM 1
  1.1.2. Focal Adhesions 3
1.2. Integrins 6
  1.2.1. Integrins structure 6
  1.2.2. Integrins are linked to the cytoskeleton 8
  1.2.3. Integrin signalling 9
1.3. Talin 10
  1.3.1. Talin structure 10
  1.3.2. Talin localisation 13
  1.3.3. Talin Function 13
  1.3.4. Talin2 17
1.4. Mouse Gene targeting 18
  1.4.1. Generation of transgenic mice from targeted mouse Embryonic Stem Cells 18
  1.4.2. Classical Knockout 20
  1.4.3. Conditional Strategy 24
  1.4.4. Genetrap approach and the Sanger Centre Project 26
1.5. Project’s aims 29

Chapter 2: Materials and Methods 30
2.1. Preparation and analysis of DNA 30
  2.1.1. Precipitation of DNA 30
2.1.2. Agarose Gel electrophoresis 30
2.1.3. Bacterial culture of DNA in DH5α 31
2.1.4. Alkaline Mini prep of plasmid DNA 31
2.1.5. Midi prep of plasmid DNA (QIAgen) 32
2.1.6. Midi Preparation of cosmid DNA 32
2.1.7. Caesium chloride plasmid DNA preparation 32
2.1.8. Quantification of nucleic acids 33
2.1.9. Sequencing of DNA 33

2.2. Polymerase Chain Reaction (PCR) 33
2.2.1. Oligonucleotides synthesis 33
2.2.2. Primer sequences 34
2.2.3. Amplification of the \textit{Tln2} targeting vector arms 35
2.2.4. Screening of \textit{Tln2} homologous recombinants 35
2.2.5. Identification of the 3'\textit{loxP} in the \textit{Tln2} targeting event 36
2.2.6. PCR from cDNA 36
2.2.7. Mouse genotyping: PCR analysis on lysed tail DNA 36

2.3. Cloning 36
2.3.1. Restriction digest of plasmid DNA 37
2.3.2. Gel Purification 37
2.3.3. Plasmid dephosphorylation 37
2.3.4. Ligation 37
2.3.5. Transformation of plasmid into DH5α 37
2.3.6. Plasmids 38

2.4. Southern blot analysis of genomic DNA 40
2.4.1. Southern blot 40
2.4.2. 5'end labelling of oligonucleotides and hybridisation 41

2.5. Gene expression analysis 41
2.5.1. RNA extraction from cultured cells 41
2.5.2. Northern blot 41
2.5.3. Random primed labelling of probe 41

2.6. Tissue Culture 42
2.6.1. Growth media and general protocols 42
2.6.2. Growing MEFs and Mitomycin C treatment 43
2.6.3. Trypsinisation 43
2.6.4. Freezing and storage 44
2.6.5. Electroporation of ES cells 44
2.6.6. Clone picking and freezing 44
2.6.7. DNA extraction for screening 45
2.6.8. Preparing ES cells for blastocyst injection 45
2.6.9. MEF transfection 46
2.6.10. Staining of ES cells with X-Gal 46
2.7. Protein analysis 47
2.7.1. Antibodies 47
2.7.2. Protein extraction from mouse tissues 47
2.7.3. Protein quantification assay 47
2.7.4. SDS PAGE gel 47
2.7.5. Western Blot analysis 48
2.7.6. Immunofluorescence 49
2.8. Animal work 50
2.8.1. Lysis of tail samples for DNA extraction 50
2.8.2. Generation of mouse embryonic fibroblast cell lines 50
2.9. Histology 51
2.9.1. Slides preparation 51
2.9.2. β-Gal staining of whole mount embryos 51
2.9.3. Fresh tissue extraction and sectioning 52

Chapter 3. Analysis of Talin2 expression 53
3.1. Introduction 53
3.2. Transcript expression in mouse and human tissues 54
3.2.1. Northern Blot analysis 54
3.2.2. Analysis of Human Kidney cDNA 57
3.2.3. Analysis of protein expression in mouse tissues 63
3.3. Analysis of talin2 expression in the mouse using a Tln2 genetrap cell line 65
3.3.1. Analysis of talin2 expression in a Tln2 genetrap ES cell line 65
3.3.2. Analysis of talin2 expression in mouse adult tissues using the Tln2 genetrap transgenic mouse 69
3.3.3. Analysis of talin2 expression in developing embryos using the Tln2 genetrap transgenic mouse line

3.4. Discussion

Chapter 4. Generation of a Tln2 conditional Knockout

4.1. Introduction

4.2. Mapping of the 5'end of the mouse Tln2 gene

4.2.1. Analysis of the cosmid clone

4.2.2. Ensembl genome database analysis

4.2.3. Generation of a map of 5'end of the Tln2

4.3. Generation of Tln2 conditional targeted ES cells

4.3.1. Construction of the Tln2 targeting vector

4.3.2. Generation of a PCR positive control vector for the screening of the Tln2 homologous recombination event

4.3.3. Generation of Tln2 conditional knockout ES cells

4.4. Generation of Tln2 conditional knockout mice

4.5. Deletion of the Tln2 conditional allele using Cre recombinase

4.5.1. Action of Cre recombinase in mouse embryonic fibroblasts (MEFs)

4.5.2. Action of Cre recombinase in mice

4.6. Discussion

Chapter 5. Discussion

References
List of figures

Figure 1.1. Cells interact with adjacent cells and the extracellular matrix (ECM) 2
Figure 1.2. Schematic representation of a focal adhesion complex 4
Figure 1.3. The Role of Rho kinase in FA assembly 7
Figure 1.4. Talin domains and ligand binding sites 11
Figure 1.5. Model for talin activation of integrins 16
Figure 1.6. Generation of transgenic mice from targeted 129Ola ES cells 19
Figure 1.7. Different strategies for classical gene targeting 21
Figure 1.8. Strategy used for Tln1 gene disruption in mice 23
Figure 1.9. Conditional gene targeting strategy based on the use of Cre-loxP
and Flp-Frt recombination systems 25
Figure 1.10. Gene alteration using a genetrap strategy 28
Figure 2.1. Restriction maps of plasmids 39
Figure 3.1. Position of the talin2 probes used for the Northern blot analysis 55
Figure 3.2. Mouse and human Northern blots hybridised with different probes
or Tln2 cDNA 56
Figure 3.3. PCR amplification of different regions of the Tln2 cDNA 59
Figure 3.4. Amplification of the 5’ end of the human Tln2 transcript in kidney 61
Figure 3.5. Alignment of the kidney ESTs with the novel human kidney cDNA 62
Figure 3.6. Analysis of talin protein expression in mouse and human lysates 64
Figure 3.7. ES Cell line RRI434: Identification of the gene trap insertion in the
mouse Tln2 gene 67
Figure 3.8. Expression of talin2- βgeo in the RRI434 ES cells 68
Figures 3.9. to 3.12 Expression of talin2-βGal in mouse tissues dissected from
Tln2GT/+ mice 70-73
Figure 3.13. Expression of talin2-bGal in heterozygous developing embryos by
wholemount staining 75
Figure 3.14. Expression of talin2-bGal in heterozygous Tln2GT/+ 10.5dpc embryos
by β-gal staining of sections 76
Figure 3.15. Expression of talin2-bGal in heterozygous Tln2GT/+ 14.5dpc embryos
by β-gal staining of sections 77
Figure 4.1. Strategy for gene conditional targeting using the Cre-triloX system 85
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.</td>
<td>Agarose gel and Southern blot analysis on single restriction digests of the Tln2 C10 cosmid clone</td>
</tr>
<tr>
<td>4.3.</td>
<td>Agarose gel and Southern blot analysis on double restriction digests of the Tln2 C10 cosmid clone</td>
</tr>
<tr>
<td>4.4.</td>
<td>Restriction map of the 5' end of Tln2</td>
</tr>
<tr>
<td>4.5.</td>
<td>Strategy adopted for the inactivation of the Tln2 gene</td>
</tr>
<tr>
<td>4.6.</td>
<td>PCR amplification of the isogenic arms of the Tln2 conditional targeting vector</td>
</tr>
<tr>
<td>4.7.</td>
<td>Cloning of the Tln2 targeting vector: Subcloning of the 3 homology arms</td>
</tr>
<tr>
<td>4.8.</td>
<td>Final cloning step of the Tln2 conditional targeting vector</td>
</tr>
<tr>
<td>4.9.</td>
<td>Identification of the homologous recombination event and generation of a PCR positive control</td>
</tr>
<tr>
<td>4.10.</td>
<td>PCR screening for ES cell homologous recombinants</td>
</tr>
<tr>
<td>4.11.</td>
<td>Genotyping of the different Tln2 alleles after exposure to Cre</td>
</tr>
<tr>
<td>4.12.</td>
<td>Cre mediated deletion of the Tln2 floxed allele in MEFs</td>
</tr>
</tbody>
</table>
List of Tables

Table 1.1: Integrin links to the actin cytoskeleton 8
Table 1.2. Comparison of the two human Talin1 (TLN1) and Talin2 (TLN2) genes and proteins 17
Table 3.1. Summary of the Northern hybridisation results 57
Table 3.2. Expression of Tln2-βgeo at different stages of development. 79
Table 4.1. Sizes of the fragments for each of the five digests 87
Table 4.2. Summary of hybridisation results of each restriction digest with each probe 87
Table 4.3: Summary of hybridisation results of the double digests with each probe 88
Table 4.4. Genotyping of the different Tln2 alleles resulting from the Cre mediated recombination. 104
Table 4.5. In vivo Cre mediated deletion of the Tln2<sup>IV+</sup> allele 108
Abstract

Talin1 is a large, ubiquitously expressed cytoskeletal protein that couples integrins to the actin cytoskeleton. Analysis of genomic and EST databases shows the presence of a second Tln gene encoding a closely related protein (75% identity in human), Tln2. Disruption of both Tln1 alleles in mouse results in early embryonic lethality. Although both talin1 and talin2 are very similar in sequence and most probably bind the same ligands, talin2 cannot compensate for the loss of talin1 in early mouse development. To further understand the role of talin2, analysis of talin2 expression at the mRNA and protein level in mouse and human has been carried out. A genetrap cell line (BayGenomics) expressing a fusion protein talin2-β-galactosidase has been used to report on talin2 expression. This cell line was used to generate transgenic mice, which allowed the study of talin2 expression at different stages of embryonic development and in various adult mouse tissues. In order to study the role of talin2 in mammalian physiology and ontogeny, a conditional knockout approach based on a Cre-triloxx system was used to inactivate Tln2 expression in mouse. The successful generation of a Tln2 conditional knockout mouse is reported.
Acknowledgments

I would firstly like to thank Prof David Critchley for giving me the chance to work on such a challenging and interesting project. He made me feel very welcome in the lab, and was always available for any advice I needed.

I would also like to thank Catrin for all her help over the past few years, especially over these last few hectic months. Her skills as a supervisor, her encouragements, her perfectionism and her generosity will always be remembered.

A big thank you to Sue Monkley, who showed great patience in promptly providing answers to my numerous questions and enquiries. Her expertise and help have been invaluable during the course of this project.

I also will hold many thoughts of my fellow lab members: Kathryn, Vicky, Ruth, Susan, Bip & many others, for all the conversations and laughs which made every day life in the lab very enjoyable.

To all the staff from the Division of the Biomedical services in the MSB, as well as Jane, Anthony and Lucy in the Transgenic Unit thank you for all their help with the mouse related work & Jenny Edwards (MRC Toxicology Unit) for her help with the histology sections.

My parents, not only passed on their scientific inheritance and showed support throughout my endeavours, but also put up with me over the past year. They never lost patience and always provided a happy, peaceful environment for me to develop this thesis.

Finally, thanks and hugs must go out to all my friends: Marie-Ange, my best friend, who understands me better than anyone else; all my friends who live far away, but thanks to new technologies have stayed in touch, always eager to learn how things were going & also to my friends from the forum who followed my ups and downs on a daily basis and kept believing in me more than I ever did.

Thank you!
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ RACE</td>
<td>5’ rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine tri phosphate</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium Persulfate</td>
</tr>
<tr>
<td>βgeo</td>
<td>β galactosidase-Neomycin reporter gene</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumine</td>
</tr>
<tr>
<td>CAM</td>
<td>Cell Adhesion Molecules</td>
</tr>
<tr>
<td>CMV</td>
<td>cytomegalovirus</td>
</tr>
<tr>
<td>CIP</td>
<td>Calf Intestinal Phosphatase</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dubelco Modified Eagle Media</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dpc</td>
<td>days post-coitum</td>
</tr>
<tr>
<td>ECM</td>
<td>Extra cellular matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethyl Diamine Tetra Acetic acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>Ethyleneglycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid</td>
</tr>
<tr>
<td>ES cells</td>
<td>Embryonic Stem cells</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed Sequence Tag</td>
</tr>
<tr>
<td>FA</td>
<td>Focal adhesion</td>
</tr>
<tr>
<td>FAK</td>
<td>Focal Adhesion Kinase</td>
</tr>
<tr>
<td>Frt site</td>
<td>Flp recombinase target site</td>
</tr>
<tr>
<td>HEK293</td>
<td>Human embryonic kidney cells</td>
</tr>
<tr>
<td>HGMP</td>
<td>Human Genome Mapping Project</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>HygR</td>
<td>Hygromycin resistance cassette</td>
</tr>
<tr>
<td>I4</td>
<td>RRI434 Tln2 ES cell line</td>
</tr>
<tr>
<td>IMS</td>
<td>Industrial Methylated Spirit</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>LB</td>
<td>Luna Bertari Medium</td>
</tr>
<tr>
<td>LIF</td>
<td>Leukemia Inhibitory Factor</td>
</tr>
<tr>
<td>MAP Kinase</td>
<td>Mitogen Activated Pathway Kinase</td>
</tr>
</tbody>
</table>
MEFs  Mouse Embryonic Fibroblasts
MLC  Myosin light chain
MLCP  Myosin light chain phosphatase
MLCK  Myosin light chain kinase
MTN  Multiple Tissue Northern Blot (Clonetech)
MW  Molecular weight
Neo<sup>R</sup>  Neomycin resistance cassette
OD  Optical Density
ORF  Open reading frame
PBS  Phosphate Buffer Saline
PCR  Polymerase chain reaction
PIP2  Phosphatidylinositol bisphosphate
PIPKIγ  Phosphatidylinositol-4'5'-kinase type Iγ
RNA  Ribonucleic Acid
ROCK  Rho associated kinase
rpm  rotation per minute
RT-PCR  Reverse transcription- Polymerase chain reaction
SDS  Sodium DodecylSulfate
TalA  Dictyostelium discoidum TalinA protein
TalB  Dictyostelium discoidum TalinB protein
TE  Tris - EDTA
TEMED  (N,N,N',N'- Tetramethylethylenediamine)
Tln1  Talin1 mouse gene
TLN1  Talin1 human gene
Tln2  Talin2 mouse gene
TLN2  Talin2 human gene
Tln2<sup>+/+</sup>  Talin2 homozygous wild type
Tln2<sup>fl/fl</sup>  Talin2 homoygous for the floxed allele
Tln2<sup>+/fl</sup>  Talin2 heterozygous for the floxed allele
Tln2<sup>++/-</sup>  Talin2 heterozygous for the null allele
Tln2<sup>-/-</sup>  Talin2 homozygous for the null allele
Tln2<sup>++/GT</sup>  Talin2 heterozygous for the genetrap allele
Tln2<sup>GT/GT</sup>  Talin2 homozygous for the genetrap allele
X-Gal  5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Chapter 1: Introduction

1.1. Interactions between cells and their environment

The ability of eukaryotic cells to interact with their environment is a key process for the development and homeostasis of a multicellular organism. These interactions occur between cells and neighbouring cells or between cells and the extra cellular matrix (ECM).

1.1.1. Cells interact with adjacent cells and the ECM

The organisation of cells into tissues and organs is very much dependent on the molecular interactions at the cellular level (Figure 1.1A). Intercellular interactions are mediated by cell adhesion molecules (CAMs). CAMs are transmembrane proteins that fall into four different families: cadherins, immunoglobulins (Ig) superfamily, integrins and selectins. E-cadherins and Ig form homophilic intercellular interactions. CAMs can interact with the same kind of CAM on an adjacent cell (homophilic binding) or to a different kind of CAM (heterophilic binding). E-Cadherins form homophilic bonds with adjacent cells. Members of the Ig family can form homophilic and heterophilic linkages. Integrins mediate heterophilic interactions with proteins from the ECM such as fibronectin or collagen. Selectins also form heterophilic interactions, but they recognise specialised sugar residues on glycoproteins and glycolipids from adjacent cells. These adhesive molecules contain multiple domains and their cytoplasmic face interacts with structural proteins that are linked to the cytoskeleton or to downstream signaling pathways (Hynes, 2002). Various factors determine different types of junctions. The binding affinity of interacting molecules, the spatial distribution of the molecule involved, and the influence determine the type of function and its properties. Interactions between cells and adjacent cells or with the ECM, not only provide some mechanical support to a tissue, but also play a role in communication between cells. Because cell-cell adhesions are associated with elements of the cytoskeleton and signaling proteins inside the cell, the surrounding of a cell and the cytoplasm have an influence on the cell properties ("outside-in" effects) and similarly cell shape and events within the cell have an influence on the cell’s environment ("inside-out" effects). The junctions that a cell establishes with its neighbouring cells and the ECM contribute to the structure of the tissue it belongs to and the function of the organ.
Figure 1.1. Cells interact with adjacent cells and the extracellular matrix (ECM). (A) Different adhesion molecules mediate different types of intercellular and cell-matrix interactions. Cell Adhesion Molecules are composed of 4 main families of proteins. Cadherins and immunoglobulins (Ig) can establish homophilic interactions with molecules on the plasma membrane of an adjacent cell. Integrins and Selectins can establish heterophilic interactions with molecules of an adjacent cell or of the extracellular matrix. (B) The small intestine epithelium is composed of a monolayer of polarised cells that establish multiple types of cell-cell and cell-matrix adhesions that contribute to the tissue’s structure and function. Tight junctions play the role of maintaining the link between adjacent cells. Gap junctions allow communication between cells by allowing small molecules to diffuse between adjacent cells. Adherens junctions, desmosomes and hemidesmosomes are tightly linked to the cytoskeleton and contribute to maintain the cell shape and the tissue structure. (source: Molecular Cell Biology)
The epithelium of the small intestine is made of a monolayer of polarised cells (Figure 1.1.B). Its role is to transport the product of digestion coming from the stomach, from the lumen across the basal lamina into the blood. The luminal side of the epithelium has microvilli that extend from the plasma membrane into the lumen. The correct orientation of the cells is maintained by a network of numerous connections between cells in the epithelial layer and between cells and the basal lamina. Three major classes of junctions are present in the intestinal epithelium. The first category of junctions has the role of maintaining the structure as a monolayer by holding adjacent cells together; they are the *tight junctions* and are present just under the microvilli. They prevent the diffusion in between cells of many substances between the lumen and the extracellular space. *Gap junctions* allow the rapid diffusion of small, water-soluble molecules between adjacent cells. The third category of junctions participates in cell-cell adhesion as well as in cell-ECM adhesion and play a role in maintaining the shape and the rigidity of the epithelial layer. They are the *adherens junctions*, and are usually located near the apical surface just below the tight junctions. They are tightly linked to the actin cytoarchitecture and play a role in controlling the cell shape. Other junctions such as desmosomes and hemi desmosomes also play a role in maintaining the integrity of epithelia.

1.1.2. Focal Adhesions (FA)

a) FA components

Specialized regions of close contact between the plasma membrane and the underlying substratum called FAs mediate adhesion to the ECM. These were first identified by electron microscopy (Abercrombie et al., 1971) as electron dense regions at the plasma membrane. At these sites bundles of actin filaments are anchored to the transmembrane protein family of integrins via a complex of numerous cytoskeletal components. They are composed of structural proteins that act as linkers between the actin cytoskeleton and the integrin receptors that are linked to proteins of the ECM. They also contain signaling molecules that mediate downstream signaling pathways involved in cell migration, cell survival, differentiation and gene expression (Figure 1.2). Although it is speculated that FA were an *in vitro* artefact in cultured cells, similar structures are observed in *in vivo* structures such as myotendinous junctions in skeletal muscle and dense plaques in smooth muscle (Burridge et al., 1987).
Integrin heterodimer

**Figure 1.2. Schematic representation of a focal adhesion complex.** Integrin receptors are linked to the actin cytoskeleton. They form complex structures called focal adhesions (FA) and are directly or indirectly linked to actin stress fibers via multiple structural proteins such as talin, α-actinin, vinculin.
b) FA formation and turnover
Cell-ECM interactions are involved in many different physiological processes including cell proliferation, suppression of apoptosis, cell migration and differentiation. FAs are a good model system to study cell adhesion because they are the type of adhesion developed by many types of cells in culture. Studies on FAs show that cells interact with the ECM via the integrin family of adhesion receptors, the cytoplasmic domains of which are linked to the actin cytoskeleton by proteins such as talin and α-actinin (Figure 1.2).

FAs are complex structures containing numerous structural and signaling proteins. They are formed at the end of actin stress fibers to which they are linked. The association of integrins with actin binding proteins such as talin, α-actinin and filamin, leads to the organisation of actin filaments into stress fibers. Similarly, the assembly of actin filaments into stress fibers promotes integrin clustering and FA assembly.

The balance between FA assembly and disassembly plays an evident role in mechanisms such as cell migration and spreading. Therefore coupling between integrin receptors and the actin cytoskeleton is tightly regulated. Calcium dependant proteolysis by calpain is one of the mechanisms by which FA turn over is mediated. Many structural and signaling proteins present in FA complexes are known to be calpain substrates (Reviewed by Lebart, 2006). Franco et al, by using a calpain resistant talin mutant in NIH3T3, showed a reduced FA turnover in the absence of talin cleavage. Vinculin, which is an actin binding protein present in FA also plays a role in FA turnover and cell motility, although it is not required for FA assembly, since vinculin-/- mouse embryonic fibroblasts assemble such structures (Xu W, 1998). Evidence has also been given that vinculin inhibits cell migration by stabilising FAs, and this role is regulated by phosphatidylinositol 4,5-bisphosphate (PIP2), which induces vinculin activation by allowing it to switch to an open conformation (Saunders et al., 2006).

c) The role of Rho kinases
Contacts between cells and the ECM enhance their spreading. Integrins regulate cell motility through the Rho family of GTP binding proteins (Nobes and Hall, 1999), particularly Rho, Rac and cdc42. Rho is able to drive the formation of large complexes by activating Rho kinase (ROCK), which in turns phosphorylates the myosin-binding
site of the myosin light chain phosphatase (MLCP) leading to its inhibition (Kawano et al., 1999). Myosin Light Chain (MLC) in its phosphorylated form is accumulated and this results in actomyosin contraction, stress fibers formation and in turn FA adhesion assembly (Figure 1.3).

Additionally Rho binds to and activates PI4(5) kinase (Chatah and Abrams, 2001), leading to an increase in PIP2 production. PIP2 interacts with vinculin resulting in the open conformation of vinculin thereby making available binding sites to talin and actin (Gilmore and Burridge, 1996) and reinforces the nascent FA complex. PIP2 is also involved in the release of actin monomers from profiling and gelsolin complexes, thereby enhancing actin polymerisation and stress fibers assembly.

Whereas Rho regulates the association of actin-myosin filaments to form stress fibers (Ridley and Hall, 1992), the other family members Rac and cdc42 regulate the formation of lamellipodia and filopodia (Nobes and Hall, 1995). Furthermore the activities of these three proteins are linked to each other. Cdc42 can activate Rac, leading to the association of filopodia and lamellipodia. Rac can activate Rho leading to the formation of FAs.

1.2. Integrins

1.2.1. Integrins structure

Integrins are a family of heterodimeric transmembrane proteins that are involved in the cell-ECM and cell-cell interactions. They are composed of two non-covalently associated subunits; the α-subunit is 120-180kDa and the β-subunit 90-110 kDa. At present, 18 α-subunits and 8 β-subunits have been described. Each subunit contains a short cytoplasmic domain (<50 residues, with the exception of the β4 subunit) that is indirectly linked to the actin cytoskeleton. Integrins recognize both cell-surface receptors and ECM. Each α-β combination has its own binding specificity and signaling properties, although some of the integrin ligands can bind several different integrins. Integrins bind to a wide range of ligands such as collagen, fibronectin and laminin.
Figure 1.3. Rho induced Focal Adhesion assembly. Rho’s main downstream effector is ROCK which inhibits the phosphorylation of MLCP, leading to an increase in MLC phosphorylation, allowing actomyosin contraction. It also activates LIMK which in turns inhibits coflin by phosphorylation. Active coflin enhances actin depolymerisation. Rho also activates PIP Kinase, leading to the production of PIP2 allowing the release of actin monomer to be released from gelsolin and profilin complexes. This results in an increase in actin polymerisation and actin filaments assembly. The formation of actin filaments along with stress fiber assembly contributes to the assembly of new focal adhesion. ROCK: Rho Kinase, MLCP: Myosin Light Chain Phosphatase, MLCK: Myosin Light Chain Kinase, PIP2: Phosphatidylinositol 4,5’ Phosphate, LIMK: LIM Kinase.
Integrin β-cytoplasmic tails are necessary and sufficient to link integrins to the actin cytoskeleton. There is less evidence that α-tails are directly linked to the cytoskeleton (Burridge and Chrzanowska-Wodnicka, 1996) although they are thought to negatively regulate binding of the α cytoplasmic tails to the cytoskeleton (Ginsberg, 1995). The interaction between the α and β tails is modulated by ligand binding to the extracellular domain (Leisner et al., 1999). Consequently, changes in the interactions with extracellular ligands may affect the integrin-cytoskeleton linkage.

