Investigations of the Ufm1 pathway and its association with a familial form of hip dysplasia

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medical and Human Sciences

2013

Michal Dudek

School of Medicine

Table of Contents

List of Figures

List of Tables

List of Supplementary Figures

Abstract

Michal Dudek, University of Manchester
PhD in the faculty of Medical and Human Sciences 2013
Investigations of the Ufm1 pathway and its association with a familial form of hip dysplasia

Beukes Hip Dysplasia (BHD) is an autosomal dominant disorder where the abnormal shape of the hip joint leads to secondary osteoarthritis. The locus of BHD has been previously mapped to 4q35 and screening of candidate genes within this region revealed a mutation in the gene encoding the Ubiquitin-fold modifier 1 specific protease 2 (Ufsp2). The mutation prevents Ufsp2 from cleaving its target, Ufm1. Ufsp2 and Ufm1 are both components of a novel ubiquitin-like protein modification system which involves Ufm1 being processed via the E1, E2 and E3 enzymes (Uba5, Ufc1 and Ufl1, respectively) and attachment to target protein(s) one of which has been identified (Ddrgk1). The aim of this study was to investigate the link between the *UFSP2* mutation and the BHD phenotype by: (i) examining the expression of components of the Ufm1 system *in vivo* and *in vitro*; (ii) development of an *in vitro* Ufm1 conjugation system; and (iii) generation and analysis of transgenic mice overexpressing the *Ufsp2* gene with the BHD mutation.

The expression of *Ufsp2* was determined by radioactive RNA *in situ* hybridisation of mouse tissue sections. The analysis revealed *Ufsp2* expression predominantly in the bone of the hip joint and in the bone and secondary ossification centres of the knee of 10 day old mice. Real-Time PCR analysis showed increased expression of components of the Ufm1 system during *in vitro* osteogenic and chondrogenic differentiation which coincided with induction of ER stress evidenced by upregulation of *Bip*. These components were also upregulated in response to chemically induced ER stress *in vitro*. Analysis of the promoter regions of Ufm1 system genes identified unfolded protein response elements in the upstream sequences of *Uba5, Ufl1, Ufm1* and *Lzap* genes and the elements in *Uba5* and *Lzap* were found to be required and responsive to ER stress using luciferase promoter assays.

A Tandem Affinity Purification method was developed for isolation and identification of Ufm1 conjugation targets from cell lines expressing modified forms of Ufm1. Mass spectrometry analysis of Ufm1 conjugates purified from HEK293T cell line identified Uba5 and Ufc1 but no new Ufm1 targets. Western blot comparison of Ufm1 conjugated proteins purified from the HEK293T and 2T3 osteoblast cell lines identified putatative Ufm1 conjugation targets and increased conjugation in osteoblasts in response to ER stress. One of these targets was identified as Ddrgk1 but the remaining putative targets remain to be identified by mass spectrometry.

Transgenic mice overexpressing the mutated *Ufsp2* gene were generated and subjected to phenotypic analysis. No significant differences were found between transgenic and wild type mice following X-ray, histological and weight analysis.

Higher expression of *Ufsp2* in bone and secondary ossification centres as well as upregulation of components of the Ufm1 system in response to ER stress suggests that the molecular pathway between the *UFSP2* mutation and the BHD phenotype may relate to abnormal ER stress responses during osteoblast differentiation. Further studies are however required to determine how the Ufm1 system modulates ER stress responses and how disruption of these processes caused by the *UFSP2* mutation causes BHD.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

I would like to thank my supervisor Gillian Wallis for her tremendous guidance over the past four years and for the invaluable help in writing this thesis. I am indebted to the members of Boot-Handford and Canfield labs, past and present, for sharing their knowledge and experience with me. I am also grateful to my lovely wife Ola for supporting me in this endeavour, proofreading the thesis and for introducing me to coffee when we first met which significantly improved my academic performance.

Abbreviations

Agc1 - aggrecan
AMP - adenosine mono phosphate
APC - anaphase promoting complex
Aspn - asporin
Atf4 - activating transcription factor 4
ATG8 - autophagy-related protein 8
ATP - adenosine tri phosphate
BBFH7 - BBF2 human homolog on chromosome 7 (CREB3L2)
BHD - Beukes hip dysplasia
BiP - heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)(HSPA5)
BLAST - Basic Local Alignment Search Tool
BMP - bone morphogenic protein
bp - base pair
bZIP - basic-leucine zipper
CAND1 - cullin-associated and neddylation-dissociate 1
CCD - cleidocranial dysplasia
Cdc - cell division cycle
CDK2 - cyclin dependent kinase
CD-RAP - cartilage-derived retinoic acid-sensitive protein
CFTR - cystic fibrosis transmembrane conductance regulator
CLS - Coffin-Lowry syndrome
CMD - camptomelic dysplasia
CMV - cytomegalovirus
Col1α1 - collagen type I alpha 1
Col2α1 - collagen type II alpha 1
COMP - cartilage oligomeric matrix protein
COPII - coat protein complex
CYLD - cylindromatosis
DDH - developmental dysplasia of the hip
Ddrgk1 - DDRGK domain containing 1
DMEM - Dulbeccos's modified Eagle's medium
DNA - deoxyrybonucleic acid
DTT - dithiothreitol
DUB - deubiquitinating enzyme
EBV - Epstein-Barr virus
ECM - extracellular matrix
eIF2 - eukaryotic initiation factor 2
ER - endoplasmic reticulum
ERAD - ER associated degradation
FAT10 - leukocyte antigen F-associated (UBD)
FGF-10 - fibroblast growth factor 10

FRZB - frizzled-related protein
Gdf5 - growth/differentiation factor 5
GDP - guanosine di phosphate
GFP - green fluorescent protein
GTP - guanosine tri phosphate
HECT - homologous to E6-AP C-Terminus
HPV - human papillomavirus
Ibsp - integrin binding bone sialoprotein
Igf1 - insulin-like growth factor 1
Ihh - indian hedgehog
IRE1α - inositol-requiring protein 1 alpha
ISG15 - interferon stimulated gene 15
JAG1 - jagged 1
JAK - Janus kinase
JAMM - JAB1/MPN/Mov34 metalloenzyme
JNK - c-Jun N-terminal kinase 1
kDa - kilo Dalton
LC - liquid chromatography
LOD - logarithm (base 10) of odds
Lzap - LXXLL/leucine-zipper-containing ARF-binding protein (Cdk5rap3
MAD2 - mitotic arrest deficient 2
MCDS - metaphyseal chondrodysplasia type Schmid
MED - multiple epiphyseal dysplasia
MEM - modified Eagle's medium
MHC - major histocompatibility complex
MJD - Machado-Joseph disease proteases
MMP13 - matrix metallopeptidase 13 (collagenase 3)
MS - mass spectrometry
MSC - mesenchymal stem cells
MW - molecular weight
N-CAM1 – neural cell adhesion molecule 1
NEDD4 - neural precursor cell expressed, developmentally down-regulated
NEM - N-ethyl maleimide
NF-κB - nuclear factor kappa B
NLS - nuclear localisation signal
NTA - Nitrilotriacetic acid
OA - osteoarthritis
OST - dolichyl-diphosphooligosaccharide protein glycosyltransferase
OUT - OTUbain proteases
PCI - proteasome component
PCNA - proliferating cell nuclear antigen
PCR - polymerase chain reaction
Pdi - protein disulphide isomerase
PERK - PRKR-like endoplasmic reticulum kinase

PLSD-T - platyspondylic lethal skeletal dysplasia, Torrance type
PML - promyelocytic leukemia
PSACH - pseudo achondroplasia
PTHrP - parathyroid hormone-related protein
RAD51 - BRCA1/BRCA2-containing complex, subunit 5 (RecA homolog)
RFP - red fluorescent protein
RING - really interesting new gene
RNA - rybonucleic acid
RT-PCR - reverse transcription polymerase chain reaction
Runx2 - runt related transcription factor 2
S1P - serine protease site-1
S2P - metalloprotease site-2 protease
SCF - Skp, cullin, F-box
SDS-PAGE - sodium dodecyl sulphate polyacrylamide gel electrophoresis
Sox9 - SRY-box containing gene 9
Spp1 - osteopontin
STAT - signal-transducer and activator of transcription protein
SUMO - small ubiquitin-related modifier
TAP - tandem affinity purification
TG - transgenic
TGF-β - transforming growth factor beta
TNF-α - tumour necrosis factor alpha
Uba5 - ubiquitin-like modifier activating enzyme 5
UBLs - ubiquitin-like proteins
UCH-L1 - ubiquitin carboxy-terminal hydrolase L1
Ufc1 - Ufm1 conjugating enzyme 1
Ufl1 - Ufm1 specific ligase 1
Ufm1 - ubiquitin-fold modifier 1
Ufsp1 - Ufm1 specific peptidase 1
Ufsp2 - Ufm1 specific peptidase 2
Ugt1 - UDP-glucosyl transferase
UPR - unfolded protein response
UPRE - unfolded protein response element
USP - ubiquitin-specific protease
VHL - von Hippel-Lindau tumor suppressor, E3 ubiquitin protein ligase
XBP1 - X-Box binding protein-1

Chapter 1. Introduction

Osteoarthritis (OA) is the most common form of human joint disease that causes significant pain and disability. It is estimated that 60%-70% of people over 65 have some degree of OA and, with an increase in the average age of populations in the western world it is becoming a major socio-economic burden [1]. The disorder is characterised by progressive loss of articular cartilage and results in the destruction of diarthrodial synovial lined joints, the only treatment for which is joint replacement for end stage disease. Hence, a major focus of OA research is identification of new targets for disease modifying therapy [2].

Most forms of OA are caused by a combination of genetic and environmental factors [3, 4]. In rare instances, and usually as a secondary consequence of an underlying form of skeletal dysplasia, OA can have a Mendelian pattern of inheritance. These heritable forms of OA are tractable by classical genetic analysis which has led to the identification of genes important for joint structure and function [5]. The functional effects of mutations in such genes are providing insights into the aetiology and pathogenesis of OA and may provide potential targets for therapy [6].The research conducted as part of this thesis investigated the consequences of a mutation that causes one such Mendelian form of OA, called Beukes hip dysplasia (BHD) [7, 8]. The research builds on the prior identification of a unique mutation within the Ufm1-specific peptidase 2 gene, *UFSP2*, which causes BHD. Thus far, the disease has been identified in only one multigenerational family. Members of this family acquire OA in childhood due to the abnormal shape of the hip joint. Other than the affected hip joint the individuals with BHD are healthy. At the time of the identification of the BHD mutation little was known about *UFSP2* other than that it was a component of a novel ubiquitin-like system, with a role in processing Ufm1 protein and cleaving Ufm1 from unknown target proteins.

The aim of the research reported here was therefore to investigate the role of the Ufm1/Ufsp2 pathway and how disruption of this pathway may lead to an OA phenotype. To inform these studies, this introduction provides an overview of joint formation, OA and its associated disorders, the current status of understanding of the features of ubiquitin–like proteins such as Ufm1, as well as, the role of the Unfolded Protein Response (UPR) in skeletal development and pathology.

1.1 Chondrogenesis

Cartilage is an avascular, supporting, and articular skeletal tissue. The cellular component consists of chondroblasts and chondrocytes, which deposit the extra-cellular matrix (ECM), and chondroclasts, that degrade the ECM. The ECM is mainly composed of glycosaminoglycans and collagen type II [9].

During embryonic development the blueprint (or anlagen) of the skeleton arises through the condensation of mesenchymal cells. In the formation of the axial and appendicular skeleton these mesenchymal cells differentiate into chondrocytes that synthesise cartilage. This cartilage has two distinct fates during development: transient cartilage which is removed and replaced by bone by the endochondral ossification process, and permanent cartilage, such as articular cartilage which remains at the distal ends of developing bones and forms the surface of synovial joints [10]. The differentiation of mesenchymal cells into chondrocytes is tightly regulated and depends on many factors and processes (see Figure 1) involving transcription factors, cell-cell and cell-matrix interactions, and mechanical stimulation [11]. For chondrogenesis to occur it is essential that the mesenchymal stem cells (MSCs) are densely packed. The condensation of MSCs is mediated by interaction with extracellular matrix (ECM) proteins, mainly fibronectin, by interaction between cells facilitated by membrane proteins (e.g. N-CAM and N-cadherin), and by paracrine stimulation with TGF-β (Transforming Growth Factor beta) and BMPs (Bone Morphogenic Proteins) [12]. JAG1-Notch signalling is also thought to play an important role in cell concentration and in priming chondrogenesis [13]. Following condensation, the cells start to express the transcription factor, Sox9, which is the earliest marker of chondrogenesis [14]. Its expression is regulated positively by TGF-β and BMPs, and negatively by β-catenin and Wnt [15]. In addition, TGF-β also induces expression of N-CAM and N-Cadherin, and assists in expression of the genes encoding collagens type II (*Col2α1*) and XI (*Col11α1* and *Col11α2*), fibronectin, aggrecan (*Agc1*) and COMP (Cartilage Oligomeric Matrix Protein). Sox9 in turn is a direct regulator of the genes encoding collagen type II, aggrecan, the α2 chain of type XI collagen and Cartilage-Derived Retinoic Acid-Sensitive Protein (CD-RAP), as well as its related transcription factors Sox5 and Sox6. These latter transcription factors are essential for expression of the gene encoding the α1 chain of collagen type IX (*Col9α1)* and act as cofactors in stimulating expression of *Agc1* (see Figure 1) [11, 16, 17].

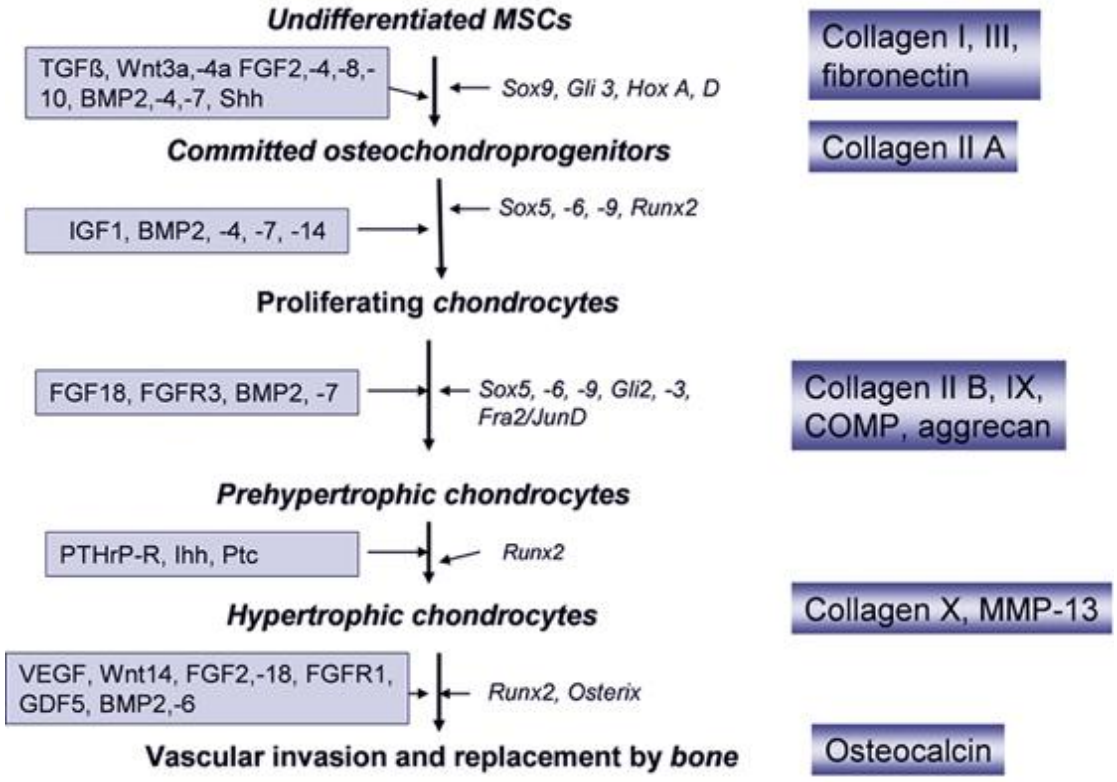


Figure 1. Chondrogenic differentiation cascade.

Chondrogenesis is regulated by an intricate network of transcription and growth factors along with the expression of ECM molecules. The figure shows different stages of MSC differentiation governed by growth factors and signalling molecules (boxes on the left) and stage specific ECM molecules (boxes on the right). Reproduced from [18].

1.2 Joint development

The first sign of joint formation is the flattening of chondrocytes across the continuous cartilaginous anlagen of future bones. This characteristic feature is called the interzone and is thought to develop from a distinct subpopulation of chondrocytes of the anlagen. The interzone forms three discrete layers: a dense intermediate cell layer and two outer cell layers. It was initially postulated that the outer layers participate in initial lengthening of bones and that the intermediate layer gives rise to articular chondrocytes and other synovial joint cells [19]. However, studies have shown that chondrocytes destined to give rise to articular chondrocytes are present on either side of the interzone in early mouse development demonstrating that not all articular chondrocytes are derived from the interzone [20]. Further, some of the joint structures such as regions of the meniscus of the developing knee joint may also arise partly from invading cells and not from the interzone [21]. At the molecular level, the cells of the interzone are

distinguishable from the anlagen chondrocytes by the absence of expression of *Matn1* (matrilin-1) and *Col2α1*, and expression of *Gdf*5 (growth/differentiation factor 5) , *Col1α1*(encoding the α1 chain of collagen type I) [20] and *Cux1* (cut homeobox 1 transcription factor) [22]. The interzone and nascent joint cells were also found to express a number of pro- and anti-chondrogenic genes like *Wnt9a, Wnt4, Gli3, CD44, Erg and Noggin* [19]. Other signalling molecules such as FGF-10, α5β1, Gdf-6 have also been implicated in joint morphogenesis. This suggests that joint development and maintenance are under the regulation of a complicated network of positive and negative stimulating factors but the precise roles of each and exact order of signalling are not fully understood.

Following joint specification and articular cartilage differentiation, cavitation must take place for the functional joint to develop. The process by which the articular surfaces of the joint separate is not fully understood but mechanical stimuli, hyaluronan synthesis, and CD44 mediated cell-cell interactions are thought to play a key role. Differential hyaluronan synthesis at the interzone and developing articular surfaces is induced by mechanical stimulation (i.e. limb movement of the embryo). The increase in hyaluronan concentration within the interzone stimulates the cells through the CD44 receptor which induces cell separation (see Figure 2) [23, 24]. Growth of articular cartilage progresses in the opposite direction to hyaluronan synthesis and is facilitated by a population of progenitor cells residing at the articular surface [25].

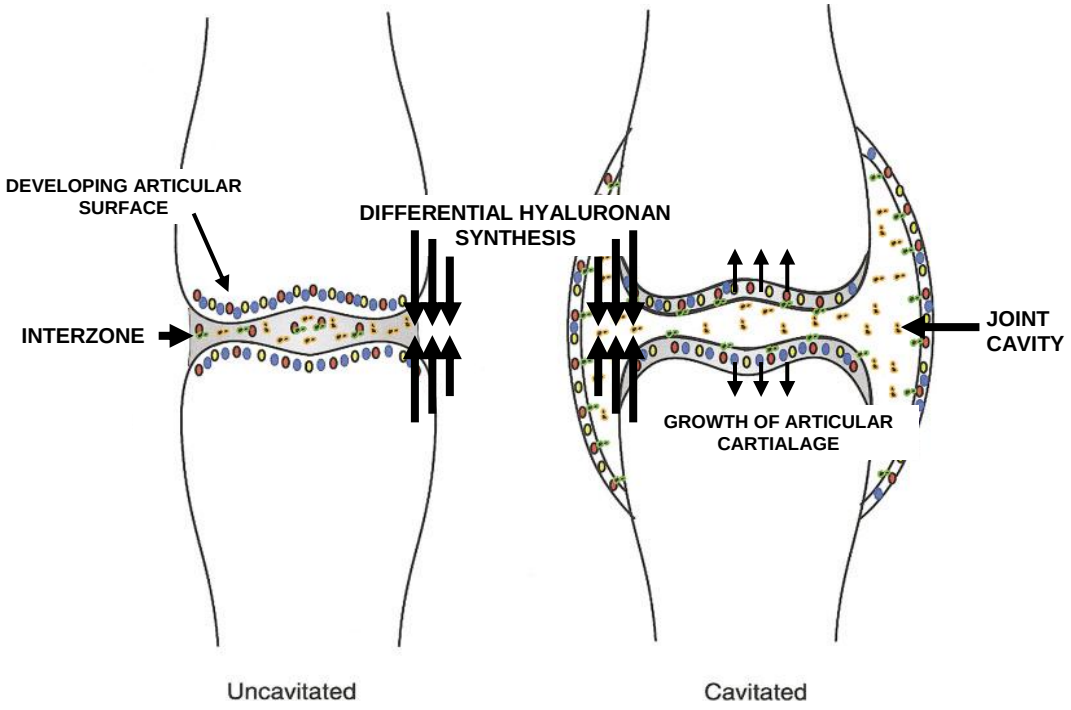


Figure 2. Joint cavitation process.

The arrows show the direction of differential hyaluronan synthesis in the interzone (thick arrows) and articular cartilage growth (thin arrows). Adapted from [23].

1.3 Endochondral ossification

The bones of the skeleton form by two distinct routes termed intramebranous ossification and endochondral ossification. Intramembraneus ossification is the process in which the flat bones of the skull and most of the mandible and clavicle are formed. This mechanism involves direct differentiation of mesenchymal cells into osteoblasts, which then proceed to form bone and the periosteum. The bones of the axial and appendicular mammalian skeleton however form via the endochondral ossification process which results in the conversion of the cartilaginous anlagen to bone with accompanying longitudinal and radial growth.

Endochondral ossification initiates when perichondrium surrounding the cartilaginous blueprint of long bones develops osteogenic potential and lays down a thin layer of bone (bone collar) (see Figure 3 - 1). Then, in the process of primary ossification centre formation, mesenchymal cells and blood vessels invade space created by chondrocytes degenerating in the middle of the diaphysis. The mesenchymal cells differentiate into osteoblasts and bone marrow cells. The osteoblasts then synthesise

irregular woven bone along the surface of the remaining calcified cartilage which is later remodelled to form the compact bone of the diaphysis (see Figure 3 - 2 and 3) [26].

In the process of endochondral ossification of long bones the cartilaginous epiphyses are separated by an ossified shaft (see Figure 3 - 4). The epiphyses contain three histologicaly different zones of cartilage: articular, epiphyseal and growth plate (see Figure 3 - 5). The growth plate is responsible for the longitudinal growth of the bones. The growth plate can be divided into zones based on the function or morphology of the chondrocytes (see Figure 4). Farthest away from the diaphysis is the resting (or reserve) zone of cartilage. Here are located stem-like cells that give rise to proliferative chondrocytes [27]. Next lays the proliferative zone where cells actively divide. The cells divide so that they align in neatly arranged parallel columns along the long axis of the bone. This spatial orientation of the columns directs the growth of the bones. As new columns appear on top of the old ones during growth, cells in the columns closer to the diaphysis stop dividing and start to enlarge and mature so that they form what is called the hypertrophic zone. In the hypertrophic zone mature chondrocytes farthest from the epiphysis degenerate and the matrix surrounding them calcifies. At this point the cells apoptose and the space left by them is invaded by blood vessels and osteoblasts [28, 29]. At one point in development the secondary ossification centre appears in the epiphyseal zone of cartilage, in the middle of the epiphysis. The centre grows in size until a thin layer of articular cartilage is left at the end of the bone that will form a part of the joint. The process of secondary ossification centre formation is not completely understood but it is thought to depend on several processes like: presence of cartilage canals (which contain mesenchymal cells that give rise to osteoblasts) and their later closure, mechanical loading, and interplay of PTHrP and IHH that drives chondrocyte hypertrophy [30]. The growth plate separates the primary and secondary ossification centres until later in development when the growth stops and the cartilage is replaced by bone [31, 32].

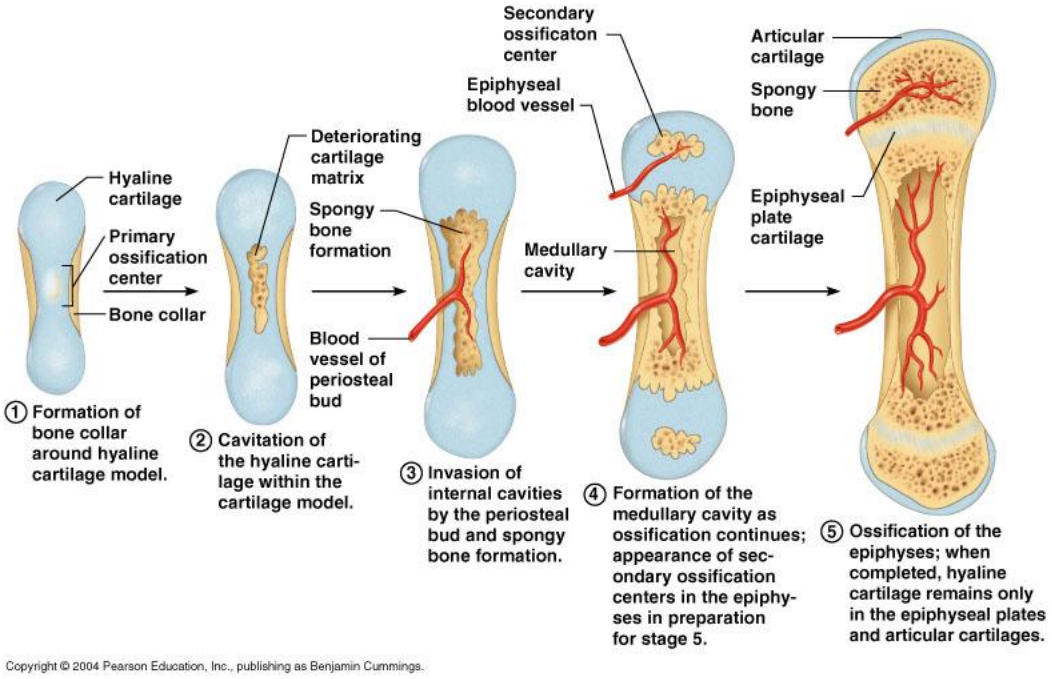


Figure 3. Endochondral ossification.

The process of endochondral ossification starts with formation of a bone collar around the cartilaginous blueprint of the future bone (1). Inside the part of the hyaline cartilage model, that will form the diaphysis, chondrocytes start to degenerate forming a cavity (2). Blood vessels invade the cavity and supply MSCs that differentiate into osteoblasts and bone marrow. Secondary ossification centres form in the epiphyses (4). The ossification process progresses until cartilage is left only in the growth plate and on the articular surface (5). Copyright by Pearson Education, Inc.

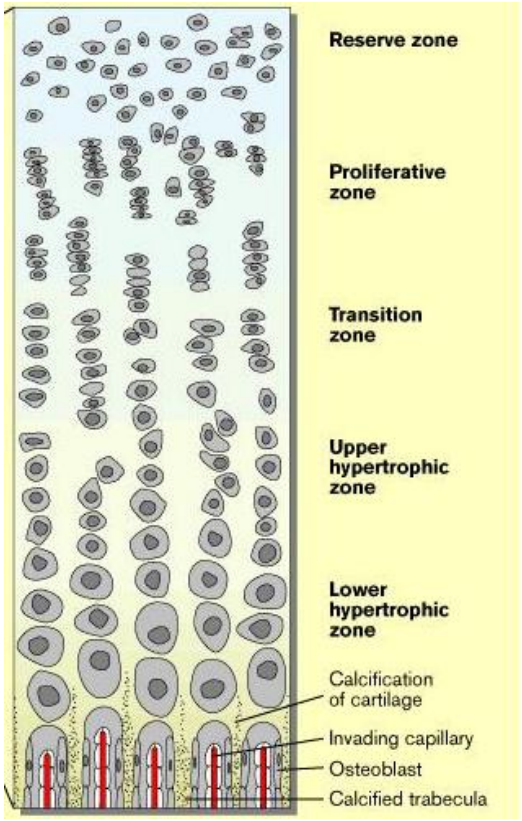


Figure 4. Organisation of the growth plate.

Distinctive zonation can be observed in the growth plate. Farthest away from the diaphysis undifferentiated stem-like cells can be found. These cells give rise to the chondrocytes in the proliferative zone. Proliferative chondrocytes form columns when they divide and expand the growth plate parallel to the long axis of the bone. The chondrocytes start to grow in size and form the hypertrophic zone which further expands the growth plate. Finally, the cartilage ECM mineralises and the cells start to apoptose which allows vascular invasion and replacement of the cartilage matrix by bone. Reproduced from [33]

1.4 Articular cartilage

Three types of permanent cartilage can be distinguished in the adult: elastic, hyaline and fibrous cartilage. The articular cartilage is a form of hyaline cartilage which covers the ends of bones within articulating joints and is not consumed by the endochondral ossification process during normal development and in healthy joints. In articular cartilage, chondrocytes comprise about 3-5% of the tissue volume and the collagen is predominantly type II. Collagens type VI, IX and XI are also present in minor quantities. Type II collagen forms complexes with type XI that arrange into large fibrils. Many other proteins like collagen type IX, COMP, matrilin-3, decorin and biglycan adhere on the surface of those fibrils and are involved in cross-linking them. The most abundant nonfibrous molecule in articular cartilage is the proteoglycan aggrecan which consists of the core protein with attached chondroitin sulphate and

keratan sulphate side chains, which give it a hairbrush-like shape. Multiple aggrecan molecules are attached to long hyaluronan chains via the link protein (Figure 5) [34-36]. In the adult articular cartilage of synovial joints, discrete zones can be observed (Figure 6). The zones are adapted to respond to the stresses experienced by the cartilage. At the surface of the joints, in the superficial zone, flattened chondrocytes are suspended in an ECM composed of a dense net of thin collagen fibrils aligned parallel to the surface, with a high concentration of decorin and biglycan, and a low concentration of aggrecan. The superficial zone chondrocytes, as well as cells of the synovial lining, secrete the superficial zone protein known as lubricin which helps in frictionless articulation of interacting joint surfaces. The high tensile properties of the superficial zone make the surface of the joint extremely resistant to the forces applied during articulation. Beneath the superficial zone is the midzone. It is characterised by the presence of round chondrocytes, high aggrecan content and thicker, randomly arranged collagen fibrils. The cells are less abundant than in the surface zone. Beneath is the deep zone with even fewer cells and the highest concentration of aggrecan. The content of collagen is lowest here but the fibrils have the largest diameter. Between the deep zone and bone is a region of calcified cartilage which serves to buffer and integrate the different properties of those two tissues. The chondrocytes of the calcified zone have a hypertrophic phenotype and synthesise collagen type X [35].

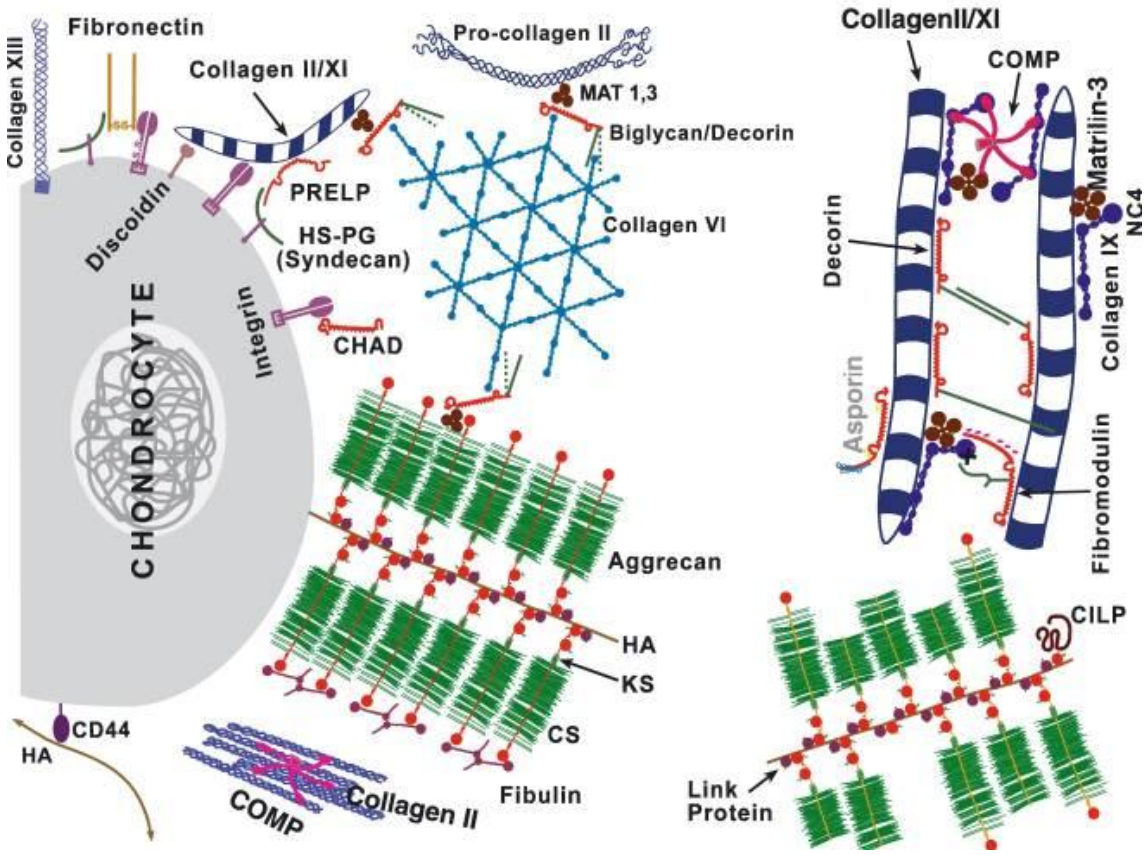


Figure 5. Organisation of the cartilage ECM.

Chondrocytes are the cells that produce the ECM of cartilage. Collagen type II and XI are arranged together in large fibrils that form complexes with other smaller proteins like collagen type IX, COMP, matrilin-3, decorin and biglycan. Aggrecan is one the most abundant large molecules in cartilage. Multiple aggrecan molecules are attached to long hyaluronan chains (HA) via the link protein. Reproduced from [37]

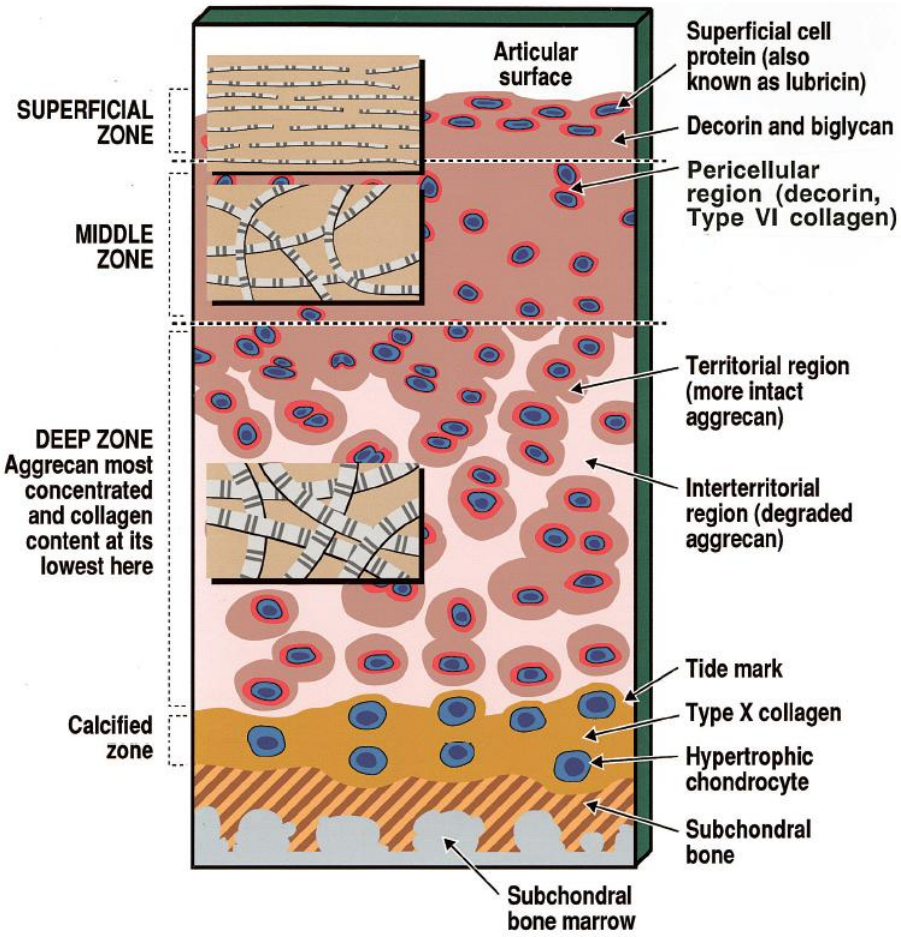


Figure 6. Graphic representation of zones distinguished within the articular cartilage of adult human joints.

The tensile properties of the superficial zone make the surface of the joint extremely resistant to the forces applied during articulation. The further away from the surface of the joint the fewer cells are present and collagen type II fibrils become thicker but more sparsely organised. Closest to the bone the cartilage is calcified which provides a buffer zone between the articular cartilage and the bone. Reproduced from [35].

1.5. Bone formation

1.5.1. Bone structure

Bone is a complex tissue that serves as support in locomotion, protection of internal organs and plays an important role in calcium homeostasis of the whole organism. Two different morphological types of bone, each having different mechanical properties, can be distinguished. Cortical (compact) bone that forms the diaphysis of long bones is thick and dense which provides maximum resistance to bending and torsion (see Figure 7 c). Cancellous (spongy, trabecular) bone located in the metaphyses and epiphyses forms an intricate mesh that is aligned in a way that provides the best support against the direction of mechanical forces applied to the particular bone. Cancellous bone forms a support for the articular cartilage and helps to absorb impact loads applied to synovial joints (see Figure 7 b) [26, 38].