### 1.2.2. Integrins are linked to the cytoskeleton

Integrins have been shown to bind to a number of cytoskeletal adaptor and signaling proteins. They generally cluster at the sites of close apposition of the plasma membrane to the ECM forming structures called FAs (Burridge and Chrzanowska-Wodnicka, 1996; Jockusch et al., 1995; Hall, 1998). The association of integrin receptors with the actin cytoskeleton occurs via direct or indirect interaction of the β cytoplasmic tails with actin binding proteins. Examples of proteins that link the integrins to the actin cytoskeleton are shown in Table 1.1.

<table>
<thead>
<tr>
<th>Linker Protein</th>
<th>Integrin</th>
<th>Other ligands</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talin (270 kDa)</td>
<td>β1A, β1D, β2, β3</td>
<td>Layilin, PECAM-1, FAK, vinculin, PIP2</td>
<td>Actin bundles / networks</td>
</tr>
<tr>
<td>60 nm, antiparallel dimer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-actinin (100 kDa)</td>
<td>β1A, β2</td>
<td>ICAM 1-3, NMDA-R etc. Zyxin vinculin etc. PI3K, PKN, PIP2</td>
<td>Actin bundles</td>
</tr>
<tr>
<td>35 nm, antiparallel dimer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filamin (&gt;250 kDa)</td>
<td>β1A, β2, β3, β7</td>
<td>Gp1b, Trio, Ral1A-GTP, Cdc42, Rac, Rho, SAPK</td>
<td>Actin cross-linking</td>
</tr>
<tr>
<td>Parallel dimer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myosin X (&gt;200 kDa dimer)</td>
<td>β5</td>
<td></td>
<td>Actin-binding, Molecular motor PH domain</td>
</tr>
<tr>
<td>Skelemin (200kDa)</td>
<td>β1, β3</td>
<td></td>
<td>Actin binding</td>
</tr>
</tbody>
</table>

Table 1.1. Integrin links to the actin cytoskeleton
In addition, in *Drosophila melanogaster* a protein called Integrin Linked Kinase has been showed to be involved in linking integrins to the actin filaments indirectly by binding an actin binding protein, actopaxin (Travis *et al*., 2003).

However each linker protein has a different structure, and supports a different type of actin cytoarchitecture. Filamin is a Y-shaped parallel dimer, and forms orthogonal actin networks (Stossel, 2001) whereas α-actinin is a short rod shaped anti-parallel dimer that promotes actin bundling (Blanchard, 1989). Therefore each linker protein between integrins and the actin apparatus is involved in different actin organization and required for different physiological contexts.

### 1.2.3. Integrin signaling

Integrins transduce signals by associating with (i) adapter proteins that connect the integrins to the cytoskeleton, (ii) cytoplasmic kinases and (iii) transmembrane growth factor receptors (Giancotti, 1997). Ligand binding to integrins leads to integrin clustering and association with cytoskeletal and signaling complexes that promote the assembly of actin filaments (outside-in signaling). The reorganization of actin filaments stimulates the clustering of integrins (inside-out signaling). As a result, integrins, ECM proteins and cytoskeletal proteins form aggregates on each side of the membrane, assembling FAs.

Integrins activate various tyrosine kinases, including FA Kinase (FAK). The activation of FAK is coupled to the assembly of FAs. FAK autophosphorylates (Y397), creating a binding site for Src or Fyn (Schaller and Parsons, 1994). The Src kinase then activates multiple FA components. Src also phosphorylates FAK (Y925) creating a binding site for the adapter Grb2, which in turn binds SOS, a guanine nucleotide exchange factor for Ras (Schlaepfer *et al*., 1997). In this way, a link is established between the growth factor activated Ras/MAPK pathway, which is involved in cell proliferation, and cell adhesion to the ECM. In the absence of attachment, the Ras mediated signaling pathway is blocked either at the level of Raf or MEK (Lin *et al*., 1997). Cell anchorage to ECM is closely related to cell survival and growth (Re *et al*., 1994).

The composition of the ECM is important in the control of cell growth. Some types of cell grow only on specific substrates. For example, myoblasts proliferate on fibronectin
but stop growing on laminin and differentiate into myotubes (Adams and Watt, 1993; Lin and Bissell, 1993). These results show that integrins promote cell growth but also regulate differentiation.

1.3. Talin

1.3.1. Talin structure

Talin is a large (270 kDa, 2541 amino acids) dimeric protein, which is localised predominantly at integrin-containing cellular junctions with the ECM (Burridge and Connell, 1983), although it is also found in some cell-cell junctions such as the junctions between T-cells and antigen-presenting cells (Monks et al., 1998). Talin has binding sites for several ligands including the cytoplasmic domains of integrin β-subunits, and the C-type lectin layilin, an integral membrane protein that co-localises with talin in membrane ruffles of migratory cells (Borowsky and Hynes, 1998). Talin also binds F-actin, a number of other cytoskeletal proteins, an isoform of phosphatidyl inositol-4'-5'-kinase type Iγ (PIPKIγ) which is implicated in the regulation of cell motility, cell proliferation and the suppression of apoptosis (Ilic et al., 1997). Therefore, talin may play a key role in coupling integrins to the actomyosin contractile apparatus and in integrin signaling pathways.

Electron microscope images of talin show an elongated molecule (60 nm) with a globular head and a flexible tail (Isenberg et al., 2002). These two regions can be separated by calpain II cleavage into an N terminal head-region (50 kDa) and a C-terminal rod-region (220kDa) (Rees et al., 1990) (Figure 1.4). The talin head contains a region of homology to the FERM domain family (Band 4, Ezrin, Radixin, Moesin) of cytoskeletal proteins, which act as adaptors between membrane proteins and the actin cytoskeleton (Takeuchi et al., 1994). (Girault et al., 1998; Schultz et al., 1998). FERM domains contain three subdomains F1-F3, and F3 domain closely resembles a phosphotyrosine binding (PTB) domain. The crystal structure of a part of the talin FERM domain bound to part of the integrin β3-subunit cytoplasmic domain has recently been solved (Garcia-Alvarez et al., 2003). An NPxY motif in the integrin β3–cytoplasmic tail binds to the talin F3 phosphotyrosine binding domain (Calderwood et al., 2002b), and this interaction has been shown to be sufficient to activate integrin
**Figure 1.4. Talin domains and ligand binding sites.** Talin is a 2541 aminoacid cytoskeletal protein composed of a globular head and a flexible tail, which can be separated by calpain II cleavage. The head domain contains binding sites to β integrin subunit, F-Actin, and signalling proteins such as PIPKγ. It also binds other components of focal adhesions such as vinculin which is also an actin binding protein.
αIIbβ3 (Calderwood et al., 2002a). The talin head also binds the cytoplasmic domain of layilin (Borowsky and Hynes, 1998) and recent studies show that layilin binds to the F3 subdomain in a similar manner to the integrin β-subunits (Barsukov and Critchley, unpublished data). Similarly, the talin head also binds to PIPKIγ via the same hydrophobic groove in F3 subdomain of the talin FERM domain, and NMR studies have shown that interactions of talin with PIPKIγ and β integrin are mutually exclusive (Barsukov et al., 2003).

The affinity of the talin head for the cytoplasmic domain of β3 integrin is increased six fold by proteolytic cleavage between the talin head and the rod domain (Yan et al., 2001) suggesting that in intact talin, the integrin binding domain is masked in some way (Sampath et al., 1998). Phosphorylation of PIPKIγ on Y661 by c-Src is important for the association of the kinase with talin and its targeting to FA (Ling et al., 2003). The interaction between talin and PIPKIγ is adhesion dependent and activates PIPKIγ, which is thought to drive talin to the plasma membrane and facilitate its incorporation to FAs (Di Paolo et al., 2002; Ling et al., 2002). The talin head domain contains an N-terminal membrane-binding domain, which inserts into lipid bilayer containing PIP2 (Seelig et al., 2000). The association of talin with PIP2 results in a conformational change of the protein, which becomes more sensitive to protease cleavage (Martel et al., 2001). Moreover, in the presence of PIP2, the affinity of talin for the β-integrin cytoplasmic domain is higher. This result suggests that PIP2 induces a conformational change in talin that makes the integrin-binding site more accessible.

The talin head also contains an actin binding site (Hemmings et al., 1996; Lee et al., 2004) and a myosin-binding site (Lin et al., 1998). The interaction between talin and myosin stimulates the ATPase activity of myosin. The authors suggested that talin might play a role in the interaction between soluble myosin and actin in membrane ruffles and then in the initiation of actomyosin contraction required for cell motility.

The talin rod contains two actin-binding sites (Hemmings et al., 1996), the most C-terminal of which (residues 2345-2541) shows homology with the yeast actin-binding protein Sla2p, and the huntingtin-interacting protein Hip1 (Kalchman et al., 1997). Interestingly, a second β-integrin binding site is also located in the C-terminal region of
the rod (Xing et al., 2001). The talin rod domain also contains binding sites to another actin binding protein, vinculin. Vinculin is a 116kDa protein, and multiple vinculin-binding sites have been identified in the talin rod. Using a yeast two hybrid screen three vinculin binding sites were initially identified as single α helices (Bass et al., 1999). More recently 8 additional vinculin binding sites have been identified using recombinant peptides. Structural analyses based on NMR and crystallography have given evidence that the presence of hydrophobic amino acid residues played a role in the vinculin-talin interactions. (Gingras et al., 2006; Patel et al., 2006).

1.3.2. Talin localisation

FAs are a model of cell-ECM interactions formed by fibroblasts in culture. However, many cell types develop structures in vivo, which are very similar to FAs and implicate the same proteins such as talin. Talin is mainly localised in FAs in cells in culture, in membrane ruffles where it colocalises with layilin (Borowsky and Hynes, 1998), and at sites of phagocytosis in macrophages (Greenberg et al., 1990) and in some intercellular junctions (Burridge and Connell, 1983). It has been shown that talin is highly concentrated in myotendinous junctions (Tidball et al., 1986), where the actomyosin cytoskeleton is in contact with the plasma membrane. Interactions between T cells and Antigen presentating cells (APC) involve a complex of receptors and cytoskeletal proteins. Experiments using immuno-fluorescent staining shows that talin is present at the sites of cell-cell contacts between T-cells and APC (Monks et al., 1998). Talin is also present in neuronal growth cones. Its inactivation by laser irradiation in these cells result in defects in filipodial motility (Sydor et al., 1996), implicating that talin might play a role in filipodia extension or retraction.

1.3.3. Talin Function

Talin’s role in cell adhesion, cell migration and in development has been studied by inhibiting its function (knockout) in Dictyostelium discoideum (Niewohner et al., 1997). In talin-null cells a defect of cell motility and in cell adhesion to a substrate was observed. These cells also showed defects in particle uptake in phagocytosis and in cytokinesis resulting in the formation of binucleated cells. In mammalian cells, microinjection of anti-talin antibodies affects cell spreading, migration, and cell motility and disrupts FAs (Bolton et al., 1997; Nuckolls et al., 1992). Moreover, down regulation of talin expression in HeLa cells using antisense RNA technology also
resulted in defects in cell spreading and the assembly of cell-ECM junctions (Albiges-Rizo et al., 1995).

In *Caenorhabditis Elegans*, talin is localised in muscle adhesion complexes called dense bodies where it is associated with PAT-3β integrin. These structures play a major role during *C. Elegans* embryogenesis (Barstead and Waterston, 1991; Moulder et al., 1996). Gonad morphogenesis is dependent on migration of a specialised gonadal cell type, the distal tip cell (DTC) (Blelloch et al., 1999; Hubbard and Greenstein, 2000). Talin knockdown using a RNA interference approach caused severe defects in gonad formation because of aberrant DTC migration (Cram et al., 2003). It also resulted in paralysis and loss of muscle filament organisation, showing that talin plays a role in cell migration as well as in muscle organisation in *C. Elegans* larval and adult tissues.

In *Drosophila Melanogaster*, the absence of talin caused failure in germband retraction, muscle detachment and a wing blister phenotype due to separation of the two cell layers (Brown et al., 2002). The authors also studied the role of talin in integrin activation in *Drosophila*. Talin is highly present in muscle attachment sites where muscles attach via the ECM to the epidermis. The *Drosophila* integrin heterodimer expressed in muscle cells is αPS2βPS (αPS2 subunit is similar to vertebrate α5, IIb, V and 8, βPS subunit is the orthologue to β1) and studies showed that the isolated head domain of talin is sufficient for its localisation at these sites (Tanentzapf and Brown, 2006). However talin mutants (R367A) in which aminoacids involved in the talin-integrin interaction did not localise to muscle attachment sites. However studies on the full-length talin showed that the R367A mutation did not affect its recruitment to these sites. The αPS2βPS integrin ligand in the ECM is a protein called tiggrin. In the absence of talin, integrin remains bound to tiggrin, although is unable to connect to cytoskeletal components (Brown et al., 2002). More recently in *Drosophila*, talin has been shown to play an integrin independent role in the suppression of cadherin expression by regulating the level of cadherin transcript in oocyte localisation (Becam et al., 2005).

The role of talin in cell adhesion in mouse has been investigated by the production of *Tln1*−/− knockout cells. Disruption of both *Tln1* alleles in mouse embryonic stem (ES) cells inhibits cell spreading on gelatin and laminin, and the ability to assemble FAs on
fibronectin (Priddle et al., 1998). The gene targeting lead to the deletion of the 3’end of the first coding exon, the adjacent intron and the 5’end of the second coding exon. However, differentiated $Tln1^{-/-}$ ES cells were able to spread on gelatin and to assemble FAs containing vinculin and paxillin. This result suggested that other cytoskeletal proteins are important in integrin-mediated cell-adhesion and that they can replace talin in certain cell types. To study the importance of talin in development, mice with a heterozygote $Tln1^{+/}$ genotype have been generated. These mice are indistinguishable from their $Tln1^{+/+}$ littersmates and they are viable and fertile. However, intercrosses of these animals generated no $Tln1^{-/-}$ mice at birth. In depth analyses of litters at different stages of pregnancy showed that the embryos die at gastrulation stage and show failure of mesoderm cell migration (Monkley et al., 2000). At 6.5 to 8.5dpc, abnormal embryos were genotyped as homozygous mutant $Tln1^{-/-}$. At 9.5dpc, no $Tln1^{-/-}$ were present in the litter. Expression of talin protein was also analysed in mutant embryos by analyses of trophoblast cells derived from 3.5dpc embryos. It appeared that there was a talin immunoreactive protein expressed in these cells and that these cells were able to assemble FAs. This could be due to a potential reinitiation of translation downstream of the deletion leading to the production of a truncated talin protein or to the upregulation of talin2.

Integrin signaling via FA kinase (FAK) and Src results in the association of PIPKγ to talin (Figure 1.5). PIPK1γ belongs to the family of proteins, which synthesises PIP2 (PIP2). Local enrichment of PIP2 plays a major role in cell migration at the leading edge of a cell as it regulates. The binding of talin to PIPK1γ results in the translocation of the talin-PIPKγ complex to the plasma membrane. Active PIPK1γ at the plasma membrane leads to the production of PIP2. The local enrichment of PIP2 enhances the affinity of talin for β integrin (Di Paolo et al., 2002; Ling et al., 2002) making the integrin-binding site in talin available (Martel et al., 2001). PIPK1γ is subsequently displaced as the binding site for on talin overlaps with the one of β integrin. Studies carried out by Tadokoro et al. (2003) using RNA interference to knock down talin expression in mammalian cells show that talin binding to the integrin β tail of αvβ3 in megakaryocytes and α5β1 in chinese hamster ovary cells is required for inside out activation of integrins.
Figure 1.5. Model for Integrin activation by talin. Signalling via integrins and FAK leads to the phosphorylation of PIPKIγ. This allows the association of talin and PIPKIγ and the complex is subsequently translocated to the plasma membrane. At the plasma membrane, activated PIPKIγ enhances a local increase in PIP2. PIP2 in turn allows talin to switch to an open conformation, thus allowing it to dimerise and bind to integrin. As the binding to β integrin and PIPKIγ are mutually exclusive, the kinase is released into the cytosol. Binding of talin to integrin β tail constitute the initial link to a nascent FA complex.
1.3.4. Talin2

Recent analysis of human and mouse EST databases have identified a second gene encoding a protein highly related but distinct from talin, now renamed talin1 (Tln1) (Monkley et al., 2001). The second talin gene was called talin2 (Tln2). Talin2 protein is very similar in size to talin1 and they show a high level of sequence identity throughout their length. The main differences between the two human genes are in the size of the genes (see Table 1.2) and their expression patterns.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosomal location</th>
<th>Size</th>
<th>Gene</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Number of exons</td>
<td>Average intron size</td>
</tr>
<tr>
<td>Tln1</td>
<td>9p13</td>
<td>≈ 30 kb</td>
<td>57</td>
<td>486 bp</td>
</tr>
<tr>
<td>Tln2</td>
<td>15q15</td>
<td>≈ 190 kb</td>
<td>57</td>
<td>3490 bp</td>
</tr>
</tbody>
</table>

Table 1.2. Comparison of the two human talin1 (Tln1) and talin2 (Tln2) genes and proteins (Monkley et al., 2001).

Given the close similarities between the two talin proteins in size and in domain structure, it is likely that talin2 binds many of the same ligands. The second talin gene does not exist in Drosophila and C. elegans, but there is a talin2 gene in mouse (Tln2), rat and cow. In Dictyostelium discoideum, two genes have been found (TalA and TalB) but knockouts of these genes give different phenotypes, which suggest that the two proteins have different functions. It seems that TalB is more specific to multicellular morphogenesis, whereas talA is important in cytokinesis and cell motility. The TalA and TalB proteins are also quite different (30% identity) (Tsuijoka et al., 1999). It has been shown that talin2 as talin1, binds to PIPKIγ by its FERM domain and is thought to play a role in PIPKIγ activation (Di Paolo et al., 2002). It also binds F-actin by the conserved C terminal I/LWEQ module and this interaction is regulated by intrasteric inhibition (Senetar et al., 2004).

In adult mouse tissues, Northern blot analyses (Monkley et al., 2001) revealed a single major transcript for talin1 (8kb) that is expressed at similar levels in various tissues, while the expression levels of Tln2 varied greatly across tissues and different sizes of transcript were observed. The fact that disruption of Tln1 is embryonic lethal in mouse
indicates that talin2 cannot compensate for the loss of talin1 during early mouse development. This could be because they are not expressed at the same stage of development or the two proteins may have different functions or both are required for normal cellular function and form a talin1/talin2 heterodimer. Using a Tln2 gene trap embryonic stem cell line, Chen and Lo (2005) have generated a transgenic mouse expressing a fusion of the N terminal region of talin2 (aminoacids 1-1295) and β galactosidase. The authors showed that the homozygous mutants survived and were healthy, thus giving evidence that the N terminal half of talin2 in mouse is sufficient for mouse development and survival. This might be due to the fact that talin1 and talin2 although very similar in size and structure, might be required at different stages of development. Talin1 deficiency leads to embryonic lethality, showing that talin2 does not compensate for talin1 in early mouse development and that it might play an important role in specific tissues later on in development or in the adult mouse.

1.4. Mouse Gene targeting

The development of transgenic techniques to alter or suppress in vivo gene expression has become a very important tool to study and better understand genetic disorders by generating animal models and to investigate gene function particularly in embryonic development. Gene targeting typically involves the introduction of a mutation in a target gene in vivo, which may result in the complete inactivation or an alteration of the gene expression.

1.4.1. Generation of transgenic mice from targeted mouse Embryonic Stem cells

The ES cells used for gene targeting are extracted from 3.5dpc embryos generally from 129Ola mouse strain. After transfection or gene alteration, they can be injected into C57BL/6 3.5dpc blastocysts. They are then transferred into the uterus of a superovulating surrogate mother where they go through embryogenesis and develop to term. When the litter is born, the blastocysts containing a mix of the blastocyst coat colour and albino coat colour giving rise to chimaeric animals. Generally C57BL/6 host blastocysts are used, giving chimaeras that are a mix of white and black coat colour.
Figure 1.6. Generation of transgenic mice from targeted 129Ola ES cells. (A) After gene alteration and isolation of correct clones, 129Ola ES cells are microinjected into 3.5dpc host blastocysts (C57BL/6). The blastocysts are then microinjected into a superovulating surrogate mother. (B) After gestation, blastocysts composed of a mixture of 129Ola injected ES cells and C57BL/6 ES cells give rise to chimaeras. The coat colour of offspring from a backcross between the chimaeras and a C57BL/6, attests whether germline transmission has been achieved from the chimaeras.
Chapter 1 – Introduction

The next step is to ensure that the injected 129Ola cells have participated in the formation of the germline cells in the embryo. The male chimaeras are backcrossed to C57Bl/6 females. If germline transmission of the 129Ola ES cells has been achieved this will result in an agouti coat colour in the offspring as opposed to a black coat colour. If the gene of interest has been mutated only on one allele, the agouti pups have a 50% chance of carrying the gene targeting event (Figure 1.6).

1.4.2. Classical knockout

The classical knockout strategy to disrupt a gene is to replace or alter a part of the gene by homologous recombination between a targeting vector transfected in ES cells and a gene of interest. The target gene can be modified in different ways. A part of the coding sequence can be deleted by inserting an antibiotic resistance gene in order to inhibit the expression of a functional protein. A reporter gene, allowing the monitoring of the gene expression, a point mutation allowing studies on a particular function of the protein can also be inserted.

Two different vectors can be used to knockout a gene, replacement and insertion vectors (Figure 1.7). Replacement vectors are designed to replace some sequence in the endogenous gene with a selectable marker or a chosen sequence such as a chosen mutation or gene alteration. They contain two regions of homology of the target locus separated by a selection marker interrupting part of the coding region. Before transfection, the vector is linearised by a single restriction digest outside the region of homology with the gene. After a double recombination event in each arm of the vector, it results in gene inactivation when the introduced sequence induces a frameshift and inserts a premature STOP codon. Insertion vectors also contain two regions of homology with the gene of interest separated by a non-homologous region. Before transfection into ES cells the vector is linearised within the homology region. Homologous recombination between the vector and the target gene results in the insertion of the complete vector at the gene locus and a partial gene duplication.
Figure 1.7. Different strategies for classical gene targeting. Exons are represented by green boxes, intronic regions by back thick lines, the positive selection marker $Neo^R$ by the light blue box and plasmid DNA from the targeting vector by a thin black line.

Targeting vectors contain two regions of homology with the target locus separated by a non homologous sequence, usually a selection marker ($Neo^R$). Before transfection into ES cells the vector is linearised by a single restriction digest. The vector is digested outside the arms of homology for replacement vectors (A). Homologous recombination in each arm of homology results in the replacement of a part of the genomic region with vector DNA. In the case of an insertion vector (B), the single restriction digest is performed within the region of homology with the gene of interest. When homologous recombination occurs, it results in the insertion of the $Neo^R$ as well as a duplication of a portion of the gene.
This strategy can be used to introduce numerous different mutations. The most frequent mutation introduced is an insertion of a selection marker within a gene of interest as shown in Figure 1.7. The purpose of such an insertion is the inactivation of the gene expression and the abolition of the production of a fully functional protein. Using such a strategy, the gene structure must be well known and the exons to be interrupted or deleted carefully chosen. A reporter gene can also be coupled to the gene disruption. This enlarges the possibilities of studies by allowing the monitoring of the mutated gene expression. Other mutations such as single nucleotide mutation or deletion of a chosen portion of the coding region can also be introduced for more targeted studies, such as genetic diseases.

However, this approach only allows for the introduced mutation to be expressed in all the tissues and at all stages of embryonic development. In the case of early embryonic lethality of the homozygous mutant animals, the scope of potential analyses in development, cultured cells or in the adult animals is reduced and in some cases impossible.

A classical approach has been undertaken to disrupt the Tln1 in the mouse (Monkley et al., 2000). A selection marker was introduced between the first and the second coding exon, leading to the deletion of codons 37 to 66. Disruption in both alleles gave rise to embryonic lethality and no homozygous mutant Tln1/^+^ survived past 9dpc of embryogenesis as described above. A detail of the gene targeting strategy adopted is shown in Figure 1.8. Therefore only the early embryonic phenotype could be analysed. This severe phenotype reduced the scope of potential analyses of Tln1 function in vivo at later stages of development and in the adult animal. Studies were also carried out on Tln1/^+^ undifferentiated ES cells. Due to their inability to spread on laminin and gelatin and their adhesion defects, they were difficult cells to keep in culture for any long-term cell biology experiments. So although the knockout approach for Tln1 was a good tool for understanding gene function in vivo, the study of genes playing an important role in embryonic development and ontogeny using this strategy is reduced.
Figure 1.8. Strategy used for Tln1 gene disruption in mice (Monkley et al., 2000). A replacement targeting vector was constructed. It led to the removal of a portion of the first and second coding exons as well as the insertion of a selection marker to allow selection for recombinants. The insertion of HygR cassette led to the introduction of a stop codon and a polyadenylation signal within the HygR cassette aiming to inhibit the production of a full length protein. However the presence of an in frame initiation codon in the third coding exon could allow reinitiation, producing a truncated protein lacking the N terminal 72 amino acids.
1.4.3. Conditional Strategy

Some genes are vital and a classical knockout often causes embryonic lethality. Methods have been developed to inactivate gene expression, or activate a mutation of interest in only selected cell types or organs. These tools are based on the use of site-specific recombination systems present in phages and yeast.

Cre recombinase of the P1 bacteriophage directs recombination between $\textit{loxP}$ (locus of cross over) recombination sites. Similarly the Flp integrase of \textit{Saccharomyces cerevisiae} mediates recombination between Frt (Flp recombination target) recombination sites. In both cases, the only requirements are the presence of the recombination enzyme and the recombination sites in target DNA. No additional cellular factors are necessary. The $\textit{loxP}$ and Frt sites are 34bp DNA sequences, comprising 13bp palindromes separated by an asymmetric 8bp spacer. The recombination enzyme catalyses DNA strand exchange between two recombination sites resulting in deletion, duplication, inversion, translocation of sequences depending on the orientation of the recombination sites and the number of molecules involved (Figure 1.8).

The approach is based on the generation of a targeted allele in ES cells containing the recombination sites, flanking a region of interest, for a conditional chosen mutation. These ES cells can then be used to generate transgenic mice by injection into blastocysts. A big range of transgenic mice expressing Cre or Flp recombinase under the control of a tissue or cell type specific promoter, or inducible are now available. These systems allowed gene targeting to be controlled temporally and spatially, and therefore the study of gene and protein function in particular physiological or pathological contexts. In some cases the null phenotype also made the generation and culture of particular cell types difficult, and the existence now of various ways of administration of Cre and Flp also facilitates \textit{in vitro} studies.

Another approach based on RNA interference (RNAi) has been developed as a conditional gene targeting strategy. RNAi is a gene silencing process that occurs at the post-transcriptional level. It is based on the use of double stranded fragments of RNA homologous to the gene to be silenced. They are recognised by the RNAi machinery,
Figure 1.9. Conditional gene targeting strategy based on the use of Cre-loxP and Flp-Frt recombination systems. To overcome the embryonic lethality that might be caused by the loss of a functional protein, conditional approaches have been developed. They are based on the use of recombination systems present in Phage P1 (Cre-loxP) or yeast (Flp-Frt). Site specific recombinases (Cre and Flp) mediate inter and intra molecular recombination events, leading to deletion (A), duplication (B), inversion (C) or integration (D) events. The loxP or Frt recombination sites are indicated by a blue triangle.
which uses the antisense strand as a template to recognise the complementary mRNA and degrade it. Long double stranded RNAs (dsRNA, >500bp) can induce gene silencing in various organisms such as worms, flies and plant. A similar approach using shorter iRNA (<30bp), small interference RNAs (siRNA), has also become a very powerful tool to study gene function in mammalian cells and organisms. Gene silencing can be achieved by delivering synthetic siRNAs, or siRNA expressing plasmids or viruses in cultured cells as well as directly into the target tissue or in the blood stream in a living animal. siRNA could therefore be used to target ES cells in culture that could be used to generate transgenic mice in which a gene of interest has been silenced. This strategy also allows gene silencing in organisms for which knockout strategies have not been developed. Mouse ES cells can be transfected with a siRNA expressing plasmid or viral vector, recombinant clones can be checked for the copy number, the degree of gene silencing before being used to generate transgenic mice. To make possible the siRNA approach for genes involved in major processes such as embryonic development which silencing leads to embryonic lethality or sterility, inducible RNAi systems have been developed. The vector contains a promoter, and reporter gene flanked by recombination sites upstream of the siRNA sequence. Without any induction, the reporter gene would be expressed and the stop codon introduced at the 3’end of the reporter gene would stop any transcription of the siRNA, as it contains no IRES sequence. On exposure of the appropriate recombination enzyme, the reporter gene would be deleted and expression of the siRNA induced. However this approach has several down sides. The use of RNAi in vitro and in vivo brings the risk of not a complete knockdown and also the targeting of another gene than the one aimed for, and therefore the analysis difficult to interpret. Applied to in vivo studies, it involves invasive procedure in live animals by intravenous, or injection into organs and therefore might have a general effect on the animal health. Although this approach has several downsides, it could be a possible way of further understanding talin1 and talin2 function in living organisms (reviewed by Sandy et al., 2005).

1.4.4. Genetrap approach: Sanger Centre project

With the completion of the human and mouse genome sequences, there has been increased interest in developing tools for genome-wide mutagenesis in the mouse, with the goal of analysing genome function in the context of the intact organism. It
principally relied on exposing animals to radiation or chemical mutagens. Another alternative is to induce mutation by randomly introducing well characterised sequence into mouse ES cells which can then be used to generate transgenic mice.

The basic principle of gene trapping is that the transgene inserted into ES contains a reporter gene that lacks components necessary for gene expression. It will only be expressed if it has been integrated into an active gene and will follow the pattern of expression of the promoter or the gene in which it is inserted. The insertion usually results in the alteration of the gene function by disrupting the sequence. Gene trap vectors contain a sequence tag that allows identification of the locus of insertion and a reporter gene, which provides a good tool to monitor gene expression.

BayGenomics (University of California, Berkeley) are part of the international gene trap consortium (IGTC, http://www.genetrap.org/index.html), which regroups various organisations who have carried out high throughput gene trapping projects and made publicly available their gene trap ES cell lines to the scientific community. They have used a βgalactosidase-Neomycin (βgeo) marker gene (Skarnes et al., 2004). It was designed such as after insertion in a host cell gene, a fusion transcript is made between the endogenous gene up to the insertion site and the β galactosidase-Neomycin (βgeo) marker (Figure 1.10). Their initial goal was to identify genes implicated in cardiovascular and pulmonary diseases. In addition to generating loss-of-function alleles, gene trap vectors offer a variety of post-insertional modification strategies for the generation of other experimental alleles. The insertion loci and the targeted gene are identified by RT-PCR by amplifying and sequencing the exons adjacent to the insertion. The exact locus of insertion can then be identified by a BLAST analysis and added to the database.
Figure 1.10. Gene alteration using a genetrap strategy. (A) The genetrap vector used to ‘trap’ Tln2 in mouse ES cells by BayGenomics was pGT01xf. It contained 1.5 kb of Mouse En2 intron 1, the splice acceptor site of mouse En2 exon 2 (SA), the fusion of β-galactosidase and neomycin transferase (β-geo ) and the SV40 polyadenylation signal (pA). It also contains various recombination sites to allow post insertional manipulations at the locus. (source: http://www.sanger.ac.uk/PostGenomics/genetrap/vectors/) (B) Gene targeting by genetrapping. Insertion of the βgeo cassette in an active gene leads to the expression of a fusion protein between the N terminal part of the protein expressed by the endogenous gene and the βgeo reporter gene. This can simultaneously generate a hypomorphic or null allele. Exons are shown in green, intronic regions by the black thin lines. The protein expressed by the endogenous gene is in blue, the βgeo reporter gene is indicated in purple.
1.5. Project's aims

1.5.1. Analysis of talin2 RNA and protein expression

(i) To determine in what tissues and at what stages of development talin2 transcripts and protein are expressed in mouse and human using Northern blot analysis as well as genomic database analysis.

(ii) To analyse talin2 expression in vivo in the developing embryo and in the adult mouse using Talin2 in a gene trap system (β galactosidase).

1.5.2. Generation of a conditional knockout allele for Tln2

(i) To establish a restriction map the 5' end of the Tln2 gene in order to design a cloning strategy for the targeting vector

(ii) To construct the targeting vector using the plasmids containing loxP sites and a selection marker

(iii) To generate Tln2\textsuperscript{fl/}\textsuperscript{+} ES cells by electroporation of the targeting vector into ES cells, screening of clones under Neo\textsuperscript{R} selection and identification of homologous recombination events

(iv) To generate Tln2\textsuperscript{fl/}\textsuperscript{-} mice and MEFs and use them for in vivo and in vitro Cre mediated recombination
Chapter 2 - Materials and Methods

All the chemicals used for the different protocols were from SIGMA unless otherwise stated. All enzymes were from New England Biolabs unless otherwise stated.

2.1. Preparation and analysis of DNA

2.1.1. Precipitation of DNA

DNA was precipitated to clear the samples before they could be used for experiments. Three volumes of 100% (v/v) ethanol of the sample to precipitate was added and 1/10th volume of 3M sodium acetate (pH5.5). The sample was then centrifuged for 10 minutes at 13,000rpm in an Eppendorf bench microfuge. The supernatant was carefully removed and 500μl of 70% ethanol (v/v) added to the pellet. The sample was centrifuged for 5 minutes in the same conditions, the supernatant was removed carefully and the pellet was left to air dry for 15 minutes, until all traces of ethanol had evaporated. The sample was finally resuspended in sterile autoclaved TE (10mM Tris-HCl pH8, 1mM EDTA pH8).

2.1.2. Agarose Gel electrophoresis

To separate and visualise the DNA, samples were loaded in wells of an agarose gel. Agarose was dissolved in 1X TAE (40mM Tris base, 20mM NaOAc, 1mM EDTA pH7.4). Its concentration varied from 0.8% to 2% (w/v) depending on the size of the DNA fragments to analyze. The agarose was melted in the microwave oven and poured in a gel mould containing a gel comb and left to set. The gel was then put in a gel tank containing 1X TAE. The samples were loaded with 6X loading dye (0.5% OrangeG, 30% glycerol (v/v)) as well as a marker (1kb+ DNA Ladder, New England Biolabs) as a size reference. The gel was run at 100V for 45 minutes to an hour until the DNA was separated enough. To visualize the DNA, the gel was stained in 5μg/ml Ethidium bromide in 1 X TAE and visualised by using the UV lamp of a Gel Doc system (Biorad) and visualised on computer using Quantity One software.
2.1.3. Bacterial culture of DNA in DH5α

a) Plasmid DNA
LB media: 10g Bactotryptone, 5g Yeast extract, 10g NaCl, up to 1 litre of sterile H₂O and autoclaved. A final concentration of 50μg/ml ampicillin was added, before the culture was inoculated.

LB agar plates: 10g Bactotryptone, 5g Yeast extract, 10g NaCl, 15g Agar up to 1 litre with sterile H₂O and autoclaved. The media was left to cool down to 50°C, the ampicillin was added up to a final concentration of 50μg/ml and the plates were poured under sterile conditions.

b) Cosmid DNA
To grow DH5α transformed with the talin2 C10 cosmid, the same LB media as for plasmid DNA was used. A final concentration of 50μg/ml kanamycin sulfate was added, before the culture was inoculated.

c) Bacterial culture
One colony was picked from an agar plate with a sterile tip and grown in 3ml LB medium, in a 15ml sterile tube overnight at 37°C in the shaker. This culture was when necessary diluted 1/1000 and grown in a larger flask for preparation of large amounts of plasmid or cosmid DNA.

2.1.4. Alkaline Mini prep of plasmid DNA
A single isolated colony was picked from an LB agar plate and grown in a 5ml culture of LB medium and 100μg/ml of ampicillin. This was centrifuged in a microfuge tube for 4 minutes at 13,000rpm. The supernatant was discarded. The pellet was resuspended in 100μl of P1 (50mM Tris-HCl, 10mM EDTA, 100μg/ml RNase-A). The cells were lysed with 200μl of P2 solution (200mM NaOH, 1% SDS) for 5 minutes. Then, 150μl of chilled solution P3 (3M potassium acetate, pH 5.5) was added, mixed and the tube was left for 5 minutes at room temperature. The tube was centrifuged at 13000rpm for 10 minutes. The supernatant was carefully removed and put in a fresh tube. The DNA was then ethanol precipitated as following the method described in section 2.1.1 and resuspended in 50μl sterile TE.
2.1.5. Midi prep of plasmid DNA (QIAgen)

The QIAgen plasmid Midiprep kit was used to extract plasmid DNA from a 100ml overnight culture containing the appropriate selection following the manufacturers' instructions. The principle of the extraction is similar to the alkaline miniprep described in section 2.1.4. using the same lysis solutions. However after all 3 solutions have been added following the same protocol, the DNA sample was purified through a QIAfilter column.

2.1.6. Midi Preparation of cosmid DNA

A single colony was inoculated in a 5 ml volume of LB media containing 50μg/ml kanamycin sulfate. This culture was then used to inoculate a 200ml culture overnight at 37°C. The midi prep protocol described above was used to purify the cosmid DNA.

2.1.7. Caesium chloride plasmid DNA preparation

A 500 ml culture of bacteria was grown overnight at 37°C, in sterile LB containing 30μg/ml of ampicillin. The culture was then centrifuged at 6,000rpm for 10 minutes. The solutions used were the same as the ones used in the alkaline mini prep protocol. The pellet was thoroughly resuspended in 10ml of P1. 20ml of freshly prepared P2 was then added, the contents mixed and the samples incubated for 5 minutes on ice. Then 15ml of P3 were added, the content mixed and the tubes were left on ice for further 5 minutes. This was centrifuged at 9,000rpm for 15 minutes. The supernatant was transferred to a clean centrifuge tube containing 50ml of ice-cold isopropanol, mixed well and centrifuged again at 9,000rpm for 15 minutes. The supernatant was removed and the pellet dried by inverting the tube on a tissue for 30 minutes to 1 hour. The pellet was resuspended in 5.5ml of TE. The DNA was transferred to a 15ml centrifuge tube and 6g of caesium chloride (CsCl), 250μl of a 10μg/ml solution of ethidium Bromide and 250μl of TE were added and the contents mixed. The tube was placed in the dark (aluminium foil) to avoid cross-reaction between the DNA and the ethidium bromide, for a few minutes to dissolve and was then centrifuged at 4,000rpm for 5 minutes. The clear red solution was transferred into two Beckman Quickseal ultracentrifuge tubes. The tubes were balanced with 1.1mg/ml of CsCl in TE and heat-sealed. A centrifugation at 100,000rpm (Beckman ultracentrifuge) for 12 to 16 hours at 20°C generated a CsCl gradient in the centrifuge tube in which super coiled plasmid DNA and genomic DNA were separated from each other. To extract the plasmid band, the tube was carefully
pierced on the top with a 27G needle to provide a vent. Then it was pierced under the plasmid band and the lower of the two bands was removed (the smaller upper band contains genomic DNA). The ethidium bromide was then extracted using water-saturated isobutanol (50% sterile H₂O, 50% isobutanol) and all the ethidium bromide had disappeared. The DNA was finally precipitated by adding an equal volume of water and 6 volumes of 100% ethanol, centrifuging at 11,000rpm for 30 minutes at 4°C and washing the pellet in 70% (v/v) ethanol. The DNA was resuspended in 500μl TE.

2.1.8. Quantification of nucleic acids
To measure the concentration of nucleic acids in samples, the samples were diluted in a 500μl of sterile water, at a 1/500 or 1/250 dilution. The samples were transferred into Quartz cuvettes and the Optic Density was measured using an eppendorf Biophotometer. The concentrations were calculated by measuring the OD at 260nm, 1OD unit corresponding to 50μg/ml for double stranded DNA, 40μg/ml for RNA and 33μg/ml for single stranded oligonucleotides.

2.1.9. Sequencing of DNA
All DNA sequencing was performed by Lark Technologies.

2.2. Polymerase Chain Reaction (PCR)
All PCR enzymes were from Roche unless otherwise stated. All PCR cycling reactions were performed in a DNA Engine DYAD™. All reactions were set up using sterile filtered tips.

2.2.1. Oligonucleotides synthesis
The oligonucleotides were synthesised and quantified by Invitrogen custom primers service. They were synthesised on a 50nm scale and dispensed in a lyophilised pellet in a 2ml microcentrifuge tubes. The pellet was resuspended in sterile H₂O to a stock concentration of 200μmol/L. When required, the primers were diluted to a working concentration of 20μmol/L.
2.2.2. Primer sequences

**Talin2 conditional knockout**

- KpnLarml.F: 5’- GGGGT ACC AG ATCCT AGCCT AT AC AGGG-3’
- KpnLarm2.F: 5’- GGGGTACCTGGGCTCTTAAGAGTTCTGGGTGGAACCT-3’
- ClaLarml.R: 5’- CCATCGATAATGGCCGACACTTCCCACCACACATC-3’
- Sal1Marm1.F: 5’- ACGCGTCGACCATTCTGTCTCCTCTGTCC-3’
- Apa1Marm1.R: 5’- TTGGGCCCCCAGTCAGGATCTCAGTATG-3’
- XhoRarm1.F: 5’- CCGCTCGAGGCTGGGAGTCAATTTTA-3’
- XhoRarm3.R: 5’- CCGCTCGAGTTAAGAGTGAGTGCCATCAGTGGGAA-3’
- gT2In2F3: 5’- GACATAAATCCAAAGGCATAAGCTGC-3’
- 3’Neorev: 5’- CTCTTGCTGCTGCTTGTGTC-3’
- gT2In3F: 5’- GCTCATGTCTGTAGTCCAGGAAAT-3’
- gT2In4R: 5’- GCCAGAGCTACATGGAGTGTCATGAA-3’
- gT2In2F4: 5’- TCCAGAAGGCGAAGGAGCAAGCGC-3’
- gT2In3R: 5’- CAAACTGAATGAAGGCCCAACAG-3’
- gT2In3R: 5’- CAAACTGAATGAAGGCCCAACAG-3’

**Talin2 cDNA Library screening**

- hT2-1F: 5’- TGCCGTCACGGAAAGGTCT-3’
- hT2-2F: 5’- TCCGAGGGTTCCAGGCACA-3’
- hT2-3F: 5’- CCACAGGACAACCCCAGATCCACGC-3’
- 3’probeF: 5’- GTGGATCCAGAAGACCCAAC-3’
- hT2-1R: 5’- CTCTGAATCCTGGTCGGCCT-3’
- hT2-2R: 5’- ACCATCACCTTGGCAGCCC-3’
- hT2-3R: 5’- TGGAGGCCACACCAACTGCG-3’
- hT2-4R: 5’- CCACAAGGTCTGGCATGCCC-3’
- hT2-6R: 5’- TTGATGGCCAGAGGAGCTGC-3’
- hT2-8R: 5’- AGCTTGCCAGTTGTGTCACCTT-3’
- SP6-2: 5’- AGGTACCGGCTCCGGGATCC-3’

**Talin2 genetrap**

- mT2Ex26F1: 5’- GCACAAAGCCTCCAGGAG-3’
- mT2Ex26F2: 5’- CAAGCACAAGGCCCTCCAG-3’
2.2.3. Amplification of the talin2 targeting vector arms

To amplify the talin2 targeting vector arms, the reactions were set up in large scales in order to obtain enough DNA for further cloning step. The 5’ arm of the talin2 targeting vector (primers KpnLarm1.F and ClaLarm1.R, product 1.2kp) and the vector to provide a PCR positive control (primers KpnLarm2.F and ClaLarm1.R) were amplified, using wild type genomic DNA extracted from embryonic stem cells. The middle arm (primers Apa1Marm1.R and Sal1Marm1.F, product 722bp) and 3’ arm (primers XhoRarm1.F and XhoRarm3.R, 2.9bb) were amplified using DNA extracted from the talin2 cosmid C10. All the reactions were set up as triplicates of 50μl reactions, containing 4.5μl of 11.1X buffer, 1 μl of each primer (stock 20μmol/L), 1μl of Expand DNA Polymerase, 1μl of DNA template and up to 50μl of sterile H2O.

11.1X PCR buffer: 334μl of Tris-HCl pH8.8, 2M; 166μl of ammonium sulfate 1M; 67μl of MgCl2 1M; 7.2μl β-mercaptoethanol 100%, 68μl EDTA 10mM pH8.0; 150μl each dNTP 100mM; 170μl BSA 10mg/ml.

For the 5’ arm, middle arm and positive control arm, the cycle was set up as follows: 94°C for 30 seconds the 35 cycles of (94°C for 15 seconds, 60°C for 30 seconds, 72°C for 1min30 seconds) followed by 72°C for 5 minutes.

For the 3’ arm, the cycle used was 94°C for 30 seconds then 35 cycles of (94°C for 15 seconds, 60°C for 30 seconds, 72°C for 4 minutes) followed by 72°C for 5 minutes.

The samples were then electrophoresed onto a 1% agarose gel. Once the size of the fragment was checked, the samples were put through a Qiagen column for purification.

2.2.4. Screening of talin2 homologous recombinants

Two primers were used (gT2In2F3 and 3’Neorev) to amplify a region between talin2 and the inserted NeoR cassette. The reactions were set up in a 10μl final volume, containing 0.9μl of 11.1X buffer, 0.5μl of each primer, 0.25μl of Taq Polymerase, 1μl of DNA sample and up to 10μl with H2O. The reactions were set up using the following
Chapter 2 - Materials and Methods

conditions: 94°C for 30 seconds followed by 35 cycles of (94°C for 15 seconds, 60°C for 30 seconds, 72°C for 30 seconds) followed by 72°C for 5 minutes. The samples were then electrophoresed onto a 1% (w/v) agarose gel.

2.2.5. Identification of the 3′loxP in the talin2 targeting event

Two primers were used (gT2In3F and 3' gT2In4R) to amplify a region across the 3′loxP site in the talin2 targeting event. The reactions were set up in a 15µl final volume, containing 1µl of each primer, 1µl of DNA sample and up to 15µl of PCR Readymix (ReadyMix PCR MasterMix, ABGene). The reactions were set up using the following cycle: 94°C for 30 seconds followed by 35 cycles of (94°C for 15 seconds, 60°C for 30 seconds, 72°C for 30 seconds) followed by 72°C for 5 minutes. The samples were then electrophoresed onto a 2% (w/v) agarose gel.

2.2.6. PCR from cDNA

For all the different primer combinations to map the talin2 kidney transcript start site, the reactions were set up in a 15µl final volume, containing 1µl of each primer, 1µl of DNA sample and up to 15µl of PCR Mastermix (ReadyMix PCR MasterMix, ABGene). The reactions were set up using the following conditions: 94°C for 30 seconds followed by 35 cycles of (94°C for 15 seconds, 60°C for 30 seconds/kb, 72°C for 30 seconds) followed by 72°C for 5 minutes. The samples were then electrophoresed onto a 2% (w/v) agarose gel.

2.2.7. Mouse genotyping: PCR analysis on lysed tail DNA

A set of primers was designed for each genotyping strategy using Vector NTI sequence analysis software. In all cases the reactions were set up in a 15µl final volume, containing 1µl of each primer, 1µl of DNA sample and up to 15µl of PCR Mastermix. The reactions were set up using the following condition: 94°C for 30 seconds followed by 35 cycles of (94°C for 15 seconds, 60°C for 30 seconds, 72°C for 30 seconds) followed by 72°C for 5 minutes. The samples were then electrophoresed onto a 2% (w/v) agarose gel.

2.3. Cloning

All the restriction and cloning enzymes were from New England Biolabs unless otherwise stated.
2.3.1. Restriction digest of plasmid DNA

Plasmid DNA was digested with restriction enzymes for cloning purposes. In a total volume of 100μl, 10μg of DNA was digested, with 10μl of appropriate buffer, 2μl of restriction enzyme and H₂O was added up to 100μl. The reaction was left in a 37°C incubator overnight. For analytical purposes, plasmid DNA was digested in a total volume of 20μl containing 2μl of the appropriate buffer, 1μg of the DNA, 1μl of the chosen restriction enzyme and sterile H₂O was added up to 20μl. All restriction digests were left in a 37°C incubator for 2 hours.

2.3.2. Gel Purification

DNA fragments were extracted from an agarose gel using the QIAgen gel purification kit following the manufacturer’s instructions. The band of DNA was carefully removed from the agarose gel under a UV lamp.