Both types of bone have the same cellular and matrix composition. Collagen type I is the major organic component with types V, VI, VIII and XII present in small quantities. Additionally, non-collagenous glycoproteins and bone specific proteoglycans like osteocalcin, osteonectin, bone sialoprotein and bone phosphoproteins are present in the matrix. Collagen type I forms thick fibrils arranged in an overlapping manner with numerous crosslinks between them. Apatite is the major mineral component of the matrix. Apatite crystals, composed of calcium and phosphate, are responsible for the rigidity of the bone as opposed to collagen which provides flexibility. Beside the structural role, the mineral component of the matrix serves as a store of calcium, phosphorus, sodium and magnesium. Combined together the organic and mineral components of bone provide the balance between resistance and flexibility required for the functioning of the skeleton [26, 39].

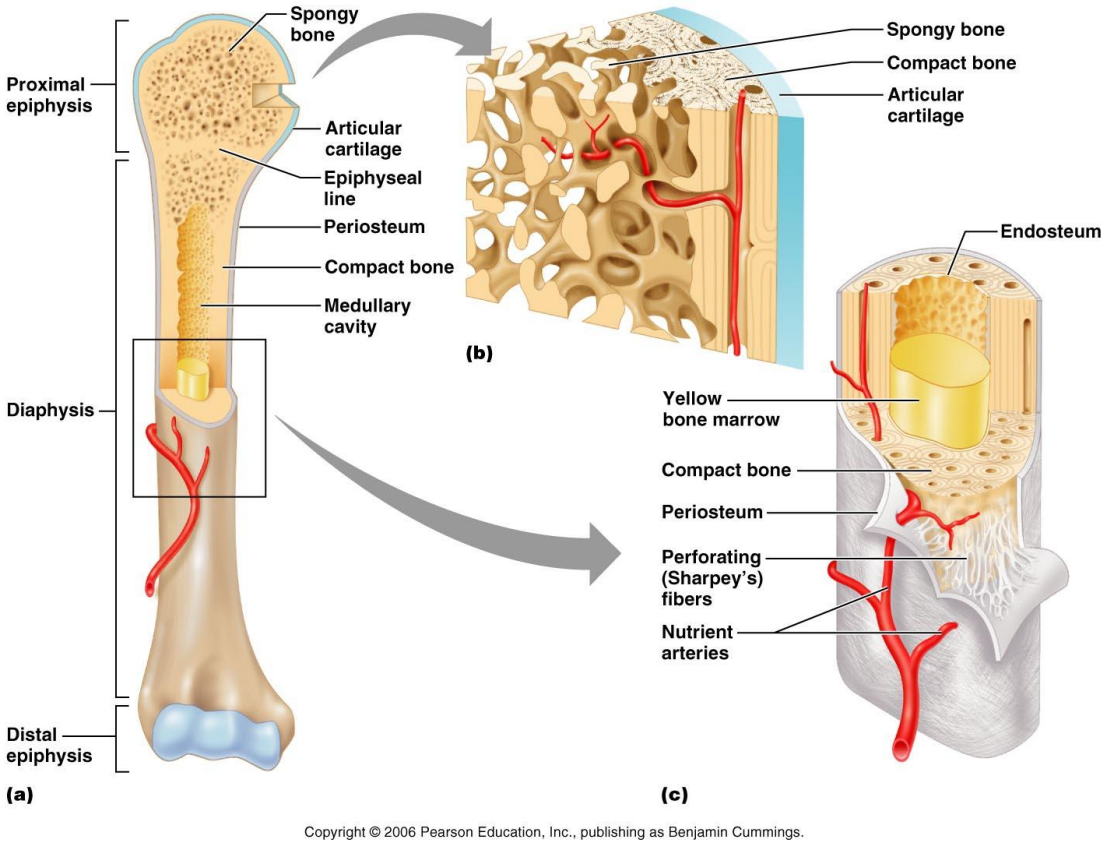


Figure 7. Bone anatomy.

Two major regions of the long bones can be distinguished: the diaphysis that forms the long shaft of the bone and the epiphyses that form the rounded ends of the bone. The diaphysis is composed of compact bone with bone marrow containing medular cavity. The diaphyses are composed mostly of cancellous bone with articular cartilage on their surface. Growth plate can be found between the diaphysis and the epiphyses and is replaced by bone later in life which is visible as the epiphyseal line. (a) Long bone with diaphysis, epiphyses and articular cartilage; (b) Cancellous bone of the epiphyses; (c) Compact bone of the diaphysis. Copyright by Pearson Education, Inc.

1.5.2 Osteoblast differentiation

Osteoblasts differentiate from mesenchymal osteoprogenitor cells located in the marrow, endosteum and periosteum. Osteoblasts and chondrocytes share a common origin. Both types of cells differentiate from mesenchymal progenitor cells that are activated by expression of the Sox9 transcription factor [40]. Committed osteochondro progenitor cells additionally express the Runx2 transcription factor which is repressed by Sox9 when the cells commit to the chondroblast lineage [41]. *Sox9* is expressed in all osteoblast progenitors. However, during differentiation into osteoblasts *Sox9* expression is downregulated and so *Runx2* is released from its repression and activates a series of osteoblast specific genes like *Osterix*, *Col1a1*, *Ibsp* (integrin binding sialoprotein**)**, *Spp1*

(osteopontin) and *Mmp13* (matrix metalloproteinase) [42-45]. No *Sox9* expression can be found in mature osteoblasts. Expression of *Osterix* marks the full commitment of progenitor cells to the osteoblast lineage. During maturation of osteoblasts *Runx2* expression decreases but the transcription factor still cooperates with Atf4 in induction of *Osteocalcin* expression which results in calcification of the matrix (Figure 8) [46, 47].

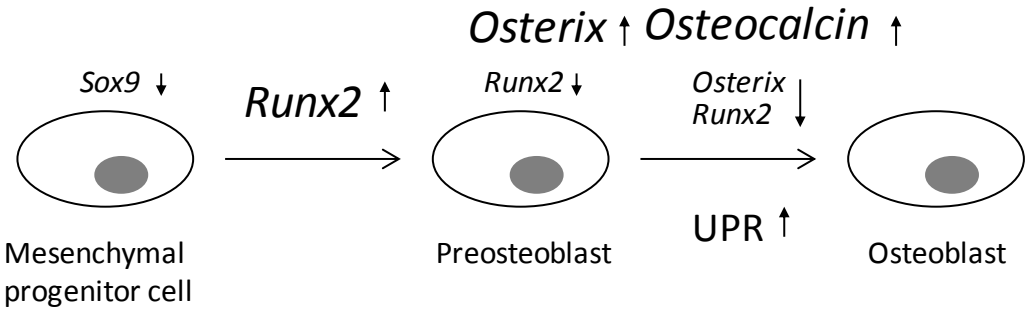


Figure 8. Simplified illustration of the osteoblast differentiation pathway.

After downregulation of *Sox9* in osteochondro progenitor cells *Runx2* activates a series of osteoblast specific genes. *Runx2* expression decreases later during the differentiation process while *Osterix* and *Osteocalcin* increases consecutively. High *Osteocalcin* expression results in calcification of the matrix. Figure based on data from [48].

1.5.3 Bone formation and remodelling

In addition to osteoblasts described in 1.5.2, three further types of cells form the cellular component of bone and are involved in bone formation and remodelling. These are bone lining cells, osteocytes and osteoclasts. Metabolically active osteoblasts have an oval polyhedral shape and their cytoplasm contains expanded rough endoplasmic reticulum (ER) and Golgi apparatus along with abundant mitochondria. By secreting the organic components of the bone matrix osteoblasts build the osteoid (unmineralised bone) [49]. During the course of their lives osteoblasts may follow one of three pathways. They may remain on the surface of the osteoid and continue to synthesise the matrix or they may start surrounding themselves with the matrix as the bone grows and become osteocytes. The osteocytes reside in pits within the mineralising matrix called lacunae. However, they maintain contact with osteoblasts and other osteocytes by cytoplasmic processes extending from their lens-shaped bodies forming canaliculi in the

bone. Osteoblasts may also decrease their activity and become bone lining cells. These cells are more flat in appearance and their cytoplasm contains fewer and smaller organelles [26, 49, 50]. In response to parathyroid hormone the lining cells degrade the osteoid surrounding the bone allowing osteoclasts access to the mineralised matrix and initiation of bone resorption. Osteoclasts have a haemopoietic origin. When stimulated osteoclast precursor cells, present in bone marrow or circulating in blood, proliferate and subsequently fuse to form large multinucleated cells. Osteoclasts on average contain 3 to 20 nuclei. Numerous mitochondria are found in the cytoplasm providing energy for bone resorption. The plasma membrane on the side facing the bone forms many folds that are described as ruffled border. Osteoclasts attach themselves to the bone like a suction cup, then acidify thus created space and secrete matrix degrading enzymes [51].

1.6 Skeletal disorders

As described previously, chondrogenesis and bone and joint development are regulated by a complex network of regulators of cellular differentiation and ECM synthesis and deposition. This network ensures the appropriate growth, morphogenesis structure and function of all elements of the skeleton. Abnormalities arising through genetic defects at any stage of skeletal formation can lead to a form of osteochondrodysplasia. As the condition BHD has features of chondrodysplasias, OA and developmental dysplasia of the hip, these conditions are described here.

1.6.1 Chondrodysplasias

Chondrodysplasias are a subset of the osteochondrodysplasias that result from abnormalities of chondrogenesis and the differentiation of chondrocytes during endochondral ossification. The chondrodysplasias are a heterogeneous group of over 150 genetic disorders caused by mutations in genes encoding, transcription factors, membrane receptors and ECM components. These disorders are characterised by abnormalities in the shape of skeletal elements and the rate and extent of bone growth.

The number and diversity of conditions classed as chondrodysplasias reflect the multiple genes that, if mutated, can lead to abnormalities of cartilage structure and function, and impaired endochondral bone formation. The classes of genes and genotype/phenotype correlations have been detailed in several comprehensive reviews [52-55] and examples of these are provided in Table 1. Mutations in genes encoding

transcription factors can cause a wide variety of phenotypes. The phenotype depends on the function of the affected gene and its role in, for example, chondrocyte differentiation or regulation of ECM protein expression. As an example, mutations in *SOX9* (see 1.1 and 1.5.2) have been found to cause camptomelic dysplasia (CMD1) [56]. Studies in mice have shown that Sox9 activity is required for mesenchymal cell condensation during establishment of the chondrogenic anlagen of the skeleton. In addition, the expression of several ECM genes such as *Col2a1*, *Col11a1* and *aggrecan* is under control of Sox9 [57]. Therefore, the phenotype of CMD1 may be an outcome of impaired chondrogenesis as well as reduced production of cartilage matrix structural proteins. Another transcription factor important for skeletal development was found to be mutated in cleidocranial dysplasia (CCD). CCD was found to be caused by haploinsufficiency of *RUNX2* [58]. *Runx2* knock-out mice show absence of bone and impaired chondrocyte maturation while heterozygotes display features similar to CCD (see 1.5.2) [59]. Membrane receptor mutations may result in either inactivation or irreversible activation of the receptor and thus lead to absence or excess signalling. As such, different mutations in the same receptor can lead to distinct phenotypes depending on how the mutation affects the function of the receptor (as for PTHrP, Table 1). Mutations of ECM components can affect both the structural integrity of the matrix but also chondrocyte differentiation due to impaired interactions and signalling between the ECM and the chondrocyte. Further, there is increasing evidence that mutations in genes encoding ECM components may lead to incorrect protein folding, assembly and trafficking that can trigger ER stress responses in the chondrocyte. Specifically, it is known that mutations of *COL9A2*, *COL9A3*, *COMP*, and *MATR3* cause Multiple Epiphyseal Dysplasia (MED) and Pseudochondroplasia (PSACH) [60, 61] and it is known that the mutated proteins, if secreted, have an effect on the mechanical properties of the ECM in addition to triggering ER stress through impaired protein folding. The triggering of ER stress has multiple consequences including altered chondrocyte differentiation, impaired deposition of the ECM and, if ER stress is prolonged, chondrocyte apoptosis. The triggering of ER stress has been proposed as a common pathogenic consequence of the range of mutations that cause MED and PSACH (see Table 1) [61]. Interestingly, BHD has been previously characterised as a form of MED based on its phenotypic characteristics which hints that BHD may have a similar pathogenic basis to MED [62].

Disease	Gene	Function	Phenotype
Campomelic Dysplasia	*SOX9*	Transcription factor	Bowing and angulation of long bones and short-limbs, short stature. Additionally, hypoplasia of all cartilage elements can be observed. [59].
Cleidocrenial Dysplasia	*RUNX2*	Transcription factor	Delayed closure of the fontanelles, hypoplastic clavicles and dental anomalies. In general, bones forming by intramembranous ossification or combination of both intramembranous and endochondral ossification are affected the most. Growth plate anomalies, predominantly in the hip, are also evident clinically. Mouse homozygous mutants of *Runx2* die after birth and are completely devoid of bone [55, 58].
Achondroplasia	*FGFR3*	Membrane receptor	Long, narrow trunk with short limbs, enlarged head with prominent forehead but hypoplastic midface. Small foramen magnum and decrease in size of the spinal canal relative to the size of the spinal cord may cause lumbar spinal stenosis [55, 63].
Metaphyseal Dysplasias	*PTHrP*	Membrane receptor	Activating mutations of parathyroid hormone related protein receptor (PTHrP) cause the Jansen form, and inactivating mutations cause the Blomstarnd form of metaphyseal dysplasia [55]. The Jansen form results in constitutive cAMP accumulation and is characterised by short limb dwarfism and agonist-independent hypercalcemia [64]. The Bloomstrand form is fatal and is characterised by osteosclerosis and advanced maturation of the skeleton [65].
Multiple Epiphyseal Dysplasia (MED)	*COMP, DTDST, MATN3, COL9A1, COL9A2, COL9A3*	ECM proteins	Delayed ossification of epiphyses and changes in their shape. Growth retardation, club feet, muscular hypotonia and brachydactyly may also occur. Joint laxity leads to premature osteoarthritis of hips and knees [66, 67] [68] [69] [70] [71] [72] [73].

Table 1. Examples of forms of chondrodysplasia that result from mutations of genes that alter chondrogenesis, chondrocyte differentiation and cartilage structure.

1.6.2 Osteoarthritis

OA is a common musculoskeletal disorder affecting approximately 60% of men and 70% of women above the age of 65 [1]. Progression of the disease causes disability and joint pain that can also lead to other diseases as a result of reduced activity. The reduction of working time, early retirement and need for hip or knee joint replacement is a significant socio-economic burden.

The disease is characterised by progressive loss of articular cartilage centred on the load-bearing areas, formation of osteophytes (bone spurs at the joint margins), changes in the subchondral bone and inflammation of the synovial membrane [74]. Radiographic features include joint space narrowing, appearance of osteophytes and subchondral cysts, and subchondral sclerosis [75]. Risk factors like age, gender, trauma, overuse, genetic background and obesity predispose to development of the disease [76]. Therefore, two types of OA are distinguished: primary, without any obvious cause, and secondary, where there is an identifiable cause such as accident or injury or an underlying disorder of joint formation [77]. As forms of chondrodysplasia may alter the

integrity or shape of the joint, OA is a common secondary consequence of these conditions.

The articular cartilage of joints is the main focus of OA pathology. In OA the articular cartilage appears macroscopically as yellowish or brownish and is typically soft. In early OA the surface is no longer smooth and as the disease progresses there is visible fibrillation and matrix loss leading to eburnation of subchondral bone in late stages [77]. On the microscopic level, in early OA, the surface of cartilage roughens and cells in the superficial zone become hypertrophic. In advanced OA cracks and fissures can be observed in the cartilage surface. Cells disappear from the superficial and intermediate zone. The tidemark between uncalcified and calcified cartilage becomes unclear and blood vessels from subchondral bone invade the calcified cartilage region. Fibrocartilaginous scar-like tissue appears in focal areas of cartilage loss [74]. The most apparent changes in molecular structure and composition of the ECM are the loss of aggrecan and loosening of the collagen type II network. The latter leads to further loss of proteoglycans and hyper hydration of the matrix. The altered physical properties of articular cartilage make it more fragile to mechanical stress which in turn amplifies the damage on the molecular level. Once the initial balance is lost, a positive feedback cycle progresses OA. This is also reflected by the expression pattern of OA chondrocytes. A mixture of anabolic (IGF-1, BMPs, TGFβ) and catabolic (TNFα, IL-1β) factors are secreted by the cells [78]. Increased collagen (type I in the superficial, and type II in deeper zones) and metalloproteases (in the superficial zone) synthesis are a sign of disturbance of the homeostasis [79]. There is also evidence that the articular chondrocytes acquire a phenotype resembling that of transient chondrocytes that are consumed during the endochondral ossification process and expresses markers characteristic of hypertrophic chondrocytes such as collagen type X [80].

1.6.3 Hip dysplasia

Developmental Dysplasia of the Hip (DDH) includes a range of pathologies, from mild acetabular dysplasia with a stable hip to dislocated hip with a dysmorphic head of the femur and acetabulum [81]. Reported prevalence of the disease depends on the geographical region as well as methods applied in clinical screening but generally it is regarded to be in the range of one to fifteen cases per 1000 live births [82, 83]. Several foetal and maternal factors are thought to increase the risk of DDH. Among the most commonly associated factors are female gender, breach positioning during birth

and a positive family history of DDH [83]. As abnormalities in the shape of the hip joint often lead to secondary OA attempts have been made to uncover the genetic factors that increase the susceptibility to DDH and subsequent OA. In familial forms, several genomic regions have been linked to susceptibility to DDH such as 13q22 [84] and 17q21 [85] but studies have thus far failed to identify any particular disease causing gene mutations. Polymorphisms in a number of genes have however putatively been associated with DDH. Studies have shown association of polymorphisms with DDH and OA in *GDF5* (a member of TGF-β superfamily also known as cartilage-derived morphogenic protein-1) [86-88], *ASPN* (asporin, a small leucin-rich repeat protein family member highly expressed in cartilage) [89, 90] and *FRZB* (a Wnt antagonist) [91]. However, the exact mechanisms whereby these polymorphisms may contribute or cause the disease remain unknown. Discovery of the region linked to the BHD phenotype was made through a similar study [8] followed by identification of the mutation in the *UFSP2* gene. Discovery of the mechanism underlying BHD may provide an insight into the molecular basis of other forms of DDH.

1.7 ER stress

ER stress has been implicated as part of the normal processes of osteoblast differentiation (see 1.5.2) and has been implicated as a contributor to the pathogenesis of some forms of chondrodysplasia (see 1.6.1) as well as in the Ufm1/Ufsp2 pathway. Hence ER stress may contribute to the pathogenesis of BHD and therefore is reviewed in this section.

1.7.1 ER stress responses

Protein production is a hallmark of all living cells. In eukaryotic cells most of the proteins destined for membranes and secretion are synthesised on the ER membrane and co-translationally translocated to the ER lumen where folding and post-translational modification take place. The ER capacity has to reflect the cell's need for new proteins. This is particularly true for multicellular organisms where the needs of cells change rapidly during development as well as are different between specialised cells in a fully developed organism. Tight regulation of this dynamic process is crucial for cell survival, organ development as well as proper functioning of the whole organism. An imbalance in this process in the form of the accumulation of misfolded or slowly folding

proteins (termed ER stress) triggers a program of actions (the Unfolded Protein Response or UPR) aimed at expanding the ER capacity to fold proteins and at the same time reducing the load of proteins entering the ER by inhibiting translation of non-essential mRNAs. Eventually, cells recover the balance between the ER load and the ability to process the entering proteins. However, in some cases (such as when mutated proteins are produced or there are changes to the cellular environment) the stress may be too high or sustained for too long which leads to ER stress induced apoptosis [92].

1.7.2 Protein folding and quality control

Membrane proteins and ones destined for secretion are synthesised on ER membrane bound ribosomes and are directed there by the Signal Recognition Particle pathway [93]. The newly synthesised peptide chain is co-translationally translocated through the Sec61 translocon complex. The nascent chain, while it still emerges from the pore, is assisted in folding by a myriad of chaperones, foldases and cofactors. Some of the best characterised factors are BiP (Grp78) and Grp94 which bind the hydrophobic sequences and prevent premature folding of the emerging chain [94]. As the protein chain enters the ER it is co-translocationally modified by addition of N-linked oligosaccharides by the OST (oligosaccharyltransferase) enzyme which are then trimmed by ER α-glucosidase I and II [95]. The glycosylated protein then interacts with calnexin and calreticulin in conjunction with Erp57 and PDI that catalyse disulphide bond formation [96]. After release from the calnexin/calreticulin complex the single remaining glucose residue is trimmed by glucosidase II to prevent immediate reassociation with the chaperone complex. Successfully folded proteins can subsequently exit the ER and enter the secretory pathway. Proteins that do not fold correctly are recognised by the ER folding sensor UGT1 which re-glucosylates the glycoprotein allowing it to re-enter the calnexin/calreticulin cycle for another round of folding [97]. Proteins that fail to achieve their native structure eventually enter the ER associated degradation (ERAD) pathway which results in their retrotranslocation to the cytoplasm and disassembly by the proteasome (Figure 9) [98].

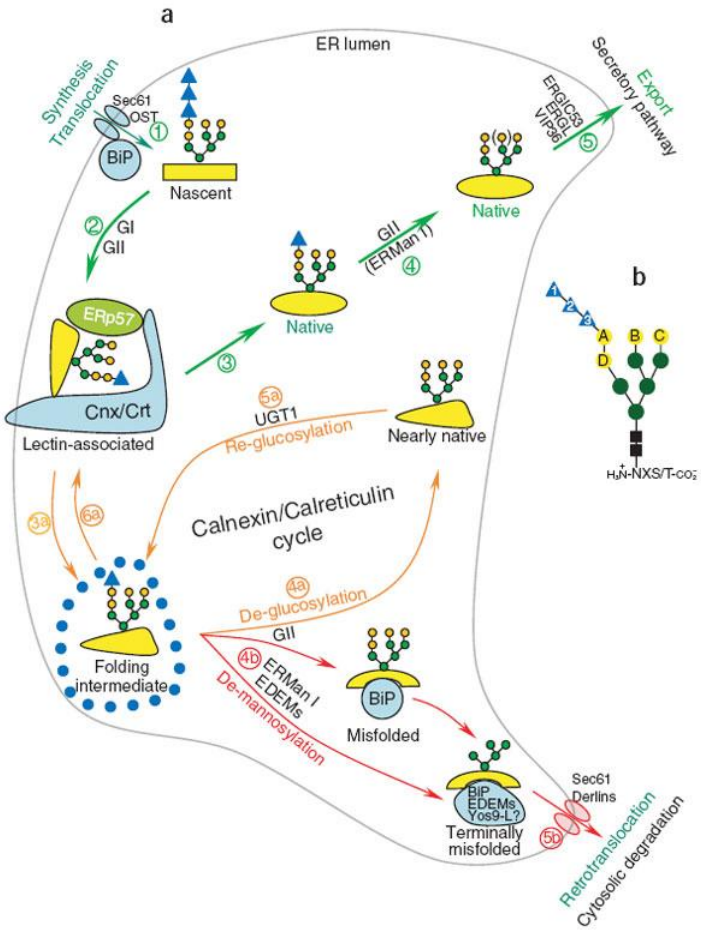


Figure 9. Protein folding and glycosylation in the ER.

a) Sec61 complex is the site of entry to the ER for newly synthesised proteins. BiP assists in folding and OST co-translocationally glycosylates the protein. Subsequently the protein is assisted in folding by the Calnexin/Calreticulin complex. Successfully folded proteins are exported from the ER and ones not folded properly are retained longer until folding is completed. Terminally misfolded proteins are degraded through the ERAD pathway. **b)** N-linked oligosaccharide structure. Blue triangles – glucose, yellow circles - (1-2)-linked mannose residues, green circles - other mannose residues, black squares - N-acetylglucosamine residues. Reproduced from [95]

Increasing amounts of unfolded proteins in the ER, caused by exceeding the folding capacity of the ER chaperones or folding-defective mutant polypeptides, trigger UPR which is aimed at restoring the balance and flow-through of newly synthesised proteins. Mammalian cells evolved three distinct pathways that sense and regulate the UPR (Figure 10).

The quickest to respond to the task is the double stranded RNA activated protein kinase (PKR)-like endoplasmic reticulum kinase (PERK) which mediates the attenuation of translation of new polypeptides. PERK senses unfolded proteins through

its ER lumen domain. In response to ER stress PERK dimerises and trans-autophosphorylates its cytoplasmic domain which results in activation of the protein kinase domain [99]. Activated PERK is then able to phosphorylate the α-subunit of eukaryotic translation initiation factor 2 (eIF2). eIF2 requires for its activity the GTP nucleotide and the phosphorylation facilitated by PERK blocks the access for eIF2B, a guanidine nucleotide exchange factor, that recycles eIF2 to its active GTP-bound form. The GDP-bound form of eIF2 is not able to initiate translation which in consequence lowers the load of the ER [100, 101]. This happens in case of most mRNAs but paradoxically this condition may also increase the efficiency of translation of some transcripts. One of these is activating transcription factor 4 (ATF4) mRNA which possesses in its 5' untranslated region a series of open reading frames that under non-stressed conditions cause early termination of translation. However, in stressed cells, high levels of eIF2α phosphorylation cause ribosomes to skip the inhibitory upstream open reading frames and allow efficient translation of ATF4 [102]. ATF4 is a transcription factor regulating genes involved in amino acid import, glutathione biosynthesis, and resistance to oxidative stress but also genes of the UPR like BiP and Grp94 chaperones [103].

Another regulatory transmembrane protein present in the ER is ATF6. In unstressed cells ATF6 remains in the ER as an inactive precursor with its ER luminal domain bound by BiP. Following depletion of BiP, due to increased amount of unfolded proteins requiring assistance, a Golgi localisation signal is uncovered and ATF6 is translocated to the Golgi apparatus [104] where serine protease site-1 (S1P) cleaves its luminal domain and the metalloprotease site-2 protease (S2P) cuts the N-terminal cytosolic domain [105]. The cytosolic domain is a transcription factor of the basic-leucine zipper (bZIP) family that when released travels to the nucleus and activates transcription of genes involved in ER quality control and other ER resident proteins and some of the ER associated degradation components [106]. Recently, a number of ATF6 related bZIP transcription factors have been characterised among which are OASIS and BBF2H7 [107, 108].

IRE1α is the third and evolutionarily oldest ER stress regulating pathway. IRE1α ER lumenal domain is similar to the one of PERK and it was shown experimentally that the domains are interchangeable and the whole protein retains its functionality [109]. In its inactive form, just like PERK, it is bound by BiP which prevents homodimerisation

of the protein kinase domain. Upon release from this repression IRE1α dimerises and trans-autophosphorylates which activates its cytosolic domain [110]. The cytosolic domain of IRE1α possesses endoribonuclease activity that catalyses unconventional splicing (on the cytosolic face of the ER membrane) of the mRNA encoding X-Box binding protein-1 (XBP1) transcription factor [111]. XBP1 (see 1.7.3 for further details) controls transcription of diverse regulatory networks among which are protein biosynthesis, folding and trafficking, UPR, ERAD, cell growth and differentiation [112]. The endoribonuclease activity of the cytosolic domain of IRE1α was also shown to play a role in degradation of specific mRNAs, based on their ER membrane localisation and the sequence they encode, thus reducing ER load [113]. In the early phase of ER stress IRE1α can also mediate JNK activation which induces autophagosome formation but following prolonged stress may lead to apoptosis [114].

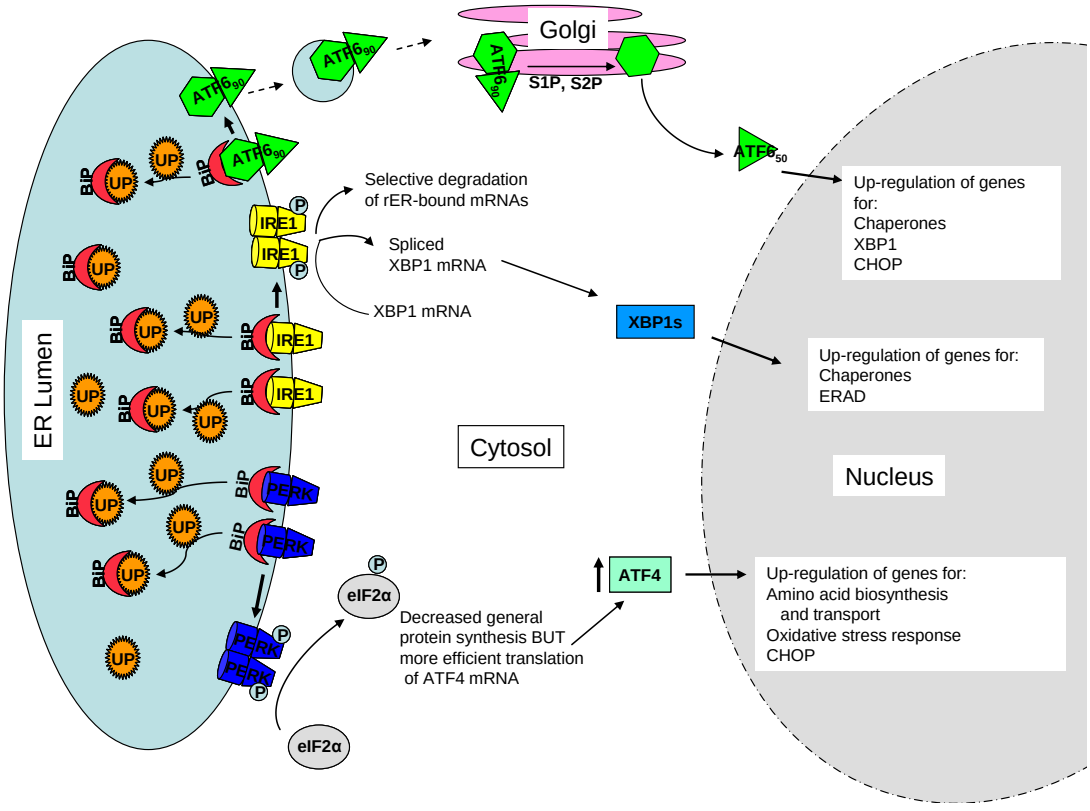


Figure 10. The unfolded protein response.

Three pathways are involved in the UPR: ATF6, IRE1 and PERK. Activation of all three pathways depends on dissociation of the BiP chaperone in response to increased concentration of unfolded proteins inside the ER. ATF6 following dissociation of BiP is transported to the Golgi apparatus where its cytosolic domain is cleaved by S2P. Then the domain migrates to the nucleus where it activates its downstream genes. IRE1 when released from BiP facilitates splicing of XBP1 mRNA which results in translation of an active transcription factor. PERK when not bound by BiP phosphorylates eIF2α which decreases general translation but increases efficiency of ATF4 translation, an active transcription factor and effector of the UPR. Reproduced from [61].

1.7.3 ER stress and skeletal development

A role of the ER stress response in skeletal development has been highlighted by gene knock-out studies in mice. For example *Perk* deficient mice have impaired osteoblast differentiation and maturation which results in severe osteopenia [115] and *Atf4* deficient mice have delayed bone formation during embryonic development and low bone mass throughout postnatal life [116]. Additionally recent studies have shown elevated ER stress response in chondrocytes of mice expressing mutated forms of COMP, collagen II and X [68, 72, 117]

The *OASIS* (old astrocyte specific induced substance) gene was first isolated from long-term cultured astrocytes [118]. Later, expression studies revealed high expression of OASIS in bone during mouse development particularly in primary osteoblasts but not primary osteoclasts [119]. OASIS is a bZIP transcription factor that structurally mimics Atf6 but has little amino acid sequence homology. It is an ER transmembrane protein with ER lumenal and cytoplasmic domains that upon activation of ER stress is cleaved by S1P and S2P proteases [120]. *Oasis* knockout mice exhibit growth retardation and osteopenia. The decreased bone density was shown to result from decreased bone formation and delay in osteoblast maturation. *Oasis* deficient osteoblasts contained enlarged ER with accumulated ECM proteins like collagen type I-α1 and osteocalcin. Interestingly, it was also shown that OASIS can bind UPRE-like (Unfolded Protein Response Element) sequence in the *Col1a1* gene promoter [121].

BBF2H7 is a transcription factor closely related to OASIS. Their sequence and mechanism of action are very similar and both are activated by ER stress with the difference that BBF2H7 is actively expressed in proliferating chondrocytes in cartilage of long bones during development [108]. *Bbf2h7* knockout mice exhibit dwarfism among other features and die shortly after birth due to respiratory problems. Studies have shown that the cartilage ECM of *Bbf2h*7 knockout mice is markedly reduced and chondrocytes show disorganised structure in the proliferative zone compared to the wild type mice. The ER of *Bbf2h7* deficient chondrocytes is abnormally enlarged with collagen type II and COMP aggregated in the lumen. *Sec23a,* a gene involved in formation of COPII vesicles and transport of proteins from ER to Golgi, was shown to be a target of BBF2H7. The accumulation of proteins in the ER, high level of ER stress and reduced ECM may be caused by defective protein secretion mechanism in B*bf2h7*

deficient mice [122]. BBF2H7 was also shown to protect the chondrocytes from ER stress induced apoptosis by activating the ATF5-ACL1 pathway [123].

XBP1, (see 1.7.2), is a transcription factor that is activated by IRE1α mediated mRNA splicing. The unspliced *XBP1* mRNA can be translated and results in a shorter less stable protein that is thought to act as a transcriptional repressor [124]. It was also found that the unspliced version may play a role in directing the mRNA to the ER membrane for splicing by IRE1α as the unspliced protein possesses a hydrophobic amino acid sequence with affinity for the ER membrane. Additionally, during translation, due to unique sequences in the *XBP1* mRNA, ribosomes may pause before terminating translation which gives time for IRE1α mediated splicing to commence in case of activation of ER stress [125]. Upon activation of ER stress IRE1α facilitates splicing of a 26 nt intron which results in a frame shift and a different N-terminal domain than the protein resulting from unspliced mRNA. Spliced XBP1 is an active transcription factor containing a bZIP domain that preferentially binds to the ACGT core sequence [112]. It was shown that the IRE1α-XBP1 pathway is required for progression of osteoblast differentiation through activation of the *Osterix* gene that possesses the ACGT sequence in its promoter region [48]. XBP1 was also found to enhance chondrocyte hypertrophy by associating with RUNX2 and regulating some of the genes essential for chondrogenesis. It was found that XBP1 activated expression of IHH and downregulated PTHrP [126].

The close interlinking of osteogenesis and ER stress was also demonstrated by a study which found that *ATF6* transcription can be directly activated by the RUNX2 transcription factor. ATF6 in turn activates *Osteocalcin* gene expression through an UPRE like TCAGCT sequence present in its promoter [127].

Further importance of ER stress in skeletal development is supported by *Perk* knockout mice which develop not only multiple skeletal dysplasias and abnormal expression of *Col1a1* but also diabetes mellitus due to loss of pancreatic alpha and beta cells. The compact bone of long bones of *Perk* deficient mice shows large perforations and discontinuities and the trabeculae of the spongy bone shows overgrowth to compensate for the lack of regular bone collar. Overall, the mice exhibited similar defects to people with Wolcott-Rallison syndrome [115]. Deletion of *Perk* was found to affect differentiation and proliferation of osteoblasts as well as their function in

deposition and mineralization of ECM. Deficiency in PERK causes retention and accumulation of procollagen type I in the ER but this does not induce the ATF6 and IRE1α downstream targets like *Bip* and *Xbp1*. Osteoblast markers like *Osteocalcin* and *Ibsp* (integrin binding bone sialoprotein) are also downregulated similarily to mice deficient in *Atf4* [116, 128, 129].

ATF4 was found to be able to induce expression of osteoblast markers in non-osteoblastic cell lines. *Osteocalcin* promoter can be bound by ATF4 in cooperation with RUNX2, possibly in the same region as ATF6 [46]. ATF4 tissue specific activity seems to be regulated at post-transcriptional level by ubiquitination [130]. ATF4 signalling was also implicated in the Coffin-Lowry Syndrome (CLS) of which characteristic features include delay in bone maturation, delayed closure of fontanelles and overall short stature. ATF4 activity was shown to be regulated by phosphorylation by RSK2, a kinase mutated in CLS [116].

A large group of skeletal disorders caused by mutations in collagen genes can be distinguished. Most of phenotypes exhibit bone and cartilage malformations but also muscular and vascular disorders are present. A missense mutation in *COL2A1* causing platyspondylic lethal skeletal dysplasia, Torrance type (PLSD-T) was shown in a mouse model to result in ER stress in chondrocytes [72]. Similarly, mutations in *COL10A1* cause metaphyseal chondrodysplasia type Schmid (MCDS). Mouse model of the disease exhibited irregular growth plate, was found to retain the protein intracellularly and the chondrocytes had elevated UPR as evidenced by increased *Bip* expression [73].

Chondrocytes and osteoblasts are a very active type of cells and specialise in secretion of proteins that form the vast ECM of cartilage and bone. Therefore, it is of no surprise that disturbances in the protein folding mechanism and secretion pathway result in multiple skeletal disorders. What is interesting though is the fact that the process of cell differentiation and formation of the skeleton is regulated at so many levels by genes involved in UPR. Furthermore, many studies have shown that certain level of UPR is a prerequisite for differentiation of skeletal cells and their normal ECM deposition. A precise balance needs to be maintained between physiological levels of UPR that allow high ER throughput and such levels that induce cell apoptosis.