2.3.3. Plasmid dephosphorylation

If a plasmid used for a cloning step had been linearised by a single enzyme, it needs to be dephosphorylated to avoid religation to itself. The phosphatase used was the Calf Intestinal Phosphatase (CIP). 1μl of CIP was added to the restriction digest reaction followed by incubation at 37°C for 1 hour. The enzyme was then heat-inactivated by incubating the reaction at 75°C for 10 minutes.

2.3.4. Ligation

The total ligation reaction was set up in a 20μl final volume. It contained the vector and the insert with a molecular ratio of 1/3 (vector/insert) with approximately 100ng of vector, 2μl of 10 X T4 DNA ligase buffer, 1μl of T4 DNA Ligase and up to 20μl of sterile H₂O. The reaction was then placed at 16°C overnight. A negative control reaction was also set containing only the vector DNA, 2μl of 10 X T4 DNA Ligase buffer, 1μl of T4 DNA ligase and up to 20μl of sterile H₂O.

2.3.5. Transformation of plasmid into DH5α

1-5μl of the ligation reaction was placed on ice with 100μl of commercial DH5α (Invitrogen), in a polypropylene tube, for 30 minutes. The reaction was then heat-shocked at 42°C for 45 seconds and put back on ice for 2 minutes. 1 ml of LB medium was then added and the tubes were placed in a 37°C shaking incubator for 1
hour. 150µl of this was plated on LB plates containing 100µg/ml ampicillin using a
glass spreader sterilised with IMS and flamed.

2.3.6. Plasmids

The \textit{ploxP(pgkNeopA)loxPSallx2} and \textit{ploxP:KS3(+)} plasmids were a gift from R.
Fassler (Max Planck Institute of Biochemistry, Martinsried, Germany). \textit{pBSIISK+} was
from Stratagene, \textit{pGEMTEasy} was from Promega. The \textit{pCrePAC} plasmid was provided
by a gift from the Institute of Cancer Research, London. The \textit{Lawrist7} plasmid was
from the Resource Centre of the German Human Genome Project (RZPD) and was the
backbone vector in which the \textit{Tln2} genomic cosmid was cloned. The restriction maps of
the plasmids used are showed on Figure 2.1.
Figure 2.1. Restriction maps of plasmids
2.4. Southern blot analysis of genomic DNA

2.4.1. Southern Blot

Cosmid DNA was digested following standard methods of restriction digest described in section 2.3.1. After digestion, the samples were ethanol precipitated and resuspended in 20μl of sterile 1X TE. They were then loaded on a 1% (w/v) agarose gel. The DNA was visualized as normal under a Gel Doc system (Biorad) after ethidium bromide staining. A picture of the gel was taken using a ruler to have a reference for the distances of the bands from the wells. If the DNA to be blotted was genomic DNA, the gel was first washed in 0.25M HCl for 10 minutes and then washed twice in 0.4M NaOH for 30 minutes to denature the DNA. A Southern blot was set using sponges flooded with transfer buffer (0.4M NaOH) in a box. A piece of gel blot (10MM Whatman) paper cut in the same size of the gel was put on top of the sponges, followed by the gel inverted and above it a Zeta-Probe GT genomic tested blotting membrane (Biorad) also cut to the size as the gel, followed by a second piece of gel blot paper. Finally, paper towels were put on top of the set to a height of 15 cm and 100g of weight was added on top. The blot was set overnight. The next morning the membrane was washed for 10 minutes in 2X SSC (1 litre of 20X SSC: 175.3g NaCl, 88.0g trisodium Citrate, pH7) and baked at 80°C for 1 hour.

The membrane was first prehybridised at 37°C for 2 hours in Church Buffer (0.5M NaHPO4, 7% SDS, 0.5mM EDTA). The labelled probe was then added and the membrane was hybridised at 42°C overnight. The membrane was then washed 3 times for 5 minutes in 6X SSC, 0.1% SDS at room temperature and exposed to an X-Ray film overnight.

2.4.2. 5'end labelling of oligonucleotides and hybridisation

A reaction containing the following components was set up: 100pmol of oligonucleotide, 5μl of 10X kinase buffer, and up to 50μl of sterile H2O. The reaction was transferred to ice, 1 μl of T4 polynucleotide kinase, and 1μl of γ-P32-ATP were added in a radiation-designated area, behind a perpex screen. The reaction was then left to incubate at 37°C for 1 hour. A stock solution containing 1μl glycogen (20 mg/ml stock) with 200μl of 0.2M EDTA pH8 was prepared, and 2μl of this mixture were added to the labelling reaction. The labelled oligonucleotide was finally precipitated by adding 5μl of 4M LiCl and 150μl of prechilled 100% ethanol, and left at -20°C for 2
hours. The sample was then centrifuged for 10 minutes at 13,000rpm. The supernatant was removed and discarded in the cin bin. The pellet was washed using 500μl 70% Ethanol, and after centrifugation at 13,000rpm for 5 minutes, the supernatant was carefully removed and the pellet was left to air dry for 5 to 10 minutes until all the ethanol had evaporated. The pellet was resuspended into 50μl of sterile H$_2$O.

**Sequences of oligonucleotides**

5'T2For  5'-AGTCATCCATGAAAATGGTGCC-3'
mT2-1F   5'-TAATCCAGTGCCGACCTGCTTC-3'
mT2 -4F  5'-TCACGATCTGCACGAGTAGG-3'

2.5. Gene expression analysis

2.5.1. RNA Extraction from cultured cells

RNA was extracted from a 10cm plate of Tln2 genetrap ES cells using the RNAeasy column (QIAgen) kit following the manufacturer's instructions.

2.5.2. Northern Blot

The commercial MTN (Multiple Tissue Northern from Clonetech) was prehybridised for 30 minutes at 68°C in the Express-Hyb (Clonetech) solution included in the kit. The membrane was put into some fresh Express-Hyb solution and the labelled probe was added. The hybridisation was left at 68°C for 1 hour. The membrane was then washed twice at room temperature in 2X SSC, 0.1%SDS for 30 minutes, followed by a last wash in 0.2%SSC, 0.1%SDS at 55°C. The membrane was then exposed to X-Ray film, which was developed after a suitable time of exposure.

The membrane was stripped by putting it in 0.5% SDS at 90°C and left until it had gone down to room temperature, and could be used for further analysis.

2.5.3. Random primed labelling of probe

The probe was synthesised by PCR or digest of plasmid DNA. After 50-100ng of probe were denatured at 95°C for 5 minutes, the Ready-to-go labelling beads (Amersham Pharmacia) were used for labelling with $\alpha$-32P dCTP and the ProbeQuant50 columns (Amersham Pharmacia) to separate the labelled DNA from the incorporated radioactivity following the manufacturers’ instructions. The probe was then denatured just before hybridisation by incubation at 95°C for 5 minutes.
2.6. Tissue Culture

All tissue culture reagents were from Invitrogen, unless otherwise stated.

2.6.1. Growth media and general protocols

All the plasticware used for cell culture was from Helena Bioscience. All cell types were grown at 37°C and in 10% CO₂.

a) Mouse Embryonic Fibroblasts (MEFs)

Dulbecco’s Modified Eagle Media (DMEM) with L-Glutamine, 1000mg/L D-Glucose, sodium pyruvate supplemented with:
- 10% v/v foetal calf serum (Sigma)
- 100U/ml penicillin, 100µg/ml streptomycin 100X
- 20mM L-Glutamine

b) Embryonic Stem (ES) Cells E14.1a

DMEM with L-Glutamine, 4500mg/L D-Glucose supplemented with:
- 15% (v/v) Stem cell tested foetal calf serum (LabTech)
- 20mM L-Glutamine
- 100U/ml penicillin, 100µg/ml streptomycin 100X
- 1mM non essential aminoacids
- 10mM sodium pyruvate
- 115µM β-mercaptoethanol
- 1ml recombinant Leukaemia Inhibitory Factor (LIF)

LIF was made by a calcium phosphate transient transfection of COS7 cells with a LIF expression vector, followed by harvesting the supernatant after 48hours. It was provided by Sarah Munson (Wellcome trust ES cell facility, University of Leicester).

ES cells were grown on mitomycin treated Mouse Embryonic Fibroblasts (MEFs) feeders.

c) Talin2 Genetrap cell lines

Glasgow MEM (G-MEM) (BHK-21) with L-Glutamine, without Tryptose Phosphate Broth supplemented with:
- 15% (v/v) Stem cell tested foetal calf serum (LabTech)
- 20mM L-Glutamine
Chapter 2 – Materials and Methods

100U/ml penicillin, 100μg/ml streptomycin 100X
1 mM non essential aminoacids
10 mM sodium pyruvate
115 μM β-mercaptoethanol
1 ml recombinant Leukaemia Inhibitory Factor (LIF)
The cells were grown on a dish coated with 0.1% (w/v) gelatin, in a 37°C incubator with 10% CO2.

2.6.2. Growing MEFs and Mitomycin C treatment

A vial of an early passage of MEFs was thawed out and plated onto a 10 cm dish with MEF media. The cells were grown and expanded onto at least 16 15 cm plates. They were then treated with \(2 \mu g/ml\) of Mitomycin C in media for 2 hours, and washed with PBS 3 times very carefully. They were trypsinised and placed into a 50ml falcon tube. The cells were then counted and frozen down in aliquots of \(10^6\) or \(2 \times 10^6\) cells. These cells were used as feeders to grow E14.1a ES cells as follows.

<table>
<thead>
<tr>
<th>Size of plate</th>
<th>Number of MEFs / well</th>
<th>(10^6) MEF vials per plate</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 cm</td>
<td>(1 \times 10^6)</td>
<td>1</td>
</tr>
<tr>
<td>6 cm</td>
<td>(5-6 \times 10^5)</td>
<td>0.5</td>
</tr>
<tr>
<td>6 well</td>
<td>(4 \times 10^5)</td>
<td>2</td>
</tr>
<tr>
<td>12 well</td>
<td>(2.5 \times 10^5)</td>
<td>3</td>
</tr>
<tr>
<td>24 well</td>
<td>(1 \times 10^5)</td>
<td>2.5</td>
</tr>
<tr>
<td>48 well</td>
<td>(4 \times 10^4)</td>
<td>2</td>
</tr>
<tr>
<td>96 well</td>
<td>(1.4 \times 10^4)</td>
<td>1.5</td>
</tr>
</tbody>
</table>

2.6.3. Trypsinisation

The cells were first washed with sterile 1X phosphate buffer saline (stock of 20X PBS: 160g NaCl₂, 4g KCl, 18g of Na₂HPO₄). Trypsin at 1mg/ml (in 1X PBS) was then added to the cells and the cells were incubated at 37°C for 5 minutes. The cells in trypsin were then transferred to a fresh tube with 5 ml of growing media to inactivate the trypsin, centrifuged at 1,300rpm (IEC Centra CL3) for 4 minutes. The supernatant was carefully removed and the pellet was resuspended and replated in the appropriate amount of media.


2.6.4. Freezing and storage

The cells to be frozen were trypsinised with 1ml of 1X trypsin EDTA as indicated above and 1ml of media was added to inactivate the trypsin. The cells were centrifuged at 1,300rpm (IEC Centra CL3) for 4 minutes and the supernatant carefully removed. The cells were resuspended in an appropriate volume of 1X freezing media (10% (v/v) DMSO, 45% (v/v) DMEM and 45% (v/v) FCS) and the cells were aliquoted in 0.5ml or 1ml aliquots. They were finally placed in a freezing container that allows slow freezing (NALGENE Cryo 1°C) at -80°C and later transferred to liquid nitrogen.

2.6.5. Electroporation of ES cells

ES cells were grown onto three 10cm dishes. The media was removed and the cells were washed with 1X PBS. The cells were then trypsinised using 5ml of 1X trypsin and incubated at 37°C for 10 minutes. The cells were then transferred to a fresh sterile tube and the plate was washed with 5ml growing media to collect the remaining cells. The cells were then counted and centrifuged at 1300 rpm for 4 minutes. The supernatant was removed and the cells were gently resuspended in PBS in a final volume of 0.8ml. 20μg of the targeting vector was linearised using a single restriction digest, ethanol precipitated and resuspended under the hood in sterile H₂O. The cells were transferred into an electroporation cuvette (BioRad, 0.4mm) and the targeting vector was added to the cells. The cells were electroporated at 0.250kV, 500μF. The cells were finally transferred to 3 10cm dishes (with MEF feeders) each containing 10ml of ES growing media and placed at 37°C overnight. The next day, the media was removed, the cells washed twice with PBS and fresh media containing G418 at 250μg/ml was added. The media was then changed everyday for 6-7 days.

2.6.6. Clone picking and freezing

The media was removed and the cells were left in PBS. Approximately 300 colonies were picked under the microscope in a 10μl volume. Each colony was transferred to a 96well dish, and 30μl of 2X trypsin was added to each well. The colonies were then incubated at 37°C for 15 minutes and transferred to a 48 well dish, on feeders with 500μl of media. The clones were then fed everyday with media containing no selection for 5-6 days. When the cells were at least 50% confluent the media was removed, and the cells were washed in 1X PBS. They were then trypsinised in a 75μl total volume, of
which 15μl was plated on a 12 well dish coated with 0.1% (w/v) gelatin for DNA extraction and 60μl was frozen down with 40μl of FCS and 100μl of 2X freezing media in a 96 well dish at -80°C.

2.6.7. DNA extraction for screening

Two different methods were used. The Gene Elute mammalian genomic DNA kit (SIGMA) was used to extract DNA from a 12 well plate following the manufacturers’ instructions. The cells were washed with 1X PBS, lysed with the lysis buffer provided by the extraction kit, and transferred to a DNA column. This was followed by two washes and centrifugations. Finally the DNA was eluted using 200μl of the provided elution solution into a fresh tube. The samples were then stored at 4°C until they were screened. Alternatively, the cells were washed with PBS and 500μl of DNA Lysis buffer (50mM Tris-HCl pH7.6, 1mM EDTA, 100mM NaCl, 0.2% (w/v) SDS) containing 100μg/ml of Proteinase K was added to the well. The cell lysate in the buffer was then removed, placed in a fresh 1.5ml eppendorf tube and incubated at 65°C for 2 hours. The DNA was then precipitated with 1ml 100% ethanol and 50μl of 3M NaOAc pH5.5. The DNA pellet was washed in 70% ethanol and resuspended in 500μl of 1X TE. If the DNA was not properly dissolved, the tubes were incubated at 37°C for 2 hours or at room temperature overnight.

2.6.8. Preparing ES cells for blastocyst injection

Positive colonies identified by screening were first thawed out a week before the injection date on two wells of a 6well dish containing feeders. The media was changed everyday until the cells were ready to be trypsinised (3-4 days). Each well was then split onto 2 other wells containing feeders. The morning of the injection the cells were washed with PBS and trypsinised in 1ml of 1X trypsin until single cell suspension, 1ml of media was added and the cells were centrifuged at 1,300rpm for 4 minutes and resuspended in 500μl of injection media. The cells were then taken to the Transgenic unit (TGU) on ice. The cells were then injected into C57BL6 bastocysts following the Home Office regulated procedures.

Injection Media: DMEM with 4500mg/L D-Glucose (15.2ml), FCS (ES cell batch tested, 4.0ml), L-Glutamine (0.4ml), sodium pyruvate (0.2ml), penicillin/streptomycin (0.2ml).
2.6.9. MEF transfection

MEFs to be transfected were expanded into 3 10 cm dishes. After washing with 1 X PBS and trypsinisation, the cells were counted and 1x10^6 cells were used for transfection. The DNA for transfection was prepared using the CsCl plasmid prep (section 2.1.7) and resuspended in sterile H₂O under sterile conditions in the fume hood. The Nucleofector Kit 6457 was used following the manufacturer’s instructions.

2.6.10. Staining of ES cells with X-Gal

All the solutions were prepared fresh on the day of the experiment, except the stock solution of 5X X-Gal base (0.4M disodium hydrogen phosphate anhydrous, 90mM sodium diHydrogen phosphate, 1H₂O, 10mM magnesium Chloride, 6H₂O).

Washing solution: Dilution of the 5X X-Gal base to 1X in H₂O.
Fixing Solution: 0.2% (w/v) glutaraldehyde, 50mM EGTA in 1X X-Gal Base.
Staining solution: 2mM MgCl₂, 5mM K₄Fe(N)₆, 5mMK₃Fe(N)₆; 1mg/ml 5-bromo-4-chloro-3-indolyl-β-galactoside (X-Gal) in 1X X-Gal base. The X-Gal was prepared in Dimethyl Formamide and stored in the dark at -20°C until needed.

The cells to be stained were split the day before (as indicated in section 2.6.3.) staining such as on the day of staining so they were about 70% confluent. The growing medium was carefully removed and the cells were washed with washing solution. They were then fixed for 10 minutes in the fixing solution at room temperature and then washed 3 times in the washing solution. The staining solution was then left overnight in the dark at 37°C. The cells were then washed 3 times and left in the fixing solution. They could be stored in the fixative at 4°C until needed.
2.7. Protein analysis

2.7.1. Antibodies

Western blot antibodies

<table>
<thead>
<tr>
<th>Antibody name</th>
<th>Protein raised against</th>
<th>Dilution</th>
<th>Secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD77</td>
<td>Talin (aa2464-2541)</td>
<td>1/5000</td>
<td>Anti mouse-HRP (1/3500)</td>
</tr>
<tr>
<td>C20 (Sigma)</td>
<td>Talin (20 Cterm aa)</td>
<td>1/1000</td>
<td>Anti goat-HRP (1/3500)</td>
</tr>
<tr>
<td>T2FP2</td>
<td>Talin2</td>
<td>1/500</td>
<td>Anti rabbit-HRP (1/3500)</td>
</tr>
<tr>
<td>F9</td>
<td>Vinculin (258-1066)</td>
<td>1/500</td>
<td>Anti mouse-HRP (1/3500)</td>
</tr>
</tbody>
</table>

2.7.2. Protein extraction from mouse tissues

The tissue was extracted from the animal, wrapped with aluminium foil and immediately transferred to liquid nitrogen to be frozen. It was then stored at -80°C until needed. When the protein was needed, the tissue was thawed in 2ml of buffer (10mN NaCl, 1mM EGTA, 50mM Tris pH8.0, 1% (w/v) SDS, 1mM AEBSF), cut into smaller pieces with a clean scalpel and homogenised twice for 30 seconds with a homogeniser. It was then transferred to a fresh microfuge tube (1.5ml), centrifuged for 5 minutes at 13,000rpm at 4°C. The supernatant was transferred to a fresh tube. The protein lysates were then aliquoted into 200μl samples and kept frozen at -20°C.

2.7.3. Protein quantification assay

The assay used to quantitate protein lysates from tissues was the BCA based assay (Reagent Kit, Pierce). The two working reagents (WR) provided in the kit were mixed (v/v 50 parts of A for 1 part of B). A standard curve was drawn using different concentrations of 2mg/ml BSA in 100μl with 2ml of WR mix. The BSA was diluted in 1% (w/v) SDS and 50mM TRIS pH8.0. BSA concentrations were measured at OD562nm and a standard curve drawn; 100μl of each protein sample was added to 2ml of the WR mix. The absorbance at OD 562nm was measured and the concentration was calculated by reference to the BSA standard curve.

2.7.4. SDS PAGE gel

The protein samples were electrophoresed on a SDS-Polyacrylamide gel using a BioRad running tank filled with 1X SDS running buffer (10X Stock: 250mM Tris-HCl, 1%
(w/v) SDS, 1.92M Glycine). 30μg of protein was loaded on a 7.5% separation gel along with high molecular weight protein markers (Sigma).

Composition of the Running gel:
- 1.875 ml of 30% (v/v) Acrylamide
- 2.8ml 1M Tris-HCl pH 8.8
- 2.625 ml H₂O
- 100μl 10% SDS
- 50μl 10% (w/v) Ammonium persulfate (APS)
- 20μl TEMED (N,N,N’N’-Tetramethylethylenediamine)

Composition of the Stacking gel:
- 66μl of 30% (v/v) Acrylamide
- 625μl 1M Tris-HCl pH 6.8
- 2.625 ml H₂O
- 50μl 10% (w/v) SDS
- 25μl 10% (w/v) APS
- 10μl TEMED

The two gel mixes were made up in parallel, but the TEMED was added just before pouring the gel. The separation gel was poured first and left to set, and then the stacking gel was poured on top of it with a comb. Once the gels were set, the plates were then put in the tank and the tank was filled with running buffer. The samples were then mixed with 6X loading buffer (for 20ml stock: 1.25 ml of Tris/HCl pH 6.8, 4ml 10% (w/v) SDS, 1ml β-Mercaptoethanol, 2ml glycerol and a few drops of bromophenol Blue) loaded into wells and electrophoresed at 100V for 30 minutes and then at 200V for 1 hour.

2.7.5. Western Blot analysis

The gels were electroblotted onto Hybond-C membranes (Amersham Biosciences) in semi dry transfer buffer (for 1 litre: 243g Glycine, 5.81g Tris-HCl, 0.375g SDS, 200ml ethanol) for 1 hour at 10V. The membrane was first stained in 0.5% (w/v) PonceauS (Sigma) to detect protein transfer. The position of the markers was labelled on the blot. The Ponceau was then washed off using 1X TBS, 0.1% (v/v) Tween20 (10X TBS stock: 60.5g Tris base, 90g NaCl, pH7.4). The non-specific sites were blocked by incubating the membrane in 5% Marvel (Skimmed milk powder) in 1X TBS-Tween for 1 hour at room temperature with shaking. The primary antibody was diluted with the appropriate concentration in the blocking buffer and incubated with the membrane overnight in the 4°C room with shaking. The blot was then washed 3 times in TBS-Tween for 10 minutes each at room temperature, with shaking. The appropriate secondary antibody diluted in blocking buffer, was then incubated for 1 hour at room temperature with the
membrane, with shaking in the dark. Finally, the blot was washed 3 times in TBS-Tween for 10 minutes each at room temperature, with shaking and in the dark. The membranes were the put on Saran paper and covered with detection solution. The detection solution was made of a 1:1 (v/v) mix of Supersignal West Pico Luminol Enhancer solution and Supersignal West Pico Stable peroxide solution (PIERCE) and left for 4 minutes. The membranes were then exposed to X-Ray (Sigma) film for different lengths of time between 30 seconds and 5 minutes to optimise the signal and developed in a Xograph (Compactx4) developer.

2.7.6. Immunofluorescence

Immuno- fluorescence antibodies

<table>
<thead>
<tr>
<th>Antibody name</th>
<th>Protein raised against</th>
<th>-dilution</th>
<th>Secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>ab1047</td>
<td>βGalactosidase</td>
<td></td>
<td>Anti mouse Texas Red</td>
</tr>
<tr>
<td>C20 (Sigma)</td>
<td>Talin (20 Cterm aa)</td>
<td>1/1000</td>
<td>Anti goat Texas Red</td>
</tr>
<tr>
<td>Phalloidin FITC</td>
<td>Actin</td>
<td>1/200</td>
<td>none</td>
</tr>
</tbody>
</table>

The cells to be analysed were split the day before onto coverslips that had been sterilised with IMS and flamed, and placed in 8 well dishes. The cells were counted and 2 x 10^5 cells were inoculated per well of an 8 well dish with 2ml of the appropriate growing medium. The following solution was prepared and used for the preparation of all the other reagents and to dilute the antibodies. PBS with magnesium and EGTA (PBSME) containing 3mM EGTA (Ethylene glycol-bis(2-aminoethylether)-N,N,N',N"-tetraacetic acid) 9mM MgCl_2 and up to 1 litre of PBS. The growing medium was carefully removed from the cells, which were then washed twice with 1ml of PBSME. The cells were then fixed for 10 minutes at room temperature in 1ml of 4% (w/v) paraformaldehyde (PFA) in PBS. The PFA was washed off and the cells were washed once in 2ml PBSME. The cells were permeabilised with 1ml of 0.2% (w/v) Triton X-100 in PBSME for 3 minutes. The triton was washed off by 3 washes in 2ml of PBSME each over 20 minutes with shaking. The cells were blocked in 1% (w/v) BSA for an hour at room temperature. In the mean time, the primary antibody was diluted in 0.2% (w/v) BSA in PBSME in a 150μl final volume and centrifuged for 5 minutes at 13,000 rpm at 4°C. After blocking in BSA, the primary antibody was applied for 20 minutes at room temperature and washed off by 3 washes in PBSME over 20 minutes with shaking. This was followed by incubation at room temperature in the dark with 150μl of
the appropriate secondary antibody coupled to a fluorophore diluted in 0.02\% (w/v) BSA in PBSME (previously centrifuged for 3 minutes at at 13,000 rpm at 4°C). The cells were washed 3 times over 20 minutes to 0.02\% (w/v) BSA with shaking and in the dark. Finally the cells were incubated with fluorescently labelled phalloidin diluted in 0.02\% (w/v) BSA for 10 minutes at room temperature and in the dark. The final wash was done in 2ml of PBSME with shaking and in the dark for 10 minutes and the coverslips were mounted on a microscope slide with 10\mu l of Antifade Kit (Prolong Antifade Kit, Molecular Probes), left at 37°C for 30 minutes and stored in the dark until observation.

The cells were visualised on a Nikon TE300 inverted fluorescence microscope using a x100 lens (Hoffman optics), the images were captured using a Hamamatsu Orca ER camera and the images were analysed using Openlab 3.1 software.

2.8. Animal work

2.8.1. Lysis of tail samples for DNA extraction

A small piece of tail was removed under local anaesthesia by the staff in the biomedical services (BMS). Then, 75\mu l of denaturation buffer (25mM NaOH, 0.2mM EDTA, pH12) was added to tissue in 0.5ml microfuge tube, followed by an incubation at 95°C for 30 minutes. The tails were then cooled down to 4°C on ice and 75\mu l of neutralising solution (40mM TrisHCl pH5) were added. The sample was then ready for PCR genotyping.

2.8.2. Generation of mouse embryonic fibroblast cell lines

The mother was culled by cervical dislocation by the staff in BMS. The mouse abdomen was then swabbed in ethanol, the abdominal wall was then cut and the uterus located at the posterior. The uterus was dissected out and each embryo (in its yolk sack) separated and transferred into cold PBS.

Each embryo was transferred to a petri dish containing cold PBS and was held on ice until they were dissected. The yolk sac was removed, the head dissected off and the liver dissected out. A little piece of the tail was removed for genotyping. The tissue was left to settle and the PBS carefully removed and the tissue was washed once more with 5ml of ice cold PBS. The PBS was removed and replaced with 2ml of 5X Trypsin. This was followed by incubation at 4°C for 6-18 hours (overnight) then at 37°C for 20 to 30
minutes. Finally, 5 ml of warm medium was added and the cells transferred to a 10 cm dish containing another 5ml of media. The mouse embryonic fibroblasts generated were grown at 37°C, 10% (v/v) CO₂ until confluent.

2.9. Histology
All the histology consumables were from RA Lamb unless otherwise stated.

2.9.1. Slides preparation
Microscope slides were racked up into metal racks. They were then soaked in 5% Decon overnight. The following day, they were washed for 30 minutes in hot running water, and washed in MilliQ water 5 times for 5 minutes. The slides were then left to dry out in the oven at 60°C for 30 minutes. This was followed by the subbing procedure, which were successive submersions for one minute in each of the following tanks, while agitating: Subbing solution (2% 3-aminopropyltriethoxysilane in acetone, 8ml in 400ml), Acetone (twice), MilliQ water (twice). Finally, the slides were dried in the oven and were ready for use.

2.9.2. β-Gal staining of whole mount embryos
The embryos were dissected out of the uterus under a Nikon SMZ-U microscope, and separated from their extraembryonic membranes and transferred into fixative straight away. The embryos aged up to 9.5dpc were dissected by Dr.Sue Monlkey. The embryos were fixed in 0.2% (v/v) Gluteraldehyde in PBS. The fixation was performed at room temperature and the fixation time depended on the size of the sample: for early embryos: 10-15 minutes, for embryos up to 13dpc, 15-30 minutes. When examining many embryos at one time, they were placed individually in wells of a 24 well tissue culture dish the fixative was carefully aspirated with a pasteur pipette. The embryos were then rinsed in a buffer containing detergent (0.1% deoxycholate, 0.02% NP40), 3 times at room temperature for 15-30 minutes each. The embryos were then stained in 1mg/ml 5-bromo-4-chloro-3-indolyl-β-D galactoside (X-Gal) in PBS containing 2mM MgCl₂, 5mM K₄Fe(N)₆, 5mM K₃Fe(N)₆ at 37°C in the dark. The staining time depended on the size of the sample. The samples were incubated at 37°C in the dark: small samples for 1-3 hours, older embryos for 4-5 hours or longer. After staining, the embryos were stored at 4°C in 70% ethanol. The embryos were then dehydrated through alcohol series (70% ethanol, 90% ethanol, 95% ethanol, 100% ethanol) and transferred
to xylene and embedded in wax. The embedding process was performed by Jenny Edwards (MRC Toxicology Unit). The samples were then sectioned using a microtome into 10μm sections.

2.9.3. Fresh tissue extraction and sectionning

The tissues were dissected from the animal straight away after culling. Each tissue was transfered into a labelled plastic holder for frozen section containing the OCT embedding solution; more OCT was added on top until the piece of tissue is covered. The sections were transferred immediately to a hexane-dry ice bath to freeze instantly. When all the tissues of interest were frozen, they were transferred to –20°C and left for 1 hour at least. The sections were cut using a cryostat (BRIGHT, model OTF). Between 5 and 10 sections of 25μm were cut for each tissue on an appropriately labelled slide. The slides were then left to dry at room temperature for an hour. Some slides were then fixed for 15 minutes, washed twice for 5 minutes and stained with X-Gal in the same conditions as for the embryo staining (section 2.9.2.) overnight at 37°C in the dark. The slides were rinsed thoroughly in water and transferred into a succession of IMS baths to be dehydrated: 70% IMS for 5 minutes twice, 90% IMS for 5 minutes twice, 100% IMS for 5 minutes twice. They were finally rinsed in xylene twice for 10 minutes in a glass container. The sections could then be mounted by adding a drop of mounting media on the section and a coverslip on top of it. The slides could then be visualised on the microscope and photographs taken using a Nikon TMSF inverted microscope and visualised using a Leica DMLB digital camera.
Chapter 3. Analysis of Talin2 expression

3.1. Introduction

Numerous studies on the biochemistry of the talin protein have demonstrated its key role in integrin mediated adhesion. It is thought to be an important player in the initiation of the link between the extracellular matrix and the actin cytoskeleton inside the cell and has been shown to be present in a variety of adhesion structures in cultured cells as well as in physiological contexts (reviewed by Critchley, 2004).

Talin1 is a 270kDa protein encoded by an 8kb transcript widely expressed in mouse tissues. Analyses of EST databases indicated the presence of a talin related gene named \textit{Tln2} (Monkley et al., 2001). The talin2 protein in human shares 75% identity and 86% similarity with talin1. The degree of similarity is higher in key regions of the proteins such as the FERM domain in the head region as well as the sites binding to $\beta$1-integrins and actin at the C terminus. Given those similarities between the two talin proteins, it is likely that talin2 will bind many of the same ligands as talin1. A second talin gene does not exist in \textit{Drosophila} and \textit{C. elegans}, but a \textit{talin2} gene exists in mouse (\textit{Tln2}), rat and cow. In \textit{Dictyostelium discoideum}, two talin genes have been identified (\textit{TalA} and \textit{TalB}) but knockouts of these genes give different phenotypes, which suggest that the two proteins have different functions (Tsujioka et al., 2004). TalB is more specific to multicellular morphogenesis, whereas TalA is important in cytokinesis and cell motility. The TalA and TalB proteins are also quite different in sequence (30% identity) (Tsujioka et al., 1999) and TalB contains an additional C-terminal domain.

In adult mouse tissues Northern blot analysis has revealed a single major transcript for \textit{Tln1} (8kb) that was expressed at similar levels in the various tissues, while the expression levels of \textit{Tln2} varied greatly across tissues and different sizes of transcript were observed (Monkley et al., 2001). The fact that disruption of the \textit{Tln1} gene was embryonic lethal in mouse indicates that \textit{Tln2} does not appear to compensate for the loss of talin1 during early mouse development. This could be because it is not expressed at the same stage or in the same cells or the two proteins may have different functions. It is also a possibility that both are required for normal cellular function and may form a talin1/talin2 heterodimer in certain tissues.
The aims of this project were:
(i) To determine in what tissues and at what stages of development talin2 transcripts are expressed in mouse and human using Northern blot analysis as well as EST database analysis.

(ii) To analyse talin2 protein expression *in vivo* in the developing embryo and in the adult mouse using talin2 gene trap mice that expresses a talin2 - β galactosidase fusion protein.

### 3.2. Transcript expression in mouse and human tissues

#### 3.2.1. Northern Blot analysis

Previous studies where a mouse tissue Northern blot (MTN) was probed with the 3' untranslated region (UTR) of the *Tln2* transcript (Monkley *et al.*, 2001) showed a difference in the expression pattern between *Tln1* and *Tln2*. *Tln2* mRNA was expressed differentially amongst the tissues analysed and showed a range of transcript sizes whereas *Tln1* transcript was expressed in all the tissues tested as a single transcript of around 8kb. *Tln2* mRNAs varied in size and in level of expression in the different mouse tissues analysed. The larger *Tln2* transcripts (8.5kb and 10kb) were expressed in heart and brain at a high level, but also at a lower level in lung, liver and kidney. In mouse kidney and testis, short transcripts of 3.9kb and 4.8kb are respectively expressed at a high level.

To confirm these results and obtain more information on the composition of the various *Tln2* transcripts expressed in mouse, the mouse MTN and a human MTN were both used for probing with various *Tln2* coding sequence probes (Figure 3.1). Based on the Ensembl database information, probes within the *Tln2* coding region were designed at the 5' end and the 3' end of the coding sequence and synthesised by PCR. The first probe was contained in a region encoding the N-terminal actin-binding site; the second probe was contained in the C-terminal actin-binding site. The probes were radioactively labelled and hybridised to the blots. The results for the different probes are shown in Figure 3.2.
Figure 3.1. Position of the talin2 probes used for the Northern blot analysis. The 5' coding probe encodes part of the FERM domain of the head region and is situated on the cDNA from positions 579 up to 1273bp. The 3' coding probe is contained in the C-terminal rod region and encodes part of the actin-binding site, from positions 7141 up to 7876bp.
Figure 3.2. Mouse and human Northern blots hybridised with different probes for Tln2 cDNA. Commercial MTN (Clonetech) containing polyA+ RNA extracted from adult mouse or human tissues were hybridised with different probes Tln2 cDNA. The blots were normalised for β-actin loading. The blots contain RNA from heart (H), brain (B), spleen (S), lung (L), liver (Lv), skeletal muscle (SM), kidney (K) and testis (T). (A) Monkley et al (2001) used the mouse MTN previously and a 3'UTR probe and this result is shown here. After stripping using 0.5% SDS, the mouse MTN was hybridised in the same conditions as before with a 5' coding probe (B) and a 3' coding probe (C). (D) Human MTN hybridised with the 3' coding probe. The blot contains RNA from heart (H), brain (B), placenta (P), Lung (L), Liver (Lv), skeletal muscle (SM), Kidney (K) and Pancreas (P).
These results confirm that Tln2 mRNA was differentially expressed in various tissues analysed. In mouse heart and brain, two high molecular weight transcripts of 8.5 and 10kb are expressed at a high level, and are detected with all three probes used, which suggest that they contain the full-length Tln2 coding sequence. These transcripts are also expressed at lower levels in other mouse and human tissues such as liver, kidney (mouse) and skeletal muscle (human). A shorter transcript of 6.8kb was also expressed in heart and at a lower level in liver that hybridises only with the 3'UTR probe but not with the 3' coding probe. Finally, in mouse and human kidney and in mouse testis much shorter transcripts are expressed at a high level. These smaller transcripts hybridise only with the 3'UTR and 3'coding probes, which suggests that these transcripts are lacking a part of the coding sequence, at the 5' end. As talin2 is a 2541 amino acid protein, these shorter transcripts cannot encode for the full-length protein. These results are summarised in Table 3.1.

<table>
<thead>
<tr>
<th>Transcript sizes</th>
<th>Tissue(s)</th>
<th>Region of the transcript</th>
</tr>
</thead>
<tbody>
<tr>
<td>10kb</td>
<td>All</td>
<td>+  +  +</td>
</tr>
<tr>
<td></td>
<td>(Predominantly heart and brain)</td>
<td></td>
</tr>
<tr>
<td>8.5kb</td>
<td>All</td>
<td>+  +  +</td>
</tr>
<tr>
<td></td>
<td>(Predominantly heart and brain)</td>
<td></td>
</tr>
<tr>
<td>6.8kb</td>
<td>All</td>
<td>-  -  +</td>
</tr>
<tr>
<td></td>
<td>(Predominantly heart and skeletal muscle)</td>
<td></td>
</tr>
<tr>
<td>4.8kb</td>
<td>Testis</td>
<td>-  +  +</td>
</tr>
<tr>
<td>3.9kb</td>
<td>Kidney</td>
<td>-  +  +</td>
</tr>
</tbody>
</table>

Table 3.1. Summary of the Northern hybridisation results using the mouse MTN

3.2.2. Analysis of Human Kidney cDNA

a) PCR Screen and identification of the potential truncated talin2 kidney transcript

The Northern blot results showed that in mouse and human kidney, a Tln2 transcript of 3.9kb was expressed. This transcript only contained a part of the full length transcript, including the 3'UTR and a 3'coding region. In order to map more precisely the smaller transcript in human kidney, plasmid DNA from a human kidney cDNA library (kindly provided by Prof. M.J.A. Tanner, University of Bristol) was screened by PCR. The
cDNA library was first diluted 1/100 and 1 µl of it used for each PCR reaction. Different pairs of primers were designed using sequence data from the *Tln2* cDNA from the Ensemble database. The amplicons starting at the 3' end of the cDNA, going upstream towards the 5' end of the transcript were used to find the kidney transcript start site. The positions of the primers and amplified regions on the human *talin2* cDNA are indicated in Figure 3.3. Six PCR amplifications were performed using ABGene PCR Master-Mix, in a 20 µl final reaction volume, in the standard cycling conditions (Figure 3.3A). The regions between nucleotides 6884-7351 bp, 5863-6363 bp, 4769-5252 bp and 4769-4869 bp were successfully amplified showing that the kidney transcript cDNA in the library contained those regions. However, regions between nucleotides 3865-5252 bp, 3865-4869 bp and 3865-4380 bp were not amplified from the library. These PCR results suggest that the *Tln2* clones contained in the library have only a part of the full *Tln2* length transcript. The clones contained the 3' half of the transcript and regions were amplified from the 3' UTR upwards to the 5' end until a region between nucleotides 4380 bp and 4769 bp. Any region upstream of that point was not amplified.

This could be due to (i) the fact that the cDNA has been synthesised using an oligo dT primer and therefore the reverse transcriptase has copied only a portion of the full length transcript or (ii) the cDNA library contains clones from short *Tln2* transcripts that exist normally in the kidney and that the full length transcript is not expressed in human kidney or is not represented in the library.

In order to precisely determine the start site of the short transcript from kidney, PCR was performed on the library DNA, using a forward primer within the vector (SP6-2 primer) and a reverse primer just 200 bp downstream of the predicted start of the putative transcript (*hT2-3R*). The PCR reaction resulted in two main products of 450 bp and 700 bp (Figure 3.4A). The DNA fragments were gel purified subcloned into pGEMT/easy and sequenced using primers within the plasmid.

The sequence of the 700 bp fragment aligned perfectly with the full-length cDNA suggesting that this fragment was amplified from a clone in which the polymerase ran off the RNA synthesising only part of the full-length cDNA strand. The second fragment of 450 bp aligned to the full-length cDNA from position 4,564 bp on the cDNA but contained a region of 80 bp that did not correspond to either any other region of the *Tln2* cDNA or any vector sequence. This novel sequence was then analysed using the
Figure 3.3. PCR amplification of different regions of the Tln2 cDNA.
(A) Map of the different PCR products of the Tln2 cDNA. The regions successfully amplified are indicated in dark blue and were:
- PCR1, primers 3' probeF and hT2-1R, (6884-7351, 468bp)
- PCR2 primer hT2-1F and hT2-2R (5863-6363bp, 501bp)
- PCR3 primers hT2-2F and hT2-3R (4769-5252bp, 484bp)
- PCR4 primers hT2-2F and hT2-3R (4769-4869bp, 100bp)
The regions not amplified are indicated in light blue were:
- PCR5 primers hT2-3F and hT2 (3865-5252bp, 1387bp)
- PCR6 primers hT2-3F and hT2-6R (3865-4869bp 1004bp)
- PCR7 primers hT2-3F and hT2-4R (3865-4380bp, 685bp)
BLAST (Basic Local Alignment Search Tool) program against the mouse Tln2 genomic sequence, and the programme aligned this sequence with an intronic region (Figure 3.5). This suggests that the transcript contains a novel exon that was contained in intron 34 and that was spliced to the next exon of the transcript i.e. exon 35.

The short transcript identified by the Northern blot analyses could therefore be the result of one of two processes. (i) The full length Tln2 transcript is subject to alternative splicing leading to the removal of some of the exons, producing a shorter transcript. (ii) The shorter Tln2 transcript was expressed by an alternate promoter, which would be located in an intron, starting transcription in a novel exon followed by the downstream exons of the gene. This latter hypothesis was supported by the fact that PCR analysis of a kidney cDNA library gave evidence for the existence of a novel exon.

b) EST and Promoter analysis

A hypothetical kidney transcript containing the novel transcribed sequence followed by exons 35 to 57, was constructed using VectorNTI software. The first 500bp of this putative transcript was used in a BLAST search in the human and mouse EST databases (Ensembl). One human EST from a pooled colon, kidney and stomach library, (clone BI517694) and one mouse EST from an adult kidney library (clone AI317453) aligned with the sequence and contained a part of the intronic sequence identified from the library screen. The mouse and human genomic sequence in the additional exon showed 80% identity. The alignment of the ESTs with the putative kidney transcript is shown in Figure 3.5. These results show that this novel exon was not an artefact of the kidney cDNA library analysed but was probably encoded in a novel Tln2 transcript.

An analysis of the genomic region between exons 34 and 36 was performed using NIX software (Nucleotide identification of unknown sequence, HGMP bioinformatics software). This software analyses the presence of potential exons, promoters, and various transcription signals based on analogy with a whole range of known consensus sequences. The program predicted no internal exon corresponding to the extra sequence identified by PCR, and no evidence of a strong transcription start or a promoter was predicted either. However the presence of two ESTs produced from two different research groups confirms that it was likely the transcript predicted from the cDNA library was expressed in mouse and human kidney (Figure 3.5).
Figure 3.4. Amplification of the 5' end of the human Tln2 transcript in kidney. (A) PCR amplification the 5' end of the kidney cDNA contained in the library was performed using a forward primer in the vector in which the library was constructed (SP6-2 primer) and a reverse primer just downstream of the region containing the start of the Tln2 transcript identified by PCR (hT2-3R). The reaction identified two fragments of 450bp and 700bp. Those fragments were gel extracted and subcloned in pGEMT-Easy and were sequenced. (B) Schematic view of the potential expression of the Tln2 shorter transcript in kidney. The green boxes represent the exons, the black thick line the intronic regions. The exon in light blue corresponds to the novel sequence identified and is labelled as 34n.
Figure 3.5. Alignment of the kidney ESTs with the novel human kidney cDNA. Three sequences were aligned: A mouse kidney EST (mT2 AI317453), the human talin2 cDNA sequence identified using the PCR screen (hT2 kidney cDNA), and a human talin2 EST (hT2 B1517694). The human novel exon is identified in green, exon 35 in blue and exon 36 in red.
3.2.3. Analysis of protein expression in mouse tissues

Given that a novel transcript was identified by the cDNA library PCR screen and Northern blot analyses, a truncated protein in kidney may be expected to be expressed. The truncated transcript size was estimated to be between 3.9 and 4.1 kb in length. Analyses using Vector NTI were performed to look for potential Open Reading Frames (ORF) in the hypothetical kidney transcript. The only potential ORF found by the computer was an ORF in frame with full-length talin2, starting at amino acid 1610 up to 2545. The potential initiating methionine would be contained in exon 36 and, the protein encoded would therefore have a length of 935 amino acids and a molecular weight of approximately 99.1 kDa. The T2FP2 antibody as well as TD77 and C20 should recognize this truncated protein as it was raised against the region between amino acids 1649 and 1954.

The laboratory possesses different antibodies (Figure 3.6A) raised against talin1 that are known to cross react with talin2. The antibodies 8D4, C20 and TD77 were raised against the talin1 protein; T2FP2 was raised against talin2 protein. Proteins were extracted from mouse heart and kidney as well as from HEK293 cells (human embryonic kidney cells) and quantified using a BCA assay as described in Materials and Methods. Proteins were resolved by a 10% SDS-PAGE gel. The gel was then electroblotted onto a nitrocellulose membrane and probed with the different antibodies (Figure 3.6B). The antibodies TD77 and C20 detected one large band in kidney and HEK293 lysates of >250 kDa, which corresponds to the full-length talin1 protein (270 kDa). In HEK 293 cells an additional high molecular weight (MW) band was detected only by C20. An additional band of around 100 kDa below the talin1 band is recognised by C20 in HEK293 lysates and by TD77 in kidney and HEK293 lysates. The 8D4 antibody also recognizes the full length Talin (270 kDa) in heart, kidney and HEK293 lysates. All the anti talin antibodies used recognised the full length talin in kidney and HEK293 lysates. TD77 is an antibody often used in the laboratory to detect talin protein and it detects a non-specific 100 kDa band suggesting that the 100 kDa band observed is not specific and not related to the shorter talin2 kidney transcript.
Figure 3.6. Analysis of talin protein expression in mouse and human lysates. (A) Map of the anti talin antibodies raised against talin1 and talin2. (B) Kidney (K) and HEK293 (293) protein lysates were used in a Western blot analysis with C20 and TD77 which were antibodies raised against a C-terminal region of talin1. (C) 8D4 is a monoclonal antibody that recognizes a N-terminal region between amino acids 400-822. The anti vinculin antibody (F9) was used as a loading control. (D) T2FP2 is a polyclonal antibody and was raised against specific for talin2. It recognizes the region between amino acids 1659-1944. (H: Heart, K: kidney) and human embryonic kidney cells (293). The markers sides are indicated on the left of each blot.
Finally, to confirm whether the 100kDa band recognized by C20 and TD77 was specific to kidney and if it was talin2, the three lysates were probed with T2FP2. The T2FP2 is a newly made anti talin2 antibody, which was raised against a region of talin2 that was divergent between talin1 and talin2. The 100kDa band was not recognised by the T2FP2 antibody suggesting that it was not a talin2 truncated protein.

These western blot results do not appear to consistently detect a truncated protein that would be specific to kidney cells or tissue and that was specific to talin2. They give no evidence of the expression of a truncated protein in kidney.

3.3. Analysis of Talin2 expression in the mouse using a Tln2 genetrap cell line

3.3.1. Analysis of Talin2 expression in a Tln2 genetrap ES cell line

Gene trapping is a high-throughput approach that is used to introduce insertional mutations across the genome in mouse ES cells. In addition to generating standard loss-of-function alleles, newer gene trap vectors offer a variety of post-insertional modification strategies for the generation of other experimental alleles. Baygenomics (http://baygenomics.ucsf.edu/) have developed a genetrap project and currently have a database of all the genes targeted and the corresponding ES cell lines are available to the scientific community. Each gene targeted is identified by 5’ Rapid amplification of cDNA ends (5’RACE) by amplifying and sequencing the exons adjacent to the β galactosidase-Neomycin (βgeo) insertion. The gene in which the insertion has occurred is then identified using a BLAST search (Ensembl database). At the time the project started, one ES cell line containing an insertion of the βgeo cassette in Tln2 was present in the BayGenomics database, and its reference was RRI434 (14). This cell lines was ordered and grown for further analysis in cell culture and for the generation of transgenic mice for analysis in vivo.

The first step was to confirm that the cell line received did indeed contain a gene trap insertion in Tln2 and that it could be used for analysis of talin2 expression. The cell line RRI434 had a βgeo insertion in intron 26 of the mouse Tln2 gene, this intron had a length of 1kb, and therefore genomic sequence from the database could be used to
design primers in the last exon before the insertion along with a primer contained in the genetrap vector (pGTOLfx) to determine the exact site of insertion. Different primers were designed in exon 26 and used in a PCR reaction with the AB Gene PCR mastermix. The reactions were setup on genomic DNA from I4 genetrap ES cells and also wt E14.1a ES cells to ensure the specificity of the primers. Two forward primers, one within exon 26 and the second one at the beginning of intron 26, were used with the same reverse primer (contained in the genetrap vector, 86r), and gave respectively fragments of 350 and 200bp. The absence of product in wt ES cells shows that the insertion was specific to the genetrap genomic DNA. These results suggest that the insertion of the β-geo marker was approximately 250bp from the splice site of exon 26. The PCR results are shown in Figure 3.7.

Once the genomic insertion of the β-geo marker had been confirmed it was necessary to ensure the expression of a correct fusion transcript between the Tln2 exons and β-geo. RNA was extracted from a plate of cultured RRI434 ES cells and was used to synthesise cDNA using random hexamers. The cDNA was used in PCR reaction using a primer in exon26 and primer at the beginning of the coding region of β-geo for amplification. Two different reactions were performed to ensure the correct splicing of the genetrap vector to Tln2. The expected fragments of the 2 reactions were 250bp (primers mT2 Ex26F3 and bgeo R1) and 360bp (primers mT2 Ex26F1 and bgeo R1) and in both cases the fragment amplified showed the expected size, confirming the proper splicing event between exon26 and the βgeo cassette.