1.8 Ubiquitin-like pathways

There are different ways in which proteins can be modified in order to alter their properties. Cellular proteins can for example be phosphorylated, acetylated, methylated, glycosylated or modified by covalent attachment of other proteins. These modifications can change the activity of the target protein or influence its interaction with various cellular mechanisms. Ubiquitin is the prototype of a family of protein modifiers. It was discovered in the mid 1970s and its role in protein degradation was characterised by Ciechanover, Hershko and Rose in the early 1980s for which they were awarded the Nobel Prize in Chemistry in 2004. Since then knowledge of the roles of ubiquitin in cellular processes has grown exponentially. Other ubiquitin-like proteins (UBLs) have been discovered further diversifying the cellular regulatory system.

1.8.1 The ubiquitin system

Depending on the mode of attachment, ubiquitin can perform different roles in cellular processes. The best characterised role of ubiquitin is targeting of proteins to 20S and 26S proteasome complex by covalent attachment of K48-linked polyubiquitin chains. Proteins such as cell-cycle regulators, transcription factors, tumour suppressors, oncoproteins and membrane proteins are regulated by ubiquitin mediated proteasomal degradation [131]. K63-linked polyubiquitination takes part in DNA damage response, vesicle trafficking and kinase activation. Proteins can also be monoubiquitinated. This process plays a role in histone regulation, endocytosis and viral budding [132].

The ubiquitin protein modification system involves a multienzyme cascade consisting of an activating enzyme (E1), a conjugating enzyme (E2), a protein ligase (E3) [133] and a deubiquitinating enzyme (DUB) (Figure 11) [134]. Ubiquitin, as well as all UBLs, are produced in an inactive pro-form with a C-terminal extension, which needs to be cleaved by a DUB to uncover a terminal glycine residue essential in conjugation reactions [135]. In budding yeast there is only one E1 enzyme, 11 E2s, and more than 20 E3 ligases, but other UBL systems are not as varied as the ubiquitin neither in yeast nor in any other organism [136].

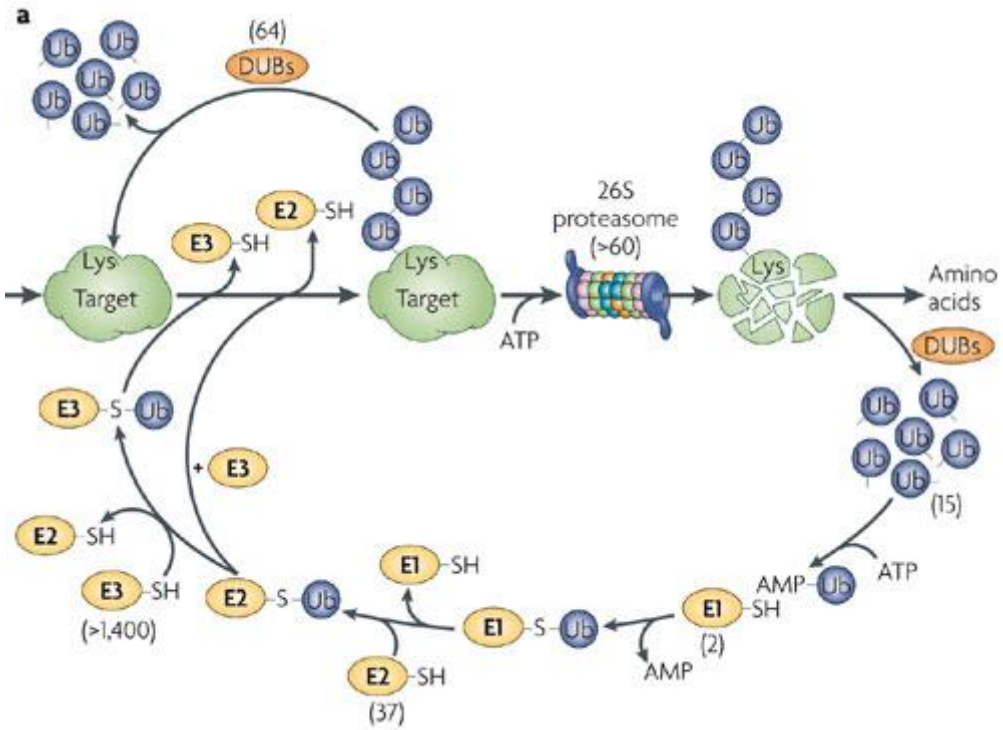


Figure 11. The ubiquitin protein modification and degradation pathway.

Ubiquitin is attached to its target proteins through an isopeptide bond facilitated by the E2 and E3 enzymes. Ubiquitin can be deconjugated by DUBs from its target and re-enter the E1-E2-E3 conjugation pathway after degradation of the target by the proteasome. Reproduced from [137].

The ubiquitin cycle starts with the binding of MgATP by E1. This step greatly increases the affinity of E1 for ubiquitin. Then, the ubiquitin is bound to E1 forming an adenylate intermediate and PPi is released. Subsequently, a thioester bond is formed between the C-terminal glycine of ubiquitin and a cysteine residue of E1, and AMP is released. A fully loaded E1 carries two ubiquitins: one as an adenylate, and second as a thioester. Next, the activated ubiquitin is transferred by transacylation to a thiol group of the E2 enzyme. From E2 ubiquitin can be transferred directly to its target with E3 acting as a bridging factor, or form a thioester with E3. Finally, the thioesters of E2 or E3 become donors for isopeptide bond formation between the lysine ε-amino group of the target and the C-terminal glycine of ubiquitin (Figure 11) [133, 138].

The number of E2 enzymes contributes to the specificity of the ubiquitin system. In general, each E2 can interact with several E3s, and each E3 can usually bind more than one E2. Many of the E2s have specific functions in cellular processes depending on which E3 they are bound to. For example, *S. cerevisiae* Ubc2/Rad6 (E2) upon binding of Ubr1 (E3) induces proteolysis of cohesin, but if it cooperates with Rad18 (E3) it

plays a role in postreplicational DNA repair. On the other hand, Cdc34 (E2) by interaction with different SCF E3s targets different proteins for degradation [133]. It is also possible for one target to be ubiquitinated by a range of E2/E3 pairs, *e.g.* the DNA polymerase processivity factor PCNA is first modified by the addition of ubiquitin by one E2/E3 pair, and elongation of the polyubiquitin chain is facilitated by a different pair [136].

All known E2s consist of a conserved ~150 amino acids catalytic core domain and different N or C-terminal extensions, which are thought to be responsible for specific interactions with E3s. According to crystal structures of E2 complexes the binding sites of E1 and E3, on the surface of E2, partially overlap and are mutually exclusive. E2s can covalently bind only one ubiquitin, but some as the UbcH5 family of E2 proteins that take part in peroxisome biogenesis, can bind another ubiquitin noncovalently in the process of polyubiquitination [139].

The final step in protein ubiquitination involves E3 enzymes. They are the most varied part of the system, and thus the most important factor in substrate specificity. Their mechanism of action, depending on the class of the enzyme, is to either bring E2 and the substrate in proximity and assist in transition of ubiquitin from E2 to the substrate, or to acquire the ubiquitin from E2 and perform the ubiquitination reaction of the substrate independently. Three main classes of E3 enzymes are distinguished: RING (Really Interesting New Gene) finger, HECT (Homologous to E6-AP C-Terminus) and U-box [131, 133, 136].

The RING domain E3s are the most abundant class. The RING finger is characterised by a short motif rich in cysteine and histidine residues, which can coordinate two zinc ions. The zinc ions are responsible for sustaining the domain conformation, which has been shown to bind directly to E2s. Their mechanism of action in general is to bring the E2 and the substrate together, and position them in a conformation favouring transition of ubiquitin from E2 to the substrate. RING E3s can bind directly to E2 and the substrate, or bind the substrate indirectly as part of a ubiquitin ligase complex [133, 136]. An example of such a multi subunit E3 ligase is the SCF (Skp, cullin, F-box) complex, in which the substrate is bound by an F-Box protein, which in turn is connected to cullin by Skp1 adaptor. The cullin serves as a scaffold for E3 and E2. An even more complicated example is the APC (Anaphase Promoting Complex) ubiquitin ligase, which is essential in degradation of proteins controlling

metaphase to anaphase transition and the exit from mitosis. The APC complex comprises of as much as 12 different subunits [140]

Another class of E3s similar to RING, but much less numerous, are the U-Box containing ubiquitin ligases. The U-Box is a 75 amino acid domain that resembles RING, but lacks the zinc binding motif. Instead, the structure of the domain is stabilised by salt bridges. The model of U-Box action is also analogous to RING in that it can become a part of a multisubunit ligase complex, and does not facilitate the ubiquitination reaction directly, but rather helps in bringing the substrate and E2 together [141].

The HECT family of ligases differs significantly from other E3s because they do not form multisubunit complexes, and they accept the ubiquitin from E2, and pass it independently to the substrate. The HECT domain consists of ~350 amino acids, always at the C-terminus of the protein, and forms two distinct surfaces: one E2 binding, and second with a catalytic cysteine residue, essential in thioester bond formation with ubiquitin. The most prominent subfamily of HECT ligases are the NEDD4s (neural precursor cell expressed developmentally down-regulated). The NEDD4 structure contains three types of domains: C2-WW-HECT. C2 domain is responsible for localisation of the ligase and can interact with phospholipids, inositol phosphates and proteins in a Ca^{2+} dependent or independent manner. Usually NEDD4 proteins contain several WW domains, which facilitate interaction with other proteins and determine substrate specificity [142, 143].

The last in the multi-enzyme cascade are the DUBs. These proteases perform various essential roles in the ubiquitin system. Since ubiquitin is translated in a pro-form or as a linear polyubiquitin, they act at the beginning of the cycle by cleaving C-terminal end of pro-form of ubiquitin, or the polyubiquitin chain, uncovering the active terminal glycine residue. DUBs recover ubiquitin from random binding with small cellular nucleophiles. They also act as antagonists of ubiquitination by clipping polyubiquitin chains, and removing ubiquitin from its target proteins. This activity has a regulatory effect. It prevents degradation of mistakenly marked proteins and also works as a recycling mechanism by cleaving polyubiquitin chains on the entry of target protein to the proteasome, and releasing ubiquitin monomers [144]. Deubiquitination is a highly regulated process and this has a consequence in the structure of DUBs. The activity of most of deubiquitinating enzymes is cryptic therefore they can act only after binding of specific substrate, or by attachment to a scaffold assisting in substrate recognition.

These interactions require involvement of specific domains responsible for binding of the scaffolds, substrates, or inhibitors. Eukaryotic DUBs are divided into five families, four of which are cysteine proteases: USP (Ubiquitin-Specific Proteases), UCH (Ubiquitin C-terminal Hydrolases), OUT (OTUbain proteases) MJD (Machado-Joseph Disease proteases), and one metalloprotease family containing JAMM (JAB1/MPN/Mov34 metalloenzyme) domain. Additionally some pathogens also encode deubiquitinating enzymes which can deregulate host cell ubiquitin homeostasis [132].

1.8.2 Ubiquitin-like proteins

Nearly two decades after the discovery of ubiquitin another similar protein was identified and named Small Ubiquitin-related Modifier (SUMO). SUMO binding of proteins is more labile than that of ubiquitin and only a small fraction of the target protein is SUMOylated and therefore it was discovered owing to its unusually stable complex with RanGAP1. SUMO proteins share only 15-20% sequence similarity with ubiquitin but their three-dimensional structure is nearly identical [135, 145]. Nevertheless, SUMO performs completely different roles than its prototype. Attachment of SUMO can mask a binding site on a protein, introduce a binding site with its own surface, change conformation of a protein by allosteric interaction, or even act as an antagonist of ubiquitin by competing for the same Lysine residue and thus preventing degradation of the target protein [146]. Although different mechanisms are possible the most common is the assistance in binding of other proteins. SUMO has been shown to promote nuclear localisation of a variety of proteins. Most notable are the PML nuclear bodies that include transcription factors, chromatin modifiers, and various genomic maintenance proteins. Additionally, attachment of SUMO to transcription factors in majority of cases has been found to inhibit their activity [147]. Another way in which SUMO can inhibit transcription is attachment to histones. This process was shown to antagonise ubiquitination. SUMOylation of histones can also promote their Rad51-dependant attachment to nuclear periphery in the event of persistent double strand DNA breaks. There is also evidence suggesting that SUMO may play a role in heterochromatin maintenance and organisation [148, 149]

NEDD8, and its yeast homologue Rub1, is yet another ubiquitin-like modifier. Its sequence is the most related to ubiquitin of all known UBLs, and it is 60% identical [150]. It is not as well characterised as SUMO and its best known function is modification of cullins in SCF complex. SUMO modification promotes their E3 activity

by inducing conformational changes [151]. Attachment of NEDD8 can also prevent inhibition of SCF by CAND1 (cullin-associated and neddylation-dissociate 1) by masking its binding site [152]. NEDD8 was also found to form polymeric chains *in vitro* but the role of this process in cells has not been identified yet [153]. Deregulation of the SCF ubiquitin ligases have been implicated in many types of tumours therefore the NEDDylation pathway is emerging as a potential anti-cancer drug target [154].

A different UBL possessing two ubiquitin-like domains, termed ISG15 (interferon stimulated gene), was found to be secreted to the media after treatment of human monocytes and lymphocytes with interferon β. In further studies increased production of interferon γ by lymphocytes, and proliferation of NK cells was observed in response to treatment with ISG15. ISG15 is also involved in a classic ubiquitin-like pathway, and is conjugated to intracellular proteins, which increases following stimulation with interferon. Later, some of the conjugates were identified as JAK and STAT proteins [155, 156]. Recent studies report important role of ISG15 in antiviral response. Experiments have shown that following virus infection majority of newly synthesized proteins are modified by attachment of ISG15. This attachment is not very specific but since during infection viruses hijack the protein synthesis machinery it is mostly the viral proteins that are targeted by ISG15. Many of the viral proteins were shown to lose their function as a result of the modification [157-159]

Several other ubiquitin like proteins were identified but their roles are not well understood. ATG8 and ATG12 UBLs were found to stimulate the development of autophagosomes [135, 152, 160]. Urm1 was found to conjugate with yeast proteins such as thioredoxin peroxidase protein Ahp1p, and to participate in thiolation of tRNA [161, 162]. FAT10 (Ubiquitin D) similarly to ISG15 possesses two UBL domains. Its gene lies in the major histocompatibility complex locus and is induced by IFN-γ and TNF-α and it is also upregulated in some forms of cancer [163]. FAT10 was also found to non-covalently bind to the spindle checkpoint protein MAD2 and promote 26S proteasome degradation of proteins covalently conjugated with single FAT10 molecule [164, 165].

1.8.3 Ubiquitin and ubiquitin-like related disorders

The cause of ubiquitin related diseases can be either the accumulation of proteins that in normal conditions should be targeted to proteasomal degradation, or rapid degradation of the targets caused by increased polyubiquitination.

The disorder cystic fibrosis is an example with increased proteasomal degradation of a target. Mutations causing this disorder are located in an ion channel CFTR (Cystic fibrosis transmembrane conductance regulator) and prevent the receptor protein from reaching the cell surface. Instead, CFTR is retained in the ER and completely degraded by the proteasome [166].

Increased protein ubiquitination can be connected with various forms of cancer. A well known tumour suppressor protein p53 is targeted for degradation by E6 protein of high risk strains of Human Papilomavirus that cause urine cervical carcinomas [167]. Another tumour suppressor p27 that inhibits CDK2 and prevents the transition of cells from G1 to S phase of the cell cycle was found to be rapidly degraded in certain colorectal, prostate and breast cancers [168]. Several cancers were found to be related to DUB mutations suggesting their important role in regulation of cellular processes. Brooke-Spiegler syndrome, familial cylindromatosis and familial trichoepithelioma were originally described as unrelated diseases until the discovery that they are connected by mutation in the *CYLD* gene, which encodes a deubiquitinating enzyme with a tumour suppressor activity. CYLD in complexes with NEMO protein, downregulates NF-κB signalling pathway by deubiquitinating TRAF2 protein, a NF-κB modulator. TRAF2 is rapidly degraded in the presence of the CYLD mutation [169].

Cancers can also result from accumulation of proteins that are not correctly directed for proteasomal degradation. Some colorectal cancers can be induced by accumulation of β-catenin, a part of signal transduction in colorectal epithelium differentiation pathway, which results from mutations in its ubiquitination recognition motif, or in the ubiquitination machinery. Other malignancies can also be caused by mutations in ubiquitinating enzymes. pVHL protein, subunit of a ubiquitin ligase complex responsible for degradation of transcription factors connected with vascularisation was found to be mutated in individuals with von Hippel-Lindau syndrome, characterised by formation of highly vascularised tumours and cysts [166]. The TNFAIP3/A20 (tumor necrosis factor, alpha-induced protein 3) is another tumour suppressor. It possesses DUB/E3 activity that was found to be downregulated in several types of lymphoma. A20 is an important factor in termination of NF-κB signalling and its mode of action is quite unique because it has both DUB and E3 domains. A20 recognises K63 linked polyubiquitinated proteins, cleaves the ubiquitin, attaches K48 linked chains, and as a consequence targets NF-κB signalling intermediates for proteasomal degradation [170]. Abnormalities in ubiquitin mediated turnover of

membrane proteins were associated with several diseases as well. The Liddle's syndrome, caused by a mutation in kidney epithelial sodium channel, is characterised by excessive Na^+ and water reabsorption which results in hypertension. The mutation renders the channel proteins unrecognizable to its ubiquitin ligase NEDD4 and causes their accumulation.

The ubiquitin proteasome pathway plays an important role in the immunological response. Foreign proteins are digested by proteasomes forming short peptides which are later presented as antigens by MHC class I molecules. Certain pathogens like HPV (human papillomavirus), EBV (Epstein-Barr virus) and CMV (cytomegalovirus) were found to develop strategies to avoid this type of response [171]. Defects in signalling related to impaired proteasomal degradation can also lead to autoimmunological diseases. An example of such disease is Incontinentia Pigmenti caused, contrary to cylindromatosis, by impaired NF-κB signalling [166, 169].

Disorders of the ubiquitin pathway have also been connected with neurodegenerative diseases. Parkinson's disease is thought to result from aggregation of α-synuclein, which is a substrate for UCH-L1, a dimeric protein that can function as an E3 and DUB. It is not clear which activity is essential in disease development because mutations in both domains were found to cause degenerative phenotype [172, 173]. Angelman disorder which is characterised by mental retardation, epilepsy and ataxia, was linked to mutation in E6-AP ubiquitin ligase. There is also evidence suggesting involvement of ubiquitin in Alzheimer disease and polyglutamine repeat neurodegenerative disorders, such as Huntington disease [174].

SUMO may also be responsible for some neurodegenerative disorders. Theoretically, abnormalities in SUMOylation may negatively influence any pathway where there is interplay between SUMO and ubiquitin. This could be the case in Huntington's disease where SUMOylation of mutant HTT protein makes it more soluble and more toxic. In *D. melanogaster* lowering SUMOylation, in contrast to ubiquitination, had a reduced neurodegenerative effect in flies expressing mutant HTT. SUMO was found in other neurodegenerative inclusions, but its role in development of pathologies remains unknown [175]. NEDD8 similarly to SUMO was found to aggregate in degenerative brain inclusions like Machado-Joseph disease, but the effect of its aggregation is also unclear [176, 177].

The network of ubiquitin and ubiquitin-like modification is vast with multiple steps at which it can malfunction. As it plays major regulatory roles in cellular

processes the effects of deregulation can be very severe. The knowledge on ubiquitin is fairly broad and its role in multiple diseases is of increasing importance. As for other UBLs, their characterisation is at an early stage but already some connections with disorders have been identified.

1.8.4 The Ufm1 system: a novel ubiquitin-like protein-conjugating system

The novel ubiquitin-like protein conjugating system, the Ufm1 system, was found by researchers when seeking to identify proteins that interacted with GATE16, a human homolog of Atg8p, by yeast two-hybrid screening [178]. One of the positive clones that they identified encoded an uncharacterised protein, the sequence of which was highly conserved among multicellular organisms. Part of the sequence exhibited homology to the ubiquitin E1 enzyme, Uba1, and other E1-like proteins of the UBL family. On the basis of sequence homology, a conserved metal binding motif and a conserved cysteine residue within the predicted active site the protein was recognised as an E1-like enzyme family member and so named Uba5.

Experiments were then conducted to confirm the interaction of Uba5 with GATE16, but GATE16 could not be pulled down when Uba5 was used as bait. The active site cysteine residue of Uba5 has been mutated to Serine allowing an O-ester instead of a thioester bond formation with its theoretical UBL protein. This type of chemical bond is stable under reducing conditions permitting easier isolation of protein complexes. In the study by Komatsu *et al.* (2004) [178] the Uba5 complex was isolated from HEK293T cells, and then analysed by nano-flow LC-MS-MS system. From this data three proteins were identified: GATE16, and two unknown proteins BM-002 and CGI-126. Since then interaction of Uba5 with GATE16 was not observed by any other study. BM-002 was found to be an 85 amino acid protein with a predicted molecular mass of 9.1 kDa, which was again conserved in multicellular organisms. The function of this protein was unknown, but the crystal structure of its *C. elegans* orthologue was available, and had been reported as possessing a ubiquitin-like fold. BM-002 shared only 16% sequence similarity with ubiquitin. However, its tertiary structure closely resembled that of ubiquitin (Figure 12). Hence it was named Ubiquitin-fold modifier 1 (Ufm1). Ufm1 possesses at its C-terminus, similar to other UBLs, a conserved glycine residue followed by a short sequence, which must be cleaved off by a deubiquitinating enzyme to generate an active form of Ufm1.

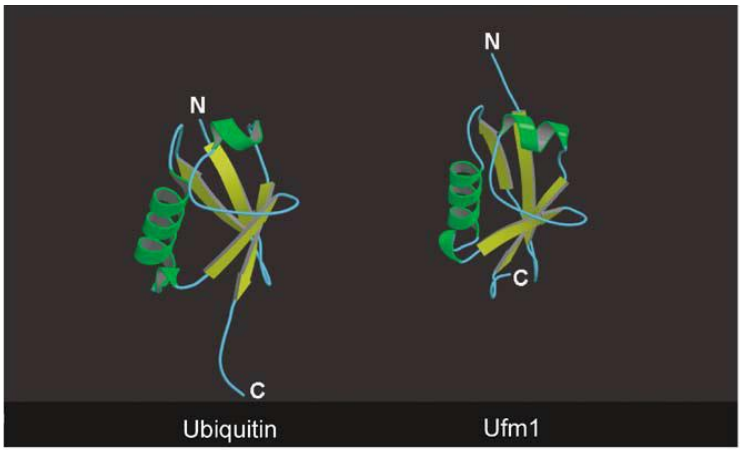


Figure 12. Structural ribbons of hsUbiquitin and hsUfm1.

Although there is little similarity between the amino acid sequences of ubiquitin and Ufm1 their tertiary structures are almost identical. Reproduced from [178].

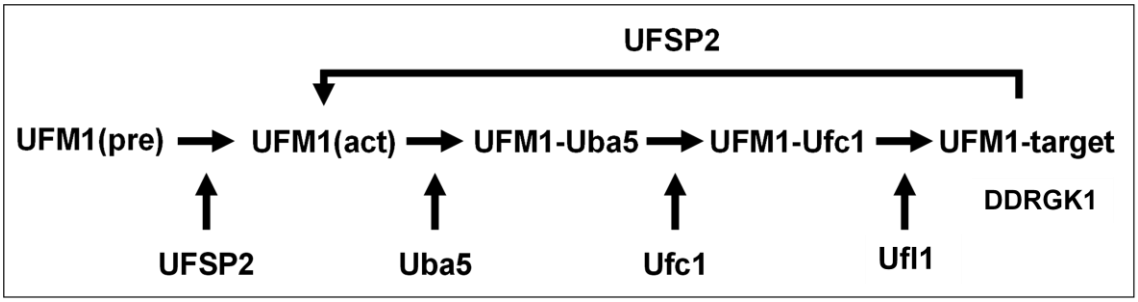


Figure 13. Schematic representation of the Ufm1 system.

Ufm1 must be activated by cleavage of the two C-terminal amino acids by Ufm1-specific protease before it can enter the E1-E2-E3 conjugation pathway. After conjugation to its target Ufm1 can be cleaved off by a Ufm1-specific protease and re-enter the pathway. Adapted from [179]

The second unknown protein identified from the LC-MS-MS spectra had a sequence of 167 amino acids, parts of which were highly conserved among various organisms. Based on the limited similarity of its active site region and interaction with Uba5 it was assumed to be E2-like enzyme and named Ufm1-conjugating enzyme 1 (Ufc1) [178].

The above experiments therefore identified three of the required components of a UBL pathway. The missing elements were an E3-like protein, a Ufm1 activating protease, and a Ufm1 deconjugating enzyme. To find proteases capable of activating Ufm1, mouse tissue extracts were fractionated, and tested for Ufm1 processing activity [180]. To do this, the positive fraction was tested against Ufm1 with a C-terminal

glycine residue modified by vinylmethylester (VME), which forms a covalent bond with the active site cysteine residue of a protease. This method led to the isolation of a complex where Ufm1 was covalently bound to a protease. The complex was further analysed by mass spectrometry, and the sequence of the protease was found to be that of a previously uncharacterised protein. BLAST searches also revealed homology with another larger protein. Both proteins possessed highly conserved cysteine and histidine residues, similarly to DUBs, that form a catalytic triad of cysteine proteases, and were therefore named Ufm1-Specific Proteases 1 & 2 (Ufsp1 & Ufsp2) [180].

A recent paper [181] has described identification of the E3 ligase of the Ufm1 system and a putative Ufm1 conjugation target (for schematic representation of the Ufm1 system see Figure 13). The conjugation target DDRGK domain containing 1 (Ddrgk1) (also known as C20orf116, UFBP1, Maxer and Dashurin) was found as a Ufc1 and Ufm1 interacting protein in pull-down assays. Pull-down assays with Ddrgk1 then identified the E3 ligase, Ufl1 (Ufm1 ligase 1, also known as KIAA0776, RCAD and NLBP). Ufl1 was found to contain a transmembrane domain that binds the protein to the ER and a possible nuclear localisation signal (NLS) [182]. The interactions of the proteins have since been confirmed by co-IP methods and *in vitro* conjugation assays [181]. The role of Ddrgk1 is not known but studies have shown that it possesses a hydrophobic N-terminal domain, responsible for anchoring it on the cytoplasmic side of the ER, followed by a NLS. Ddrgk1 also possesses a PCI (Proteasome, COP9, Initiation factor 3) domain and the Ufm1 attachment site was found to be within that domain. PCI domains were first identified in some subunits of the proteasome, COP9 and eIF3 and were found to be responsible for protein-protein interaction [183]. Studies have shown that deletion of the transmembrane domain of Ddrgk1 resulted in its nuclear localisation as opposed to the ER localisation of the full length protein. In both cases Ufm1 was found to co-localise with Ddrgk1 [184]. Overexpression of Ddrgk1 caused amplification of the ER and its re-organisation around the nucleus [181, 185]. Additionally, Ddrgk1 and Ufl1 were shown to interact with a putative tumour suppressor Lzap (C53, Cdk5rap3) and modulate its stability. Overexpression of Ufl1 led to increased stability of Lzap and Ddrgk1. On the other hand its knockdown caused rapid degradation of both Ddrgk1 and Ufl1 that was facilitated by the proteasomal pathway [185, 186]. Lzap and Ufl1 were found to bind NF-κB p65/RelA protein and inhibit TNF-α induced NF-κB activation. Interestingly, knockdown of Lzap or Ufl1 increased *Mmp9* expression. Mmp9 is a key regulator of the remodelling of skeletal tissues [182, 185, 186]. Ufm1

was also found to interact with Lzap but no covalent binding has been detected so far. Search of the BioGPS Database (http://biogps.org) shows ubiquitous expression of the components of the Ufm1 system at varying levels in different tissues with pancreatic islets, thyroid and pineal gland exhibiting expression above average level. The exception to this is *UFSP1,* the expression of which is uniform across all tissues (see Supplementary Figures 6 and 7).

The possibility of the Ufm1 system being involved in ER stress was first implicated in two studies which looked at genes upregulated during ER stress induced during ischemic heart disease and type 2 diabetes [187, 188]. Attempts to elucidate the role of the Ufm1 system in ER stress yielded contradicting results. Lemaire *et al.* (2011) [184] reported an increase in gene expression and protein levels of Ufm1 and Ddrgk1 in response to ER stressors but a decrease in conjugation of Ufm1 to Ddrgk1. Cycloheximide, a potent inhibitor of translation, on the other hand increased the level of conjugated Ufm1. The same study also reported increase in ER stress induced apoptosis in cell cultures where *Ufm1* and *Ddrgk1* expression was silenced [184]. Conversely, Zhang *et al.* (2012) [189] reported an increase in Ufm1 conjugation in response to ER stress. Silencing of the Ufm1 system components was found to elicit ER stress response in two independent studies contrary to that of the Lemaire *et al.* (2011) study which did not detect any increase in expression of ER stress markers *Bip* and *Chop*. The picture becomes even more complex in the light of the Hertel *et al.* (2013) [190] study that reported increased survival rates of nematodes under ER stress conditions if the *Uba5* was deleted and the Tatsumi *et al.* (2011) [191] study that found that *Uba5* is required for erythroid differentiation in mice and deletion of the gene is embryonically lethal. Nevertheless, the emerging consensus is that the majority of the Ufm1 system components localise to the ER (with the exception of *Leishmania donovani* parasite where they were observed to localise to the mitochondria [192, 193]) and are upregulated during UPR induced by disease (diabetes, ischemic heart disease) or chemical stressors.

1.9 Background to the current study

The disorder BHD was first identified by an orthopaedic surgeon in the Bloemfontein region of South Africa who was intrigued that many of the patients presenting for hip replacement surgery at an early age had the family name "Beukes" (*personal communication*). He and Professor Beighton from the University of Cape Town traced the family history of these patients and found that they were all part of a large, multigenerational family of Dutch origin [7]. Further investigations of this family identified that the BHD phenotype was inherited as an autosomal dominant trait. The affected family members were found to present with hip pain in early childhood and have radiographic evidence of OA by early adulthood which progressed to the need for hip replacement. Characteristic features of the BHD may be evident in children of less than 2 years of age and include delayed secondary ossification centre formation, broadening and shortening of the femoral neck, flattened irregular capital epiphysis and overgrowth of the greater trochanter. Evidence of OA in early adulthood include the formation of subchondral sclerosis, narrowing of the joint space and appearance of cysts and osteophytes (Figure 14) [7]. Other than for abnormalities of the hip joint affected individuals were healthy, of normal stature and had no other associated skeletal abnormalities. Initial linkage studies focused on cartilage extracellular matrix genes but no linkage of the disease was detected to any of the genes screened. Further investigations of this family employed genomewide screen with a panel of 290 markers with an average spacing of 11 cM which mapped the BHD locus to a region on Chromosome 4q35 [7, 8].

Work conducted in our laboratory by former PhD student Chris Watson [179] refined the linked region and screened genes within this region for mutations. Exons of genes within the 11cM locus were screened for presence of SNPs polymorphic in the Beukes family. Genotyping of these SNPs using restriction fragment length polymorphism (RFLP) methodology identified recombination events and allowed narrowing the interval of the linked allele. Genes within the refined region were then sequenced and analysed in search of mutations potentially causing BHD. One mutation was identified within exon 8 of the *UFSP2* gene (see 1.8.4). This mutation was found in all affected family members and linkage analysis between the mutation and the BHD phenotype generated a LOD score of 10.4. This mutation leads to the substitution of

Tyrosine 290 to histidine in the encoded protein. The mutation is in a conserved region in close proximity to the cysteine box which is a part of the catalytic triad of the protease (Figure 15). The crystal structure of mouse Ufsp2 predicted that the BHD mutation would alter the structure of the active site of Ufsp2 leading to its inactivation [194]. This was confirmed by *in vitro* studies which showed that introduction of the BHD mutation into Ufsp2 prevents its cleavage of Ufm1 (i.e. Ufsp2 is not able to cleave the Ser-Cys C-terminal residues from Ufm1 which would expose the active glycine required for Ufm1 to attach to its substrates via the E1 – E3 catalytic cascade, see 1.8.4) [179, 194]. These *in vitro* studies showed that the mutation had a functional effect, but as at that time little was known about the Ufsp2/Ufm1 system, this information did not provide an insight into how the mutation might lead to the BHD phenotype.

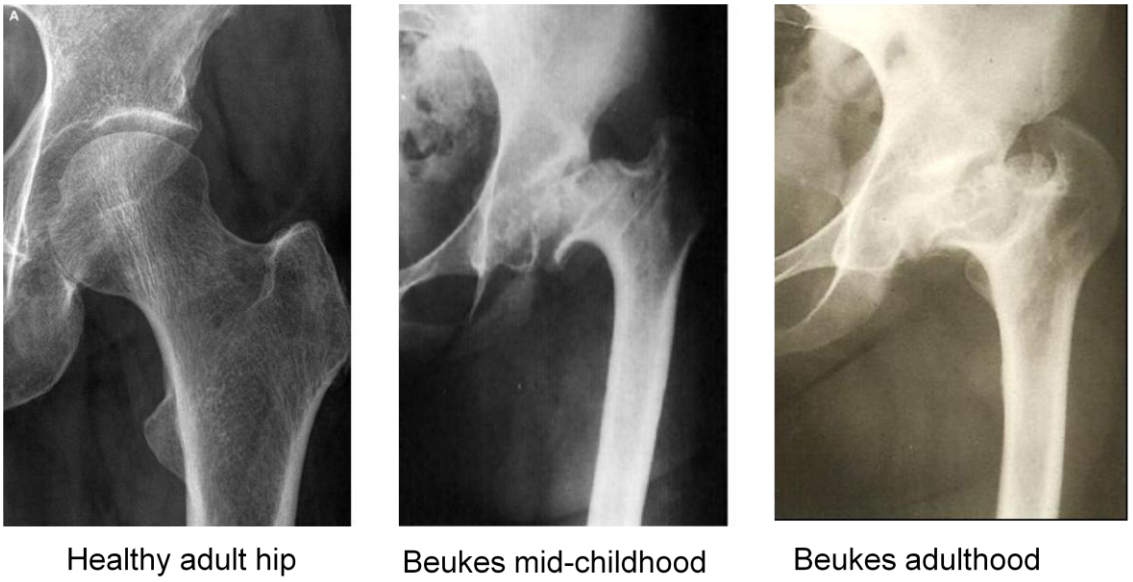


Figure 14. Radiographs of a healthy adult hip joint and hip joints of BHD patients.

Flattening of the capital femoral epiphysis, broadening and shortening of the femoral neck and overgrowth of the greater trochanter are evident in the radiographs of the BHD patients. Signs of OA are also visible in the adult BHD hip. Adapted from [7]

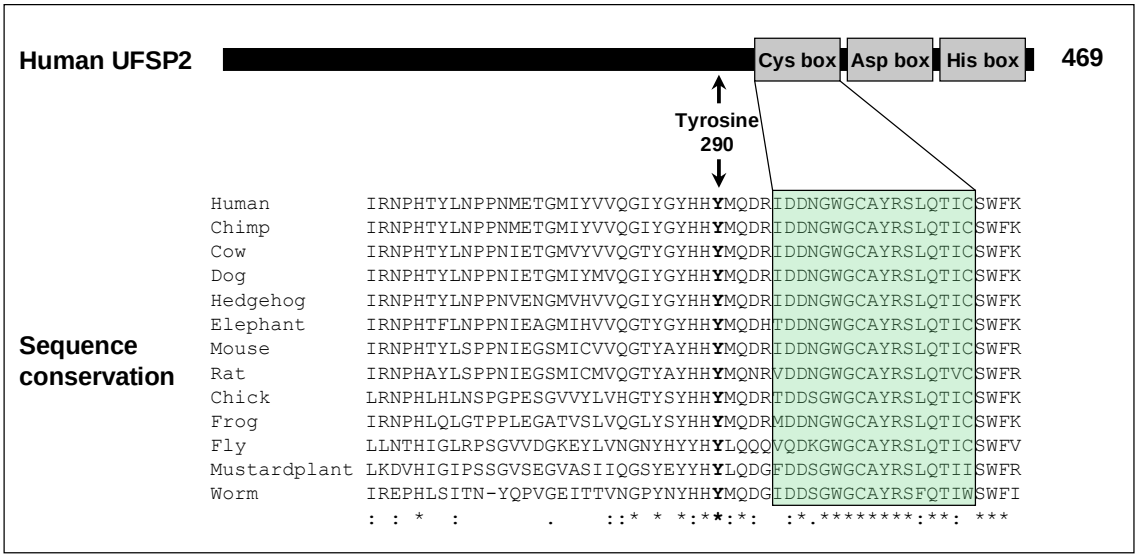


Figure 15. Schematic diagram of the Ufsp2 protein and alignment of its partial sequence.

Tyrosine 290 mutated in BHD patients is in close proximity to the cysteine box of the catalytic triad and is highly conserved across species. Reproduced from [179]

1.10 Hypothesis and aims of the study

The finding of strong linkage between the *UFSP2* mutation and the BHD phenotype and the *in vitro* evidence that this mutation prevented the activation of Ufm1 led to the hypothesis:

That impaired activation of the Ufm1 system and/or removal of Ufm1 from its target proteins might have an effect on hip joint development and thereby cause the BHD phenotype.