To analyse the expression and the localisation of the talin2-βgeo fusion protein in cultured cells at the protein level, the I4 ES cells were stained with X-gal. Positive β-gal staining was determined (Figure 3.8) but the localisation within the cells could not be precisely identified. To study the localisation in the cells more precisely, the cells were stained by immunofluorescence with an antibody raised against β-galactosidase (ABGene) along with a marker for actin filaments. The staining showed that the talin2-βgeo fusion protein is localised along actin filaments, but staining within the cytoplasm in clusters was also observed.
Figure 3.7. ES Cell line RRI434: Identification of the gene trap insertion in the mouse *Tln2* gene. (A) The genetraps vector containing the β-galactosidase-Neomycin (β-Geo) cassette was inserted in intron 26 of the *Tln2* gene. (B) Amplification of genomic DNA from I4 ES cells. Different PCR reactions were performed to locate the insertion of the marker in the intron, using primers within exon26 and a reverse primer in the genetraps vector. Primers mT2Ex26F2 and 86r gave a 350bp product (1). Primers mT2In26F3 and 86r gave a 250 bp product (2). The PCRs were performed on ES cell DNA from the RRI434 cell line (I4) as well as on wt genomic DNA (wt) and a negative control containing no DNA (-ve) sample was used to check for potential contamination in the reaction. (C) RNA was extracted from cultured I4 ES cells and cDNA synthesised using random primers. Two different RT-PCR reactions were performed using primers mT2 Ex26F3 (1) and mT2 Ex26F1 (2) along with reverse primer bgeoRl located in the β-geo coding sequence. The expected fragments were of 360 bp (1) and 250bp (2). The exons are represented by green boxes, introns by black thick lines, and the βgeo selection marker by a purple box.
Figure 3.8. Expression of talin2- βgeo in the RRI434 ES cells. (A) Expression of talin2- βgeo in the RRI434 ES cells in culture. The expression was detected by staining with X-gal overnight. (B) Expression of talin2- βgeo in RRI434 ES cells in cultured cells using immunofluorescence with an antibody raised against β-gal along with a Texas-Red coupled Phalloidin.
3.3.2. Analysis of Talin2 expression in mouse adult tissues using the Tln2 genetrap transgenic mouse

The Northern blot analyses showed that the Tln2 transcript is expressed at different levels in different mouse and human tissues. The highest level of RNA expression was observed particularly in heart and brain, skeletal muscle (human) and kidney and testis. It was important to gain more information about the cell specific Tln2 expression. This became possible using the genetrap ES cell line.

Once the RRI434 ES cell line was confirmed as having an insertion of βgeo in intron 26 of Tln2, the cells were grown and injected into blastocysts to generate chimaeras and heterozygote mice containing the Tln2-βgeo mutation were obtained. Two sessions of injections were performed and out of the 6 surrogate mothers in which blastocysts were transferred, only one chimaera was born. Germline transmission was achieved from this chimaera, and heterozygous animals were born from a mating of the chimaera to C57BL/6 mice. These animals were used for analysis of expression in adult tissues and embryos. Various tissues were dissected out of Tln2<sup>GT+/+</sup> animals, fixed and embedded in OCT. Sections were cut using a cryostat and these were fixed and stained for β-galactosidase expression overnight and mounted for observation.

The following tissues were analysed: brain, heart, lung, skin, skeletal muscle, uterus, testis, skin, stomach, intestine (Figures 3.9 to 3.12). In each case the same tissue from a wild type animal was stained (data not shown) to confirm that the staining was specific for the Tln2-βgeo heterozygous animals. These results show that the level of expression of the talin2-βgeo was high particularly in heart, brain, skeletal muscle, skin, uterus and kidney. A more localised expression was observed in the gut, lung and testis.

The staining results show that in the brain two main regions are stained and express talin2-βgeo. The staining was observed in areas of dense neuronal bodies such as the Purkinje cells in the cerebellum and the pyramidal cells in the hippocampus. In the heart, the staining was widely spread and was localised at the periphery of muscle fibres. This suggests that talin2-βgeo could be localised at the cell periphery, in the extracellular matrix or at the plasma membrane. The staining was also clear in smooth
Figure 3.9. Expression of talin2-βGal in mouse tissues dissected from \( Tln2^{GT/+} \) mice. Adult tissues were dissected from a 4 weeks old \( Tln2^{GT/+} \) animal. The samples were fixed in glutaraldehyde for 30 minutes, embedded, frozen in a dry ice-methanol bath and stored at -20°C. 20μm sections were cut and stained overnight in X-Gal (A) Brain section showing the hippocampus region (B) Brain section showing the cerebellum (C) Lung section at 100x magnification (D) Lung section at 400x magnification. Scale bars: 50μm for (A) and (B) and 100μm for (C) and (D). h: hippocampus; pc: Purkinje cells; a: arterioles
Figure 3.10. Expression of talin2-βGal in mouse tissues dissected from Tln2\(^{GT/4}\) mice. Adult tissues were dissected from a 4 weeks old Tln2\(^{GT/4}\) animal. The samples were fixed in glutaraldehyde for 30 minutes, embedded, frozen in a dry ice-methanol bath and stored at -20°C. 20μm sections were cut and stained overnight in X-Gal (A) Heart section at 100x magnification. (B) Heart section at 400x magnification (C) Skeletal muscle at 100x magnification. (D) Skeletal muscle at 400x magnification. Scale bars: 50μm.
Figure 3.11. Expression of talin2-βGal in mouse tissues dissected from $Tln2^{GT/+}$ mice. Adult tissues were dissected from a 4 weeks old $Tln2^{GT/+}$ animal. The samples were fixed in glutaraldehyde for 30 minutes, embedded, frozen in a dry ice-methanol bath and stored at -20°C. 20μm sections were cut and stained overnight in X-Gal (A) Testis section at 100x magnification (B) Testis section at 400x magnification (C) Kidney section at 100x magnification (D) Kidney section at 400x magnification. Scale bars: 50μm. sp: differentiating sperm cells; st: seminiferous tubules; hl: Henle’s loop.
Figure 3.12. Expression of talin2-βGal in mouse tissues dissected from $Tln^{2Gf/s}$ mice. Adult tissues were dissected from a 4 weeks old $Tln^{2Gf/s}$ animal. The samples were fixed in glutaraldehyde for 30 minutes, embedded, frozen in a dry ice-methanol bath and stored at -20°C. 20μm sections were cut and stained overnight in X-Gal (A) Stomach section at 100x magnification (B) Stomach section at 400x magnification (C) Small intestine at 100x magnification (D) Small intestine section at 400x magnification. Scale bars: 100μm for (A) and (B); 200μm for (C) and (D). sm: smooth muscle.
Chapter 3 – Talin2 expression

muscle of various tissues such as the lung (arterioles), and the gut (smooth muscle layer of the stomach). Some staining was also observed in the kidney, in the medulla and in the Henle's loop. The Henle's loop is composed of a single epithelial layer with a role in permeability. It was impossible to detect the exact localisation of the staining in the cells. The talin2-βgeo was also expressed in testis, in the differentiating sperm cells.

3.3.3. Analysis of Talin2 expression in developing embryos using the Tln2 genetrap transgenic mouse line

When the heterozygous animals were mature enough for mating, the males were used to mate to MF1 females. The aim of that is to generate and analyse embryos at different times of development. When the timed matings were set up, a plug date was obtained, allowing the pregnancy to be dated and the embryos to be harvested at a chosen date for expression analysis. Embryos were harvested at several stages of development and two different approaches were used for the analysis. The smaller embryos (7.5dpc, 8.5dpc, 9.5dpc) were fixed and stained as wholemounts, and photographed as such. In older embryos (10.5dpc, and 14.5dpc) the embryos were embedded, sectioned and then stained in order to obtain an expression pattern in finer detail. The staining results are shown in Figure 3.13 to 3.15.

In early embryos at 7.5dpc expression was localised in the node (Figure 3.13A). The node is an organising pole in the embryo and is a structure composed of two cell layers. At 9dpc the expression was widely spread throughout the embryo in various structures with a higher intensity in the cardiac region (Figure 3.13B). At 9.5dpc the staining shows a wide expression but a few structures are more strongly stained. These are a region of the telencephalon, which develops as a part of the brain, the heart and the otic vesicle, which develops into the adult ear (Figure 3.13C). Embryos at later stages were then analysed. At 10.5dpc, the staining was widely expressed (Figure 3.13D). In order to detect the strong staining within specific tissues and organs, after wholemount staining overnight, the embryos were embedded into paraffin, sectioned and the slides observed under the microscope. In embryos from 10.5dpc, the staining in the heart became very obvious. The somites and the surface ectoderm were also stained (Figure
Figure 3.13. Expression of talin2-βGal in heterozygous developing embryos by wholemount staining. Embryos were harvested at different stages of development and the heterozygotes were fixed and stained overnight in X-Gal. (A) 7dpc (left handside) and 7.5dpc (right handside) (scale 200μm); (B) 9dpc (scale 200μm); (C) 9.5dpc (scale 200μm) (D) 10.5dpc (scale 200μm) Tln2^{G146} embryo on the left, wt embryo on the right. n: node; ov: otic vesicle; h: heart; t: telencephalon.
Figure 3.14. Expression of talin2-βGal in heterozygous Tln2<sup>GT/+</sup> 10.5dpc embryos by β-gal staining of sections. Embryos were harvested at different stages of development and they were stained wholemount overnight in X-Gal, frozen sections (20μm) were cut and photographed. Scalebar: 500μm. Section staining: 10.5dpc Embryo. Enlargements of the brain (A), heart (B), limb bud (C) and somites (D).
Figure 3.15. Expression of talin2-βGal in heterozygous Tln2<sup>GT/+</sup> 14.5dpc embryos by β-gal staining of sections. Embryos were harvested at different stages of development and the heterozygotes were stained overnight in X-Gal, frozen sections (20μm) were cut, stained with X-Gal and photographed. Scalebar: 500μm. Enlargements of the heart (A), somites (B) and the surface ectoderm (C). h: heart; S: somites; se: surface ectoderm
3.14). The somites later develop into the cartilage of the vertebrae and ribs, the muscles of the rib cage, limbs, and back, and the dermis of the dorsal skin. Later embryos, at 14.5dpc showed a similar pattern of expression, showing expression of talin2-βgeo in the heart, in the somites and the developing skin (Figure 3.15).

<table>
<thead>
<tr>
<th>Age</th>
<th>Expression sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5dpc</td>
<td>Node</td>
</tr>
<tr>
<td>9dpc</td>
<td>Wide spread</td>
</tr>
<tr>
<td>9.5dpc</td>
<td>Wide spread</td>
</tr>
<tr>
<td>10.5dpc</td>
<td>Heart, Limb bud, brain, Somites</td>
</tr>
<tr>
<td>14.5dpc</td>
<td>Heart, skin, vertebrae</td>
</tr>
</tbody>
</table>

Table 3.2. Expression of Tln2-βgeo at different stages of development.

3.4. Discussion

To gain an insight into the expression profile of talin2 in mouse tissues and embryos, the expression of Tln2 RNA was assessed using Northern blot analysis using cDNA probes on mRNAs from different mouse and human tissues. Analysis at the protein level was performed using Tln2 gene trap ES cells and mice.

Northern blots containing various mouse or human tissues were hybridised with different probes designed for different parts of the Tln2 coding sequence. This analysis showed firstly that Tln2 mRNA was expressed in various mouse tissues. Transcripts of 8.5kb and 10kb were detected in all the tissues analysed, particularly in heart and brain. Smaller transcripts were detected in mouse kidney and testis with the 3' coding probe and 3'UTR probe but not with a 5'coding probe. These transcripts are 3.9kb in kidney and 4.8kb in testis and are not long enough to code for a full-length talin2 protein. A 3.9kb transcript was also expressed in human kidney. A 6.8kb transcript was also detected only by the 3'UTR probe in heart, brain and skeletal muscle. It did not contain the 5' and 3' coding regions used. This could be due to the fact that this Tln2 transcript resulted from an alternative splicing event in the gene. More analysis using a wider range of probes within the coding region of Tln2 would be necessary to identify the regions contained in each of
the transcripts. The data clearly suggest that there are different talin2 mRNAs in mouse and human tissues and that the exon profile of these varies between tissues.

This expression pattern can be explained by either (i) alternate splicing event or (ii) alternate promoter usage or a combination of both. In order to distinguish these possibilities in the case of the 3.9kb transcript, a human kidney cDNA library was screened by PCR, using different primers in the Tln2 coding sequence. A human kidney cDNA was identified that contains, at its 5’ end, a novel sequence of 80bp from intron 34 adjacent to exon 35 followed by exons 36 to 57. This novel sequence was found not to be an artefact of this library as it was also detected in 2 ESTs, contained in the mouse and human genome databases. Computer analysis was performed to search for potential promoters upstream of intron 34, but there was no evidence of the presence of any transcription start or promoter consensus sequence in the region. These results support the view that differential Tln2 transcripts exist in kidney but they do not distinguish between splicing or alternate promoter usage. In order to detect the presence of an internal alternate promoter, it would be necessary to subclone different parts of this region upstream of a reporter gene and check for any promoter activity. To identify the various transcripts start sites, 5’RACE analyses could be carried out.

If the novel transcript was transcribed from a promoter within intron 35 and translated using an ATG in exon 36, it would lead to the production of a 100kDa protein. To test if this was the case, protein analysis on mouse heart and brain and human kidney cells were performed by Western blot analysis using an antibody for talin1. The talin2 antibody (T2FP2) used was generated in D.R. Critchley’s lab and was raised against a region between amino acids 1659 and 1944 of talin2 but it is not currently known whether it can distinguish between talin1 and talin2. So although it does not detect a truncated protein that would be specific to kidney, more analysis with a reliable talin2 antibody would help determine whether a shorter functional protein is expressed in some cell types. This could also be due to the fact that the novel transcript identified does not lead to the production of a functional protein but has a regulatory role.

To investigate how the talin2 protein is expressed in the mouse during embryonic development and in the adult tissues, a Tln2 genetrap cell line was used. BayGenomics, now expanded at the Sanger Centre, have set up a large-scale mutagenesis project, using
Chapter 3 – Talin2 expression

genetrap technology in the aim of functional annotation of the mouse genome (Skarnes et al., 2004). The trapping strategy is based on the insertion of a reporter gene (βgeo) such that the insertion simultaneously reports on expression of the endogenous gene in which it is inserted. The insertion may also lead to gene disruption. Their database contained an ES cell line (RRI434) in which the insertion had occurred in the Tln2 gene. X-gal staining of ES cells and immunolocalisation along stress fibres showed the β galactosidase expression in these cells. The localisation of the β galactosidase staining along actin stress fibres is probably due of the expression of the talin2-βgeo fusion protein but might not reflect the localisation of full-length talin2 but this remains unproven. The talin2-βgeo fusion does not appear to localise to FAs. The head domain of talin is sufficient to target talin to FAs (Yan et al., 2001). However talin2 is expressed as a fusion protein with the β galactosidase which could have an effect on the conformation of the protein such as masking binding sites for β integrin tail or actin. This hypothesis could be tested by coimmunostaining of βgal with a FA protein such as vinculin. The use of a consistent specific talin2 antibody directed against the talin2 head would confirm this intracellular localisation.

Mice were generated from the ES cells and analyses were performed in embryos and adult tissues. Results and analysis in adult tissues are consistent with the Northern blot analyses. The expression was the highest in brain and heart. It was expressed throughout the brain but with region of dense neuronal bodies such as the Purkinje cells of the cerebellum and the Pyramidal cells of the hippocampus. The expression in smooth muscle layer was observed in various tissues such as the gut, uterus and the lung, as well as in skeletal muscle. The staining also shows there was expression in testis and kidney. Therefore talin2 is expressed widely throughout the adult mouse.

The staining of the embryos at different stages showed that talin2 expression is specific throughout embryonic development and is specialised in specific structures. The node is an organising and patterning structure, composed of two cell layers, which are intimately associated. The ventral cell layer of the node is known as the notochordal plate, in which each cell is distinguished by the presence of a single motile cilium. This cell layer later develops into a continuous epithelium. From the staining analyses carried out, the individual cells in which talin2-βgeo is expressed could not be identified. In
order to determine in which cells the talin2-βgeo fusion is expressed, sectioning through 7.5dpc embryos and staining using a specific antibody for talin2 would be required. The data collectively suggests a role for talin2 in brain development and function, as well as in smooth and cardiac muscle. Talin2-βgeo is also expressed in somites, which later develop in skeletal muscle structures and the skin.

As the Northern analyses suggested, the main organs in which talin2 is expressed are the heart, the brain and skeletal muscle. However if smaller transcripts expressed in kidney and testis led to the expression of a functional protein, this would not be reflected in the βgal expression. If these smaller transcripts were expressed from downstream promoters, the gene trap insertion would not affect their expression and the X-gal staining would not reflect their expression. So although the genetrap approach gives a lot of information on where talin2 is expressed in development and in adult tissues, the fact that the gene has a complex pattern of mRNA expression further analyses using specific antibodies particularly raised against the C terminal region would complete these data by giving information on potential truncated proteins that might fulfil a role in specific tissues.

Chen and Lo (2005) carried out analyses using mice generated from the same ES cell line. They observed that the homozygous mutants survived until birth showing that the N terminal 1295 aminoacids were sufficient for mouse embryonic development or that talin2 is specifically important in particular tissues. It could also be due to the fact that talin1 and talin2 are highly similar and therefore in the absence of talin2, talin1 can compensate for it. The knockout of both alleles using gene disruption or a siRNA approach could help determine if the two proteins function as a dimer and are both required for normal tissue function or if in certain tissues they can compensate for each other.
Chapter 4. Generation of a *Tln2* Conditional Knockout

4.1. Introduction

Analysis of human and mouse EST databases led to the discovery of a gene encoding a protein highly related to talin1, named talin2 (Monkley *et al.*, 2001). With the completion of the sequencing of the genomes of several species, a talin homolog was also be predicted in numerous other species based on BLAST alignments. It has been identified in Chimpanzee (*Pan troglodytes*), several mammals such as dog (*Canis Familiaris*), Aurochs (*Bos tauris*), small mammal species such as mouse (*Mus Musculus*), rat (*Rattus Norvegicus*), chicken (*Gallus Gallus*). The talin (thus renamed talin1) homolog has been named as talin2 in human and mouse.

Knockout studies have shown that *Tln1*−/− mice die during embryogenesis due to failure to complete gastrulation (Monkley *et al.*, 2000). This suggests that its homolog talin2 cannot compensate for its loss during early mouse development. As the two proteins are very similar (75% identity in human), it is very likely that talin2 binds the same ligands, but this remains to be demonstrated. In order to better understand the specificity of talin2 function in comparison to talin1 and to further investigate its role *in vivo* and *in vitro*, it was decided to adopt a conditional knockout approach in mouse.

In recent years, the development of site-specific recombinase systems has facilitated functional approaches to study the role of genes *in vivo* and *in vitro*, but also to develop good models for human diseases in mouse. In comparison to a classical approach for gene targeting it allows more specific studies in a whole range of physiological contexts in cultured cells as well as in mouse tissues. A classical knockout strategy was adopted for *Tln1* (Monkley *et al.*, 2000), led to the inactivation of the gene and was a good model to study the gene function, but had several disadvantages as mentioned in the introduction (Chapter 1). Early embryonic lethality or severe phenotypes reduce greatly the scope of potential analyses. Any studies in the adult mouse are made impossible in the case of an embryonic lethality. If the phenotype affects multiple tissues, any study of the gene function in particular cell lineages is also prevented.
Chapter 4 - Generation of a Tln2 conditional knockout

A system based on the Cre recombinase in combination with a trilox allele was chosen to specifically target Tln2 in the mouse. The Cre-loxP system has been widely used and therefore many modes of administrations of Cre recombinase in mouse as well as for cell culture studies are now available. This system makes the inactivation of a gene possible in almost all mouse tissues as well as in different cultured cell lines. Two reagents are necessary for this approach (i) an allele containing recombination sequences (loxP sites) flanking the region to be deleted ('floxed') (ii) a way of delivering the active recombination enzyme (Cre recombinase) either in vitro or in vivo.

The first step in designing a conditional knockout strategy is the choice of the target sites on the gene. The location of the loxP sites and the selection marker must be such that they avoid any interference with the gene function. It is also important to ensure that the positioning of the loxP sites will generate a null allele upon Cre mediated recombination. In order to achieve these goals it is important to have some good information about the gene structure, to know where the coding exons are and the position of potential regulatory elements. The obvious strategy would be to aim for the whole coding region to be flanked by loxP sites. However this would be difficult if the gene is large, or if the gene structure is not completely known. In most cases the target sites chosen are flanking a region that is essential for the normal expression of the gene or that upon deletion results in a frameshift leading to the translation of a truncated, non-functional protein.

When the sites of targeting are chosen a targeting vector needs to be constructed. After transfection into ES cells, homologous recombination event will lead to the insertion of loxP sites in the wild type allele. This targeting vector needs to include several essential components. In the case of the trilox system it has three arms of DNA sequence isogenic to that of the ES cells to be targeted. The length of these regions of homology should be as big as possible and avoid repetitive sequences. Two of those regions must flank the region to be floxed. It must also contain a positive selection gene cassette (Neomycin resistance gene, Neo\(^R\)), which is placed in between two arms of homology to allow selection for integration. In some cases a negative selection gene (such as thymidine kinase) can be placed at the end of a homology arm to select against random integration. Three recombination loxP sites must be placed in the same orientation, two of them in between the arms and a third one to flank the selection marker. A standard targeting
strategy is shown in Figure 4.1. The targeting vector constructed in this project used two plasmid constructs, one containing a \textit{loxP} site and a second one containing the neomycin resistance gene \((\text{Neo}^R)\) flanked by two \textit{loxP} sites. The maps of these plasmids are indicated in Materials and Methods (section 2.3.6).

The targeting of the allele is performed on cultured ES cells by transfection of the targeting vector plasmid DNA. Upon addition of the targeting vector, homologous recombination can occur between the isogenic arms of the vector and the genomic DNA resulting in the insertion of the \textit{loxP} sites and the \text{Neo}^R cassette at the chosen locations in the gene and the generation of a floxed allele. After growing the ES cells under selection, some ES cell clones can be isolated for screening. In the majority of the clones, the targeting vector will randomly integrate in the ES cell genome. Therefore it is necessary to design a screening strategy to identify homologous recombinants. Two methods, PCR and Southern blotting, are commonly used to distinguish between the floxed (fl) and the wild type (wt) alleles and are also then used to distinguish between different alleles generated by Cre mediated recombination.

The aims of this project were to

(i) Establish a restriction map the 5' end of the \textit{Tln2} gene in order to design a cloning strategy for the targeting vector

(ii) Construct the targeting vector using the plasmids containing the \textit{loxP} sites and the selection marker.

(iii) Generate \textit{Tln2}^{fl/+} ES cells by electroporation of the targeting vector in ES cells, screening of clones under \text{Neo}^R selection and identification of homologous recombinants.

(iv) Generate \textit{Tln2}^{fl/+} mice and MEFs and use them for an \textit{in vivo} Cre mediated recombination.

These animals and cells will allow the study of the role of talin2 \textit{in vitro} as well as \textit{in vivo}. Comparison of the role of talin1 and talin2 in cultured cells will help to determine their specific role in integrin mediated cell adhesion and motility.
Gene of interest

Targeting vector

Floxed allele

Expression of Cre Recombinase

Type I

Type II: "floxed" allele

Type III:

Figure 4.1. Strategy for gene conditional targeting using the Cre-trilo system. Exons are represented by green boxes, the loxP sites by a blue triangle, intronic regions by back lines, a positive selection marker such as Neo by the grey box with +, a negative selection marker such as the thymidine kinase gene by the grey box with -. By homologous recombination on each side of the targeting vector, the loxP sites and the positive selection marker are inserted at the chosen locations in the gene of interest. After this step, expression of Cre will generate three types of alleles due to recombination. The type I deletion corresponds to the deletion of the whole floxed region, the type II deletion is the floxed allele from which the selection marker has been removed and only the chosen region is left flanked by loxP sites. In the type III deletion the positive selection marker has not been deleted, but the genomic DNA has.
4.2. Mapping of the 5' end of the mouse Tln2 gene

In order to design a cloning strategy for generating a targeting vector, an accurate restriction map of the 5' end of Tln2 gene was required. This was achieved by (i) establishing a restriction map using a genomic cosmid clone C10 (129Ola strain) containing the 5' region of Tln2 gene by Southern blot analyses and (ii) analysing sequence data from the Ensembl mouse genomic database (C57BL/6 strain).