The overall aim of this study was therefore to investigate the link between the *UFSP2* mutation and the BHD phenotype. At the start of this research however little was known about the Ufm1/Ufsp2 system. From limited published literature (see 1.8.4) there was evidence that components of the system were conserved across species but that expression of the system did not appear to be tissue specific. Further only one putative target for Ufm1/Ufsp2 system had been identified and its expression did not appear to be tissue specific. Therefore in order for the *UFSP2* mutation to have a tissue and/or joint site specific effect, two possibilities were considered: i) that the functional consequences of the *UFSP2* mutation may have a greater effect on the development of the hip joint than on other tissues and/or ii), that if there were further targets of the Ufm1/Ufsp2 system, that these may be tissue and/or joint site specific. To investigate these possibilities three specific lines of investigation were followed to:

i. Identify whether the Ufm1/Ufsp2 system was differentially regulated during hip joint development by investigations of gene expression *in vivo* using mouse tissues from different developmental stages and *in vitro* using cell differentiation assays.
ii. Identify targets of the Ufm1/Ufsp2 system through the development of an *in vitro* Ufm1 conjugation system.
iii. Determine whether expression of mutated Ufsp2 had a tissue specific effect through development and analysis of a transgenic mouse model.

Chapter 2. Materials and Methods

2.1 Materials

Materials and their suppliers are listed in Appendix 1

2.2 Methods

Buffers components are listed in Appendix 2

2.2.1 SDS-PAGE

6x Laemmli buffer was added to protein samples. Samples were then denatured by heating at 95ºC for 10 minutes and loaded on NuPage® Novex 10% Bis-Tris polyacrylamide gel (Invitrogen). The gel was run in 1x MOPS buffer in XCell SureLock® Mini-Cell tank at 180 volts for approximately 1 hour. Protein sizes were estimated by comparing with protein standard (Biorad, Fermentas)

2.2.2 Western blotting

Proteins resolved by SDS-PAGE were transferred on PVDF or nitrocellulose membrane (GE Healthcare) using Bio-Rad electrotransfer system (1 hour at 100V). All incubations were performed on a platform rocker. The membrane was incubated overnight at 4ºC in blocking buffer. Old blocking buffer was discarded and the membrane was incubated at room temperature with primary antibody diluted 1:5000 in blocking buffer. The membrane was washed 3 times for 10 minutes with PBST and incubated with secondary antibody conjugated with HRP diluted 1:5000 in blocking buffer. The signal was detected using ECL reagent and photographic film.

2.2.3 Cell culture

HEK293T, **C3H 10T$^{1}/_{2}$** and **C2C12** cells were maintained in Dulbecco's Modified Eagle's Medium - high glucose (Sigma) supplemented with 10% FBS, 1x non-essential aminoacids and 1x penicillin/streptomycin.

2T3 osteoblast cells were cultured in Minimum Essential Medium Eagle Alpha Modification supplemented with 10% (v/v) FBS, 2 mM L-glutamine and 1x penicillin/streptomycin.

All cells were cultured in tissue culture flasks or dishes (Falcon) in 37ºC and 5% CO_2 atmosphere and passaged or harvested before reaching confluency.

2.2.4 Passaging

Cells were passaged by trypsinisation. The culture medium was removed and the cells were washed with PBS. Trypsin-EDTA was added, cells were incubated in 37ºC for 3 minutes and equal amount of medium was added to neutralise trypsine. Cells were then centrifuged in falcon tubes in Boeco C-28 centrifuge at 1600 rpm for 3 minutes. The supernatant was discarded and cells were resuspended in culture medium.

2.2.5 Cell freezing/recovery

Cells were trypsinised as previously but the pellet was resuspended in foetal bovine serum with 10% DMSO, aliquoted in cryovials (Falcon) and frozen in -80ºC before transferring to liquid nitrogen. When required cells were thawed in 37ºC and diluted 1:10 with culture medium, centrifuged in Boeco C-28 centrifuge at 1600 rpm for 3 minutes. The supernatant was discarded and cells were resuspended in culture medium and transferred to a culture flask.

2.2.6 Calcium phosphate transfection

Cells were seeded at $1x10^6$ cells per well of a 6 well plate 24h before transfection. On the day of transfection for each well to be transfected 5µg of plasmid DNA was diluted in sterile water to a total volume of 263µL and 37µL of 2M $CaCl_2$ was added. The DNA/ $CaCl_2$ solution was then mixed with 300µL of 2x HEPES-buffered saline and incubated for 30 minutes at room temperature. During that time the culture medium was removed from cells and 1.5ml of fresh medium was added to the wells. 2.1µl of 25mM chloroquine was added to the incubating solution and the whole mixture was subsequently added to the cells. The cells were then returned to 37ºC and incubated with the solution for 5 hours. After this period cells were washed 2 times with PBS and 2ml of fresh medium was added.

2.2.7 Transfection using lipid based reagent

Cells were transfected using FuGENE® HD Transfection Reagent (Promega, UK) according to the manufacturer's protocol with the reagent to DNA ratio of 3:1. The protocol is available online at http://www.promega.com/techserv/tools/fugenehdtool/.

2.2.8 Lentiviral transduction

HEK293T cells were seeded in a 6 well plate at $1x10^6$ cells per well 1 day before transfection. Cells were transfected with three vectors using previously described calcium phosphate method. 5µg of the smallest plasmid and equimolar amounts of the other two plasmids were used. For example, 5µg of pMD2-G (envelope proteins, 5824bp), 9.1µg of psPAX2 (GAG and POL, 10703bp) and 7.7µg of p201-iEP-Ufm1ΔC2 (vector with gene of interest, GFP and puromycin resistance, 8709bp) was used per each well. 24 hours post transfection a plate with cells to be transduced was prepared. The cells were seeded at density of $5x10^6$ cells per well. 48 hours post transfection the conditioned media from the virus producing cells was collected and centrifuged at 1600 rpm for 3 minutes in Boeco C-28 centrifuge. The supernatant was then filtered using a syringe driven 0.44µm filter unit. The filtered supernatant was mixed 1:1 with fresh media and 2ml was applied on cells to be transduced. 1.5ml of fresh media was added to the virus producing cells and both plates were returned to the incubator. 72 hours post transfection the procedure was repeated but the virus producing cells were discarded after collecting the conditioned media. 24 hours after adding the last conditioned media the transduced cells were washed and fresh media was added. 48 hours after adding the last conditioned media the transduced cells were trypsinised, transferred to a T75 flask and the transduced cells were selected on the basis of their antibiotic resistance (2µg/ml puromycin in case of HEK293T cells and the p201-iEP vector).

2.2.9 Preparation of competent *E. coli* cells

A single colony of DH5α cells was picked from an LB agar plate and inoculated into 5ml of LB media. The culture was grown overnight in a shaker set to 200 rpm and 37°C. 1ml of the overnight culture was added to 100ml of pre-warmed LB and incubated in the shaker until the culture reached A_{600} OD of 0.3. The cells were then centrifuged for 10 minutes at 1500 rpm at 4°C in Sigma 3-16PK centrifuge. The

supernatant was discarded and the cells resuspended in 40ml of ice cold 50mM $CaCl_2$. After 40 minutes of incubation on ice the cells were centrifuged at 1250 rpm at 4°C for 10 minutes. The supernatant was discarded and the cell pellet was resuspended in 5ml of ice cold 50mM $CaCl_2$ 15% glycerol (v/v). The cells were incubated on ice for 2 hours, transferred to 1.5ml eppendorf tubes in 50 µl aliquots and frozen in -80°C.

2.2.10 Transformation of competent *E. coli* cells

Cells were removed from -80°C and thawed on ice. 100ng of plasmid DNA was added to the cells and incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 1 minute and incubated on ice for additional 5 minutes. 950 µl of LB media was added to the tube and incubated for 1 hour in a shaker at 37°C. After this time serial dilutions were plated on selective LB agar plates.

2.2.11 Polymerase chain reaction (PCR)

PCR was used for genotyping or to amplify DNA for cloning. Reactions were performed in 0.2 ml tubes containing reagent mixture shown in Table 2, in a PCR machine set to run a cycling program shown in Table 3.

Reagent	Volume
DNA (1-100ng)	1.00µl
PCR buffer (10X)	5.00µl
$MgCl_2$ 50mM	3.00µl
dNTPs 10mM	2.50µl
Forward primer (10µM)	1.00µl
Reverse primer (10µM)	1.00µl
Polymerase	0.50µl
dH_2O	36.00µl
Total	50.00µl

Table 2. PCR reagent mixture composition.

Step	Temperature	Time	# of cycles
Initial Denaturation	95°C	3 minutes	1
Denaturation	95°C	15 seconds	35
Annealing	Variable	15 seconds	
Extension	72°C	Variable	
Final extension	72°C	5 minutes	1
Hold	4°C	Forever	1

Table 3. PCR thermo-cycler program.

2.2.12 Agarose gel electrophoresis

1% or 2% agarose gels were prepared by melting appropriate amounts (w/v) of agarose in TBE buffer by heating in microwave. Completely melted agarose was cooled down to approximately 50°C and ethidium bromide was added to a final concentration of 0.01% before pouring into a casting tray. The gel was then run in TBE buffer and the DNA visualised by UV light in a transiluminator. To estimate the size of DNA bands the sample was compared with DNA marker (HyperLadderTM range, Bioline Ltd., UK).

2.2.13 DNA extraction from an agarose gel

The DNA in gel was visualised by a handheld UV source and the appropriate DNA band was excised from the gel with a scalpel. The extraction was performed using QIAquick® gel extraction kit (Qiagen Ltd., UK) according to the manufacturer's protocol.

2.2.14 Plasmid DNA purification

Plasmid DNA was purified using QIAprep Spin Miniprep Kit or QIAGEN Plasmid Maxi Kit according to the manufacturer's protocols.

2.2.15 Mouse genomic DNA extraction

Mouse tissue biopsies were incubated in 200 µl of gDNA extraction buffer in 1.5 ml tubes at 55°C overnight. The samples were centrifuged at 13000 RPM for 10 minutes in a standard tabletop centrifuge. The supernatant was transferred to a new 1.5 ml tube and 300 µl of Phenol-Chloroform-Isoamyl alcohol mixture was added. The samples were incubated with shaking at room temperature for 10 minutes, centrifuged for 1 minute and the bottom fraction (Phenol mixture) was aspirated with a pipette and discarded. Samples were centrifuged again for 1 minute and the supernatant was carefully transferred to a new 0.5 ml tube with a pipette. 98% ethanol equal to approximately 2.5 volumes of the supernatant was added. Samples were mixed by inverting and the genomic DNA visible as white fluff was collected with a pipette tip and transferred to a new tube (alternatively the DNA can be centrifuged and the

supernatant decanted). The genomic DNA was air dried for a couple of minutes and dissolved in TE buffer.

2.2.16 mRNA extraction

mRNA was extracted from cultured cells or mouse tissue samples using RNeasy Mini Kit (Qiagen Ltd., UK) according to the manufacturer's protocol.

2.2.17 Generation of cDNA by reverse transcription PCR

100 ng of mRNA was DNase treated using Ambion® DNA-*free*™ DNase Treatment & Removal Reagents (Life Technologies, UK) according to manufacturer's protocol and used as template in reverse transcription reaction. The mRNA was reverse transcribed using TaqMan® Reverse Transcription Reagents (Life Technologies, UK) in a reaction summarised in Table 4.

Reagent	Volume
Reaction buffer (10X)	2.00µl
RNA (1µg)	Variable
$MgCl_2$ 25mM	4.00µl
dNTPs 10mM	4.00µl
RNase Inhibitor 20U/µl	0.50µl
Oligo dT 50µM	0.75µl
Reverse Transcriptase 50U/µl	0.50µl
dH_2O	Variable
Total	20.00µl

Cycling conditions
10 min at 25°C
30 min at 48°C
5 min at 95°C

Table 4. Reverse transcription reaction conditions.

2.2.18 Tandem affinity purification (TAP)

Preparation of cell lysates: Culture media was removed and the cells were washed with ice cold PBS with protease inhibitors. PBS was removed and Gu-HCl denaturing lysis buffer was added (0.4 ml per well of a 6 well plate or 4 ml per T75 flask). The cell lysates were collected in to tubes and were allowed to lyse further for 1 hour at 4°C on a rotary shaker set to low rpm. After the incubation the lysate was passed 4 times through 21G, 23G, and 25G needles to reduce the viscosity. The lysate was then centrifuged at 4°C and 13000 rpm for 30 minutes and decanted to a new tube avoiding the carryover of the cell debris collected at the bottom of the tube.

Affinity purification: Ni-NTA resin (5Prime) was directly added to the cleared cell lysate. The resin was incubated with the lysate for 4h or overnight at 4°C on a rotary shaker set to low rpm. On the next day the resin was briefly centrifuged and the supernatant (flow-through) was kept for subsequent analysis. The resin was washed with lysis buffer and then with Urea buffer A (twice) and Urea buffer B (twice). The resin was washed with neutral Buffer W before eluting with 1 resin volume of Elution Buffer (small scale) or 3 times with 1 resin volume of Elution Buffer (large scale).

The elution fractions were combined and applied on Strep-Tactin column (IBA) with bed volume 5 times of Ni-NTA resin. The flow-through was collected and applied on the column again. The column was then washed with 5 column volumes of buffer W and bound proteins were eluted with 3 column volumes of buffer E.

2.2.19 Radioactive in situ hybridisation

Probe synthesis: DNA template for the synthesis of RNA probe was prepared by linearisation of plasmids containing probe sequence with a restriction enzyme cutting at the 5' end of RNA polymerase transcription site generating antisense RNA product. The template DNA was then separated on agarose gel and extracted using QIAquick Gel Extraction Kit. The radio-labelled probe was synthesised using Riboprobe® System (Promega) in a reaction described in Table 5a for 40 min at 37°C. After this time 1µl of RNA polymerase was added and the reaction was incubated at 37°C for an additional 1 hour.

Table 5a. RNA transcription reaction

Reagent	Volume
DEPC H_2O	6.7 µl
5X reaction buffer	5.0 µl
1 M DTT	0.5 µl
10 mM GTP	1.2 µl
10 mM ATP	1.2 µl
10 mM CTP	1.2 µl
35S UTP	3.5 µl
50 µM UTP	1.0 µl
RNasin (40 U/µl)	0.5 µl
DNA template (2-3 µg)	3.2 µl
Either T7 or T3 RNA polymerase	1.0 µl
Total	**25.0 µl**

Table 5b. DNase treatment

Reagent	Volume
RNasin (40 U/µl)	0.5 µl
tRNA (10 mg/ml)	1.0 µl
1 M DTT	0.5 µl
RNase free DNase	0.5 µl
Total	**2.5 µl**

Table 5c. RNA precipitation

Reagent	Volume
DEPC H_2O	160 µl
1 M DTT	4 µl
5 M NaCl	4 µl
3 M NaOAc (pH 5.2)	20 µl
EtOH (100%)	400 µl
Total	**588 µl**

Table 5. Preparation of radioactive RNA probes for *in situ* hybridisation.

Following transcription the DNA template was digested with DNase by addition of reaction mixture described in Table 5b and incubation at 37°C for 10 minutes. The RNA probe was then precipitated overnight at -80°C by addition of reagent mixture described in the Table 5c. The precipitated probe was centrifuged in Sigma 3-16PK centrifuge at 4°C, 13000 rpm for 10 minutes. The supernatant was removed to a separate tube and the presence of precipitated 35S labelled RNA was confirmed with Geiger counter. The pellet was washed with 50 µl of 10 mM DTT diluted in 100% EtOH and centrifuged at 4°C, 13000 rpm for 5 minutes. The supernatant was removed and the pellet was checked with Geiger counter. The washing step was repeated and the pellet was then air dried for 5 minutes. After drying the RNA pellet was resuspended in 50 µl of 50 mM DTT diluted in DEPC H_2O. Probes longer than 300 bp were hydrolysed by addition of hydrolysis buffer and incubation at 60°C for time depending on the length of probe calculated using following formula:

$$\frac{\text{Transcript length (kb)} - 0.3}{0.033 \times \text{Transcript length (kb)}} = \text{Time in minutes}$$

The hydrolysis reaction was stopped with 50 µl of neutralising buffer and precipitated for 4 hours at -80°C by adding precipitation mixture from Table 5c. The washing steps were repeated as previously and the probe was resuspended in 25 µl of 50 mM DTT diluted in DEPC H_2O and stored in -80°C.

Preparation of tissue sections: Mouse tissues were fixed overnight in 4% PFA in PBS, processed in Shandon Citadel 2000 tissue processor and paraffin embedded. Tissues from mice older than newborn were decalcified for three days in decalcification solution (20% EDTA, 4% PFA in PBS pH 7.4) at 4°C before embedding. Slides were prepared by sectioning paraffin blocks on Microm HM 355S microtome with feed set to 5 and drying them overnight in 50°C. The sections were then de-waxed in xylene and rehydrated by passing through a graded ethanol series. The slides were incubated in 2% HCl for 20 minutes, washed in 2x SSC and incubated in proteinase K buffer for 10 minutes at 37°C. The proteinase was then inactivated by dipping the slides in 2mg/ml glycine for 2 minutes at room temperature. Slides were washed twice in PBS, incubated for 20 minutes in 4% PFA-PBS and washed in PBS again. The slides were then incubated in triethanolamine (triethanolamine 4.65 ml, 875 µl acetic anhydride and 345

ml H_2O) for 10 minutes at room temperature and washed in PBS followed by wash in H_2O and dehydrating through the graded ethanol series.

Hybridisation: The probe was measured using a scintillation counter and the amount of probe to be used per slide was calculated according to the following formula:

$$\frac{4{,}000{,}000}{\text{Scintilation count}} \times (\text{number of slides} + 1) = \text{amount of probe in } \mu l$$

The required amount of probe was added to the hybridisation buffer (100 µl of buffer per slide) and heated to 95°C. The hybridisation mix was then applied on the slides and the slides were covered with glass cover slips. The slides were then placed in hybrydisation chamber containing 50% formamide and incubated overnight at 55°C. On the next day cover slips were removed and slides were incubated in following wash solutions and buffers: Wash solution twice at 55°C for 20 min and once at 65°C for 20 minutes; RNase buffer twice at 37°C for 15 minutes; RNase A at 37°C for 30 minutes; RNase buffer at 37°C for 15 minutes; Wash solution twice at 65°C for 20 min; SSC/DTT buffer at 5°C for 20 min and then 5 minutes at room temperature; ammonium acetate / ethanol for 2 minutes at room temperature. The slides were then dehydrated through the graded ethanol series, air dried and exposed to autoradiography film overnight to confirm the hybridisation worked before covering the slides with photoreactive emulsion.

Developing: After confirming the hybridisation with autoradiography film the slides were covered with K5 nuclear emulsion (Ilford, UK) prepared as 50:50 solution with 2% glycerol, dried and stored in dark for 2 weeks before developing. The slides were developed in D19 developer (Kodak) and fixed using Unifix (Kodak). After developing slides were H&E stained using Thermo Shandon Linistain GLX stainer.

2.2.20 Osteogenic differentiation

Subconfluent 2T3 cells were trypsinized and plated on gelatine coated 12 well plates at a density of $1.6x10^4$ cells/cm^2. On reaching confluency (usually day 3) the cells were cultured in MEM (alpha modification) containing 5% FBS (v/v), 2 mM L-glutamine, 1x penicillin/streptomycin, 100 µg/ml L-ascorbic acid 2-phosphate, 5 mM β-glycerophosphate and 10 ng/ml rhBMP-2. Media was changed every 2 days. The cultures were usually maintained until day 12.

2.2.21 Chondrogenic differentiation in micromass cultures

C3H $10T^1/_2$ cells were allowed to reach confluency in T75 flasks. Cells were trypsinized and centrifuged in 15 ml falcon tubes. The cell pellet was resuspended in Ham's F12 (10% FBS) medium to a density of 10^7 cells/ml. Cells were deposited at the bottom of a 6 well plate in three separate 10µl drops per each well. The plate was placed in the incubator for approximately 1 hour to let the cells adhere. After 1 hour Ham's F12 medium with 100 ng/ml rhBMP-2 was added to the well. The medium was changed to fresh medium with the addition of rhBMP-2 every three days.

2.2.22 Qantitative real-time PCR

Real-time PCR reactions were performed using Abi Prism 7000 (Applied Biosystems) and SYBR® green PCR mastermix. Reactions were performed in total 12.5µl volume with 6.25µl SYBR® green, 1.5µl of cDNA generated as described in 2.2.17, 1.5µl of each primer (10pmol/µl) and 1.75µl of dH_2O. The thermocycling conditions were as shown in Table 6. Analysis was done using the formula $2^{-\Delta CT}$ where ΔCT is the difference between the CT value for the gene being measured and the gene to which relative expression is being compared. Expression was measured relative to *Actb* (β-actin) or *Gapdh* as indicated in the results.

Temperature	Time	Number of cycles
50°C	2 minutes	1
95°C	10 minutes	1
95°C	15 seconds	40
60°C	1 minute	
72°C	15 seconds	
95°C	15 seconds	1
60°C	20 seconds	1
95°C	15 seconds	1

Table 6. Real-time PCR cycling conditions.

2.2.23 Luciferase assay

2T3 cells were plated in 96 well solid white microplate (Fisher, UK) at a density of 80% per well in 100 µl of growth medium. Cells were transfected using FuGene reagent in a 3:1 DNA to reagent ratio according to the manufacturer's protocol. Each well was co-transfected with two plasmids: the pGL3-basic luciferase vector carrying

the tested genomic sequence and a control plasmid carrying *Renilla* luciferase gene under the control of the CMV promoter. The plasmids were co-transfected in 2:1 ration (ng/ng). All transfections were done in triplicate. 24h post transfection the medium was changed to fresh medium with or without 2μM thapsigargin. After 14h of incubation the media was removed and 25μl of PBS was added to each well followed by 25μl of Dual-Glo Luciferase Assay Reagent. The pGL3 firefly luciferase luminescence was measured after 10 minutes of incubation in a luminometer. Next, 25μl Dual-Glo Stop & Glo reagent was added, incubated for 10 minutes and *Renilla* luciferase luminescence was measured as previously. The values were calculated as ratio of firefly:*Renilla* luminescence for each well.

2.2.24 Colloidal Blue staining

Staining of SDS-PAGE gels was performed using Colloidal Blue Staining Kit (Invitrogen) according to manufacturer's protocol. Briefly, for one gel staining solution was prepared (55ml deionised water, 20ml methanol, 5ml stainer B, 20ml stainer A) and the gel incubated in the solution with shaking for 3 hours. Staining solution was removed and the gel washed until disappearance of the background.

2.2.25 Preparation of tissue sections for histology

Mouse tissues were fixed overnight in 4% PFA in PBS, processed in Shandon Citadel 2000 tissue processor and paraffin embedded. Tissues from mice older than newborn were decalcified for three days in decalcification solution (20% EDTA, 4% PFA in PBS pH 7.4) at 4°C before embedding. Slides were prepared by sectioning paraffin blocks on Microm HM 355S microtome with feed set to 5 and drying them overnight in 50°C. The slides were then H&E stained using Thermo Shandon Linistain GLX stainer.

2.2.26 Mass spectrometry

Digestion: bands of interest were excised from the gel and dehydrated using acetonitrile followed by vacuum centrifugation. Dried gel pieces were reduced with 10 mM dithiothreitol and alkylated with 55 mM iodoacetamide. Gel pieces were then washed alternately with 25 mM ammonium bicarbonate followed by acetonitrile. This was

repeated, and the gel pieces were dried by vacuum centrifugation. Samples were digested with trypsin overnight at 37 °C. The samples were extracted in one wash of 20mM ammonium bicarbonate, and two of 50% acetonitrile, 5% formic acid. The extract was then dried by vacuum centrifuge to 20µl.

Mass Spectrometry: digested samples were analysed by LC-MS/MS using an Ultimate 3000 (LC-Packings, Dionex, Amsterdam, The Netherlands) coupled to a HCT Ultra ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany). Peptides were concentrated on a pre-column (5 mm x 300 µm i.d, LC-Packings). The peptides were then separated using a gradient from 98% A (0.1% FA in water) and 1% B (0.1% FA in acetonitrile) to 50% B, in 40 min at 200 nL min^{-1}, using a C18 PepMap column (150 mm x 75 µm i.d, LC-Packings).

Data Analysis: data produced was searched using Mascot (Matrix Science UK), against the IPI database with taxonomy of *Homo sapiens* selected. Data were validated using Scaffold (Proteome Software, Portland, OR).

Chapter 3. Expression of the Ufm1 system

3.1 Introduction

The hip joint is a complex organ composed of multiple tissues including bone, articular cartilage, the synovial membrane and surrounding it the fibrous capsule, and ligaments. As described in 1.6.1, mutations of ECM components, transcription factors and regulatory proteins are known to disrupt processes of chondrogenesis and endochondral ossficiation leading to forms of chondrodysplasia and secondary OA through a range of different mechanisms. In addition, the effects of such mutations may not be limited to cartilage and hence the secondary OA may not necessarily be limited to abnormal cartilage structure and function. For example, one of the symptoms of PSACH is joint laxity and it has been shown that it is the presence of COMP in ligaments that contributes to their strength and is required for resistance to loading. It has therefore been suggested that the development of early onset OA associated with PSACH may be the consequence of altered ligament as opposed to cartilage function [195]. In addition, several ER stress related regulatory genes such as *ATF4*, *ATF6* and *PERK* have been found to be indispensable for osteoblast growth, differentiation and endochondral ossification (see 1.7.3) [116, 128, 130, 196]. Further, several studies have also shown the importance of muscle contractions in joint formation and endochondral ossification during the embryonic stages of development [197, 198].

With so many developmental mechanisms and so many potential tissues that might be affected by abnormalities in the Ufm1 system, the first step in our investigations of the physiological role of the Ufm1/Ufsp2 system was to investigate the expression of *Ufsp2* in the developing mouse hip joint. Thereafter, expression of the Ufm1/Ufsp2 system was investigated during osteogenic and chondrogenic differentiation and during ER stress.

3.2 Results

3.2.1 Investigation of *Ufsp2* gene expression in the hip joint during mouse development

To determine whether the *Ufsp2* gene was expressed in the hip region at different stages of mouse development, radioactive *in situ* hybridisation was performed on tissue sections from E12.5, E14.5, newborn and 10 day old mice. The *Ufsp2* cDNA probe was obtained from Chris Watson (University of Manchester) and the *Col1a1* and *Col2a1* cDNA probes were obtained from Gareth Hyde (University of Manchester). Antisense mRNA was transcribed and labelled with ^{35}S UTP and mouse tissues were processed as described in 2.2.19. The *in situ* hybridisation procedure was performed first with the *Ufsp2* and *Col2a1* probes on embryonic and newborn tissue sections. The *Col2a1* probe which was used as a positive control showed characteristic staining of the cartilaginous elements of both embryonic and new born mice (Figure 16 and Figure 17). No *Ufsp2* expression was however detected in serial tissue sections of either embryonic or newborn mice. *In situ* hybridisation was then performed on tissue sections from 10 day old mice using probes for *Ufsp2*, *Col2a1* and *Col1a1* (Figure 18). Characteristic *Col2a1* expression was detected in cartilage and *Col1a1* expression was detected in bone and ligaments. In the hip region the *Ufsp2* expression pattern resembled that of *Col1a1* in that it localised to the bone and ligaments but the level of *Ufsp2* expression was much lower. *Ufsp2* expression was also visible in the muscle but not in the cartilage of the femoral head or the acetabulum. In the knee, (Figure 19 A) *Ufsp2* expression was detected in the bone and muscles, the secondary ossification centre and, possibly with weaker expression in part of the proliferative zone of cartilage. There was no expression of *Ufsp2* in the resting cartilage or the hypertrophic zone. *Ufsp2* expression was also detected in the vertebrae (Figure 19 B) where again the signal was in the bone, fibrocartilage and ligament but also appeared to be in a zone of cartilage immediately adjacent to the fibrocartilage and therefore overlapped partially with both *Col1a1* and *Col2a1* expression.

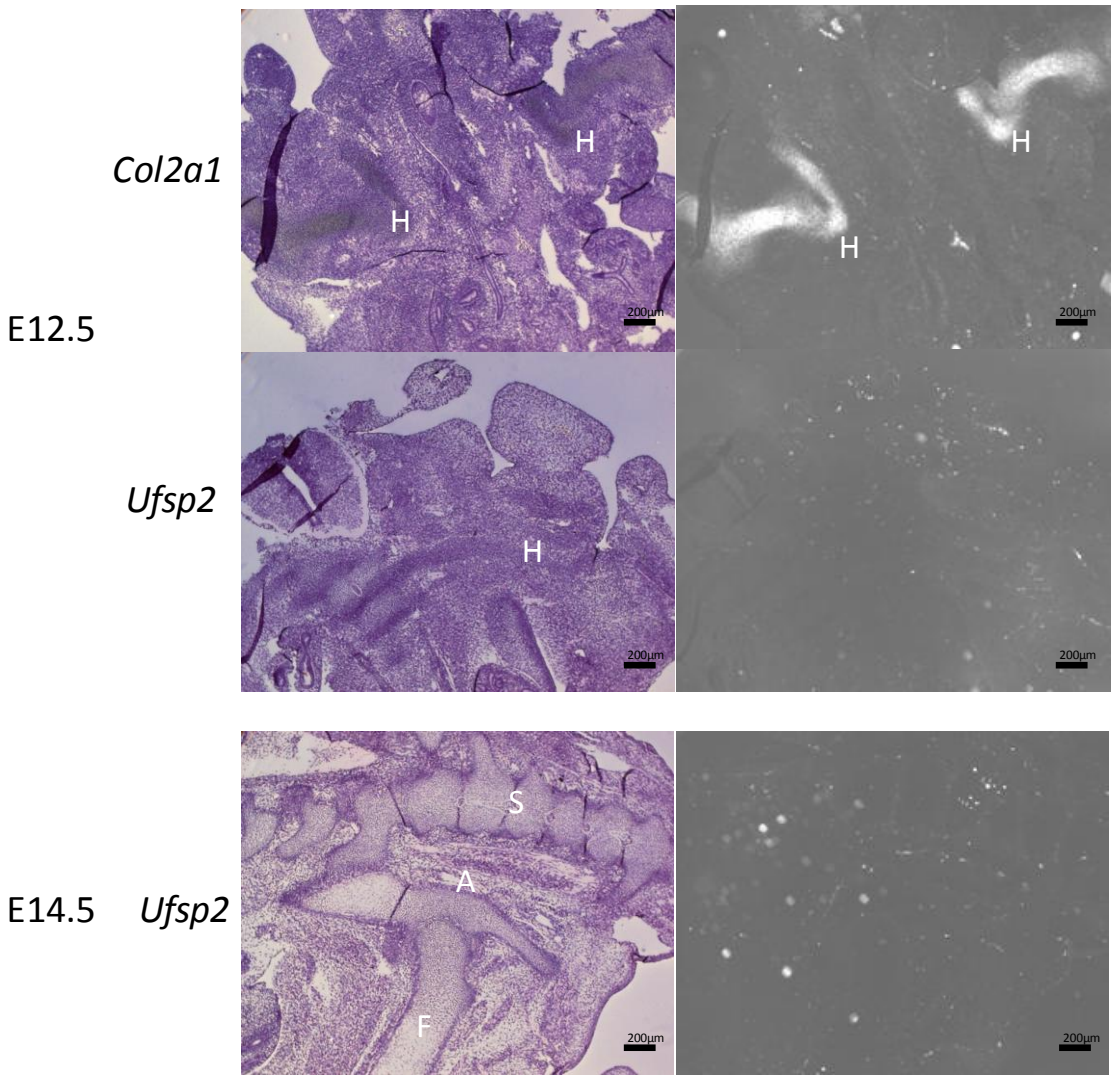


Figure 16. Radioactive *in situ* hybridisation on E12.5 and E14.5 mouse tissue sections.

Mouse tissue sections from 12.5 and 14.5 day embryos were stained for the expression of *Col2α1* and *Ufsp2*. Panels on the left show H&E staining and panels on the right are images taken in dark-field showing the RNA hybridisation signal. *Col2α1* expression is visible in the cartilage of the presumptive hip joint. No *Ufsp2* expression was detected. H – hip; S – spine; F – femur; A - acetabulum. Scale bar 200 µm.

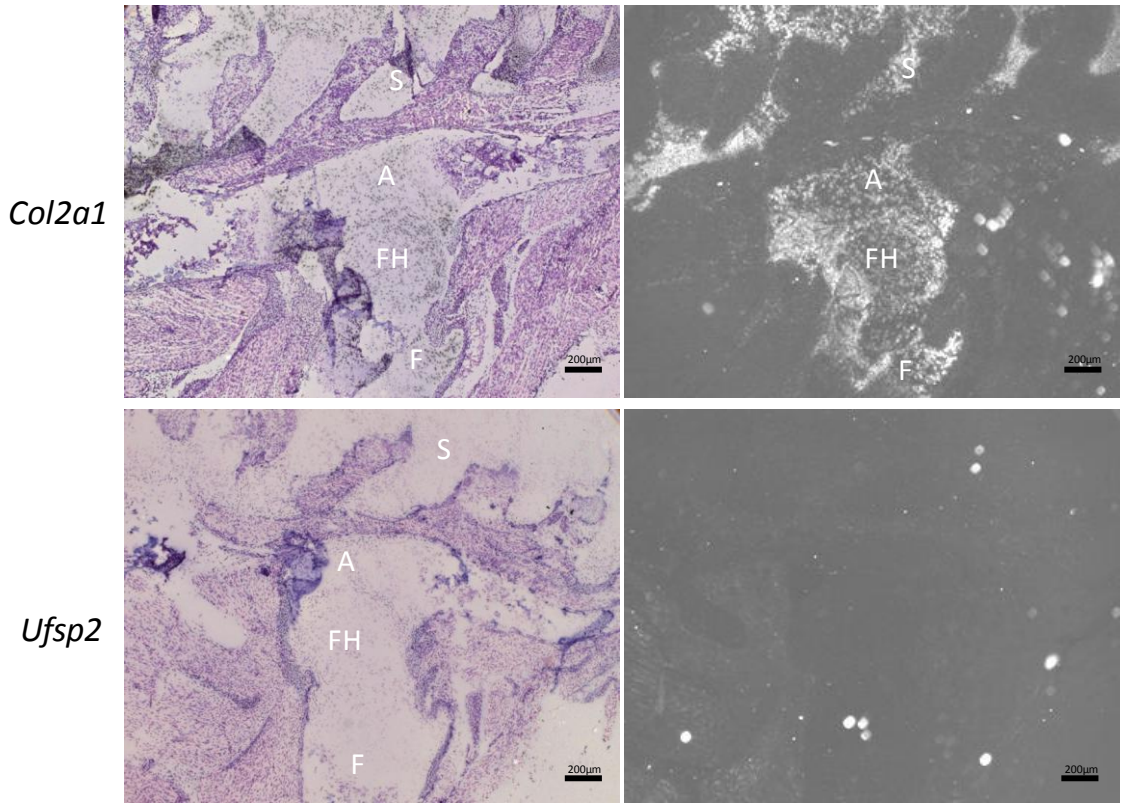


Figure 17. Radioactive *in situ* hybridisation on new born mouse tissue sections.

Mouse tissue sections from newborn mice were stained for the expression of *Col2α1* and *Ufsp2*. Panels on the left show H&E staining and panels on the right are images taken in dark-field showing the RNA hybridisation signal. *Col2α1* expression is visible in the cartilage of the presumptive hip joint and the spine. No *Ufsp2* expression was detected. A – acetabulum; FH – femoral head; F – femur; S – spine. Scale bar 200 µm.

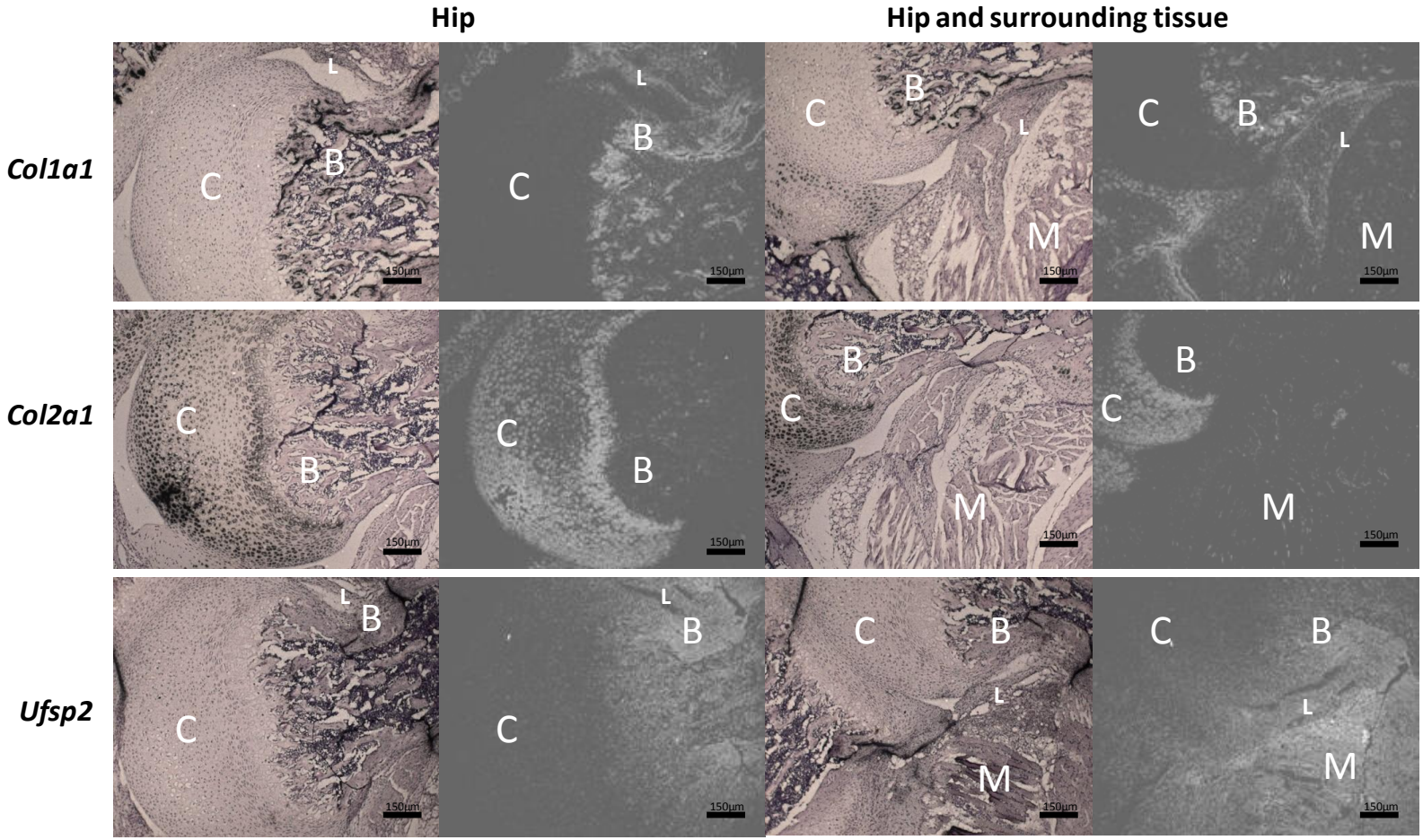


Figure 18. Radioactive *in situ* hybridisation on 10 day old mouse tissue sections with *Col1a1*, *Col2a1* and *Ufsp2* probes.

Col1a1 expression is visible in bone and fibrocartilage of ligaments. *Col2a1* is visible in the cartilage of the hip joint. *Ufsp2* expression is visible in the bone and soft tissue surrounding the joint but not in the cartilage. C – cartilage; B – bone; L – ligament; M – muscle; Images were taken in two modes: bright-field – visible H&E staining and dark-field – only radioactive probe signal visible. Scale bar 150 µm.

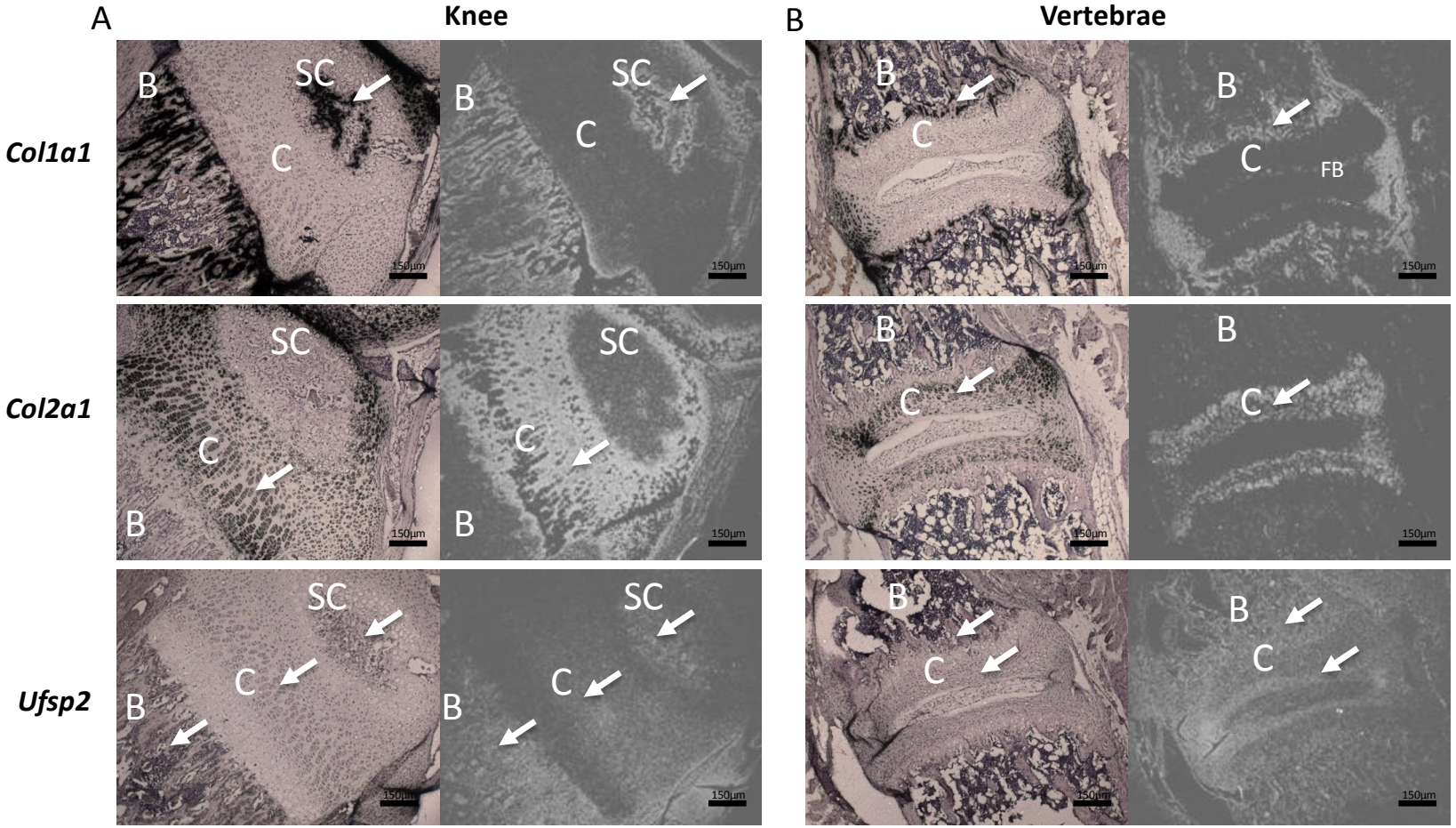


Figure 19. Radioactive *in situ* hybridisation on 10 day old mouse tissue sections with *Col1a1, Col2a1* and *Ufsp2* probes.

A. *Col1a1* and *Col2a1* expression is visible in bone and cartilage, respectively. *Ufsp2* expression is visible in bone, muscles, secondary ossification centre and proliferative zone of cartilage in the knee. **B.** In the vertebrae *Col1a1* is expressed in the bone, fibrocartilage and ligament; *Col2a1* in the cartilage; *Ufsp2* in bone, fibrocartilage and ligament. C – cartilage; B – bone; SC – secondary ossification centre; FB – fibrocartilage. Arrows point to the respective areas of expression in bright-field and dark-field images. Scale bar 150 µm.

3.2.2 Expression of the Ufm1 system during osteogenic differentiation

Based on the findings of the *in situ* hybridisation experiments it appeared that *Ufsp2* was consistently expressed by osteoblasts and it was therefore hypothesised that the Ufm1 system may have a role in either osteogenic differentiation or osteoblast function. In view of this, *in vitro* analysis of the expression of the Ufm1 system during the process of osteogenic differentiation was carried out using the 2T3 mouse osteoblast cell line [199]. This cell line was chosen as it can be induced to differentiate and mineralise following addition of an osteogenic culture medium containing rhBMP-2 [200].

To confirm that the 2T3 cell line could be induced to mineralise following addition of osteogenic medium (αMEM, ascorbic acid, β glycerol phosphate, rhBMP-2), the cells were cultured and induced to differentiate as described in 2.2.20. Non induced cultures served as controls. As is visible in Figure 20, there was evidence of mineralisation in the induced cultures from day 8.

To determine whether the Ufm1 system was differentially regulated during *in vitro* differentiation of the 2T3 cell line, mRNA was extracted as described in 2.2.16. The mRNA was extracted from induced and parallel uninduced control cultures (each condition plated in triplicate) following 2, 4, 6, 8, 10 and 12 days of culture. cDNA was generated from the extracted mRNA as described in 2.2.17 and the expression level of selected genes determined using quantitative real-time PCR as described in 2.2.22. The data shown in Figure 21 was obtained from two separate experiments.

The relative expression of *Runx2*, *Osterix*, *Osteocalcin* and *Col1a1* were measured in the 2T3 culture system as these are markers of osteogenic differentiation (see 1.5.2). *Bip* expression was also measured as its expression is known to increase during osteogenic differentiation as well as during ER stress (see 1.7.3) As is evident in Figure 21 the 2T3 cell line undergoes late osteoblast differentiation (see Figure 8) when induced with the osteogenic medium. Hence, *Runx2* expression decreases while *Osterix* is upregulated and the increase in *Osterix* is followed by the increase in *Osteocalcin* expression.

The relative expression of the components of the Ufm1 system including *Ufm1*, *Uba5* (E1), *Ufl1* (E3) and *Ufsp2*, as well as *Ddrgk1* (putative target) and *Lzap* (system associated gene) (see 1.8.4) were also measured in the 2T3 culture system. *Uba5* (E1),

Ufl1 (E3), *Ddrgk1* and *Lzap* were all significantly upregulated ($p<0.001$) in the induced cultures relative to the control at day 8. *Ufsp2* was also upregulated at day 8 but to a less significant degree ($p<0.01$) and *Ufm1* expression started to increase at day 6 ($p<0.01$) to peak at day 8 and remained at this level on days 10 and 12. Thus, in the induced cultures, *Uba5* (E1), *Ufl1* (E3), *Ddrgk1* and *Lzap* upregulation coincided with elevated ER stress as evidenced by upregulation of *Bip*. The uninduced cultures were likely undergoing spontaneous mineralisation from day 10 as evidenced by the increase in *Osteocalcin* expression but this increase was not mirrored to the same extent by *Bip* or the components of the Ufm1 system.

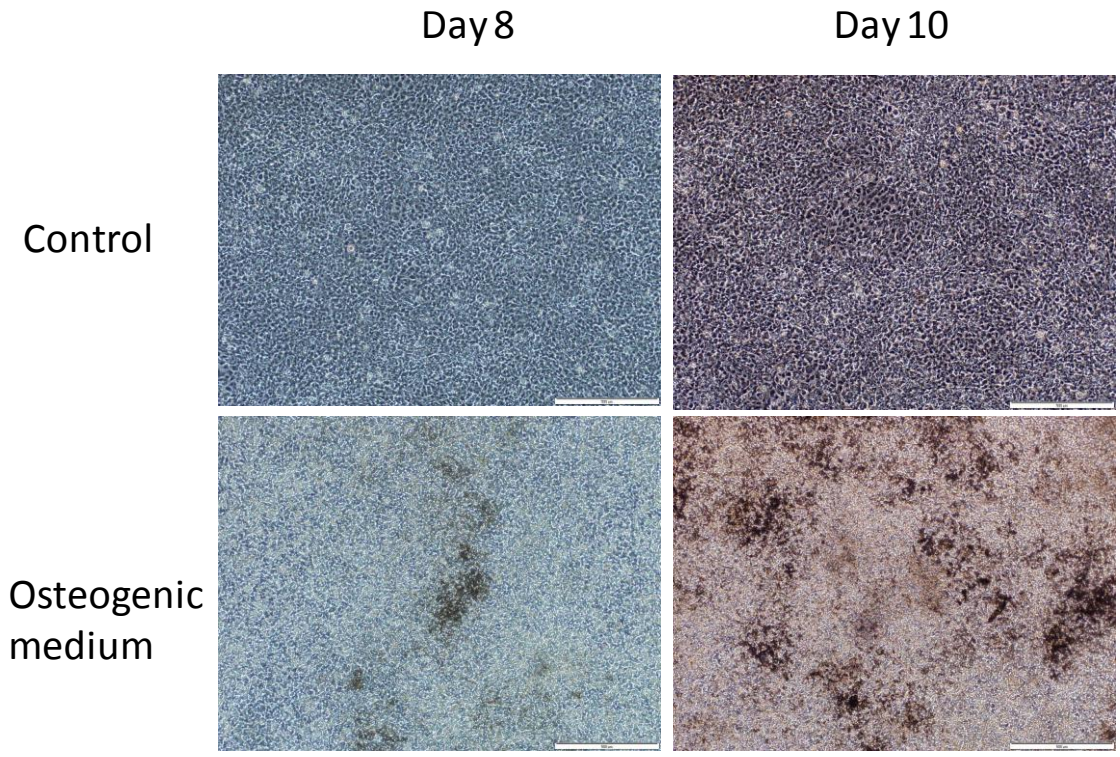
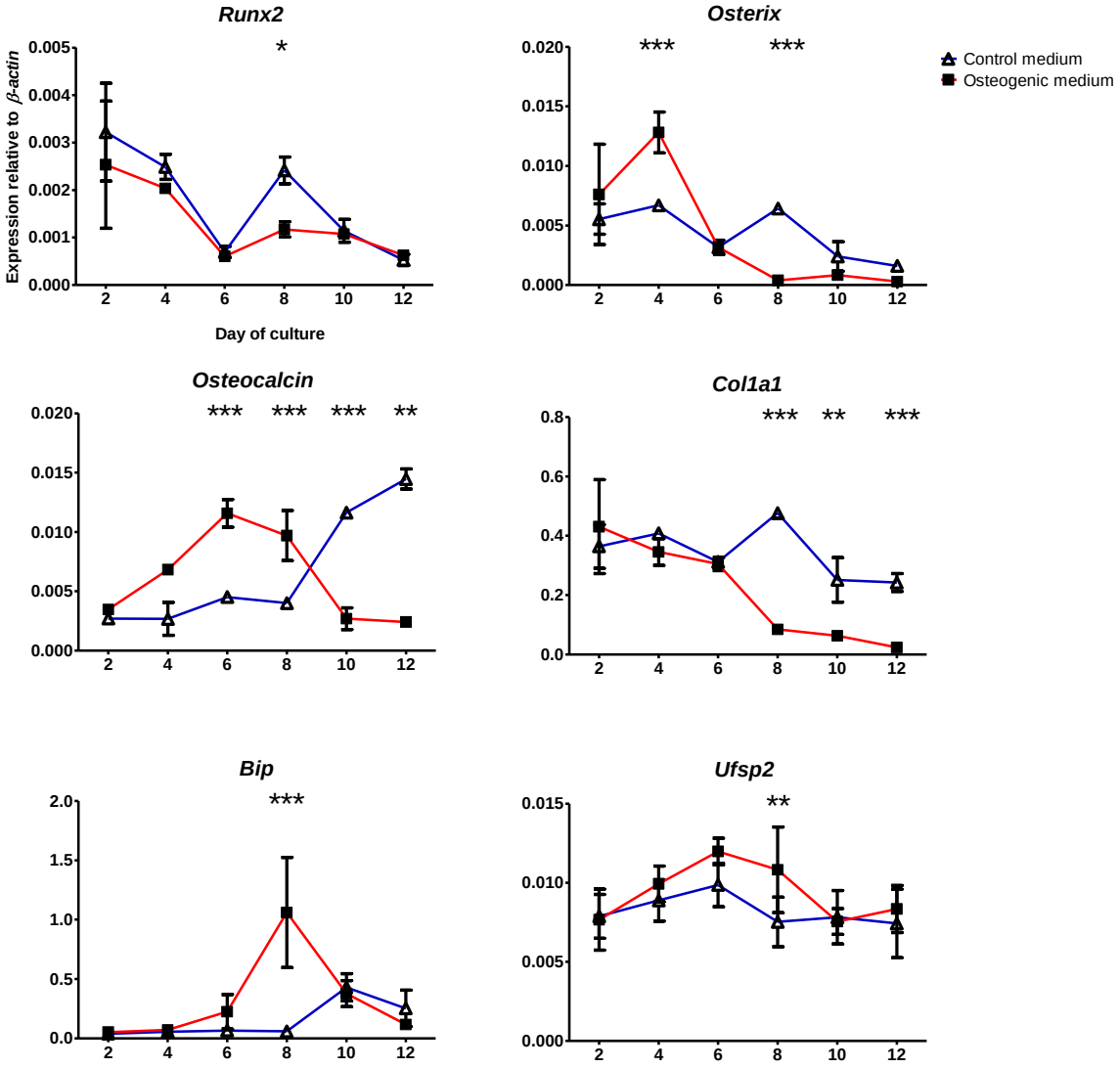


Figure 20. Osteogenic induction of 2T3 cells.

2T3 osteoblasts were cultured in control or osteogenic medium. Images were taken at days 8 and 10 of culture where dark patches of mineralisation were visible in the induced cultures without staining. Scale bar 500µm.



* **Figure 21** continued on the next page.

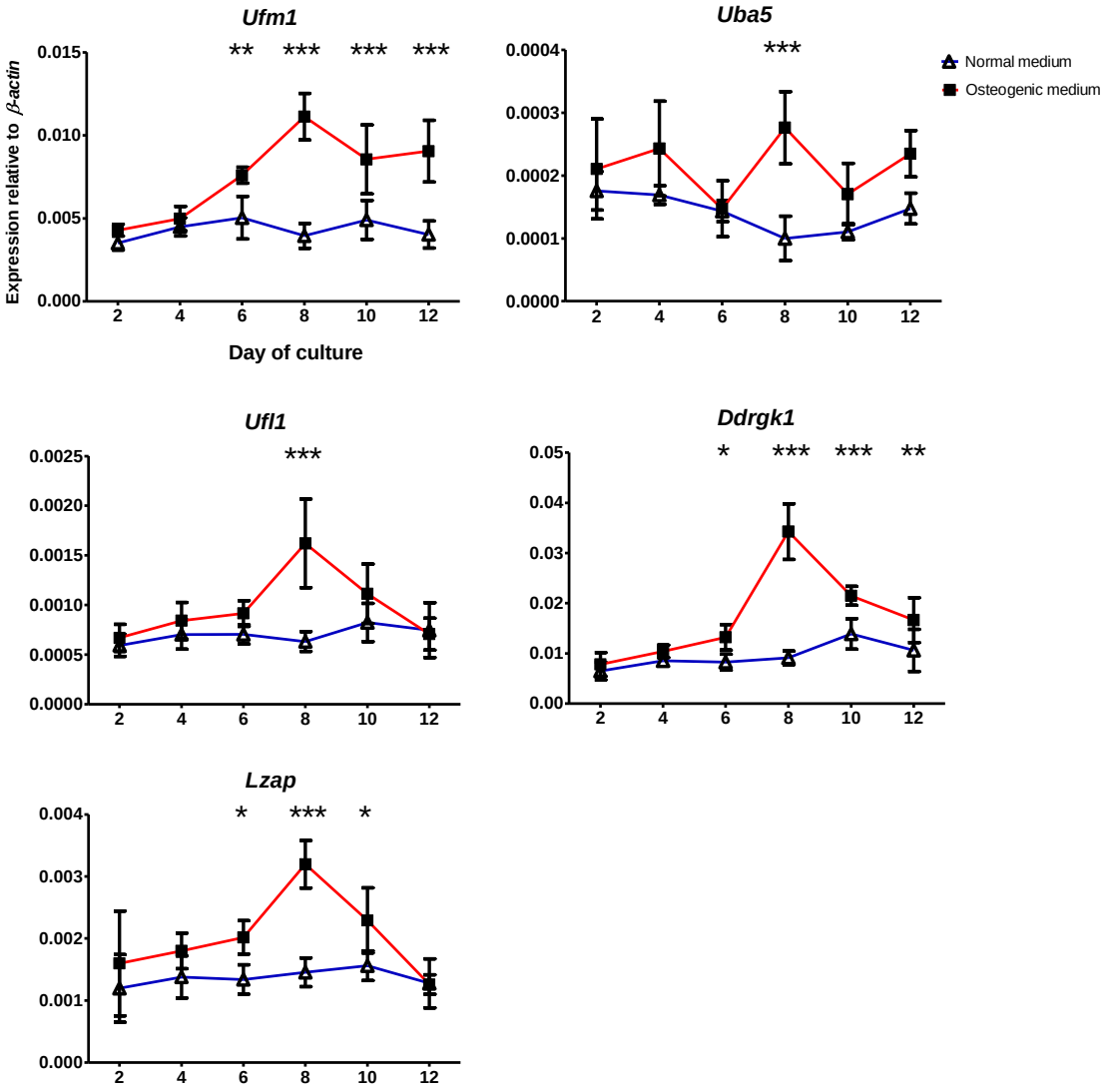


Figure 21. Expression of Ufm1 pathway genes during osteogenic differentiation of 2T3 osteoblasts.

Expression of the Ufm1 pathway was followed during 12 day culture of 2T3 osteoblasts. Cells were cultured in control or osteogenic medium and the expression of the genes indicated was determined using real-time PCR. Expression was quantified relative to *β-actin*. Values are the mean +/- standard deviation of two independent experiments performed in triplicate. Statistical analysis was performed using 1-way ANOVA with Bonferroni's post test. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$

3.2.3 Expression of the Ufm1 system during chondrogenic differentiation

The *in situ* hybridisation experiments found some evidence of expression of *Ufsp2* in cartilage as well as in bone (Figure 19) and therefore the expression of the Ufm1 system during chondrogenic differentiation was investigated. Micromass cultures of the C3H $10T^{1}/_{2}$ cell line [201] were prepared according to the method described in 2.2.21 and cultured in the presence of chondrogenic media (Ham's F12 + rhBMP-2). mRNA was extracted and transcribed to cDNA on the days 0, 3, 6, 9 of culture and the expression of selected genes was determined using quantitative real-time PCR as described previously.

One of the hallmarks of chondrogenic differentiation is secretion of large quantities of matrix of which collagen type II is the major component. *Col2a1* expression was therefore measured and shown to increase with time in micromass culture (Figure 22) indicating successful induction of chondrogenesis. The *Col2a1* expression was mirrored by the expression of *Bip*. All of the tested genes of the Ufm1 system showed significant increase of expression by the end of the culture compared to the level at day 0. This increased expression coincided with the increased expression of *Bip*.

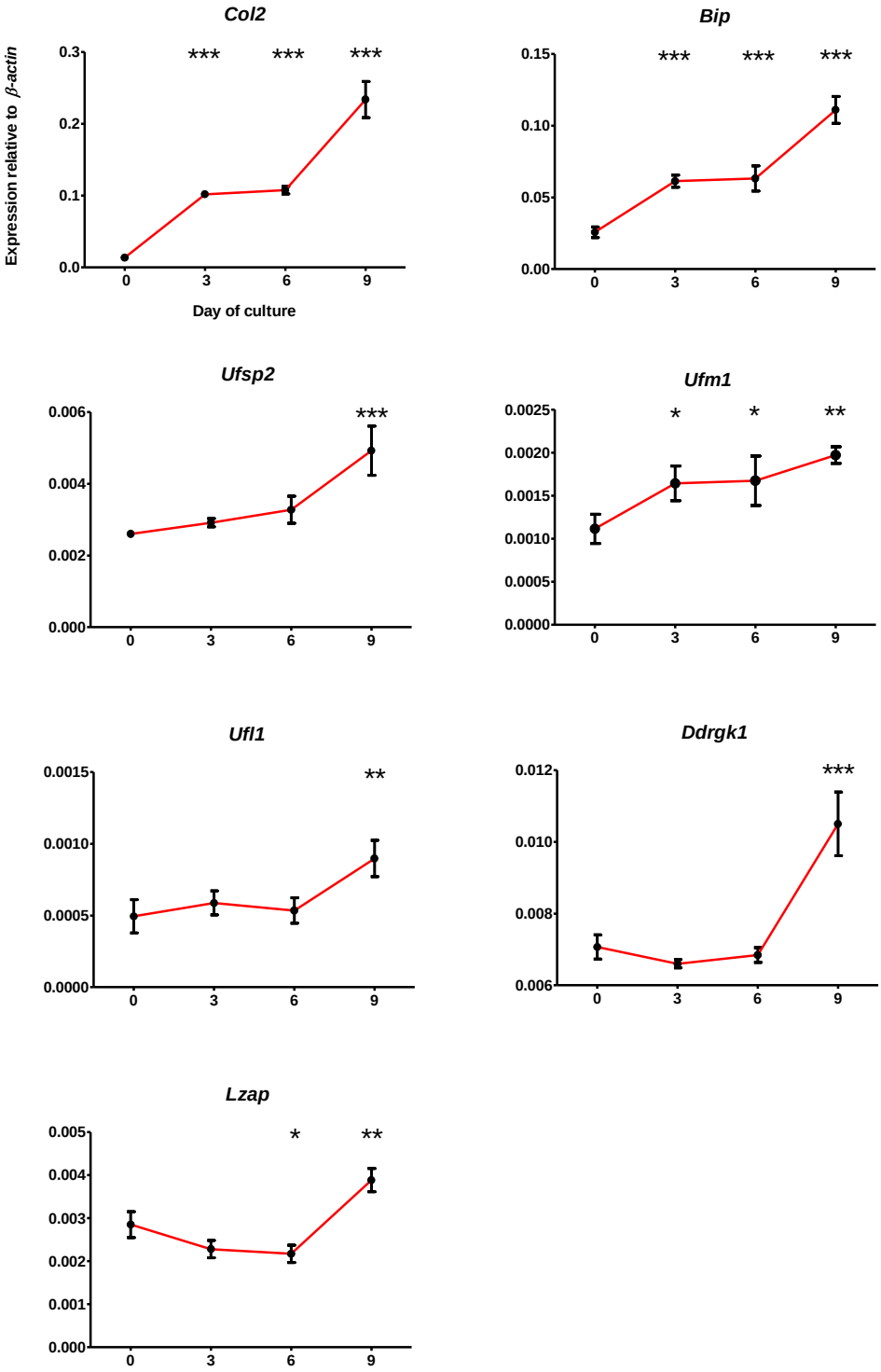


Figure 22. Expression of Ufm1 pathway genes during chondrogenic differentiation of C3H 10T$^1/_2$ cell line.

C3H 10T$^1/_2$ cells were induced to differentiate in micromass cultures in the presence of chondrogenic media. The expression of the genes indicated was determined using real-time PCR. Expression was quantified relative to *β-actin*. Values are the mean +/- standard deviation of one experiment performed in triplicate. Statistical analysis was performed using 1-way ANOVA with Bonferroni's post test. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$.

3.2.4 Expression of the Ufm1 system in response to ER stress

Upregulation of the Ufm1 system during osteogenic and chondrogenic differentiation was shown to coincide with increased expression of *Bip* which is known to be elevated during ER stress (see 1.7.2). Therefore, it was hypothesised that ER stress may be sufficient to influence the expression of the system. Three different cell lines were chosen to test the response of the Ufm1 system to ER stress: 2T3 (osteoblasts), C3H $10T^{1}/_{2}$ (embryonic fibroblasts) and C2C12 (myoblasts [202]). Cells were cultured as described previously (see 2.2.3). Subconfluent cultures were passaged 1:3, and on the following day the medium was replaced with fresh medium containing DTT at a final concentration of 2mM. mRNA was extracted at 2, 4, and 6 hour time points and cDNA prepared as previously described (see 2.2.16 and 2.2.17). The expression of Ufm1 system genes *Ufsp2*, *Ufm1*, *Ufl1*, *Ddrgk1* and *Lzap* was quantified using real-time PCR. ER stress was induced in all three cell lines in response to DTT as verified by elevated expression of *Bip* (Figure 23).

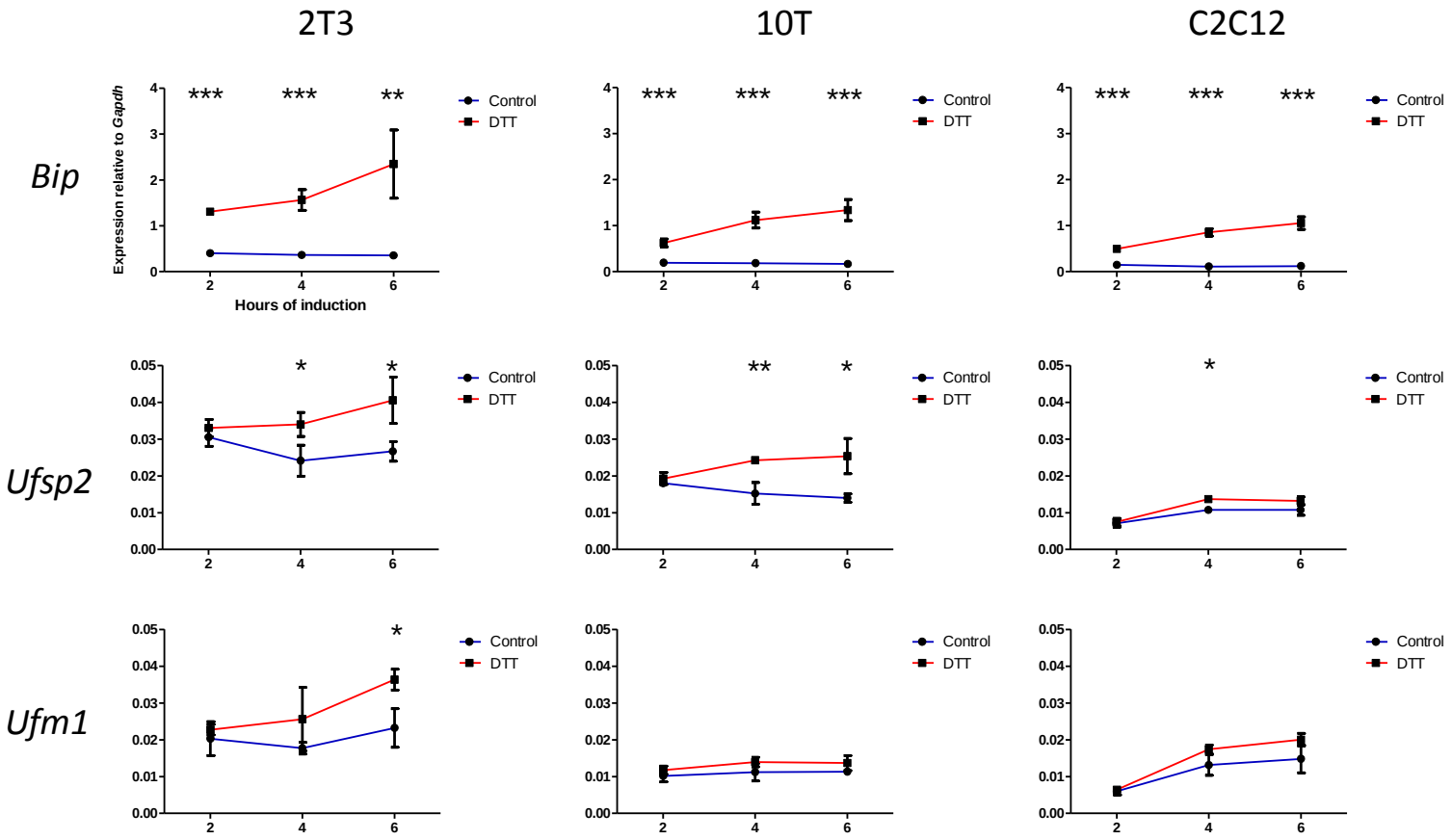
The expression of all the genes tested was upregulated in the DTT treated cultures (although not in every cell line) as compared to untreated controls. In the C212 cell line expression of *Ufl1*, *Ddrgk1* and *Lzap* was upregulated. In the C3H $10T^{1}/_{2}$ cell line all genes except *Ufm1* were upregulated. In the 2T3 osteoblast cell line all genes were significantly upregulated. Interestingly, *Ddrgk1* and *Lzap* genes were the quickest to respond to the stimulus with significantly elevated expression after 2 hours of treatment. In contrast, *Ufsp2*, *Ufm1* and *Ufl1* took 4 to 6 hours for their expression to reach significant difference as compared to untreated cells.

The specificity of the effect of ER stress on the expression of the Ufm1 system in the 2T3 osteoblast cell line was then tested by inducing ER stress with chemical agents acting on different properties of the ER. The agents used in the study were: DTT (which acts by reducing the disulphide bridges therefore directly affecting protein folding), tunicamycin (inhibitor of N-glycosylation of proteins) and thapsigargin (inhibitor of sarco/endoplasmic reticulum Ca^{2+} ATPase that affects the ER calcium balance). ER stress was induced by treating the cell cultures with 2μg/ml tunicamycin for 14 hours, 2μM thapsigargin for 14 hours and 2mM DTT for 6 hours.

As shown in Figure 24, the level of *Bip* expression was significantly higher in all induced cultures compared to the uninduced control cultures confirming that ER stress

had been induced by all three agents. All of Ufm1 system genes tested were upregulated significantly in response to all three agents, other than for, *Ufsp2* which was not upregulated significantly in response to tunicamycin and *Ufm1* which was not upregulated significantly in response to thapsigargin (Figure 24).

To rule out the possibility that induction of ER stress with chemical agents may influence the expression of *Gapdh*, the experiment inducing ER stress in 2T3 cells with thapsigargin and tunicamycin was also quantified with *β-actin* as a reference gene and the results are compared in Figure 25. While expression of *Ufm1* in response to thapsigargin and expression of *Ufsp2* in response to tunicamycin was not significantly upregulated when quantified relative to *Gapdh*, all genes were significantly upregulated under both conditions when quantified relative to *β-actin*.



* **Figure 23** continued on the next page.

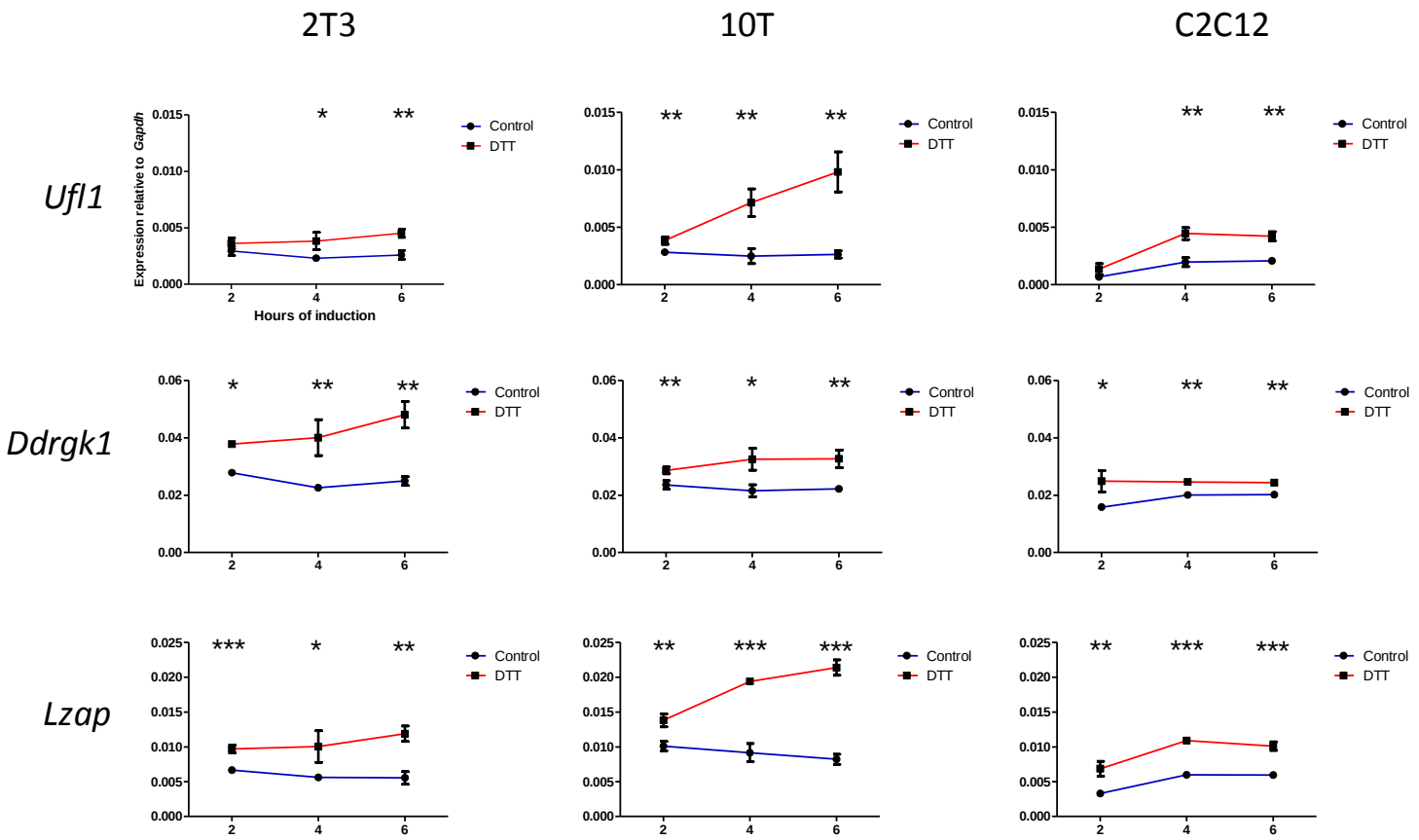


Figure 23. Expression of Ufm1 pathway following induction of ER stress with DTT.

The effect of ER stress on the expression of the genes of the Ufm1 pathway was tested in three cell lines (2T3 osteoblasts, C3H $10T^{1}/_{2}$ fibroblasts (10T), and C2C12 myoblasts). ER stress was induced by treating the cell cultures with 2mM DTT for 2, 4 and 6 hours. Untreated cultures served as controls. Expression was quantified relative to Gapdh. Values are the mean +/- standard deviation of one experiment performed in triplicate. Statistical analysis was performed using the unpaired t test. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$.