4.2.1. Analysis of the cosmid clone

A clone was obtained from a 129Ola mouse genomic library from the Resource Centre of the German Human Genome Project (RZPD). This particular library was chosen because the ES cells used for gene targeting and generation of mice are also derived from a 129Ola mouse and polymorphic differences between a targeting vector and ES cells can reduce the homologous recombination frequency. A positive clone was identified (C10) by PCR using primers contained in the first coding exon of Tln2 gene and the corresponding cosmid clone obtained.

An approach based on Southern blotting was chosen to establish a restriction map of the genomic region contained in the clone. The clone was digested by different restriction enzymes (single and double digests) that were then used in a Southern blotting experiment using 5' labelled oligonucleotides contained in different coding exons. According to initial restriction digests and to sequencing data, the insert appeared to be cloned between the Bam HI (1) and the Sfi I (301) sites on the map shown in Materials and Methods (section 2.3.6). A restriction digest with Sfi I gave two fragments (data not shown). The first one was approximately 5kb in size and corresponded to the size of Lawrist7, the cosmid vector in which the library was cloned. The second fragment was of a much larger size and potentially contained the genomic insert. This suggested that Sfi I digested the cosmid such as to release the genomic insert.

The C10 clone was digested with multiple restriction enzymes. Five different enzymes, which did not give too many fragments and did not cut within the Lawrist 7 vector, were chosen to generate a restriction map. The digests were run on a 0.8% agarose gel (Figure 4.2A); the sizes of the various digest fragments were calculated and are presented in Table
4.1. The total insert size for each restriction digest was calculated and for all digests the average was 38.1kb.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Fragment size in kb</th>
<th>Insert Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bam HI</td>
<td>13.1 10.1 6.1 4.6 4.3 4.1 2</td>
<td>39</td>
</tr>
<tr>
<td>Hind III</td>
<td>9.2 8.3 7.9 4.1 3.5 2.9 2.7 1.3 1.2 0.7 0.6</td>
<td>37.1</td>
</tr>
<tr>
<td>Sma I</td>
<td>13.7 13.1 6.1 5.0 4.1 2.0</td>
<td>38.7</td>
</tr>
<tr>
<td>Sst I</td>
<td>7.9 7.7 6.4 6.3 5.3 4.5 2.9 1.9 0.8 0.6</td>
<td>39</td>
</tr>
<tr>
<td>Xba I</td>
<td>9.6 8.7 6.1 5.8 2.2 2.2 1.8 1.7 1.4 1.2 0.9 0.6</td>
<td>36.9</td>
</tr>
</tbody>
</table>

**Table 4.1. Sizes of the fragments for each of the five digests.**

The estimated insert size was calculated by adding all the fragments size and taking out 5.3kb corresponding to the vector size for each digest and then an average between the 5 restriction digests reactions calculated.

In order to map the genomic insert in the C10 clone, gels of the above digests were blotted and then hybridised with different \([\gamma^{32}P]ATP\) labelled oligonucleotides specific to exons 3, 4 and 5 (respectively primers 5'T2For, mT2-1F and mT2-4F), which are the first 3 coding exons of \(Tln2\) (Figure 4.2). After hybridisation, for some of the digests, multiple bands were present. In those, one band in each track was significantly more intense than the others and was taken to be the fragment that was most specifically hybridised to the probe. The other bands were due to non specific hybridisation of the probes on the cosmid. The bands hybridising to the different probes are summarised in Table 4.2.

<table>
<thead>
<tr>
<th>Hybridisation Probe used</th>
<th>Bam HI</th>
<th>Hind III</th>
<th>Sma I</th>
<th>Sst I</th>
<th>Xba I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon3</td>
<td>2kb</td>
<td>9.2kb</td>
<td>4.5kb</td>
<td></td>
<td>9.6kb</td>
</tr>
<tr>
<td>Exon4</td>
<td>8.3kb</td>
<td>1.3kb</td>
<td>4.5kb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon5</td>
<td>3.5kb</td>
<td>7.8kb</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.2. Summary of hybridisation results of each restriction digest with each probe.**
Figure 4.2. Agarose gel and Southern blot analysis on single restriction digests of the Tln2 C10 cosmid clone. (A) The genomic clone C10 containing the 5' end of Tln2 was used to establish a restriction map of the region. The C10 DNA was digested with 5 different restriction enzymes: Bam HI (B), Hind III (H), Sma I (Sm), Sst I (Ss) and Xba I (X) and electrophoresed on a 0.8% agarose gel and photographed. (B) The digests were used in a Southern blot analysis using 5' labelled oligonucleotides as probes for the first 3 coding exons (exon3, 4 and 5). The markers (M) in the first lane correspond to 1KB+ markers with bands between 12kb and 100bp.
Double digests of the C10 cosmid were then performed. These digests were electrophoresed on an agarose gel, blotted onto a nitrocellulose membrane and probed using 5't radiolabelled primers for exons 3 and 4. The gel electrophoresis and the hybridisation results are shown in Figure 4.3. The results of the hybridisations are summarised in Table 4.3.

<table>
<thead>
<tr>
<th>Probe</th>
<th>Sst I</th>
<th>Sst I + Bam HI</th>
<th>Sst I + Not I</th>
<th>Sst I + Sfi I</th>
<th>Sst I + Smal</th>
<th>Sst I + Xba I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon3</td>
<td>6.4kb</td>
<td>1.6kb</td>
<td>6.4kb</td>
<td>6.4kb</td>
<td>6.4kb</td>
<td>4kb</td>
</tr>
<tr>
<td>Exon4</td>
<td>6.4kb</td>
<td>2.9kb</td>
<td>6.4kb</td>
<td>6.4kb</td>
<td>6.4kb</td>
<td>4kb</td>
</tr>
<tr>
<td>Probe</td>
<td>Xba I</td>
<td>Xba I + Bam HI</td>
<td>Xba I + Not I</td>
<td>Xba I + Sfi I</td>
<td>Xba I + Smal</td>
<td>Xba I + Sst I</td>
</tr>
<tr>
<td>Exon3</td>
<td>8.7kb</td>
<td>1.1kb</td>
<td>8.7kb</td>
<td>8.7kb</td>
<td>8.7kb</td>
<td>4kb</td>
</tr>
<tr>
<td>Exon4</td>
<td>8.7kb</td>
<td>8kb</td>
<td>8.7kb</td>
<td>8.7kb</td>
<td>8.7kb</td>
<td>4kb</td>
</tr>
</tbody>
</table>

Table 4.3: Summary of hybridisation results of the double digests with each probe.

All of these data allowed the generation of a restriction map of the 5' end of \textit{Tln2}, as shown in Figure 4.4.

4.2.2. Ensembl genome database analysis

At the start of the project, the mouse genomic region containing \textit{Tln2} was not available. When it was made available, the Ensembl mouse genome database (http://www.ensembl.org/Mus_musculus) was used to complete and confirm the restriction map established with the Southern blot analyses. It also provided sequence information on the \textit{Tln2} gene. A \textit{Tln2} sequence from a 5'EST was used to search in the mouse genome using BLAST. Sequence alignments can provide good information to compare novel sequences to known characterised genes. The BLAST tool provided a method for rapid searching of nucleotide and protein databases. The mouse \textit{Tln2} gene is located on chromosome 9 between positions 67.3Mb – 67.6Mb (Ensembl v35), on the negative strand. The genomic sequence obtained is from the C57BL/6 mouse strain.
Figure 4.3. Agarose gel and Southern blot analysis on double restriction digests of the *Tln2* C10 cosmid clone. Double digests were performed on the cosmid DNA and were used in a Southern blot analysis using the probes for exon 3 and 4. For each set, the DNA gel and the southern blot results are shown. The digests were set up as follows: (A) 1) *Sst* I + *Bam* HI; 2) *Sst* I + *Not* I; 3) *Sst* I + *Sfi* I; 4) *Sst* I; 5) *Sst* I + *Sma* I; 6) *Sst* I + *Xba* I (B) 1) *Xba* I + *Bam* HI 2) *Xba* I + *Not* I; 3) *Xba* I + *Sfi* I; 4) *Xba* I; 5) *Xba* I + *Sma* I; 6) *Xba* I + *Sst* I
4.2.3. Generation of a map of 5'end of the \textit{Tln2}

In order to check that the 129Ola cosmid clone and the Ensembl database (C57BL/6 mouse strain) sequence were similar, the C10 cosmid clone was sequenced using primers in exons 3, 4 and 5 (first 3 coding exons) and in the introns between these exons. Primers to the cosmid vector were also used for sequencing in order to determine where in the genomic sequence the cosmid C10 insert started and ended.

The sequence data was then aligned to the Ensembl sequence using Contig Express software. Contig Express is alignment software, which allows alignments between multiple sequences from various sources. The alignments showed that the C57BL/6 sequence from the Ensembl database and the 129Ola cosmid sequence were highly similar and that the restriction maps of both sequences were very close. The C10 clone contained the first 11 coding exons. In addition, the location of the restriction sites between the cosmid and the Ensembl sequence were compared. In nearly all cases the two restriction maps were identical. The combination of the Southern blots results and the Ensembl restriction map is shown in Figure 4.4.

4.3. Generation of \textit{Tln2} conditional targeted ES cells

4.3.1. Construction of the \textit{Tln2} targeting vector

It was decided to adopt a conditional targeting strategy to allow the deletion of \textit{Tln2} to be spatially induced in different tissues or temporally regulated at a particular stage of development. The strategy chosen was based on the use of Cre recombinase and \textit{loxP} sites. In the chosen approach, it contains three regions of isogeny with \textit{Tln2} called arms (5'arm, middle arm and 3'arm) as shown in Figure 4.5. \textit{LoxP} sites were inserted between the arms allowing eventual removal of the \textit{loxP} flanked region by a Cre mediated recombination. An antibiotic resistance gene (\textit{Neo}\textsuperscript{R}) was also inserted to allow a selection of ES cells possessing the targeting vector. Following homologous recombination between the targeting vector and the \textit{Tln2} genomic sequence, the vector sequence can replace the genomic DNA of the targeted cells and therefore inserts the 3 \textit{loxP} sites and the \textit{Neo}\textsuperscript{R} cassette into the gene. The \textit{Neo}\textsuperscript{R} gene was also flanked by two
Figure 4.4. Restriction map of the 5' end of Tln2. Both the Ensembl sequence data and the Southern blot analysis allowed a restriction map of the 5' end of the gene to be established. (A) A restriction map based on genomic DNA sequence from the Ensembl database imported in VectorNTI was established. Sequence information was used to determine the positions of the restriction sites. The Tln2 gene is located on Chromosome 9, on the minus strand on positions 67.3Mb up to 67.5Mb. (B) Restriction fragments identified by the Southern blot analyses on the single and double restriction digests. Each fragment of the cosmid clone that was identified by Southern blot analysis is noted with its length the enzymes from which it was generated and the probes by which they were identified.
Figure 4.5. Strategy adopted for the inactivation of the Tln2 gene. (A) Homologous recombination between the Tln2 gene and the conditional targeting vector results in the insertion of three loxP recombination sites, two of which are located upstream of exon 3 and flank the NeoR selection cassette and the third downstream of exon 3 therefore flanking a genomic region containing the first coding exon. (B) Exposure of the Tln2 floxed alleles to Cre results in, three possible recombinated alleles. The type I deletion gives rise to the deletion of the whole region flanked by recombination sites, the type II deletion corresponds to the floxed allele, removing only the NeoR cassette and the type III deletion gives rise to the deletion of the middle arm containing exon 3, leaving the NeoR cassette. Exons are represented by a green box in which the exon number is placed, loxP sites by a blue triangle, intronic regions by thick black lines and the NeoR by a grey thick arrow indicating the orientation of the gene transcription in relation with the Tln2 gene.
Chapter 4 – Generation of a Tln2 conditional knockout

loxP sites. This allowed the removal of the selection marker after homologous recombination in ES cells (Figure 4.5). In order for the homologous recombination to be as efficient as possible in the desired areas, the floxed region had to be as small as possible in order to increase the chances of homologous recombination in the 5’ and the 3’ arms. Therefore a region flanking exon3 only was chosen to be floxed as it is the first coding exon and contains the translation initiation ATG codon.

Two plasmids, one of which contained a loxP site (ploxP) and the other one containing a NeoR cassette flanked by two loxP sites (ploxPneoloxP), were used for the construction of the targeting vector. Their maps are indicated in Materials and Methods (section 2.3.6). Based on the restriction map established from the cosmid mapping, the genomic database information and the restriction map of pLoxP and pLoxPneoloxP plasmids, a cloning strategy for the construction of the targeting vector was established.

Sequencing data from the cosmid gave the information that the cosmid clone only contained 800bp of sequence upstream of exon3. Therefore, this clone could not be used to amplify the 5’ arm and genomic DNA extracted from E14.1a ES cells was used as a template. PCR amplification of the middle and the 3’ arm were performed using the cosmid DNA as a template. The database sequence was used to design three sets of primers to amplify the 3 arms of the targeting vector. The primers were designed and a few nucleotides corresponding to the restriction site chosen for cloning were added to the primer sequence.

The homology arms were amplified using Expand DNA polymerase, which had a proof reading activity, to ensure the fidelity of the amplification. The PCR amplification results are shown in Figure 4.6. When each arm was cloned, it was sent for sequencing to ensure there were no errors in the sequence (data not shown).

a) Subcloning of the 5’ arm
The 5’ arm was 1.2kb long and contained a part of the second intron. It was amplified from ES cell genomic DNA using primers KpnLarm1.F and ClaLarm1.R which contained a restriction site sequence for respectively Kpn I and Cla I. These two enzymes were then used to digest the PCR fragment. The digested fragment was ligated
Figure 4.6. PCR amplification of the isogenic arms of the *Tln2* conditional targeting vector. Three regions in the *Tln2* gene were chosen to be amplified by PCR to obtain the isogenic arms of the *Tln2* conditional targeting vector. The 5' arm (1.2kb) was amplified from E14.1a genomic DNA using primers KpnLarm1.F and ClaLarm1.R. The middle arm (722bp) and the 3' arm were amplified from cosmid C10 DNA using respectively primer pairs SalMarml.F-ApaMarml.R and XhoRarm1.F-XhoRarm3.R. The PCR reactions were set up using a proof reading enzyme (Expand DNA polymerase), performed at 60°C annealing temperature for 35 cycles. For each reaction a negative control containing no DNA sample was run and loaded, along with 1KB+ DNA Ladder (Invitrogen) as a size marker (M).
Figure 4.7. Cloning of the *Tln2* targeting vector: Subcloning of the 3 homology arms. (A) The 5' arm (1.2kb) was amplified from ES cell genomic DNA using primers KpnLarm1.F and ClaLarm1.R which contained a restriction site sequence for respectively Kpn I and Cla I. The fragment was then digested by Kpn I and Cla I and subcloned into the Kpn I and Cla I restriction sites of ploxP/neoP. The resulting construct *pneoP/T2Larm* was 6115bp. (B) The middle arm (722bp) was amplified from the cosmid C10 DNA using primers Sal1MArm1.F and Apa1MArm1.R. The PCR fragment was then digested by Apa I and Sal I and subcloned into the Apa I and Sal I sites of pBSIISK+ and then into the Sst I and Sal I sites of ploxP. The resulting construct *ploxP/T2MA* was 3726bp. (C) The 3' arm (3.9kb) was amplified using primers XhoRarm1.F and XhoRarm3.R. The PCR fragment was then digested by Xho I. It was first subcloned into pBSIISK+ and then into the Xho I sites of the ploxP/T2MA shown in (B). The resulting construct *ploxP/T2MAloxPRA* was 7626bp.
Chapter 4 – Generation of a Tln2 conditional knockout

into ploxPneoloxP also digested with Kpn I and Cla I. The ligation was transformed into DH5α and plated onto a LB agar plate containing 100μg/ml of ampicillin. Positive clones were picked, DNA extracted by mini prep and checked by restriction digest using appropriate enzymes (Figure 4.7A). The resulting plasmid was called pneolox/T2Larm.

b) Subcloning of the middle arm
The middle arm was 722bp long and contained the first coding exon (exon3). It was amplified from the cosmid C10 DNA using primers Sal1Marml.F and Apa1Marml.R, which contained a restriction site for respectively Apa I and Sal I. The fragment was amplified then digested using these two enzymes. Following the same protocol as for the 5’ arm, it was ligated and transformed into the Apa I and Sal I sites of pBSIISK+. This step was performed in order to expand the number of restriction sites available for the next cloning steps. This was followed by a ligation reaction and a transformation step into the Apa I and Sst I restriction sites of the ploxP plasmid (Figure 4.7B). The resulting plasmid was called ploxP/T2MA.

c) Subcloning of the 3’ arm
The 3’ arm was 3.9kb in length and contained the second and third coding exons (exons 4 and 5). It was amplified using primers XhoRarm1.F and XhoRarm3.R. The PCR fragment was then digested by Xho I and purified for cloning. It was first subcloned into pBSIISK+ and then into the Xho I sites of the ploxP/T2MA plasmid previously constructed (Figure 4.7C). The resulting plasmid containing the middle and 3’arm separated by a loxP site was called pT2/MAloxPRA.

d) Final cloning step
The last cloning step joined the 3 homology arms together by ligating the Eco RI fragment of the plasmid containing the middle arm and the 3’arm (pT2/MAloxPRA) into the Eco RI site of the construct containing the 5’arm (pneolox/T2LA). Once positive recombinants were identified, one was chosen and thoroughly checked by further restriction digests (Figure 4.7). The size of the final plasmid (pT2CTV) was 10.3kb and it was linearised by a Not I and Bam HI restriction digests. Eco RI cuts the construct into 2 fragments separating the backbone vector containing the 5’arm (4.2kb) from the insert containing the middle arm and the 3’arm (6.2kb). Other restriction
Figure 4.8. Final cloning step of the \textit{Tln2} conditional targeting vector. (A) The 6.1 \textit{Eco} RI fragment of the plasmid containing both middle arm and 3’arm (pLoxP/T2MAloxPRA) was digested and subcloned in the \textit{Eco} RI site of the pneolox/T2Larm plasmid. The resulting plasmid containing the three isogenic arms, the \textit{loxP} sites at the right positions and the Neo\textsuperscript{R} cassette was 10.3 kb in length and named pT2CTV. (B) The \textit{Tln2} conditional targeting vector was tested to ensure the correct orientation and the right position of the different elements. Different restriction digests were performed: (1) \textit{Not} I, (2) \textit{Bam} HI, (3) \textit{Eco} RI, (4) \textit{Bam} HI and \textit{Eco} RI, (5) \textit{Eco} RV, (6) \textit{Cla} I, (7) \textit{Sst} I. In the first lane (M) was loaded a size marker 1KB+ DNA ladder (Invitrogen).
digests using *Bam* HI + *Eco* RI, *Eco* RV, *Cla* I and *Sst* I allowed to check the correct orientation of the arms in the targeting vector and confirmed that the construct was correctly organised.

4.3.2. Generation of a PCR positive control vector for the screening of the *Tln2* homologous recombination event

Homologous recombination events were identified by screening of *Neo* \(^R\) clones using the strategy shown in Figure 4.9A. This PCR used a forward primer upstream of the targeted region and a reverse primer within the *Neo* \(^R\) gene. A PCR product was present only if homologous recombination occurred in ES cells. If there was no PCR product, it was assumed that the vector had integrated randomly in the genome. In order for this strategy to be validated, a positive control for the PCR screen had to be generated. A plasmid containing a genomic region amplified by PCR which is 150bp bigger than the left arm amplified for the complete targeting vector allowed the design of a primer outside the targeted region. This control vector was constructed using the same cloning strategy as the first cloning step of the targeting vector. A genomic region was amplified using the same cycling conditions as for the amplification of the vector arms. Primers *KpnLarm2.F* and *ClaLarm1.R* were used to amplify from E14.1a ES cell genomic DNA and gave rise to a 1.35kb product. The PCR fragment was digested by *Kpn* I and *Cla* I and the fragment was ligated into *ploxPneoloxP* plasmid and digested with the same enzymes. The ligation was transformed into DH5\(\alpha\) cells and positive colonies were screened. The resulting construct is shown in Figure 4.9.

4.3.3. Generation of *Tln2* conditional knockout ES cells

After the *Tln2* targeting vector had been thoroughly checked by restriction digest and sequencing, it was used in a transfection experiment on ES cells in order to generate conditional knockout ES clones.

E14.1a ES cells at an early passage (P3 from the harvesting) were electroporated with the *Tln2* conditional targeting vector (pT2CTV). Two experiments were carried out and for each of them 5 x 10\(^7\) ES cells were electroporated with 40\(\mu\)g of DNA linearised by digestion with *Not* I, which cuts in a unique site that is located downstream of the 3’arm (Figure 4.10). For each experiment the transfected cells were plated onto four 10cm
**Figure 4.9. Identification of the homologous recombination event and generation of a PCR positive control.** (A) A PCR strategy to identify the homologous recombination event in ES cells was established; using a forward primer upstream of the targeted region in the second Tln2 intron (gT2In2F3) and a reverse primer in the NeoR (3'neorev). This 1.5kb fragment was only amplified in homologous recombinant ES clones. Random integrants did not give rise to a PCR product using these primers. (B) To construct a positive control for PCR screening, a longer version of the 5’ arm (1.35kb), extended 150bp upstream was amplified from ES cell genomic DNA using primers KpnLarm1.2 and ClaLarm1.R. The PCR fragment was then digested by Kpn I and Cla I and subcloned into the Kpn I and Cla I restriction sites of ploxPneoloxP. The resulting construct pT2PCR+ve was 6295bp long.
Figure 4.10. PCR screening for ES cell homologous recombinants. After electroporation of the Tln2 targeting vector into ES cells a strategy based on PCR was used to identify homologous recombination events. (A) Two different PCR strategies were used to genotype ES clones. The primers previously designed (gT2In2F3 and 3’Neorev) gave rise to a 1.5kb product and specifically identified the homologous recombination event. Primers gT2In4R and gT2In3F gave a 377bp (loxP PCR) from the wt allele and a 431bp from the floxed allele. (B) A total of 600 clones were screened (data not shown) and two homologous recombinant clones (#261 and 219) were identified. The two PCRs described in (A) were performed on genomic DNA from each clone as well as on genomic DNA from wt E14.1a ES cells. In the top photo is shown the PCR, which gave rise to a 1.5kb product on both clones #219 and #261 but no amplification from the wt DNA. The 3’ loxP PCR shown on the bottom picture gave a double band with clones #261 and #219 corresponding to the wt and floxed alleles, and gave a single band of 377bp with the wt DNA indicating the presence of a loxP site in the targeted clones.
dishes containing MEF feeders and fresh ES media. The next day, the media was changed to a media containing 250μg/ml of G418. A week later, 300 clones were picked, grown, frozen down and the genomic DNA extracted for a PCR screen. Overall 600 clones were picked and screened, out of which 2 homologous recombinants were identified (clones #219 and #261). All the clones were used in a PCR reaction using primers gT2In2F3 (forward primer upstream of the targeting region in Tln2) and 3'neorev (reverse primer in the NeoR cassette). The PCR amplification was performed using Taq Polymerase in a 10μl final reaction and gave an amplification product of 1.5kb.

The two homologous recombinants identified were thawed out, expanded and their genotype checked a second time by PCR in order to ensure the presence of the targeting event. Two different reactions were performed in order to check these clones; the first one checked for the homologous recombination event as for the ES cells screen. A second PCR was used in order to check for the presence of the 3'loxP site (gT2In4R and gT2In3F) on these same 2 clones. Amplification from the wt allele resulted in a 337bp band whereas the floxed allele resulted in a 431bp band. The results are shown Figure 4.10 and they confirm that in the clones #261 and #219 contain the correct targeting event.

4.4. Generation of Tln2 conditional knockout mice

The harvesting of the blastocysts, the transfer into a surrogate mother and the breeding of the chimaeras were performed by the Transgenic Unit, Division of Biomedical services under the Home Office regulations.

The clones #219 and #261 in which homologous recombination of the Tln2 targeting vector had occurred were thawed out at the lowest passage as possible. The morning of the injection, the ES cells were trypsinised and resuspended in injection media and taken on ice to the Transgenic Unit, Division of Biomedical services. The blastocysts were obtained from mating of two C57BL/6 mice, and harvested at 3.5dpc. Between 8 and 10 ES cells were injected per blastocyst and depending on the amount of blastocysts harvested, 10 blastocysts injected were transferred to surrogate superovulating female.
Chapter 4 – Generation of a Tln2 conditional knockout

One male chimaera per cell line was born from the injection sessions. Each chimaera was mated to two C57BL/6 females and the presence of agouti pups in the litter showed germline transmission of the injected ES cells. For the clone #261, germline transmission was achieved in the first litter. Overall, a total of 40 agouti pups were born in 4 litters from 2 different mothers. The agouti pups came from the injected Tln2fl ES cells and therefore can inherit either the Tln2 wt or fl allele. In order to identify the presence of Tln2 floxed allele all these agouti animals were genotyped by PCR. A PCR reaction was performed on DNA extracted from tail samples using the Gene Elute mammalian genomic DNA kit. This kit was used to obtain good quality DNA on which the 5' long range PCR could be performed successfully. The standard lysis of tail samples gave a quality of DNA on which allowed only short PCR amplifications to be performed. Therefore, it was important to check that Tln2 was correctly targeted in the mice that would be used to start a colony and used for further analysis. The two PCR strategies described in section 4.3.3 were used. For the clone #261 a total of 20 Tln2fl animals were born and those animals were used for further analyses.

4.5. Deletion of the Tln2 conditional allele using Cre recombinase

4.5.1. Action of Cre recombinase in mouse embryonic fibroblasts (MEFs)

The efficiency of the Cre recombinase mediated recombination on the Tln2 floxed allele was tested in MEFs generated by interbreeding heterozygous Tln2fl animals. Intercrosses were set up between a male and a female Tln2fl in order to generate MEF cell lines. A litter was harvested from a 13.5dpc pregnant female as indicated in Materials and Methods. All the embryos were genotyped and used to generate an individual cell line. Three cell lines only were kept for further analysis: #1.1 Tln2fl, #1.2 Tln2+/+ and #1.6 Tln2fl. These were grown and expanded onto 2 x 10cm dishes for transfection with a plasmid transiently expressing the Cre recombinase (pCrePAC, map in Materials and Methods section 2.3.6). The transfection was performed using the Nucleofector® kit (Amaxa) using 10μg of plasmid DNA. A total of 1 x 10⁶ cells at 70-80% confluence were used for the experiment. After transfection, the cells were replated onto 2 x 10cm dishes and left for 24 hours. Then the MEFs were lysed for DNA extraction using the GeneElute Mammalian DNA extraction kit for PCR. A PCR strategy was established to identify the different possible alleles resulting from the Cre
recombinase mediated deletion (Figure 4.11). Four primer sets were designed using the sequence data from the Ensembl database.

The position of the primers is shown in Figure 4.11. Two forward primers were used, one upstream of the $\text{Neo}^R$ cassette in intron2 (gT2In2F4, (a) ) and the second one in the $\text{neo}^R$ cassette (pgk3’, (b) ). Two reverse primers were designed, one was located in the middle arm upstream of exon3 (gT2In3R, (c) ) and the second one was located downstream of the 3$'$ $\text{loxP}$ site, within intron4 (gT2In4R, (d) ). Three different pairs of primers were used to identify the different alleles and are shown in Table 4.4.

<table>
<thead>
<tr>
<th>Allele</th>
<th>Floxed</th>
<th>Wt</th>
<th>$\Delta\text{Neo}$</th>
<th>$\Delta\text{ex3}$</th>
<th>$\Delta\text{Neo}+\Delta\text{ex3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) + (d)</td>
<td>3.3kb</td>
<td>1.2kb</td>
<td>1.4kb</td>
<td>2.5kb</td>
<td>569bp</td>
</tr>
<tr>
<td>(a) + (c)</td>
<td>2.4kb</td>
<td>325bp</td>
<td>443bp</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(b) + (d)</td>
<td>1.5kb</td>
<td>-</td>
<td>-</td>
<td>692bp</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.4. Genotyping of the different $\text{Tln2}$ alleles resulting from the Cre mediated recombination. The different primers combinations allowed the identification of the different Cre recombined alleles. The fragment sizes indicated in bold correspond to the strategy used for each allele.

The wt allele was identified by a 325bp band using primers gT2In2F4 (a) and gT2In3R (c). The allele in which the Cre recombinase had deleted the $\text{Neo}^R$ cassette ($\Delta\text{Neo}$) gave a 443bp band using primers gT2In2F4 and gT2In3R. The allele in which only exon 3 has been deleted ($\Delta\text{ex3}$) gave rise to a 692bp band using primers pgk3’ (b) and gT2In4R (d). The allele containing the deletion of the whole floxed region including the $\text{Neo}^R$ cassette and exon3 ($\Delta\text{Neo} + \Delta\text{ex3}$, null) gave rise to a 569bp fragment using primers gT2In2F4 + gT2In4R. Each primer pair was used in a separate reaction using the ABGene PCR Mastermix® in a 15µl reaction. The PCR conditions were the standard 60° annealing temperature and a 45 second extension at 72° for 35 cycles.

For each MEF cell line generated, DNA was extracted 24 hours after transfection and used as a template in a PCR reaction. Cre mediated deletion was assessed by PCR amplification using the different primer sets described above. DNA from ES
Figure 4.11. Genotyping of the different *Tln2* alleles after exposure to Cre. Upon Cre mediated recombination of a heterozygous *Tln2* \(^{fl/+}\) animal, five different alleles can be identified using primers located in different areas of the allele. (a) gT2In1F4 (b) pgk3' (c) gT2In3F (d) gT2In4R.
Figure 4.12. Cre mediated deletion of the Tln2 floxed allele in MEFs. Three different MEF cell lines were generated from 13.5dpc embryos from Tln2\textsuperscript{1/l+} intercrosses. These were genotyped by PCR as Tln2\textsuperscript{1/l+} (1), Tln2\textsuperscript{2+/2} (2) and Tln2\textsuperscript{2/2} (3). They were used in a transient transfection experiment using the pCrePac plasmid. PCR was performed 24 hours after transfection in order to identify the different alleles resulting from the Cre mediated recombination. The top picture represents PCRs with primers (a) and (d), the middle picture with primers (a) and (c) and the bottom picture with primers (b) and (d).
Chapter 4 - Generation of a Tln2 conditional knockout

$Tln2^{0/+}$ cells was also used as a positive control for each of the reactions. The PCR results are shown in Figure 4.12. In the #1.1 $Tln2^{0/+}$ cell line, 24 hours after transfection, the $Tln2$ fl, wt, Δex3 and null were detected showing that the deletion had not occurred in all cells. Only the deletion of the Neo$^R$ cassette was not generated. In the #1.2 $Tln2^{+/+}$ cell line, 24 hours after transfection, only the $Tln2$ wt allele was detected confirming the specificity of Cre recombinase on the $Tln2$ floxed allele. In the #1.6 $Tln2^{0/0}$ cell line, 24 hours after transfection, all the different $Tln2$ alleles (fl, wt, Δex3, ΔNeo and null) were detected showing that although the deletion had not occurred in all the cells the tri/ox strategy chosen to target the gene was working. The wt allele in the cell line #1.3 after treatment with Cre was present although it was initially genotyped as $Tln2^{0/0}$. This might be due to an eventual contamination during the Cre transfection. But this experiment gave evidence that the loxP site orientation and location were correct.

4.5.2. Action of Cre recombinase in mice

a) Crossing of $Tln2^{0/+}$ animals to CMV-Cre mice

A Cre expressing transgenic mouse colony that ubiquitously expresses the recombinase in all cell lineages was available. In those mice, the Cre expression cassette is under the control of the human cytomegalovirus (CMV) promoter (Schwenk et al., 1995). The CMV promoter provides a strong and constitutive expression in all cell lineages including germline cells. The transgene is located on the X chromosome and so from a mating between a male CMV-Cre mouse and a female $Tln2^{+/-}$, all the female pups in the litter will inherit the CMV-Cre from their father.

Four matings between a $Tln2^{0/+}$ and a CMV-Cre animal were set up and the litters were left to birth. From these matings different genotypes can occur at the time of fertilization. Two alleles can be present at the $Tln2$ locus: floxed (fl) and wt (+), two alleles can be present at the CMV-Cre locus (CMV). So four possible genotypes could occur: $Tln2^{0/+}Cre^{+/+}$, $Tln2^{0/+}Cre^{CMV/-}$, $Tln2^{+/+}Cre^{CMV/+}$ and $Tln2^{+/+}Cre^{+/+}$. In order for the deletion to occur at the $Tln2$ locus, the presence of the floxed allele and of the CMV-Cre allele was necessary.

At 3 weeks of age the offspring were weaned and a tail sample was taken, lysed and DNA was used for genotyping. All 40 offspring, which were born from those matings,
Chapter 4 - Generation of a Tln2 conditional knockout

were tested for the presence of the Cre allele (data not shown) and PCR genotyping for the different Tln2 alleles was performed on these animals. Each of the PCRs described in section 4.5.1. were used in the same cycling and reaction conditions as stated before, on all the offspring expressing that were identified as Cre positive. A summary of the results is shown in Table 4.5.

<table>
<thead>
<tr>
<th>sex</th>
<th>Animal ID</th>
<th>wt</th>
<th>fl</th>
<th>Δneo</th>
<th>Δex3</th>
<th>Null</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tln2 locus</td>
</tr>
<tr>
<td>F</td>
<td>2CD1.5</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
<td>+/-m</td>
</tr>
<tr>
<td>F</td>
<td>2CD1.6</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>+</td>
<td>+/-m</td>
</tr>
<tr>
<td>F</td>
<td>2CD1.7</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td>+</td>
<td>+/-m</td>
</tr>
<tr>
<td>F</td>
<td>2CD1.8</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>+/-m</td>
</tr>
<tr>
<td>F</td>
<td>2CD2.7</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
<td>+/-m</td>
</tr>
<tr>
<td>F</td>
<td>2CD3.2</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>+</td>
<td>+/-m</td>
</tr>
<tr>
<td>M</td>
<td>2CD4.1</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>M</td>
<td>2CD4.2</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>M</td>
<td>2CD4.4</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Table 4.5. In vivo Cre mediated deletion of the Tln2<sup>fl/wt</sup> allele.

The animals in which only deletion of exon 3 and Neo<sup>R</sup> was detected is noted as +/-, and the animals in which the - allele as well as the other Cre deleted alleles were present are noted as +/-m.

A total of eight animals containing both Tln2<sup>fl/wt</sup> and expressing Cre recombinase were identified; 6 of these were female and 2 were male. Four of the females were mosaic at the Tln2 locus (noted +/-m), as they contained the Tln2 wt and floxed allele as well as alleles resulting from the Cre mediated recombination Δex3 and - (Δex3 + ΔNeo). In the four other pups (2 males and 2 females), the PCR genotyping detected only the presence of the null and the wt allele, suggesting that a complete deletion of the Tln2 floxed allele had occurred in all the tail cells of these animals. It is noteworthy that the deletion of only the Neo<sup>R</sup> cassette was not present in any of the mosaic animals. The Tln2<sup>+/-</sup> animals were used to start up a breeding colony by backcrossing onto the C57BL/6 inbred background. The backcrosses generated more Tln2<sup>+/-</sup> animals that could be used for phenotypic and functional analyses.

b) Generation and analysis of Tln2<sup>+/</sup> mice

In order to determine the effect of the deletion of the first coding exon of Tln2, some intercrosses between Tln2<sup>+/</sup> animals were set up. Before any further analysis, it was important to determine whether the phenotype was embryonic lethal, 7 litters were set up and left to give birth. The animals were genotyped after weaning to determine
whether the $Tln2^{+/}$ animals survive during embryonic development up to birth. A total of 45 pups were born, all were genotyped after weaning. Two PCRs were used to identify the two possible alleles at this stage: gT2In2F4 + gT2In4R for the null allele (569bp), and pgk3' + gT2In3R (566bp) for the wt allele. The two reactions were performed in a separate reaction tube in the same conditions as described in Materials and Methods using 15μl final volume with 1μl of tail sample DNA. The genotyping result showed that the $TlnT1$ animals survived to birth. Out of 45 pups born in total, 27 $Tln2^{+/-}$ (20%) were present, 9 $Tln2^{+/-}$ (60%), 7 $Tln2^{-/-}$ (15.6%) and 2 animals for which the genotype could not be determined (4.4%). If the survival rate was 100% the proportions of the different genotypes would be $Tln2^{+/-}$: 50%, $Tln2^{+/-}$: 25% and $Tln2^{-/-}$: 25%. This would indicate that $Tln2^{+/}$ are obtained at lower frequency than expected.

4.6. Discussion
A conditional knockout approach based on the use of the Cre recombinase with a tri-lox system was chosen in order to better understand the function of talin2 in mouse cells and animals. This approach was chosen because it has been very widely used and also a large range of Cre administration modes are available, in vivo as well as for cell culture. The principle of this technology is to insert recombination sites (loxP) flanking a region of interest. These sites can be recognised by the Cre recombinase, which will mediate the recombination and deletion of the fragment of DNA flanked by those sites. For $Tln2$, the region chosen was exon3, which is the first coding exon containing the initiating ATG codon. This was chosen because the removal of the first coding exon would remove the start codon for translation initiation and hopefully result in the inability to express talin2.

(i) Mapping of the 5' end of $Tln2$ and generation of the targeting vector
The first step was to construct a targeting vector containing 3 arms of homology with $Tln2$ gene, the loxP sites correctly positioned and a selection marker. In order to construct that vector using the plasmids available ploxP and ploxPneoloxP, it was important to have more information on the gene structure and also establish a restriction map of the 5' end of $Tln2$. A genomic cosmid clone screened from a 129ola genomic library containing the first coding exon was used to establish a map of the gene. It was digested by various enzymes and the digests were used in Southern blot analyses using
5' radiolabelled oligonucleotides located in different exons. The results of those experiments were used to establish a restriction map of the region of Tln2 containing the first three coding exons. This map was confirmed when the sequencing of the genomic region containing Tln2 was completed and published in the Ensembl database. The combination of both sets of data allowed a detailed map of the region to be established and therefore a cloning strategy could be determined. Interestingly the data from both the genomic clone mapping (129Ola strain) and the Ensembl sequence database (C57BL/6 strain) were very similar, although they came from two different mouse strains.

The targeting vector was constructed using the ploxP and ploxPneoloxP plasmids. The three homology arms were generated by PCR with primers designed based on sequence data available from the http://www.ensembl.org. Restriction sites sequences were added to the 5' end of the primers in order to be able to clone the resulting PCR fragments according to the chosen strategy. The 5' arm was 1.2 kb in length and contained a portion of intron 2; the middle arm was 0.7 kb in length and contained exon 3; the 3' arm was 3.9 kb in size and contained exon 4 and 5. The genomic DNA from the cosmid was used as a template to amplify those isogenic regions except for the 5' arm that was amplified from wt E14.1a ES cells genomic DNA. The different arms of the targeting vector were successfully assembled with the loxP sites and the NeoR cassette. The final construct was finally checked by multiple restriction digests to confirm the correct orientation of the different elements in it.

(ii) ES cells targeting and generation of Tln2<sup>0/+</sup> mice

The targeting vector was linearised by a Not I digest and electroporated into ES cells. The transfectants were selected by G418 application. The transfection experiment was performed twice and a total of 600 G418 resistant clones were picked and screened by PCR for homologous recombination. Two homologous recombinants were identified (clones #261 and #219, one in each experiment), expanded and further checked by PCR. Both clones were injected into blastocysts. Each of them generated a chimaera in which germline transmission of the Tln2<sup>0/+</sup> ES cells was achieved. Agouti pups in the litter of mating between a chimaera and C57BL/6 female originated from germ cells generated from the Tln2<sup>0/+</sup> ES cells and therefore were genotyped to check for the presence of the
fl or wt allele at the \textit{Tln2} locus. From those matings, 20 \textit{Tln2}^{\text{fl+/+}} animals were born and could be used to start a colony and further analysis.

(iii) Application of Cre recombinase in MEFs and \textit{in vivo}

Intercrosses between \textit{Tln2}^{\text{fl+/+}} animals were set up and MEFs were harvested at 14.5dpc. Three cell lines (\textit{Tln2}^{+/+}, \textit{Tln2}^{\text{fl+/+}} \textit{Tln2}^{\text{fl+/fl}}) were used for an \textit{in vitro} application of Cre using the pCrepac plasmid. The transfection was performed using the Nucleofector™ system (Amaza), 24 hours after transfection DNA was extracted from the cells and used for PCR amplification. The PCRs showed that the Cre mediated recombination was effective on the \textit{Tln2} floxed allele and did not affect the wt \textit{Tln2} gene. This confirms that the location and the position of the \textit{loxP} sites allow the deletion of the floxed region to occur properly and therefore can be used for further \textit{in vitro} or \textit{in vivo} analyses.

The \textit{Tln2} floxed allele generated is a powerful tool to inactivate \textit{Tln2} expression and therefore study the role of \textit{Tln2} in mouse development and adult tissues. A mouse strain present in which the Cre recombinase is under the control of the CMV promoter was used to generate \textit{loxP} site deletions \textit{in vivo}. A total of four matings between a \textit{Tln2}^{\text{fl+/+}} animal and a CMV-Cre expressing mouse were set up. Out of all the offspring, four \textit{Tln2}^{+/+} animals in which the complete deletion of the allele (removal of \textit{NeoR} and of the middle arm containing exon3) had occurred were generated including three females and two males. These matings also gave rise to mosaic animals. These contained a heterogeneous population of cells in which all different possible deletions of the \textit{Tln2} allele had occurred. Interestingly no animals contained the deletion of the \textit{NeoR} cassette. The Cre strain used is X-linked. The X chromosome inactivation in females starts at the implantation of the embryo and occurs randomly with respect to the parental origin of the chromosome, which results in cellular mosaicism (Migeon, 1994).

The next step was to analyse whether the \textit{Tln2}^{+/+} mice had a phenotype and particularly if the homozygous \textit{Tln2}^{+/+} were embryonic lethal or not. Intercrosses between \textit{Tln2}^{+/+} animals were set up and the litters were left to birth. The pups were genotyped and the result showed that there were some \textit{Tln2}^{+/+} animals that had survived to birth. The proportion of \textit{Tln2}^{+/+}, \textit{Tln2}^{+/+} and \textit{Tln2}^{+/+} animals were 20%, 60% and 15.6% (4.4% of animals in which no accurate genotype was given). Therefore the \textit{Tln2}^{+/+} mice were obtained at slightly lower frequency than expected if the animals were viable. This
could be due to the fact that some animals die during embryogenesis but most do survive. More matings would be required in order to determine whether there is some degree of embryonic lethality or not. The \( Tln2^{fl/+} \) animals that were used for in vivo Cre deletion were a mixed strain between 129Ola and C57BL/6. The phenotype could be further studied once pure colonies on either 129Ola or C57BL6 background have been established. These colonies would be established by successive backcrosses to 129Ola or C57BL6 animals after a minimum of 10 generations.

The removal of the first coding exon might have a different effect depending on the mouse strain, which might influence the regulation of the expression of \( Tln2 \). The lack of apparent phenotype could have more explanations. (i) The loss of the first coding exon only removes a small portion of the coding sequence and therefore a full length transcript of 8.5kb a potential translation initiation site can be used downstream on the mRNA. However that hypothesis has been tested by analysis of the in-frame ATG codons in good Kozak environment. The following methionine codons in a decent Kozak consensus sequence (6/10) are situated at positions 2302, 2677 and 3310, corresponding to amino acids 683, 808 and 1019 on the protein. (ii) It could also be due to the fact talin1 and talin2 share a degree of identity of 75%, that talin1 is ubiquitously expressed and might therefore compensate for the loss of Talin2 during the embryonic development and in adult tissues (iii) It could also be explained by the fact that talin2 has a function in specific tissues such as brain or heart where it is expressed, that is not sufficient to cause embryonic lethality. It may give however subtle neuronal disfunction that could be explored by further analyses. Further investigations of the mice are required to explore each of these possibilities.
Chapter 5. Discussion

Talin is a large cytoskeletal protein composed of a globular head and a flexible tail (Isenberg et al., 2002) localised in structures called focal adhesions (FA) where the actin filaments link to the extracellular matrix via the integrin receptors. FA assembly and disassembly play key roles in physiological processes such as cell migration and cell spreading and are tightly regulated. Talin plays a key role in the assembly of FA by initiating the coupling of the integrin receptor subunits to actin filaments. It has been shown to bind to β1 and β3 integrin subunits via its FERM domain, in the head region. Its role in vivo has been extensively studied in several species using knockdown strategies. In D. Discodeum, C. Elegans, D. Melanogaster and M. Musculus talin is a major component of integrin mediated cell matrix adhesion and the absence of talin leads to severe phenotype in cell attachment and spreading.

The conclusions from all these knockdown studies have been questioned since the recent discovery of a second talin gene, that codes for a talin homolog protein (talin2) present in human, mouse, rat and chicken. The two proteins share a high degree of high homology (75% in human) throughout their length and have been shown to bind the same ligands. Analysis using Northern blot studies on mouse tissues revealed that talin2 expression varied amongst mouse tissues and that short talin2 mRNAs were expressed in specific tissues such as kidney and testis (Monkley et al., 2001). In order to further understand the function of talin2 in vivo, two different approaches have been undertaken in this thesis.

Firstly, further expression analyses using Northern blots were carried out. The results confirmed that Tln2 transcript sizes vary. Large Tln2 transcripts potentially encoding for the full-length protein are expressed in most of the tissues analysed, although at a higher level in heart and brain. Shorter transcripts were expressed in mouse kidney and testis, and these only contain part of the Tln2 coding sequence, but lacking the 5' end, which encodes the N terminal actin binding site. The shorter transcript in kidney was also observed in human. This expression pattern could be the result of alternative splicing,
expression from an alternate promoter or a combination of both. In order to determine which hypothesis was the right one, a human kidney cDNA library was screened. PCR and sequencing identified a novel transcript, which could be expressed in mouse and human kidney. A novel exon was identified, and it was confirmed by the presence of ESTs present in the mouse and human databases. Protein analyses did not give evidence that this shorter transcript was translated into a functional shorter talin2 protein. However due to the absence of reliable antibodies specific to talin2, this possibility cannot be excluded.

A Tln2 genetrap mouse line was used to carry out analysis of the expression of talin2 in vivo. Talin2 is expressed as a fusion protein with the reporter gene β galactosidase. The data confirm that talin2 is expressed in various adult tissues such as heart and brain but also particularly in smooth muscle. In the developing embryo, it is highly expressed in heart and part of the brain. The genetrap approach was a good tool to gain a better understanding of talin2 expression in mouse tissues. However the fact that talin2 is expressed as a fusion protein containing only the N terminal region of the protein means that subcellular localisation might not reflect that of the wild type talin2 protein. If the shorter transcripts identified by the mRNA expression analysis lead to the expression of a talin2 functional protein, its expression would also not be reflected by the genetrap construct.

A conditional knockout approach based on the use of a Cre-tri/ox system to target Tln2 was also undertaken. The 5' end of the gene was mapped, a targeting vector as constructed and the gene was targeted in mouse ES cells. Transgenic mice in which the first coding exon of Tln2 was flanked by loxP sites were successfully generated. Cre deletion in MEFs confirmed that the loxP sites inserted could be correctly recognised by Cre recombinase. Cre deletion was performed in vivo, by crossing Tln2+/+ animals to a transgenic mouse expressing Cre under the control of CMV promoter. Homozygous Tln2−/− animals were generated and viable in preliminary crosses and no direct phenotype was observed. Tln2−/+ animals were obtained at a slightly lower frequency than expected. The surviving Tln2−/+ had no apparent phenotype. The absence of phenotype may reflect the true null phenotype for Tln2 and would be supported by the
genetrap data published by (Chen and Lo, 2005) showing that the \( Tln2^{GT/GT} \) mice also survive post natally. However the RNA expression analyses suggest that we cannot exclude the expression of truncated proteins in either of these mouse strains, which would not have been knocked down by the \( Tln2 \) conditional allele or the genetrap allele.

In the targeting strategy used only the first coding exon was removed and although no close downstream ATG initiating codon in good Kozak environment was found downstream, a downstream potential reinitiation cannot be excluded. If this is the true null phenotype, these data would suggest compensation of talin2 by talin1 in the tissues expressing talin2. Moreover, if as the mRNA expression data suggest, \( Tln2 \) expression pattern is complex and the gene undergoes alternative splicing events or is expressed under the control of alternate promoters, more than one talin2 protein might be expressed in mouse. It is a possibility that the major talin2 protein \textit{in vivo} is expressed from a downstream promoter or does not contain the first coding exon, meaning that the deletion of the first coding exon only will have no effect on its function.

In order to further investigate how \( Tln2 \) is expressed and its function in mammals, more detailed analysis on the structure of the transcripts expressed would be required, determining their exon profile and their tissue specificity. The generation of reliable specific talin2 antibodies, raised against different regions of the protein would give more information on whether shorter talin2 proteins are expressed in certain tissues. The conditional knockout approach could be used but a deletion of the complete coding sequence, leading to the inhibition of all expression from the \( Tln2 \) gene would be required.

These different approaches would help determine whether talin2 plays a similar role as talin1 in integrin mediated cell adhesion. It could also bring information on whether talin2 plays a different role in cell adhesion, or if it acts in synergy with talin1 and both proteins are required for normal tissue/cell physiology.
References


talin gene arrests mouse development at the gastrulation stage. Developmental Dynamics 219, 560-574.