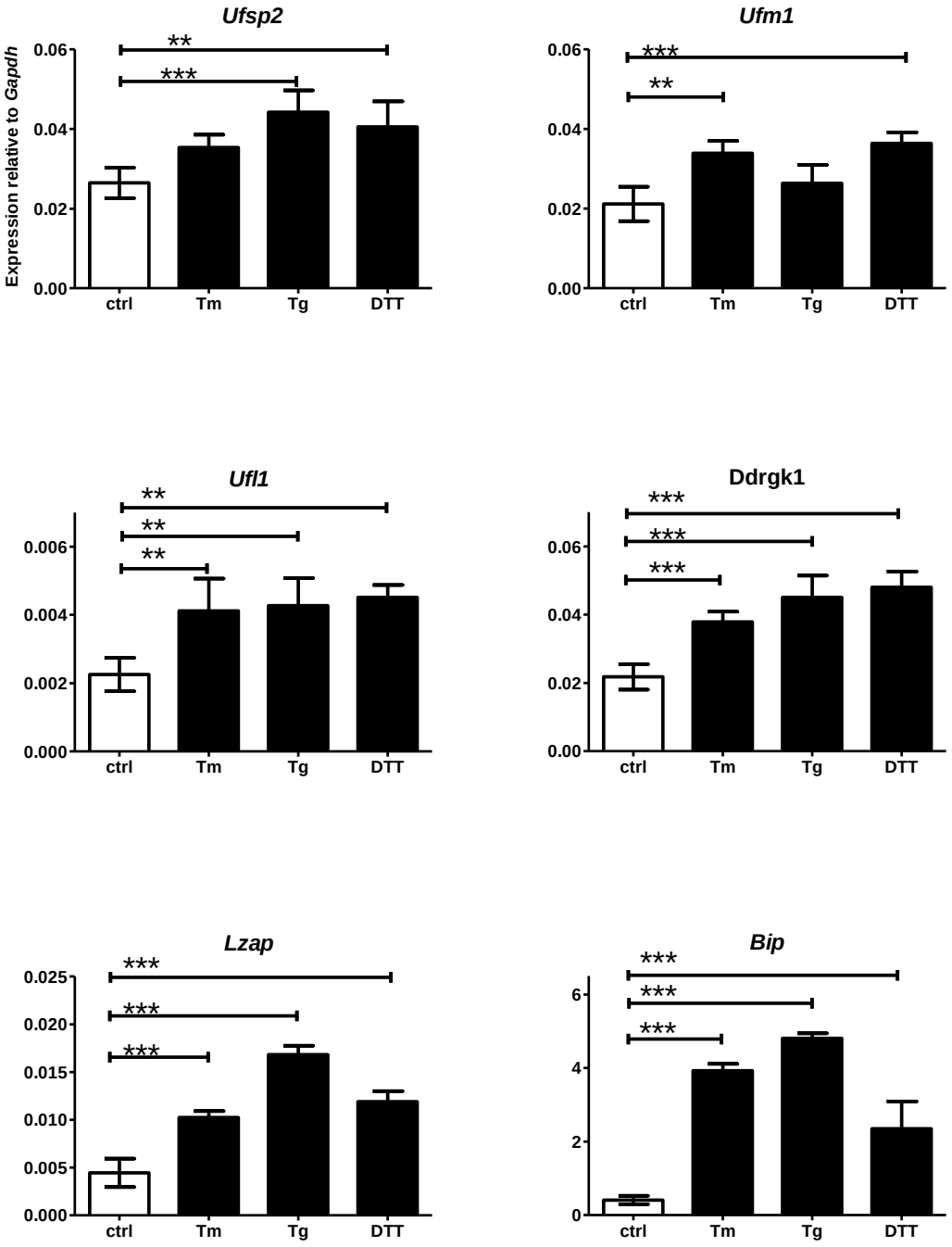


Figure 24. Expression of Ufm1 pathway in 2T3 cells following induction of ER stress with tunicamycin, thapsigargin and DTT.

The effect of ER stress on the expression of the genes of the Ufm1 pathway was tested in 2T3 cells with three chemicals: tunicamycin (Tm), thapsigargin (Tg) and dithiothreitol (DTT). ER stress was induced by treating the cell cultures with 2µg/ml tunicamycin for 14 hours, 2µM thapsigargin for 14 hours and 2mM DTT 6 hours. Untreated cultures served as controls (ctrl). Expression was quantified relative to *Gapdh*. Values are the mean +/- standard deviation of one experiment performed in triplicate. Statistical analysis was performed using the unpaired t test. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$.

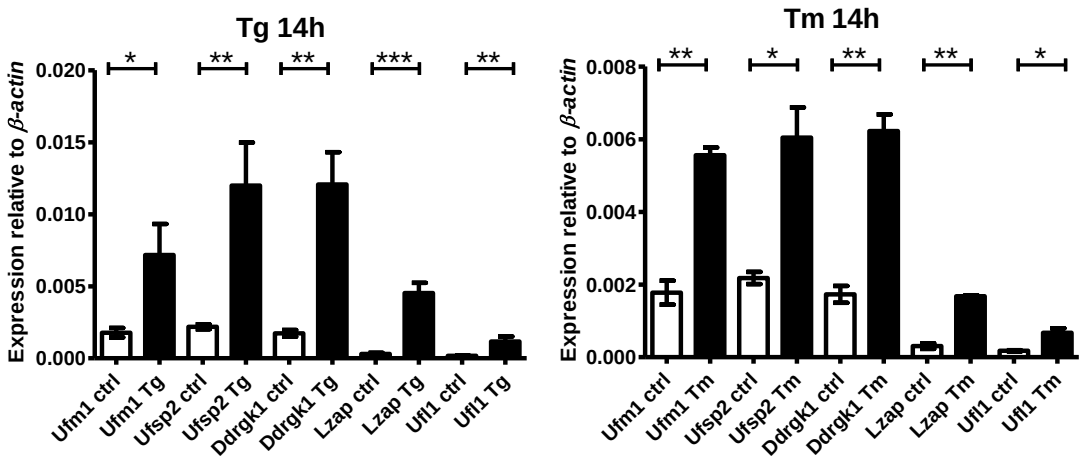


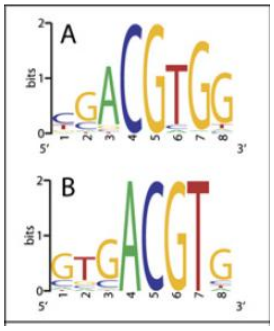
Figure 25. ER stress in 2T3 osteoblasts induced by thapsigargin and tunicamycin.

2T3 osteoblasts were treated with 2µM thapsigargin (Tg) and 2µg/ml tunicamycin for 14 hours. Untreated cultures served as controls (ctrl). Expression of each gene was quantified relative to β-actin using real-time PCR. Values are the mean +/- standard deviation of one experiment performed in triplicate. Statistical analysis was performed using the unpaired t test. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$.

3.2.5 Analysis of the promoter regions of the Ufm1 system genes

As genes encoding key components of the Ufm1 system had been found to be upregulated during ER stress, the upstream genomic sequences of these genes were analysed for the presence of known regulatory motifs. Upstream sequences of *Ufsp2, Ufm1, Uba5, Ufc1, Ufl1, Ddrk1, Lzap (Cdk5rap3)* from rat (*Rattus norvegicus*), mouse (*Mus musculus*), human and cow (*Bos taurus*) were obtained from the NCBI database. Upstream sequences of each gene from different organisms were aligned using ClustalW2 - Multiple Sequence Alignment available online at http://www.ebi.ac.uk/Tools/msa/clustalw2/. The alignments were screened for highly conserved regions and the conserved regions were analysed for the presence of known regulatory sequences. This analysis revealed the presence of the Unfolded Protein Response Element (UPRE) in the promoter regions of 4 of the 7 known Ufm1 system genes. The UPRE GACGTG sequence potentially bound by the spliced form of XBP1(S) [112] was found to be present and highly conserved in the promoter regions of *Ufm1, Uba5, Ufl1* and *Lzap* (Figure 26 and Supplementary Figures 2, 3, 4 and 5).

A



B

Ufm1

```
                                        -85       -80
                                         |         |
rat          ----------------GGAGGGCGACGTGGGCCGCGAAGGACCCCACTG----CCTGCCAT -56
mouse        ----------------GGAGGGCGACGTGGGCCGCGAAGGCCCCCGCTG----CCGGCCAT -52
human        AGGTCCCCAGATTGCAGAGGGAGACGTGGACGTGAGTGGAGCGGGGCGGTCCCCAGCACA -64
cow          AGGACCCCAGCTTGCGGAAGGAGACGTGGACCTGAGTTGGAAGAGACG----CTAGGGTG -58
                       ** ** ******* *       *        *    *  *
```

Uba5

```
                                                                -167      -162
                                                                  |         |
rat          CGGCGGCGCACCGGAAGTGGCGTGC-GGAAGTCCGGTGGGCATGCCGGAGACGTGTCTTA -148
mouse        CGGCGGCGCACCGGAAGTGGCACGCCGGAAGCCCGGTGGGCGTGCCGGAGACGTGTCTTA -157
human        CTGCGACGCACCGGAAGCGGCTCCGAGGAAGGCCTGTGGGAGTCTCGGAGACGTGTCTGT -156
cow          CAGCAACGCACCGGAAGTGGCCCGCTGGGAGACCCGCAGGAGCGGCGGAGACGTGTCGGG -160
             * **  *********** ***     ** ** ** *  **     ************
```

Ufl1

```
                                                      -43       -38
                                                       |         |
rat          CGGCCCGTCTCAGGGTGC--CAGTGGAGGGCTGAAGGCTGACGTGTCTGCCGTTTCCCGG -23
mouse        CGGTCCGTCCCGGGGTGC--CTGCGGCAGGCTGAAGGCTGACGTGTCTGCCGTTTCCCTG -23
human        CGGCCCGTTCCGCCTCTCTTCTCCCACCGCCTGTCGGCTGACGTGTCTGCAGTTCCTCCG -24
cow          CGGCCCGTTCCGCCTCGG--CTCCCGCGACCTTTGGGCTGACGTGTCTGCGGTTTGTCGG -23
             *** ****  *         *          **   *************** ***   * *
```

Lzap

```
                                  -112      -107
                                    |         |
rat          TCCGCTGAAGACGGCGACGTGGCTGAAGGAGGCCGGTTTGTTTCTGAATTGAGGCTCAGG -68
mouse        TCCGCTGAAGACGGCGACGTGGCTGAAGGAGGCCGGTTTGTTTCGGAACTGAGGCTCAAG -68
human        TGTACGGAAGGCGGCGACGTGGCCGAAGGATGCCCGTTTGTGTCTAAACGGAGGCTCGGC -73
cow          TCTGAGGAAGGCGGTGACGTGGCCAGAGAAGACCCATCTGTGCCTAACCCGAGGCTCAGC -72
             *     **** *** ********   ** *  **  * ***  *  *   *******
```

Figure 26. Identification of Unfolded Protein Response Elements (UPRE) in the promoter regions of *Ufm1, Uba5, Ufl1* and *Lzap* genes.

A. UPRE consensus sequence bound by the spliced form of *Xbp1* derived from chromatin immuno-precipitation studies by Acosta-Alvear et al (2007). Size of the letter symbolises frequency of occurrence of this particular base at this position of the consensus sequence **B.** Analysis of promoter regions of the Ufm1 system genes from four different species. The UPRE sequences are highlighted yellow and the numbers denote the position of the UPRE in respect to the ATG codon of the **mouse** gene. Asterisks denote conserved sequences. UPRE sequences were found in promoter regions of *Ufm1*, *Uba5*, *Ufl1* and *Lzap* genes.

To test whether the UPRE sequence played a role in the regulation of *Ufm1, Uba5, Ufl1* and *Lzap* under conditions of ER stress, approximately 1.5 kb of the mouse upstream promoter regions of the above genes were cloned into the pGL3-Basic luciferase vector (Promega, UK) (Supplementary Figure 1). The promoter sequences were amplified with primers including appropriate restriction enzyme cloning sites (Appendix 3). The *Ufl1* promoter sequence was not cloned as the sequence failed to amplify after multiple attempts possibly due to low complexity of the genomic region. Mutant versions of the promoter regions (where the UPRE sequence GACGTG was mutated to GTAATG) were generated by PCR with overlapping primers including the mutation and again cloned in to the pGL3-Basic luciferase vector. 2T3 cells were transfected with the promoter-pGL3 vectors and 24 hours post transfection ER stress was induced by addition of thapsigargin (2μM final concentration for 14 hours). The luciferase assay was performed according to the method described in 2.2.23.

Expression of the luciferase gene driven by the *Ufm1* promoter region was noticeably reduced following treatment with thapsigargin. Mutation of the UPRE site significantly decreased the promoter activity. Compared to the native sequence the expression of luciferase was reduced by 48%. Moreover, treatment with thapsigargin of cells transfected with the mutated promoter reduced the activity by 76% compared to the native untreated sequence (Figure 27).

Thapsigargin treatment of cells transfected with the *Uba5* and *Lzap* promoter luciferase vectors had the opposite effect on their activity compared to *Ufm1*. Following ER stress the expression of luciferase driven by *Uba5* and *Lzap* upstream regions increased by approximately 56% and 21%, respectively. When the UPRE sequence was mutated the luciferase signal markedly decreased. Mutation of the *Uba5* UPRE region reduced the luciferase signal by 66% but there was no significant change to the level of luciferase signal following thapsigargin induced ER stress. The mutation of the *Lzap* UPRE sequence decreased the luciferase signal by approximately 80% but again there was no significant change to the level of luciferase signal following thapsigargin induced ER stress (Figure 27).

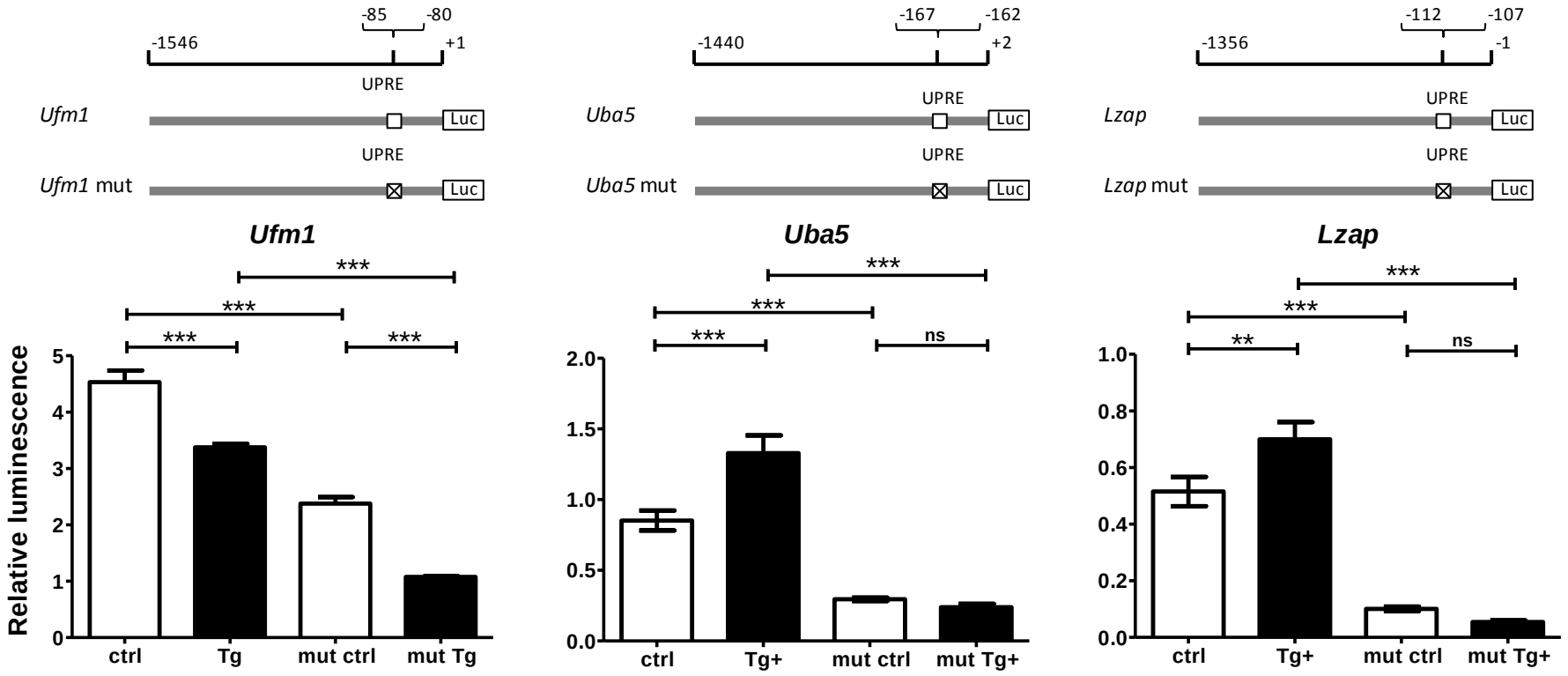


Figure 27. Luciferase gene reporter assay of promoter regions of *Ufm1, Uba5* and *Lzap*.

The ruler at the top shows the length of the genomic sequence and position of the UPRE in respect to the ATG codon. UPRE sequences (GACGTG) are marked with open squares and mutated UPRE sequences (GTAATG) are marked with crossed squares. Firefly luciferase expression driven by the native and mutated (mut) *Ufm1*, *Uba5* and *Lzap* promoter regions under normal (ctrl) and ER stress conditions induced with thapsigargin (Tg) was determined as a ratio of firefly to control Renilla luciferase luminescence. Values are the mean +/- standard deviation of one experiment performed in triplicate. Statistical analysis was performed using 1-way ANOVA with Bonferroni's post test. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$.

3.3 Discussion

As a first step towards understanding how the mutation in *Ufsp2* leads to the BHD phenotype, its expression was determined by radioactive mRNA *in situ* hybridisation against sections of embryonic and postnatal mouse hip joints. These experiments detected *Ufsp2* expression predominantly in the bone and ligaments of 10 day old mice (Figure 18). No *Ufsp2* expression was detected however in the developing hip joint or surrounding tissues in prenatal and newborn tissue sections (Figure 16 and Figure 17). Whether *Ufsp2* expression is absent in these tissues at early stages of development or whether its expression is below the limits of detection of the procedure cannot be distinguished by these experiments. However, it does appear that *Ufsp2* expression is higher at day 10 of mouse development than at earlier stages and is higher in the bone and the ligaments of the hip joint. A similar pattern of *Ufsp2* expression was found in the knee and vertebrae of 10 day old mice (Figure 19). In the knee, *Ufsp2* expression was detected in the bone but there was also some evidence of expression in the proliferative zone of cartilage where chondrocytes exhibit the highest expression of *Col2a1* (Figure 19). Interestingly, the secondary ossification centre was visible in the knee and *Ufsp2* expression was detected in the osteoblasts within the middle of that region but not in the chondrocytes of the surrounding ossifying cartilage zone. The vertebrae exhibited an interesting pattern of *Ufsp2* expression where most of the signal was localised to the bone and ligaments but also to the fibrocartilage surrounding the intervertebral disc overlapping a zone of *Col1a1* and *Col2a1* expression. For the most part, however, *Ufsp2* expression in the vertebrae colocalised with *Col1a1* expression.

From the described experiments and from the expression profile of *Ufsp2* reported in the BioGPS database (Supplementary Figure 7) *Ufsp2* expression is not limited to just the hip joint and so its expression pattern alone does not explain why the BHD phenotype is limited to the hip joint. However, from these experiments, *Ufsp2* expression does appear to increase postnatally within tissues of the hip joint and it is expressed during secondary centre ossification (as shown in the knee joint). In human development, the secondary centre of ossification of the proximal femoral epiphyses appears at ~8 months of age and ossification is not complete until 16 years of age. In contrast, in the knee joint, the secondary centre of ossification of the distal femur is present at birth and that of the proximal tibia forms around 1-3 months with ossification being complete by 13-15 years in girls and 15-18 years in boys [203-205]. Thus, early

ossification of the proximal femoral head coincides with the time that infants start to walk and hence the hip joint is load bearing. The impact of a delay or alteration in secondary centre ossification (which may be caused by Ufsp2 haploinsufficiency) may therefore have a greater effect on the hip joint than on the knee joint. Such effects of the Ufsp2 mutation can however only be explored fully in an *in vivo* model system (see Chapter 5).

For *in vitro* investigations, the *in vivo* expression studies did however indicate cell culture systems that may be appropriate for studies of the role of the Ufm1/Ufsp2 system. As *Ufsp2* expression observed in the hip joint co-localised with *Col1a1* expression in osteoblasts, patterns of expression of the components of the Ufm1 system were determined relative to markers of osteogenesis in the 2T3 cell line. The 2T3 cell line is an immortalised osteoblast cell line isolated from the calvarial bone of transgenic mice overexpressing the SV40-T antigen under control of the BMP-2 promoter [199]. Hence the cells are already committed to an osteoblastic lineage but can be stimulated to undergo mineralisation following addition of rhBMP-2 (Figure 20). During the time course (see Figure 22) following stimulation with rhBMP2, *Runx2* expression decreases, *Osterix* expression increases (to day 4) followed by increased expression of *Osteocalcin* (to day 6). This is an expression pattern characteristic of late osteogenic differentiation (see 1.5.2). Untreated cells are also known to express BMP-2 in culture [199] and so increase in *Osterix* followed by *Osteocalcin* occurs in untreated cells but at a later stage than that of rhBMP treated cultures (Figure 21). The expression of *Bip* was also determined in these experiments as it is upregulated during UPR (see 1.7.2). Moderate levels of ER stress are known to play an important role in osteogenic differentiation (see 1.7.3) and there is evidence that the Ufm1 pathway may be related to the ER stress that is induced during ischemic heart disease, myositis and type 2 diabetes (see 1.8.4 and [187, 188, 206]). As is evident in Figure 21, ER stress was indeed induced during the differentiation experiment with a significant increase in *Bip* expression at day 8. When comparing the expression of the components of the Ufm1 system as well as *Ddrgk1* (putative target) and *Lzap* (system associated gene) with the markers of osteoblast differentiation and mineralisation, there was no apparent correlation. An exception to this was a possible correlation between the expression pattern of *Ufsp2* and *Osteocalcin*. In contrast, all components of the Ufm1 system and *Ddrgk1* and *Lzap* were significantly upregulated relative to the control at day 8 which correlates with the expression of *Bip*.

Thus in this culture system it appears that the Ufm1 system may be related to ER stress as opposed to osteogenic differentiation.

The relationship between differentiation, ER stress and expression of the Ufm1 system was also explored during *in vitro* chondrogenesis using the C3H $10T^1/_2$ model system (Figure 22). In this system it was not possible to distinguish between whether expression correlated with differentiation (as measured by increased expression of *Col2a1*) or ER stress (as measured by increased expression of *Bip*) as *Col2a1* and *Bip* had the same expression patterns. Interestingly, however it did appear that during the 9 day *in vitro* chondrogenic differentiation assay changes in *Ufm1* expression closely followed that of both *Bip* and *Col2a1*. The remaining Ufm1 system genes were also upregulated but this upregulation occurred between days 6 to 9. In the chondrogenic assay it appears that increasing levels of *Col2a1* may cause ER stress and this in turn causes the increased expression of the components of the Ufm1 system. In contrast, in the osteoblast culture system expression of *Bip* did not correlate with expression of *Col1a1* expression and so, it was not clear what triggered ER stress in this culture system. The fact that *Bip* expression rapidly increased following prolonged high *Osteocalcin* expression (the most abundant non-collagenous bone ECM protein [207]) and appearance of highly mineralised cell matrix (Figure 20) may harbour a clue and would be interesting to investigate further.

Since in both the osteogenic and chondrogenic assays upregulation of the Ufm1 system was observed to coincide with elevated UPR the link with ER stress was investigated further. Three cell lines (2T3, C3H $10T^1/_2$, C2C12) were treated with DTT to induce ER stress. Following treatment of the 2T3 cell line with DTT all the Ufm1 system genes tested, including *Ufsp2, Ufm1, Ufl1, Ddrgk1* and *Lzap,* were significantly upregulated. In the C3H $10T^1/_2$ cell line all genes except *Ufm1* were upregulated and only *Ufl1, Ddrgk1* and *Lzap* were significantly upregulated in the C2C12 cell line. All three cell lines were treated with the same concentration of DTT but their response to ER stress in the form of *Bip* upregulation differed. Different cell lines have different sensitivity to chemical agents which may partly explain why all Ufm1 genes were upregulated in the 2T3 cell line and only some in the others. 2T3, the most responsive cell line, was chosen for the next experiment which tested whether chemicals affecting different pathways in the ER would elicit Ufm1 system response to a different degree. *Ufl1, Ddrgk1* and *Lzap,* were significantly upregulated in response to all three ER

stressors (tunicamycin, thapsigargin and DTT) whereas *Ufsp2* was not upregulated to a significant level by tunicamycin and *Ufm1* was not induced by thapsigargin. Because ER stress is known to influence the expression of many genes [112], to rule out the possibility that the treatment affected *Gapdh* expression, expression of the Ufm1 system following treatment with thapsigargin and tunicamycin was also measured using *β-actin* as a reference gene. The findings were similar to those using *Gapdh* as a reference gene but when quantified relative to *β-actin* all genes were significantly upregulated.

The outcome of the described experiments suggests that upregulation of the Ufm1 system is not related to a specific type of ER stress, be it in response to BMP-2 or chemical stressors, but is possibly part of a more general stress response. This is in line with experiments performed by Lemaire *et al.* (2011) [184] who showed upregulation of *Ufm1, Ddrgk1* and *Ufl1* in rat pancreatic islets insulinoma cells upon stimulation with cyclopiazonic acid and thapsigargin (both influencing the ER Ca^{2+} balance), and brefeldin A (an inhibitor of ER to Golgi transport). All three ER stressors upregulated the expression of Ufm1 system genes to a similar degree. In contrast, a more recent report by Zhang *et al.* (2012) [189] suggested that the Ufm1 system may be involved in ER to Golgi transport. In their experiments brefeldin A treatment of mouse embryonic fibroblasts (MEFs) showed the highest upregulation of *Ufm1, Uba5, Ufl1* and *Lzap* compared to thapsigargin and tunicamycin. In addition in their experiments, using *Xbp1*$^{-/-}$ MEFs, brefeldin A treatment failed to induce Ufm1 system gene expression. This suggested that expression of Ufm1 system genes may be controlled by XBP1.

To better understand the mechanism of Ufm1 system upregulation during ER stress, the promoter regions of Ufm1 system genes were screened for sequences conserved among species. The conserved sequences were then analysed for presence of known regulatory motifs. This analysis revealed UPRE sequences in the promoter regions of *Ufm1, Uba5, Ufl1* and *Lzap* that may be bound by XBP1 [112]. Luciferase gene reporter assay performed on ~1.5 kb mouse promoter regions of *Ufm1, Uba5* and *Lzap* showed that these fragments were sufficient to drive the expression of the *Luc* gene however only *Uba5* and *Lzap* were upregulated following treatment with thapsigargin. Thapsigargin treatment downregulated *Luc* gene expression driven by the *Ufm1* promoter region. Mutation of the UPRE sequence showed decrease in *Luc* gene expression driven by all three promoter regions. Moreover, thapsigargin treatment failed to induce upregulation of *Luc* gene expression driven by the mutated *Uba5* and *Lzap*

promoter regions. This signifies the importance of the GACGTG sequence in the promoter region for continuous stable expression as well as upregulation of these genes.

Interpretation of the *Ufm1* promoter assay may be more complex as in addition to the GACGTG sequence located at (-85) to (-80) upstream of the ATG codon which was included in the promoter construct assayed, another similar sequence exists downstream within the second exon at (+108) to (+113) (Figure 26 B and Supplementary Figure 2). This site was not included in the promoter construct and may be the site that is responsive to ER stress. This requires further investigation. Further, Zhang *et al.* (2012) [189] also reported existence of an XBP1 binding site at (-67) to (-54) in the human promoter of *UFM1*. However, the binding sequence that they identified in their paper (AGGGAGCCGTGGA) does not match any known XBP1 consensus binding sequence and this sequence does not exist in the human *UFM1* promoter sequence available in the NCBI database. The mouse *Ufm1* promoter UPRE sequence reported here located at (-85) to (-80) corresponds to the position (-101) to (-96) of the human *UFM1* promoter. In the human *UFM1* promoter this sequence is followed by a further GACGTG sequence (-95) to (-90) not present in other mammalian species. This sequence differs significantly from the putative XBP1 binding site published by Zhang *et al.* (2012) and is more likely to be the actual XBP1 binding site. The findings published by Zhang *et al.* (2012) are therefore questionable. In support of the data reported here is that the *Ufm1, Uba5, Ufl1* and *Lzap* promoter regions were among genomic fragments pulled down using XBP1 in a chromatin immuno-precipitation study published by Acosta-Alvear *et al.* (2007) [112]. The same study established GACGTG as one of the consensus sequences bound by XBP1. Proof that XBP1 or other bZip transcription factors binding similar sequences are responsible for upregulating these genes under ER stress condition however requires further investigation.

Many chondrodysplasias result from the inability of osteoblasts or chondrocytes to deal with increasing load in the ER, be it due to mutations in the ER machinery or mutations of ECM structural proteins resulting in their misfolding. Many of those with these disorders develop OA as a secondary consequence. In light of the increasing evidence for the importance of ER stress in skeletal development and the findings reported here it is feasible that the link between the *UFSP2* mutation and the BHD

phenotype may be related to impaired ER stress responses in the osteoblasts and/or chondrocytes of the developing hip joint.

Chapter 4. Ufm1 biochemistry

4.1 Introduction

As described in 1.8.4 our knowledge of the components of the Ufm1 system and our understanding of its function are still at an early stage. Of the components of the system, thus far the Uba5 (E1), Ufc1 (E2), Ufl1 (E3), Ufsp1 and 2 (proteases), Ddrgk1 (putative target) and Lzap (associated protein) have been identified and were found to have variable but not tissue specific patterns of expression. Of these, Ufl1, Ddrgk1 and Ufsp2 have been localised primarily to the ER membrane [178, 180, 181, 185]. This subcellular localisation is consistent with a role for the Ufm1 system in the UPR and in response to ER stress as outlined in 1.8.4 and explored further in Chapter 3. The precise role of the Ufm1 system and the effect of Ufm1 addition or removal from its protein targets have not been established. As has been the case in defining the role of the ubiquitin, SUMO and ubiquitin-like protein modification systems (see 1.8.1 and 1.8.2) the key to deciphering the role of the Ufm1 system is finding its modification targets. Experiments can then be designed to examine the effect of the attachment or removal of Ufm1 from these target proteins, specifically in investigating the link between the *UFSP2* mutation and the BHD phenotype, and how presence of mutated Ufsp2 may affect these processes. Only three studies so far have attempted identification of targets of Ufm1 modification in mammalian cells [181, 184, 186]. These studies elucidated interactions between different components of the system but only Ddrgk1 was identified as a potential Ufm1 target. This contrasts with the ubiquitin and SUMO systems which are known to modify a broad range of targets and that the modification of these targets may play different roles depending on the mode of attachment (see 1.8.1). Further, the identified targets have differed depending on the culture conditions and cell lines used. As the Ufm1 system is a ubiquitin-like system and studies to identify targets have been limited in number and scope, it cannot therefore be assumed that Ddrgk1 is its only target. In addition, to establish (i) the effect/s of Ufm1 modification of its targets, (ii) the downstream consequences of this modification and (iii) the effect of the *UFSP2* mutation on these processes, a suitable cell culture assay is required. For these reasons in this Chapter experiments are described to identify potential targets for Ufm1 modification using a range of approaches, purification and culture systems.

4.2 Results

4.2 Development of a purification method for isolation of Ufm1 conjugation targets

Isolation of Ufm1 conjugated targets was attempted first using an available Ufm1 expression construct, StrepUfm1 p201.iEP (Supplementary Figure 8). This construct was derived from the mouse Ufm1 cDNA sequence (for alignment of the human and mouse cDNA sequences of Ufm1 see Supplementary Figure 9) with the StrepTagII sequence added in frame to the 5' end of the Ufm1 sequence (minus the ATG start codon) in order to generate a fusion protein. This sequence was cloned into the p201-iEP vector in which transgene expression is driven by the constitutive CMV promoter (Supplementary Figure 10). In p201-iEP the transgene sequence is followed by an internal ribosome entry site (IRES) which allows translation of a GFP reporter and puromycin resistance gene (PAC) from the same mRNA strand. The vector can be used for transient transfection as well as lentiviral transduction and generation of stable cell lines by antibiotic selection (Methods 2.2.6, 2.2.7 and 2.2.8)

HEK293T cells were transfected with the StrepUfm1 p201.iEP and an RFP/GFP p201.iEP control vector using the calcium phosphate method (2.2.6). Transfection was confirmed based on the expression of the GFP reporter at 24 hours (Figure 28) and the cultures harvested 48 hours post transfection. Cells were lysed in Buffer W by freeze-thawing. Proteins from the cell lysate were then purified on Strep-Tactin resin followed by their separation by SDS PAGE and detection by western blotting using an anti-Ufm1 antibody (2.2.18). Expression of StrepUfm1 was detected as a single band of molecular weight approximately 10kDa, both in the lysate and in purified fractions. No higher molecular weight bands of Ufm1 conjugated targets were detected (Figure 29). The absence of conjugated proteins may have been caused by the lysis conditions not preserving the Ufm1-target protein bond as the StrepTagII system requires native lysis conditions to allow binding of the tag to the column. In these conditions activity of Ufm1 specific proteases may prevent isolation of conjugates similarly to early experiments with SUMO [208, 209]. Therefore a new strategy was devised for the isolation of conjugation targets using tandem affinity purification (TAP).

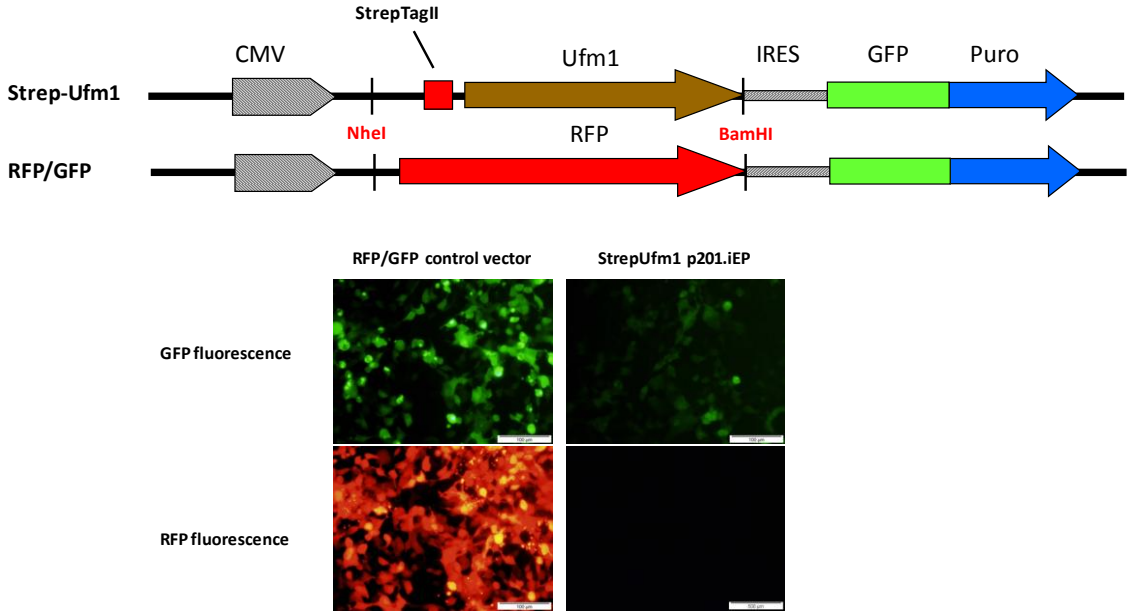


Figure 28. HEK293T cells transfected with StrepUfm1 p201.iEP and a control RFP/GFP p201.iEP vector.

The GFP fluorescence following transfection with StrepUfm1 p201.iEP confirmed expression of the transgene. Scale bar 100µm.

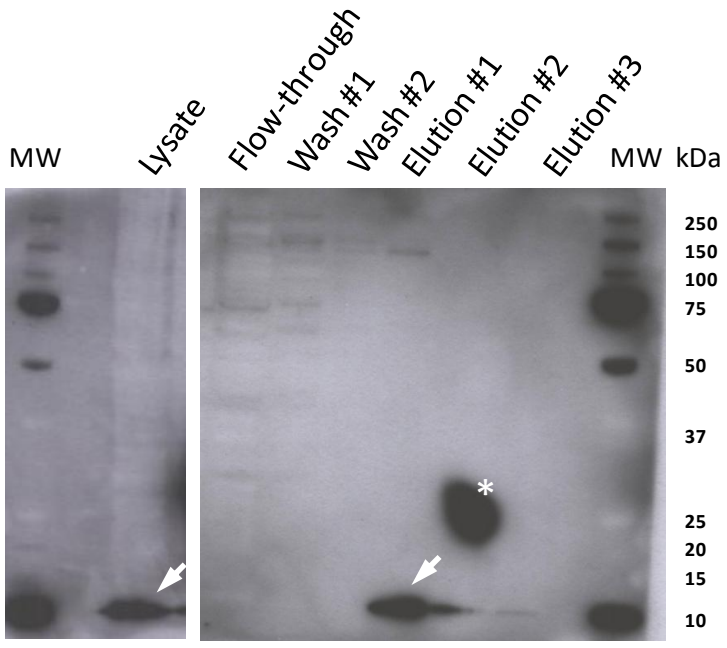


Figure 29. Western blot of proteins purified from cultures transfected with the StrepUfm1 construct.

Following transient transfection of the HEK293T cell line with StrepUfm1 p201.iEP, proteins were purified using Strep-Tactin resin and detected by the αUfm1 antibody. Arrows show unconjugated StrepUfm1. * - an artefact on the nitrocellulose membrane. MW – protein molecular weight marker.

The TAP strategy involved isolation of Ufm1 conjugated proteins via a His-tag (as Ni-NTA resin used for His-tag purification can tolerate denaturing and reducing lysis conditions) followed by additional purification via StrepTagII (the StrepTagII system has higher specificity than the His-tag system). For this purpose a new construct, Ufm1ΔC2, was generated using the cloning strategy outlined in Supplementary Figure 11. The Ufm1ΔC2 construct (Figure 30) had the last 2 C-terminal amino acids of Ufm1 (serine and cysteine) deleted thereby exposing the active glycine residue as in the activated form of Ufm1 (see 1.9.4). The His-tag was introduced between the StrepTagII and the Ufm1 sequence, and the construct was cloned into the p201.iEP vector as before. This activated form of Ufm1 should attach to target proteins in the same way as native Ufm1.

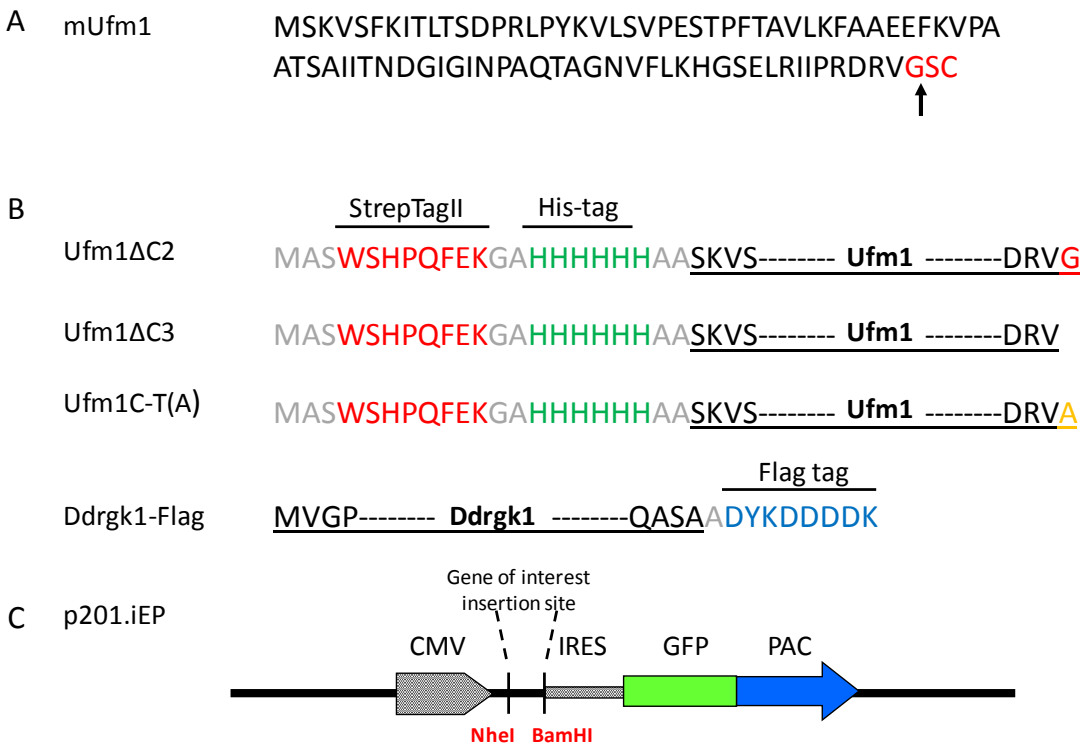


Figure 30. Sequences of Ufm1 expression constructs.

A. Amino acid sequence of the mUfm1 precursor. The arrow indicates the glycine which is exposed prior to Ufm1 attachment via this residue to target proteins. The two remaining amino acids are cleaved off during activation of Ufm1 by the Ufm1-specific proteases. **B.** The sequence of the three Ufm1 expression constructs (including the StrepTagII and His-tag sequences at the N-terminus and altered amino acid sequences at the C-terminus) and the sequence of Ddrgk1 (including the C-terminal Flag tag). **C.** Basic features of the p201.iEP expression vector used for overexpression of the Ufm1 variants. CMV – cytomegalovirus immediate-early promoter, *Nhe*I-*Bam*HI restriction sites used for cloning the Ufm1 variants, IRES – internal rybosome entry site, GFP – green fluorescent protein reporter gene, PAC - Puromycin N-acetyl-tranferase (antibiotic resistance gene)

A T75 flask of subconfluent HEK293T cells was transiently transfected with the Ufm1ΔC2 construct as previously described. A parallel transfection with the RFP/GFP vector served as a control. Cells were then lysed with 6M guanidine buffer and proteins purified on the Ni-NTA resin only (see 2.2.18). Western blotting of the purified sample using the anti-Ufm1 antibody showed a high concentration of proteins purified from the Ufm1ΔC2 transfected cells in the range of 10-50 kDa that were not present in the RFP/GFP transfected cells. Individual protein bands could not be distinguished due to the high signal (Figure 31). To aid further purification of potential Ufm1 conjugated proteins, the whole TAP procedure including purification on the Strep-Tactin resin, was tested. Cell cultures were transfected, lysed and the lysate purified on Ni-NTA resin as previously. The eluate was then applied to the Strep-Tactin column and purified (2.2.18). Samples from specific steps during the procedure were analysed by western blotting (Figure 32 A). The second purification step eliminated all the contamination visible in the His-tag purified sample. The unconjugated Ufm1 (~10 kDa) is clearly visible in both purification steps but the additional bands originating from potential Ufm1 conjugates following Ni-NTA purification are not visible following Strep-Tactin purification. This may be due to dilution of the sample as it was eluted in 4 steps. This experiment was repeated with cells from 10 T75 flasks and the final eluate was concentrated on a spin column, resolved on SDS-PAGE gel and stained with the Colloidal Blue Staining Kit (2.2.24) (Figure 32 B). The gel showed only the unconjugated Ufm1 of approximately 10 kDa similarly to the diluted Strep-Tactin fractions. This suggests that the proportion of modified proteins to free Ufm1 is very low and a large number of trasnfected cells would be needed to yield the concentration of conjugated proteins required for mass spectrometry analysis.

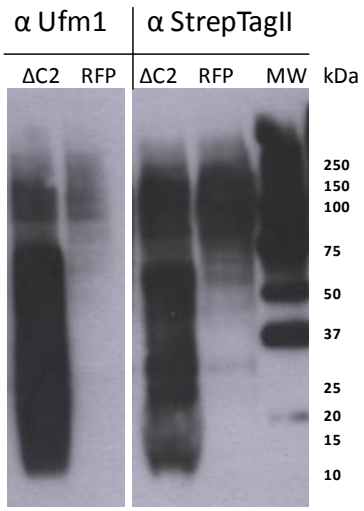


Figure 31. Western blot of proteins purified from cultures transfected with the Ufm1ΔC2 or RFP/GFP vectors.

Proteins were purified from transiently transfected HEK293T cells using Ni-NTA resin and detected using either the αStrepTagII or αUfm1 antibody. Cells transfected with the RFP/GFP vector served as a ontrol. MW – protein molecular weight marker.

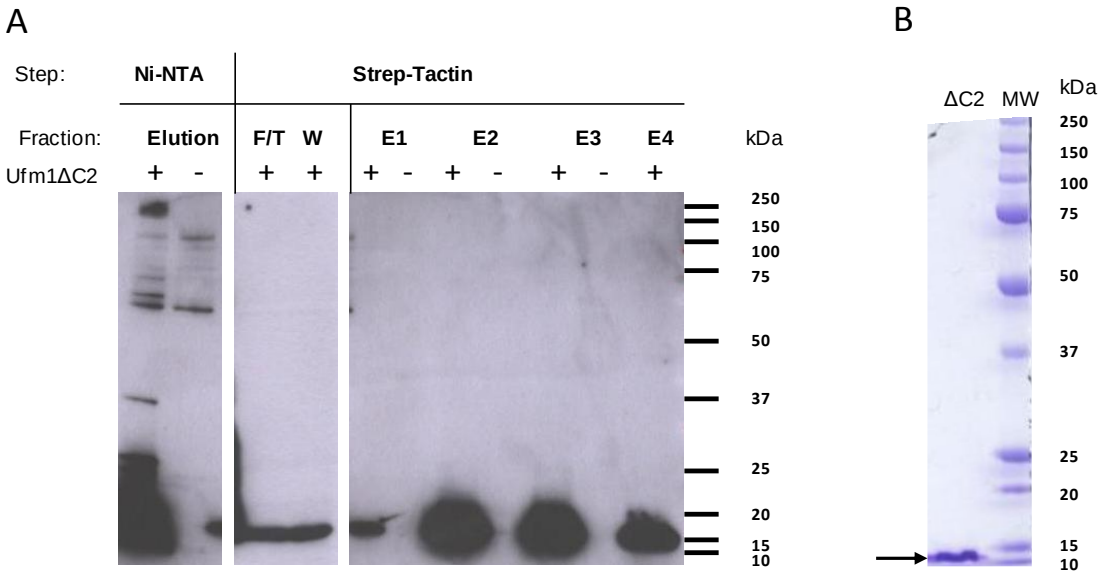


Figure 32. Western blot and SDS-PAGE of samples taken from the TAP procedure.

A. Subconfluent T75 flask of HEK293T cells was transfected with Ufm1ΔC2 (+). T75 flask transfected with the RFP/GFP vector served as a negative control (-). Proteins purified from the cultures were detected by western blotting using the anti-Ufm1 antibody. F/T – flow through, W- wash fraction, E1-E4 elution fractions **B.** Ten T75 flasks were transfected with Ufm1ΔC2. Proteins obtained by TAP were resolved by SDS-PAGE and the gel was stained with Colloidal Blue. Arrow - Ufm1ΔC2

4.2.3 Large scale purification and mass spectrometry analysis of Ufm1 conjugated proteins

Large scale transient transfection is costly and inefficient therefore a stably transduced cell line was established using the Ufm1ΔC2 vector in a lentiviral system. Lentiviral particles were generated as described in 2.2.8. HEK293T cells were transduced and selected using Puromycin. Details of the selection procedure are available in Supplementary Figures 12 and 13. The stably transduced cell line was tested for the expression of Ufm1 using the TAP procedure described previously. Figure 33 shows the first purification step (Ni-NTA) and elutions from the second step (Strep-Tactin). Two additional Ufm1 conjugated bands are visible in the final samples but the intensity of these diminishes with consecutive elutions confirming the low ratio of conjugates to free Ufm1.

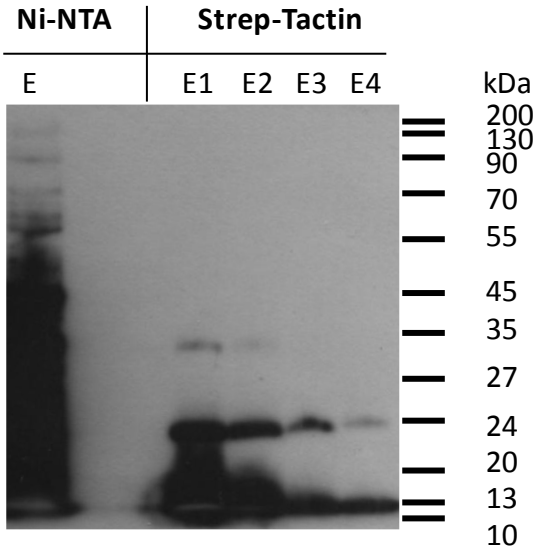


Figure 33. Western blot of TAP fractions from stably transduced cells.

HEK293T cells were stably transduced with the Ufm1ΔC2 expression vector. E – elution from Ni-NTA resin; E1 - E4 - successive elutions from the Strep-Tactin resin.

The stably transduced Ufm1ΔC2 cell line was expanded. For large scale purification, the cells were cultured in 90 150 cm^2 Petri dishes in the presence of puromycin until near confluency and harvested with 6M Gu-HCl buffer. Proteins from the cell lysate were purified by TAP. A small sample of the purified material was tested by western blotting (Figure 34 B) and the rest resolved on an SDS-PAGE gel and

stained with Colloidal Blue (Figure 34 C). Five Bands corresponding to Bands detected by western blot were excised from the gel and the proteins in the Bands were analysed by mass spectrometry (Methods 2.2.26). Of these Bands, Bands 1, 3 and 5 appeared to be absent in the negative control as shown Figure 34 A. The evidence that Bands 2 and 4 contained Ufm1 conjugated proteins was less clear but these bands were clearly visible in the preparative gel (Figure 34 C) and so their analysis was pursued. A further Band at ~95 kDa was a potential candidate from the gel in Figure 34 A but was not pursued as its concentration was low.

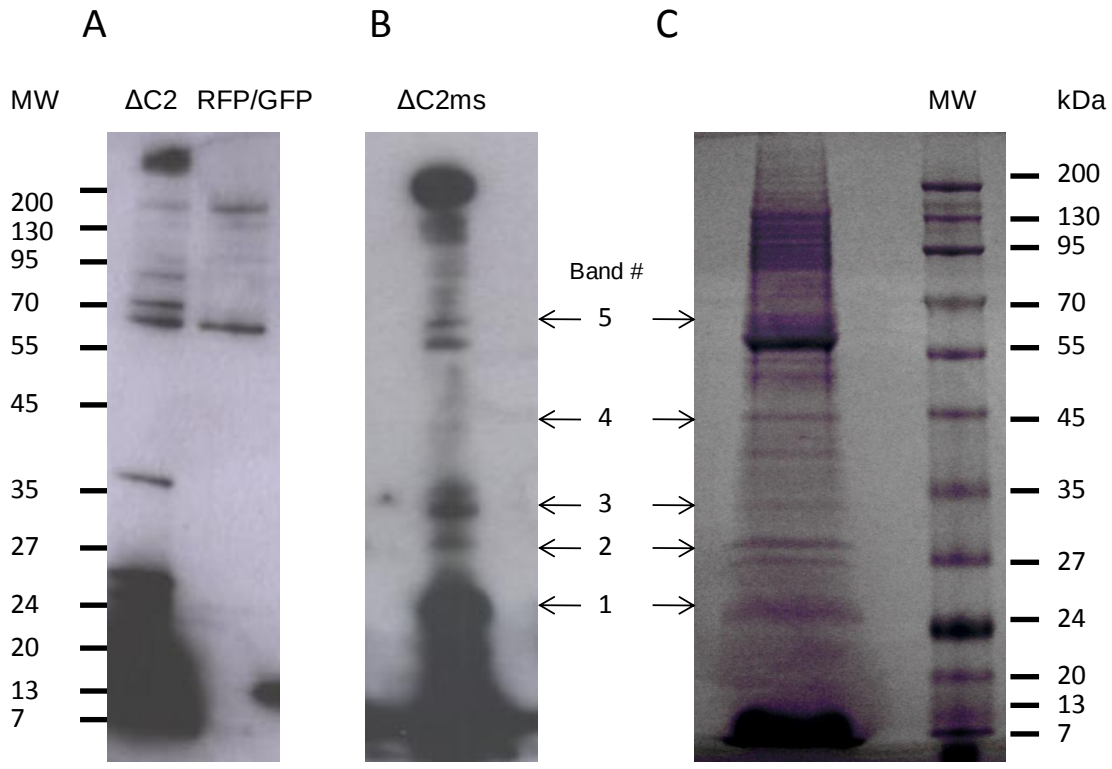


Figure 34. Preparation of Ufm1 conjugated proteins for mass spectrometry.

A. Western blot of Ni-NTA purified proteins as in Figure 32 **B.** Western blot of a 1:100 sample (ΔC2ms) of the large scale purification of Ufm1ΔC2 conjugated proteins detected by anti-Ufm1 antibody. On the left a western blot with negative control from Figure 32 is shown as a comparison **C.** Colloidal Blue stained SDS-PAGE preparative gel of the large scale TAP. Arrows indicate the protein bands cut out of the gel for identification by mass spectrometry and their corresponding bands on the western blot. MW – molecular weight marker.

The proteins identified in Bands 1-5 are shown in Table 7. In interpreting this information the molecular weights (MW) of the proteins were considered. Proteins conjugated with Ufm1 would be expected to have a MW of ~10 kDa greater than their unconjugated counterpart (i.e. the MW of the protein plus 10 kDa should approximate to the MW of the Band). In addition, proteins containing multiple histidine residues close together might have been enriched on the Ni-NTA column and not completely removed in the second purification step and so be contaminants. No proteins were identified in Band 1 (approximately 25 kDa). In Band 2 (approximately 27 kDa) peptides from 4 different proteins were identified. Two of these proteins (Nos. 2 & 19) had a similar MW to Band 2. No. 9 was keratin (with a high MW and therefore considered to be contamination) and No. 8 was identified as Ufm1 from 1 peptide. In Band 3, peptides from two proteins (Nos. 8 & 12) were identified: Ufm1 and Ufc1. Ufc1 is the E2 conjugating enzyme of the Ufm1 system (see 1.9.4) which has a MW of 19 kDa but was found in Band 3 excised at approximately 32 kDa. The 10 kDa MW shift is consistent with Ufc1 conjugation with Ufm1 as is the presence of Ufm1 peptides within Band 3. In Band 4, 6 proteins were found. Three of these proteins (Nos. 5, 9 & 10) were forms of keratin and one was actin (No. 6) (all of which are common contaminants when purifying proteins for mass spectrometry), one (No. 16) had the MW corresponding to that of Band 4 and one was Ufm1 (No. 8). In Band 5 (cut from region of 60kDa MW), 13 proteins were identified of which 3 were keratin (Nos. 5, 9 & 10), 3 were histidine rich (Nos. 4, 15 & 18) and many had MWs similar to that of Band 5 (Nos. 3, 11, 13, 14 & 17). One of the identified proteins (No. 7) in Band 5 was Uba5 the E1 enzyme of the system which is 45 kDa suggesting it must have been co-purified due to interaction with Ufm1. The MW of protein No. 1 (54 kDa) is consistent with it being conjugated with Ufm1. However it is a less likely candidate than Uba5 as being the Ufm1 conjugated protein represented in Band 5. Nevertheless, it is possible that there is more than one conjugated protein within Band 5.

The identification of Ufm1 conjugation to Uba5 and Ufc1 demonstrated that activated Ufm1 synthesised following expression of the Ufm1ΔC2 construct is appropriately attached to key components of the Ufm1 modification cascade. Furthermore, this finding demonstrated that Ufm1 conjugated proteins can be purified from HEK293T cells using the TAP procedure.

No	Identified Proteins (19)	Accession Number	Molecular Weight	Band 1 ~25 kDa	Band 2 ~27 kDa	Band 3 ~32 kDa	Band 4 45 kDa	Band 5 ~60 kDa	Comments
1	Non-POU domain-containing octamer-binding protein	IPI00304596 (+1)	54 kDa					10	Potential binding partner as found in band cut at mw ~65
2	Carbonic anhydrase 2	IPI00218414	29 kDa		6				
3	heterogeneous nuclear ribonucleoprotein L isoform a	IPI00027834	64 kDa					8	
4	SF3A2 protein (Fragment)	IPI00017341	51 kDa					6	Histidine rich
5	Keratin, type II cytoskeletal 1	IPI00220327	66 kDa				3	6	
6	Actin, cytoplasmic 1	IPI00021439 (+4)	42 kDa				6		
7	Ubiquitin-activating enzyme E1 domain-containing protein 1	IPI00015736	45 kDa					5	Uba5 (E1)
8	Ubiquitin-fold modifier 1 precursor	IPI00010207 (+1)	9 kDa		1	2	2		Ufm1
9	Keratin, type I cytoskeletal 10	IPI00009865	60 kDa		1		3	1	
10	Keratin, type I cytoskeletal 9	IPI00019359	62 kDa				2	3	
11	Isoform 1 of WD repeat-containing protein 1	IPI00746165 (+1)	66 kDa					5	
12	Ufm1-conjugating enzyme 1	IPI00294495	19 kDa			2			Ufc1 (E2)
13	Catalase	IPI00465436	60 kDa					4	
14	Isoform M2 of Pyruvate kinase isozymes M1/M2	IPI00479186 (+1)	58 kDa					4	
15	Transcriptional repressor protein YY1	IPI00014513 (+1)	45 kDa					3	10 and 3 His stretch
16	Fructose-bisphosphate aldolase A	IPI00465439 (+1)	39 kDa				3		
17	insulin-like growth factor 2 mRNA binding protein 2 isoform a	IPI00179713 (+5)	66 kDa					3	
18	E3 ubiquitin-protein ligase Hakai	IPI00290429 (+2)	55 kDa					3	Histidine rich
19	Platelet-activating factor acetylhydrolase IB subunit gamma	IPI00014808	26 kDa		2				

- Ufm1 conjugate - Possible Ufm1 conjugate - contaminant

Table 7. Proteins identified by mass spectrometry. Numbers in the columns for Bands 1-5 represent the number of unique peptides identified.

4.2.4 Testing of chemical agents potentially inducing conjugation of Ufm1 in the HEK293T cell line

Although Ufm1 conjugation to Uba5 and Ufc1 was identified (Figure 34 and Table 7) no new protein targets for Ufm1 were isolated from the HEK293T cells. A reason for this may be that, as outlined in Chapter 3, the system may only be fully activated under specific stress conditions. To investigate the possibility of a stimulus activating the Ufm1 system in the HEK293T cell line, four culture conditions known to induce cellular stress were tested: addition of H_2O_2 (200 µM) [210], DTT (2 mM), N-ethylmaleimide (20 mM) and heat shock (42°C) [211]. In these experiments stably transfected HEK293T cell lines were treated for four hours in the case of chemical agents and for two hours plus two hours of recovery in the case of heat shock. For these experiments, a second expression construct, Ufm1ΔC3 was generated (see Supplementary Figure 11), where the last three C-terminal amino acids were deleted from Ufm1 (see Figure 30 B). As the C-terminal glycine residue that is essential for the conjugation of Ufm1 to its target proteins was not present in the encoded protein, experiments with this construct served as a negative control. The stable cell line HEK293T Ufm1ΔC3 was established as described for HEK293T Ufm1ΔC2 (see 4.2.3) and one 150 cm^2 dish of each cell line per each condition were treated with the agents as described previously. The cells were harvested in denaturing and reducing conditions, proteins purified on Ni-NTA resin and conjugates detected by western blotting with the Ufm1 antibody as described in 4.2.3. As shown in Figure 35 A, a major band was detected at a MW of approximately 30 kDa from the Ufm1ΔC2 cell line that was not present in the samples from the negative control cell line, Ufm1ΔC3, when the western blot was exposed for 1 minute. A further two bands at the MW of ~70 and ~95 kDa were detected from the Ufm1ΔC2 cell line when the western blot was exposed for 10 minutes (Figure 35 B). It is likely that the ~30 kDa band is Ufc1 (corresponding to Band 3 and as identified by mass spectrometry, see 4.2.3) and that the ~70 kDa band is Uba5 (corresponding to Band 5 and as identified by mass spectrometry, see 4.2.3). The ~95 kDa band was as observed previously (see Figure 34 A) but its concentration precluded identification by mass spectrometry and so its identity remains unidentified. None of the treatment conditions induced the appearance of bands that were not in the untreated control and the relative amounts of bound to unbound Ufm1 were very low. As in Figure 35 A, H_2O_2 and heat shock did not appear to affect the intensity of the conjugated

proteins. DTT treatment appeared to slightly reduce the intensity of the ~30 kDa band, possibly due to the cell death that was observed following treatment. NEM treatment reduced the intensity of the ~30 kDa band. This may be due to observed cell death but also because NEM covalently binds cysteines [212]. A cysteine residue forms a part of the active site of the Ufc1 therefore NEM can block the Ufm1 binding (lane 9). The band in lane 9 is also visibly shifted up.

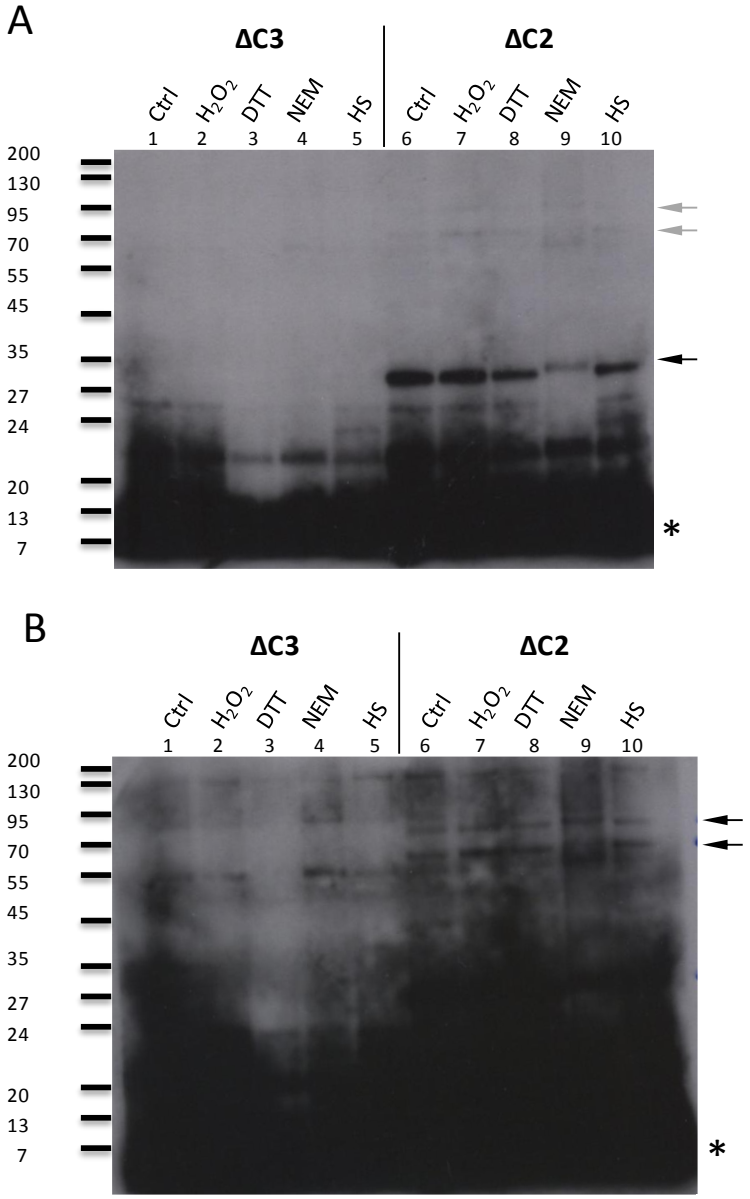


Figure 35. Treatment of stably transduced HEK293T cells under various stress conditions.

A. HEK293T cells stably transduced with Ufm1ΔC2 (ΔC2) and Ufm1ΔC3 (ΔC3) expression vectors were treated with H_2O_2, DTT, N-ethylmaleimide (NEM) and subjected to heat shock (HS). Untreated cells were used as control (Ctrl) The proteins were isolated by purification on an Ni-NTA resin and detected by western blotting with the anti-Ufm1 antibody. WB exposed for one minute **B.** The same western blot exposed for ten minutes. MW – molecular weight, * unconjugated Ufm1, ← - Ufm1 conjugate

4.2.5 Expression of the Ufm1 system in response to ER stress in HEK293T cells

Experiments in Chapter 3 showed upregulation of the Ufm1 system gene expression in response to ER stress in the 2T3 osteoblast cell line but as described in 4.2.4 DTT treatment of the HEK293T cell line did not increase Ufm1 conjugation. HEK293T response to DTT induced ER stress was therefore measured on the gene expression level. HEK293T cell cultures were induced with 2 mM DTT and mRNA was extracted after 2 and 4 hours (after 6 hours of treatment extensive cell death was observed therefore this time point was not included in the analysis) and cDNA was prepared as described in 2.2.16 and 2.2.17. The expression of *BIP* and Ufm1 system genes *UFSP2, UFM1, UFL1, DDRGK1* and *LZAP* was quantified using real-time PCR as described in 2.2.22.

DTT treatment induced ER stress in the cell line as verified by elevated expression of *BIP*. Expression of *UFL1* and *LZAP* was significantly upregulated after 4 hours of treatment ($p<0.01$ and $p<0.001$, respectively) but expression of *UFSP2, UFM1* and *DDRGK1*did not change significantly compared to the untreated control (Figure 36).

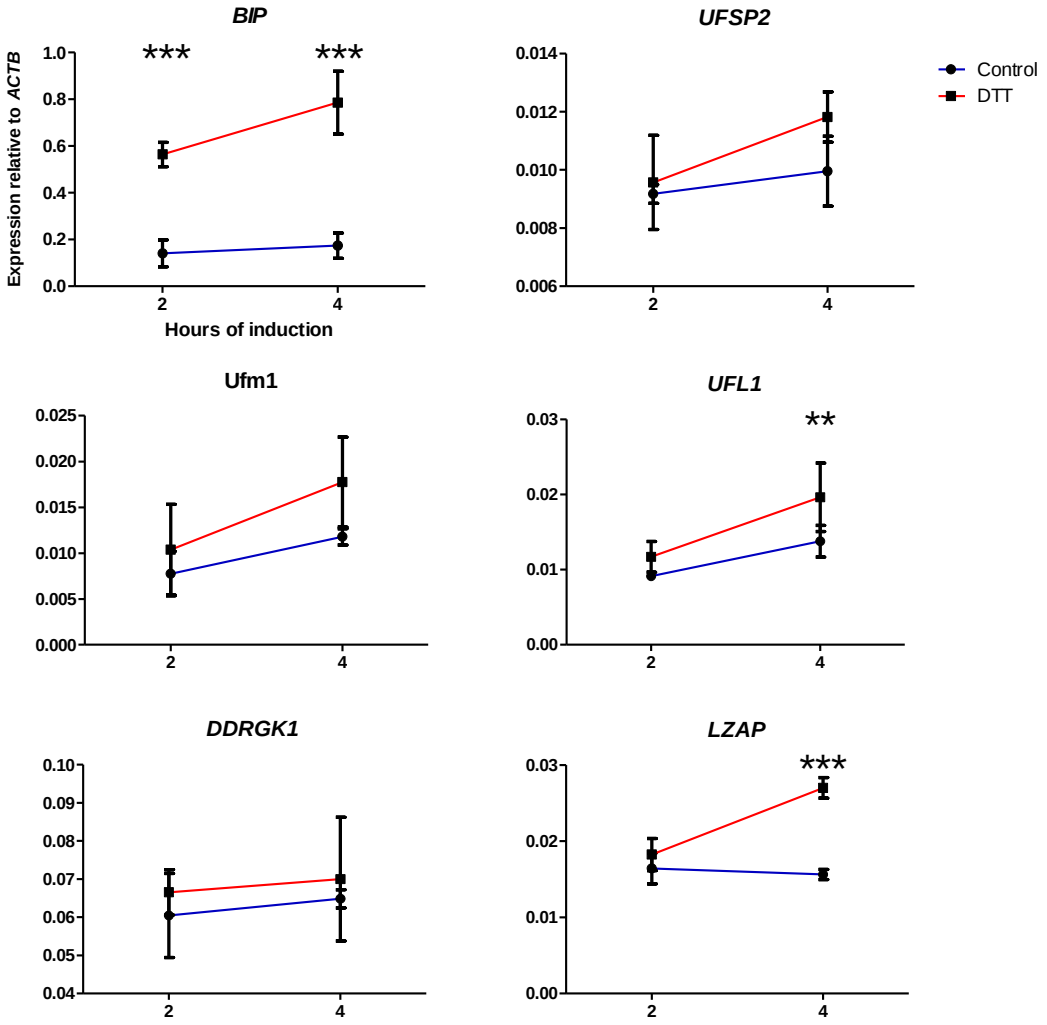


Figure 36. Ufm1 system gene expression in HEK293T cells following DTT treatment.

The effect of ER stress on the expression of the genes of the Ufm1 pathway was tested in the HEK293T cell line. ER stress was induced by treating the cell cultures with 2mM DTT for 2 and 4 hours. Untreated cultures served as controls. Expression was quantified relative to *ACTB* (β-actin) Values are the mean +/- standard deviation of one experiment performed in triplicate. Statistical analysis was performed using the unpaired t test. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$.

4.2.6 Comparison of the Ufm1 conjugated proteins in HEK293T cells versus the 2T3 osteoblast cell line

As described in 4.2.3 and 4.2.4, no new protein targets for Ufm1 were identified in the HEK293T cell line despite using stably transduced cell lines, activated Ufm1 and subjecting the cells to stress conditions. Reasons for this may include that i) the percentage of the target that is modified at any given time is very low, making it difficult to isolate conjugated targets in quantities needed for mass spectrometry, ii) that the system may not be fully active in the HEK293T cell line and iii), as indicated in Chapter 3, that the system may be activated under specific conditions such as ER stress and more so in specific cell lines than in others.

To address (i) and increase the potential of identifying target proteins, a new Ufm1 construct was generated (Figure 30) and included in further experiments. The Ufm1C-T(A) vector was designed to express a form of Ufm1 that, once attached to its target protein, cannot be removed. Ufm1C-T(A) was generated by replacing the C-terminal glycine of Ufm1ΔC2 with alanine. According to published studies [181] this modification should allow conjugation of Ufm1 to its target proteins in the usual way but the resulting bond would not be recognised by the Ufm1-specific proteases thereby blocking deconjugation. It was predicted that expression of Ufm1C-T(A) would lead to an enriched pool of Ufm1 modified target proteins and increase the chance of their identification by mass spectrometry. The above form of Ufm1 was tagged at its N-terminal end with both StrepTagII and a His-tag and the expression construct was cloned into the lentiviral expression vector, p201-iEP as before. For the amino acid sequence and for details on the cloning procedure used to generate this construct, see Figure 30 and Supplementary Figure 11.

To address (ii) and (iii) above, the 2T3 osteoblast cell line was also included in subsequent experiments. As shown in Chapter 3, genes of the Ufm1 system are activated during osteogenic differentiation of 2T3 cells stimulated by BMP-2 as well as when the cells are subject to ER stress conditions.

The Ufm1C-T(A), Ufm1ΔC2 and Ufm1ΔC3 expression constructs were transiently transfected into the HEK293T and 2T3 cell lines in parallel. Aliquots of the transfected cells were lysed with 1x Laemmli buffer without any further purification, samples were resolved on an SDS-PAGE gel and visualised by western blotting. As shown in Figure 37, the level of transfection and/or expression of the constructs was greater in the HEK293T cell line than in 2T3. In addition, the Ufm1C-T(A) construct in

the HEK293T cells produced a band not visible previously in neat lysate samples. Simultaneously, the majority of the culture was lysed in Gu-HCl buffer and the proteins were purified on Ni-NTA resin (the second Strep-Tactin purification step was not performed in routine experiments to reduce time and reagent usage) and visualised by western blotting (Figure 38). Again the majority of the purified protein could be seen as unconjugated Ufm1 at the bottom of the western blot for all three vectors While in Figure 37 expression of the constructs was visibly higher in the HEK293T cell line, after purification the ratio between the conjugated Ufm1 signal and free Ufm1 was shown to be higher in the 2T3 cells (Figure 38). This increased ratio is favourable for isolation of conjugates and subsequent identification of Ufm1 targets by mass spectrometry. In the 2T3 cell line, bands were detected in the Ufm1ΔC2 and Ufm1C-T(A) samples but not in Ufm1ΔC3 (negative control) at the MWs of approximately 27, 30, 40, 45, and 60 kDa (indicated by arrows in Figure 38). Although the same bands were present in both Ufm1ΔC2 and Ufm1C-T(A) lanes, the intensity of the bands differed between the samples. The bands at the MW of approximately 30 and 60 kDa most likely correspond to the bands identified previously as Ufc1 and Uba5 respectively. The bands of an estimated MW of 27, 40 and 45 kDa were not detected previously in stably transduced cells. It is possible that in the established cell line used in the mass spectrometry experiment the prolonged overexpression of Ufm1ΔC2 may have influenced the balance of the pathway and reduced conjugation of Ufm1 to its targets. It was not clear whether the increased intensity of the 30 and 60 kDa bands was due to overall improvement of conjugation or due to slower processing and accumulation of Ufm1C-T(A) as enzyme bound form because intensity of the new band at 40 kDa did not seem to increase. Some of the bands appeared as doublets (at 27 and 30 kDa and at 40 and 45 kDa) which could be that different isoforms of the same protein have been purified or that they indeed represent distinct proteins.

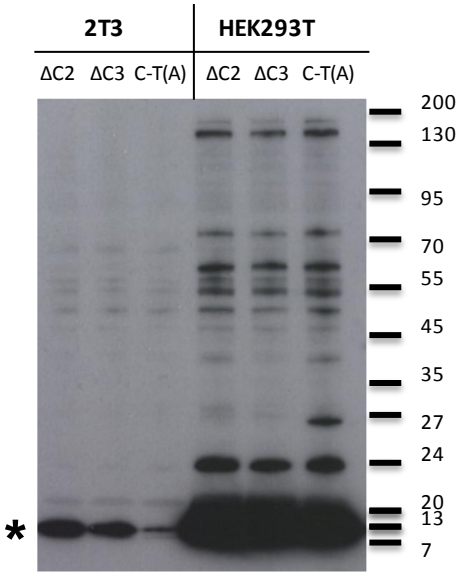


Figure 37. Comparison of the level of expression of Ufm1 in 2T3 osteoblasts and HEK293T cells.

The cells were transfected with Ufm1ΔC2, Ufm1ΔC3 and Ufm1C-T(A) expression vectors. Cells were lysed in SDS-PAGE sample buffer and the lysate analysed by western blotting without prior purification. Proteins were detected using the anti-Ufm1 antibody. * unconjugated Ufm1.

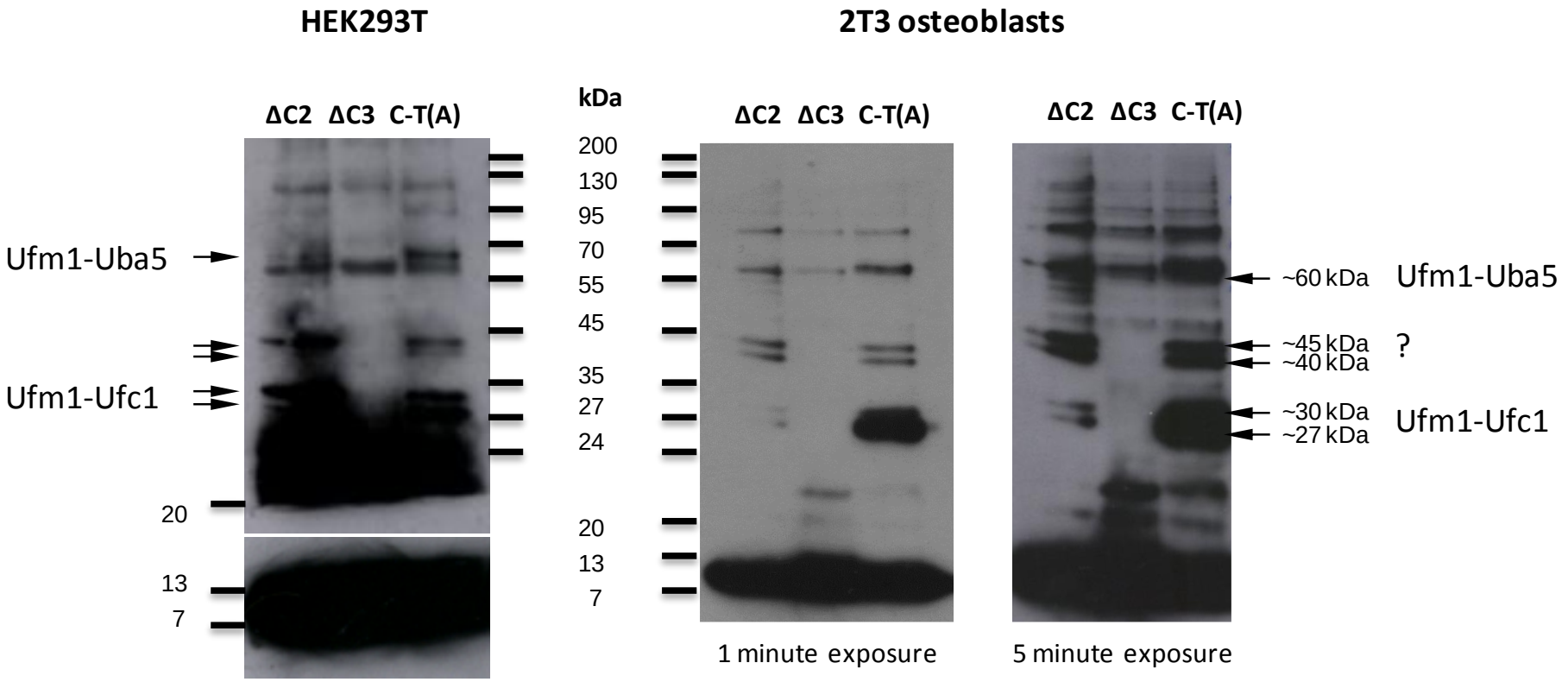


Figure 38. Comparison of proteins purified from 2T3 and HEK293T cells.

2T3 osteoblasts and HEK293T cells were transfected with Ufm1ΔC2, Ufm1ΔC3 and Ufm1C-T(A) expression vectors. The proteins were isolated by purification on an Ni-NTA resin and detected by western blotting with the anti-Ufm1 antibody. Arrows point to bands not present in the negative (ΔC3) control.

4.2.7 Effect of ER stress on Ufm1 conjugation

In Chapter 3 it was shown that the expression of the Ufm1 system is upregulated during osteogenic differentiation and during ER stress in 2T3 cells. It was therefore explored whether ER stress could lead to increased conjugation of Ufm1 to its protein targets in the 2T3 cell line.

2T3 cells were transfected as previously with Ufm1ΔC2, Ufm1ΔC3 and Ufm1C-T(A) expression vectors. ER stress was induced with thapsigargin 32 hours post transfection and the cells were harvested 48 hours post transfection (16 hours of induction). Proteins from cell cultures were purified and detected as previously described. The results are shown in Figure 39. In the cells transfected with Ufm1ΔC2 there is a visible difference between uninduced and induced cells with the latter having more intensive bands specific to Ufm1 conjugates. In contrast, induction of ER stress had the opposite effect on conjugation or stability of conjugated proteins in the case of Ufm1C-T(A). The bands at 27, 40 and 45 kDa are less intense following induction of ER stress and the band at 55kDa (the Ufm1-Uba5 conjugate) is absent.

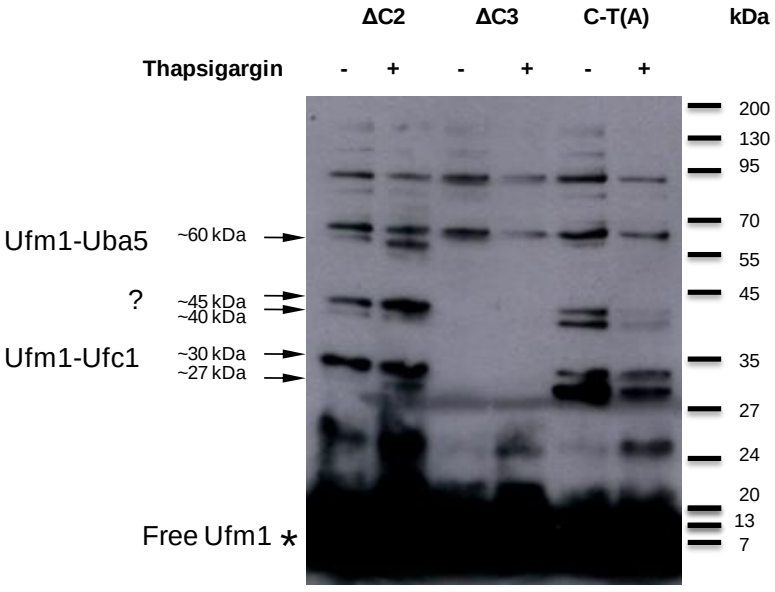


Figure 39. Induction of ER stress in 2T3 cells.

2T3 osteoblasts were transfected with 3 tagged variants of Ufm1 (ΔC2 – active, ΔC3 – inactive and C-T(A) – non-deconjugateble). The proteins were isolated isolated by purification on an Ni-NTA resin from non stressed cells (-) and cells induced with thapsigargin for 14h (+) and detected by the anti-Ufm1 antibody. Putative Uba5 and Ufc1 bands are marked → as is the possible conjugation target.

4.2.8 Analysis of Ufm1 conjugation to its putative target Ddrgk1

As discussed in Chapter 1.9.4, a potential Ufm1 conjugation target, DDRGK domain containing 1 (Ddrgk1 or Ufbp1), has been identified. To determine whether Ddrgk1 was conjugated by Ufm1 in the 2T3 cell line and whether any of the putative targets for Ufm1 conjugation identified in 4.2.3 – 4.2.7 represented Ddrgk1, the following experiments were carried out. For these experiments a cDNA clone of mouse *Ddrgk1* was obtained from Source Bioscience and the gene was amplified using primers designed to add an in-frame Flag-tag sequence at the C-terminus. The N-terminus was avoided because of potential localisation signal sequence that may be removed during maturation of the protein [184]. The construct was subsequently cloned into the p201-iEP expression vector as described previously. A diagrammatic representation of the construct, Ddrgk1-Flag, is shown in Figure 30. Expression of Ddrgk1-Flag was confirmed by transfection of the vector into 2T3 cells followed by western blot analysis of the unpurified lysate with the anti-Flag antibody (Figure 40 A) with untransfected cells serving as a control. The anti-Flag antibody detected a single band at the MW of approximately 40 kDa which is consistent with the predicted size of the Ddrgk1 protein (36 kDa) plus addition of the Flag-tag sequence (1 kDa). Subsequently, Ddrgk1 was co-transfected into 2T3 cells with the Ufm1ΔC2, Ufm1ΔC3 and Ufm1C-T(A) expression vectors. The transfected cells were lysed and proteins were purified by the His-tag present on the Ufm1 (Ni-NTA resin). The purified material was then detected using the anti-Flag antibody which should detect only overexpressed Ddrgk1. The same transfection mixtures but without Ddrgk1 were used on another set of cultures and proteins were purified in the same way as previously but the material was detected with the anti-Ufm1 antibody as a control for expression and the purification procedure (Figure 40 B). The anti-Flag antibody detected a distinct band at a MW of approximately 45 kDa in lanes with proteins purified from Ufm1ΔC2 and Ufm1C-T(A) transfected cultures but no band was present in the inactive Ufm1ΔC3 transfected sample (Figure 40 B lanes marked αFlag). The purified Ddrgk1 band appeared to correspond in size to the unidentified 45 kDa band detected by the Ufm1 antibody in samples purified from Ufm1 transfected cultures in the absence of Ddrgk1 overexpression. The same band increased in intensity in ER stressed cells transfected by Ufm1ΔC2 vector but was not present in Ufm1C-T(A) transfected cells following ER stress (Figure 39).

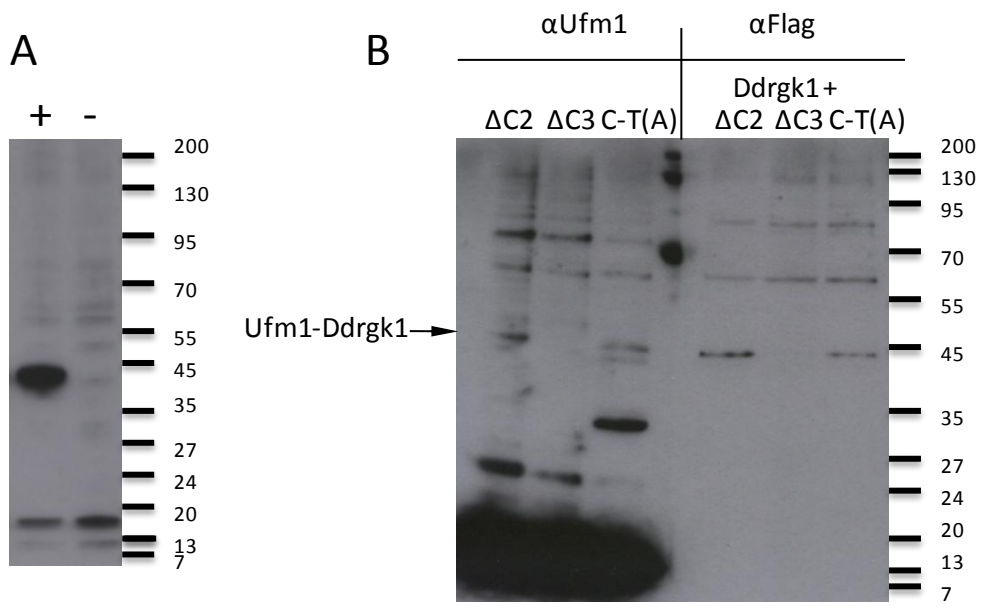


Figure 40. Co-purification of overexpressed Ddrgk1 using Ufm1 as bait.

A. 2T3 cells transfected with the Ddrgk1 expression vector (+) and untransfected cells (-) were lysed and subjected to western blot with anti-Flag antibody **B.** 2T3 cells were transfected with 3 tagged variants of Ufm1 only (left), or co-transfected with Flag-tagged Ddrgk1 (right). Proteins were isolated from the cell lysate using His-tag present on Ufm1 and detected by anti-Ufm1 antibody (left) or anti-Flag antibody (right) present on Ddrgk1.

4.3 Discussion

There is a wide variety of methods for purification and identification of targets of ubiquitin-like protein modification. One of the most popular methods is overexpression of a tagged version of the ubiquitin-like protein which is conjugated inside the cell using the same cellular machinery as the native one. Our first attempt at isolation of Ufm1 conjugation targets involved expression of Strep tagged Ufm1 (Figure 28) and purification of tagged products from HEK293T cells. The StrepTagII purification system is based on the affinity of the synthetic octapeptide WSHPQFEK Strep-tag to Strep-Tactin, which is a modified streptavidin [213]. The specificity of binding is very high but since the binding of the StrepTagII to Strep-Tactin is similar to a substrate being bound by an enzyme it depends on the Strep-Tactin being in its native conformation which precludes the use of denaturing and reducing agents. Theoretically this would allow study of the action of the Ufsp2 protease and its mutated form on the isolated targets. However this approach was not successful in producing candidate conjugation targets and only the free StrepUfm1 could be detected (Figure 29). It is possible that the native lysis conditions required by the Strep-Tactin to bind the

StrepTagII did not preserve the Ufm1-target bond even in the presence of protease inhibitors. For this reason a 6 histidine tag (His-tag) was incorporated into the constructs. The His-tag purification system relies on the affinity of histidine to nickel ions which are immobilised on a sepharose resin (the Ni-NTA resin) [214]. His-tag purification allows stringent lysis conditions including high concentration of denaturing agents such as urea and guanidine hydrochloride as well as reducing agents such as β-mercaptoethanol. This ensures purification of only covalently linked proteins. The downside of the use of the His-tag system however is that it is less specific as many eukaryotic proteins possess multiple histidine stretches and so are co-purified. This problem is eliminated if His-tag purification is followed by purification with the use of a second tag such as the StrepTagII. Combining both tags in TAP system allows the use of harsh lysis conditions in the first purification step involving the His-tag. After the change of the buffer to a milder one histidine rich contaminants are eliminated by using the StrepTagII system. Proteins purified using this method can then be analysed by mass spectrometry with a high chance of identification.

Initially, the Ni-NTA purification step was tested using lysates from HEK293T cells overexpresing the Ufm1ΔC2 vector (Figure 30). This vector was designed to circumvent the activation step of the Ufm1 precursor by the Ufm1 specific proteases as the two carboxy-terminal amino acids were deleted from the expressed Ufm1 thereby exposing the active glycine for entry into the E1-E2-E3 conjugation pathway. The Ni-NTA purification produced Ufm1 bound proteins in the range of 10-60 kDa that could be detected by the Ufm1 antibody (Figure 31). However, when the full TAP procedure was tested, the final elutions showed only the ~10 kDa free StrepHisUfm1ΔC2 band possibly due to the higher dilution factor compared to the His-tag fraction (Figure 32 A). Scaling up the experiment by tenfold led to the visualisation of free Ufm1ΔC2 on a Colloidal Blue stained SDS PAGE gel but there was no evidence of Ufm1 conjugated proteins. In this system, the free Ufm1 appeared to be in great abundance compared to any potential conjugated forms. To allow large scale purification, a stably transduced Ufm1ΔC2 cell line was established. Prior to scale up, lysate from the Ufm1ΔC2 cell line was subjected to the TAP procedure conjugation products were detected (Figure 33). The cultures and purification procedure were then scaled up and bands from the SDS PAGE gel suspected to harbour Ufm1 bound proteins were analysed by mass spectrometry. Three proteins (Ufm1, Uba5 and Ufc1) were identified confirming that the Ufm1 system is expressed in the HEK293T cell line and that Ufm1ΔC2 is correctly

processed through the Ufm1 conjugation system (Figure 13). However, no target proteins were found and Ddrgk1 was not identified. The lack of identification of Ufm1 modification targets in the cell line may be due to the Ufm1 system (i) not being fully active in this cell line and/or (ii) not being fully active under normal physiological conditions.

This possibility was investigated by treating the stable cell line used in the mass spectrometry experiment with various stress agents like H_2O_2 (oxidative stress), DTT (ER stress), N-ethylmaleimide (inhibitor of cysteine proteases) and heat shock (heat shock response). None of the conditions resulted in appearance of new Ufm1 bound bands on the western blot (Figure 35). However, a ~30 kDa band corresponding to that identified as Ufc1 by mass spectrometry diminished upon treatment with N-ethylmaleimide. Ufc1 possesses a cysteine residue in the active site that forms a thioester bond with the carboxy-terminal glycine of Ufm1. N-ethylmaleimide may covalently bind the cysteine of Ufc1 blocking the active site which would explain the reduced intensity of the band and the increased MW. Human Ufc1 has three cysteines therefore small portion of Ufc1 molecules could bind Ufm1 before being inactivated by N-ethylmaleimide but have the other cysteines modified resulting in the appearance of a higher MW band than that of untreated cells. This result supports the mass spectrometry analysis that the ~30 kDa band is Ufm1 bound Ufc1.

The absence of putative Ufm1 protein targets on western blots from the HEK293T cell lysates obtained using different purification methods and conditions, and lack of upregulation of *UFSP2* or *UFM1* in response to ER stress (Figure 36), suggested that this cell line may not be the optimum for further experimentation. Based on the results presented in Chapter 3 showing that the Ufm1 system was expressed in the 2T3 osteoblast cell line and upregulated following ER stress, this cell line was used for further biochemical studies. Ufm1 conjugation in the 2T3 osteoblasts and HEK293T cells was compared using the Ufm1ΔC2 (the activated form), Ufm1ΔC3 (the unconjugatable form) and Ufm1C-T(A) (the non-deconjugatable form) vectors (see Figure 30 for a diagrammatic representation of these vectors and 4.2.6 for their description). Although the HEK293T cells expressed more Ufm1 (Figure 37), the 2T3 cell line showed a much higher proportion of Ufc1 and Uba5 bound Ufm1 to free Ufm1 (Figure 38). Additionally, possible conjugation targets in the form of a double band at the MW of ~40-45 kDa and an extra band at the MW of ~27-30 kDa were visible that

were not present in previous experiments. These bands were only faintly visible in the HEK293T cell line.

Experiments described in Chapter 3 showed upregulation of the Ufm1 system in response to ER stress in the 2T3 cell line on the mRNA level. In 4.2.7 the effect of ER stress was investigated on Ufm1 target conjugation. Treatment of the 2T3 cell line with thapsigargin resulted in increased intensity of the Ufm1 antibody reactive bands purified from cells transfected with the Ufm1ΔC2 vector (Figure 39). This effect is contrary to the reports by Lemaire *et al.* (2011) [184] where a reduction in Ufm1 conjugation was observed in response to thapsigargin and an increase in response to cycloheximide (an inhibitor of translation). In contrast, Zhang *et al.* (2012) [189] reported a slight increase in Ufm1 conjugation in response to ER stress induced by thapsigargin. It is noteworthy that in their studies they observed similar western blot band patterns to that of Figure 39, without evidence of high MW bands as have been found in the case of other ubiquitin-like proteins. Interestingly, the thapsigargin treatment had an opposite effect on the 27, 40, 45 and 60 kDa bands purified from cells transfected with the Ufm1C-T(A) vector as their intensity decreased. Although we still do not understand the dynamics or the effects of Ufm1 conjugation it does appear that only a low level of conjugated protein targets exist under normal physiological condition and that these levels appear to increase upon induction of ER stress. Zhang *et al.* (2012) [189] have however suggested that the Ufm1 system may be involved in vesicle trafficking. One could hypothesize that the conjugation of Ufm1 to the target is very dynamic with only low number of molecules conjugated at one time and prolonged modification, as in case of the irreversible binding of Ufm1C-T(A) vector, could lead to degradation of the target protein under stress conditions.

So far only Ddrgk1 is known to be conjugated by Ufm1. It is a 315 amino acid protein with a predicted molecular weight of 36 kDa which upon Ufm1 conjugation would increase in MW to approximately 45 kDa, which approximates to the size of the unidentified 45 kDa band in Figures 38 and 39. This possibility was examined by co-expression of the Ufm1 vectors and Flag-tagged Ddrgk1 in the 2T3 cell line. The proteins were purified using the His-tag present on the Ufm1 proteins and detected by the anti-Flag-tag antibody (Figure 40). A single band at the MW of ~45 kDa was detected in cells co-transfected with Ddrgk1 and Ufm1ΔC2 or Ufm1C-T(A) whereas no band was detected in cells co-transfected with Ddrgk1 and the Ufm1ΔC3. This suggests

that Ddrgk1 may be the band observed in previous experiments at the molecular weight of ~45 kDa.

In conclusion, the TAP method has been developed and validated for the purification of Ufm1 conjugation targets. The method was successfully tested in HEK293T cells resulting in purification and identification by mass spectrometry of Uba5-Ufm1 and Ufc1-Ufm1 conjugates. However, no new conjugation targets were identified. Investigation of the Ufm1 system in the 2T3 osteoblast cell line revealed three new unidentified bands together at the molecular weight of ~27-30 and ~40-45 kDa the intensity of which increased following treatment with thapsigargin in cells transfected with Ufm1ΔC2 but decreased in cells transfected with the non-deconjugateble Ufm1C-T(A). Co-expression of Ddrgk1, the only known Ufm1 conjugation target, and tagged Ufm1 proteins allowed co-purification of Ddrgk1 as a Ufm1 conjugate and its detection at the MW of ~45 kDa. This suggests that one of the previously detected bands may be the Ddrgk1-Ufm1 conjugate. The experiments in this Chapter suggest that, for large scale purification of Ufm1 protein targets, the 2T3 osteoblast cell line transfected with the Ufm1ΔC2 vector could be used and the proteins could be purified following induction of ER stress to increase the conjugation rate.

Chapter 5. Generation and characterisation of *Ufsp2*^{BHD} transgenic mouse lines

5.1 Introduction

Mouse models have proven to be a powerful tool for the studying of human genetic diseases as 99% of the genes present in humans have orthologues in mice and numerous gene mutations causing pathologies in humans have been found to produce similar phenotypes in mice (for detailed information see http://www.informatics.jax.org/). Mouse models developed for heritable conditions have been pivotal to the understanding of the pathogenic basis of these conditions as well as in developing and testing of new drug candidates [215] (also see exemplars discussed in1.8.3).

Several different methods exist for generation of mouse models. Early research used radiation and chemically induced mutagenesis for generation of random mutations followed by the screening of mouse lines for phenotypes related to certain diseases. Random mutagenesis however introduced many unknowns in the experiments and this led to the development and adoption of more targeted or regulated approaches that are now widely used. Methods include generation of transgenic mice through:

- random integration of gene copies under the control of constitutive, inducible or tissue specific promoters
- gene knock-outs utilising homologous recombination to completely inactivate the gene of interest
- gene knock-ins utilising homologous recombination to introduce a particular mutation allowing studies of its effect on the organism conditional/inducible gene targeting through induced translocation or large deletions/inversions using *lox*P sites and Cre recombinase system [216].

Although targeted knock-out and knock-in methods constitute a more precise way to study the effects of a mutation their generation is costly and time consuming. Generation of transgenic mice where the mutated gene is randomly inserted into the mouse genome offers a quicker and less expensive alternative provided the mutation is predicted to have a gain of function or dominant negative effect.

As BHD is an autosomal dominant disorder, there are three potential mechanisms whereby the mutation in the *UFSP2* gene could elicit an effect. One of the mechanisms may be haploinsufficiency which could lead to lower rate of Ufm1 activation and/or removal from its target proteins and may influence processes that require rapid Ufm1 turnover. Haploinsufficiency for a protease is however a rare event. There is evidence however that haploinsufficiency of proteolytic components of other ubiquitin pathways influences their function. Such an example is the *Fat facets* gene involved in *Drosophila* eye formation and its mammalian homologue *Fam* involved in TGFβ signalling. The *faf* gene encodes a deubiquitinating enzyme that prevents proteasomal degradation of its targets by trimming the polyubiquitin chains in *Drosophila* or removes inhibitory monoubiquitination of Smad4 and allows TGFβ signalling in mammals. Haploinsufficiency of the enzyme results in higher degradation rate of the targets and disturbs formation of the *Drosophila* eye [217-219]. Previous studies have also examined the deconjugation of SUMO from its target proteins by members of the family of SUMO-specific proteases (SENPs) and have found that these deSUMOylation enzymes are involved in the regulation of SUMO-dependent pathways by balancing the SUMOylation status of target proteins. Alterations in the expression levels of specific SENPs has led to either decreased levels or accumulation of specific SUMO-conjugated proteins with a range of phenotypic outcomes including tumorigenesis and metastasis [220-223]. The mutation in the *UFSP2* gene may also result in a dominant negative effect if, for example, the mutated protein could be misfolded and induce ER stress similar to that caused by mutated ECM components (see Chapter 1.5). This is however unlikely as the level of Ufsp2 protein synthesis is far lower than that of the ECM components (see Figures 18 and 19) and as the study of Ha *et al.* (2011) showed that the overall structure of the mutated Ufsp2 was undisturbed but that the spatial organisation of the catalytic centre of the enzyme was affected causing its inactivation [194]. A further possibility is that the presence of the mutated protein could prevent cleavage of the target by binding to the target and blocking the access for the active non-mutated proteases. Of particular relevance in terms of the potential BHD mutation mechanism is that a catalytically inactive form of the SUMO-specific protease, SMT31P1 causes the accumulation of SUMO conjugated proteins in a dominant negative manner [224]. Whether the *UFSP2* mutation has dominant negative effect can be tested by generating a transgenic mouse line that expressed the mouse *Ufsp2* gene with the equivalent BHD mutation under the regulation of the constitutively active

CMV promoter and examine the phenotype of the mice generated. To test whether levels of expression of *Ufsp2* may have a phenotypic effect, overexpression of the native protease could address the question of an excess of Ufm1 processing (assuming that Ufsp2 activity is not regulated by additional factors) or serve as a negative control for the mutated *Ufsp2*. For these reasons expression constructs were prepared that contained either the native (WT) or mutated (BHD) forms of the mouse *Ufsp2* gene and which included a Flag-tag (to allow the transgenic Ufsp2 to be distinguished from the endogenous Ufsp2) for generation of transgenic mice by pronuclear injection. The generation of these mice and investigations to determine whether the resultant mice have a hip joint phenotype are described in this Chapter.

5.2 Results

5.2.1 Cloning of the *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag expression constructs

The cloning procedure to generate the *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag expression constructs is diagrammatically represented in Figure 41. The *mUfsp2*^{WT} and *mUfsp2*^{BHD} cDNA sequences were amplified by PCR using cDNA clones previously prepared by a former PhD student, Christopher Watson, as templates [179]. Both of the primers used for amplification incorporated restriction sites to aid subsequent cloning (Figure 41 A, Supplementary Figure 17). In addition, the reverse primer caused deletion of the endogenous termination codon, contained the Flag-tag sequence followed by a termination codon which, during amplification, were added in frame to the 3' end of the *mUfsp2* $^{WT/BHD}$ cDNA (Supplementary Figure 14). The Flag-tag was incorporated to allow the synthesis of the transgenic Ufsp2 protein to be distinguished from that of the endogenous Ufsp2. The PCR products (predicted size of 1418 bp) were resolved by agarose gel electrophoresis, excised from the gel, purified and subsequently ligated into the pGEM T-Easy vector (Figure 42 A and B; the pGEM-T Easy restriction map is presented in Supplementary Figure 15). Confirmation of the sequence of both the *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag (as shown in Supplementary Figure 16) was obtained by sequence analysis. The *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag cDNAs were subcloned into pCI-neo expression vector making use of the restriction sites incorporated by the primers (Figure 41 C and Figure 42 C). The pCI-neo vector contains a CMV promoter to drive expression of the inserted transgene and a polyadenylation

site 3' to the inserted transgene. Additionally, the expression of the transgene is enhanced by the presence of a chimeric intron (see Supplementary Figure 17). Expression of the *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag cDNAs was tested by transfecting the respective pCI-neo clones into the 2T3 mouse osteoblast cell line. Cell lysates were collected 2 days post transfection and transgene expression was detected by western blotting using anti-Ufsp2 and anti-Flag antibodies. The anti-Flag antibody detected a single protein at the expected molecular weight of approximately 60 kDa for both mUfsp2^{WT}-Flag and mUfsp2^{BHD}-Flag. The anti-Ufsp2 antibody detected the same band as well as two further background bands (Figure 43).

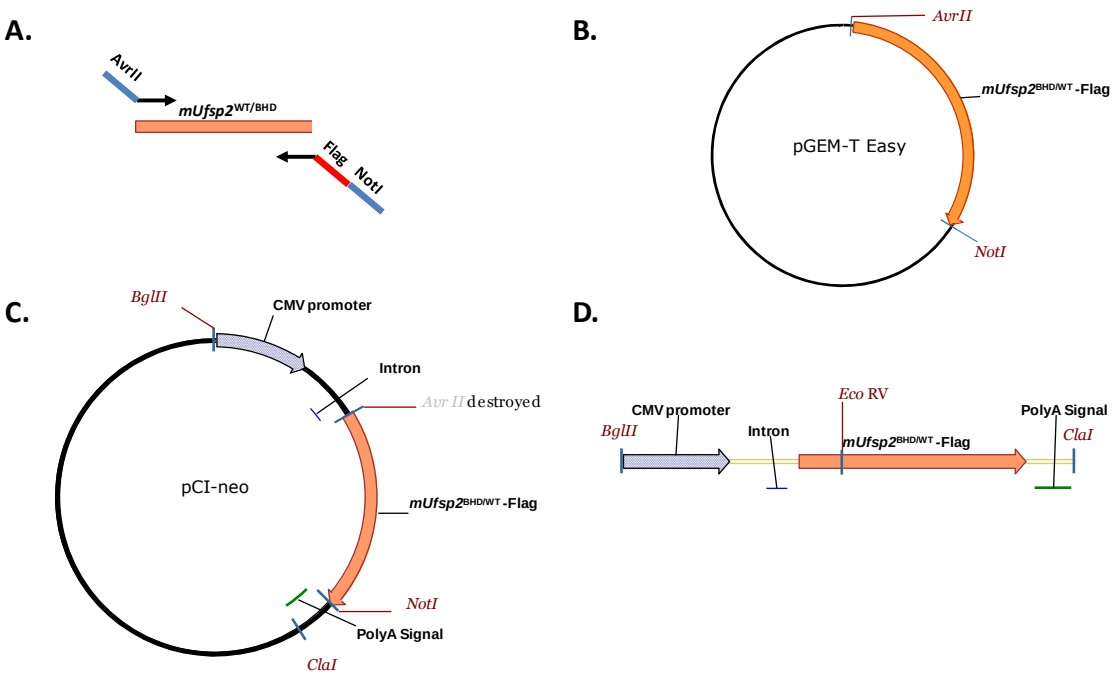


Figure 41. Schematic representation of the cloning procedure to generate the *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag expression constructs.

A. The *mUfsp2*^{WT} and *mUfsp2*^{BHD} cDNAs were amplified with primers incorporating restriction sites and a 3' Flag-tag. **B.** The *mUfsp2*$^{BHD/WT}$ –Flag PCR products were ligated into the pGEM-T Easy vector. **C.** The *mUfsp2*$^{BHD/WT}$ –Flag inserts were cloned into the pCI-neo expression vector using the incorporated restriction sites. **D.** The *mUfsp2*$^{BHD/WT}$ –Flag inserts plus associated regulatory sequences were excised from pCI-neo using the restrictions enzymes *Bgl*I and *Cla*I and purified for pronuclear injection.

5.2.2 Generation of *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag transgenic mice

The *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag inserts plus associated regulatory sequences were excised from pCI-neo, (Figure 42 D) purified (Figure 42 E) and sent to the transgenic facility at the University of Oulu in Finland for pronuclear injection.

Three rounds of pronuclear injection were performed, one with the *mUfsp2*^{WT}-Flag construct and two with the *mUfsp2*^{BHD}-Flag construct (summarised in Table 8). Ear biopsies were taken from the resultant pups two weeks after birth and sent to Manchester for genotyping. Genomic DNA was extracted from the biopsies according to the method described in 2.2.15 and the samples were screened for the presence of the transgene by PCR (see 2.2.11) using Ufsp2gntp2_For and _Rev (Appendix 3) primers designed to anneal within the chimeric intron region and the *mUfsp2*$^{WT/BHD}$-Flag insert of the transgenic construct, respectively. These primers were chosen to amplify the transgene but not the endogenous *Ufsp2* gene (Supplementary Figure 17). Five mice tested positive for the *mUfsp2*^{BHD}-Flag transgene and four for the *mUfsp2*^{WT}-Flag transgene and these mice were sent to Manchester. The mice were re-genotyped after their transfer to the Manchester animal house. As summarised in Table 8, the re-genotyping confirmed that five mice were positive for the *mUfsp2*^{BHD}-Flag transgene but only one of the four mice was transgenic for *mUfsp2*^{WT}-Flag (Figure 44). It is not clear why three of the mice had originally appeared positive for *mUfsp2*^{WT}-Flag but were clearly negative for the transgene following their transfer to Manchester and this issue was not resolved. Subsequent breeding was only carried out with the mice that were confirmed to carry the transgene.

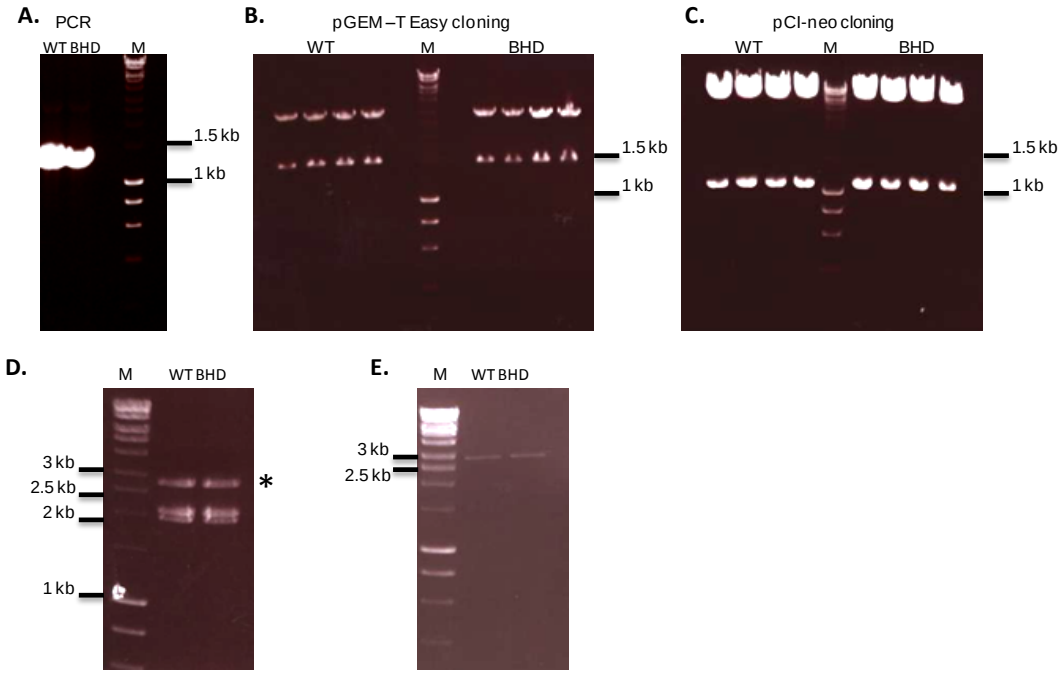


Figure 42. Amplification and cloning of *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag.

The cloning procedure was followed as outlined in Figure 41. **A.** PCR products generated following amplification of *mUfsp2*^{WT} and *mUfsp2*^{BHD} cDNAs with the AvrII_mUfsp2 and mUfsp2FlagNotRev primers. PCR products were of the predicted size (1418 bp) **B.** PCR products were cloned into the pGEM T-Easy vector and confirmed by restriction digestion with *Eco*RI. **C.** The *mUfsp2*$^{WT/BHD}$-Flag inserts were cloned into pCI-neo and excised by restriction digestion with *Eco*RV and *Not*I. **D.** The *mUfsp2*$^{WT/BHD}$-Flag inserts plus associated regulatory sequences were excised by restriction digestion with *Bgl*II and *Cla*I (the relevant 2787 bp fragment is indicated by the *). **E.** An aliquot of the *mUfsp2*$^{WT/BHD}$-Flag inserts prepared for pronuclear injection. All gels are 1% (w/v) agarose. M = Hyperladder I DNA marker.

#	Total number of pups born per injection	Construct	Number of transgenic pups per injection	Lines
1	20 pups	*mUfsp2*^{WT}-Flag	1 TG (3 false positives)	1716 (f)
2	12 pups	*mUfsp2*^{BHD}-Flag	2 TG	2036 (m) 2037 (f)
3	6 pups	*mUfsp2*^{BHD}-Flag	3 TG	2550 (m) 2552 (f) 2553 (f)

Table 8. Rounds of pronuclear injection and number of transgenic mice generated per round.

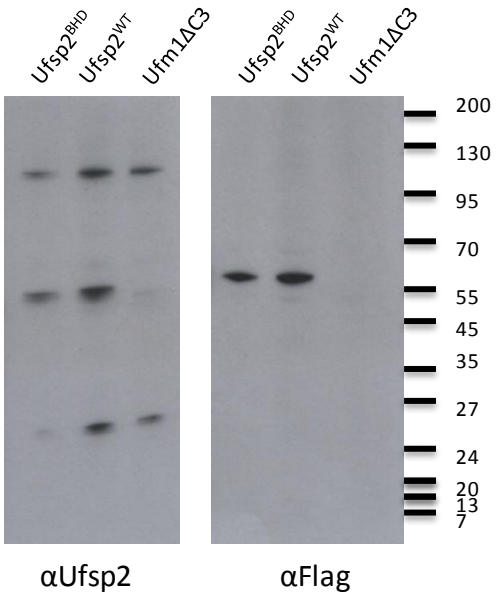


Figure 43. Detection of mUfsp2^{WT}-Flag and mUfsp2^{BHD}-Flag synthesis by western blotting.

mUfsp2^{WT}-Flag and *mUfsp2*^{BHD}-Flag pCI-neo constructs and a Ufm1ΔC3 (negative control) construct were expressed in the 2T3 mouse cell line. Cell lysates were resolved by SDS-PAGE (10%) and western blotting performed using anti-Ufsp2 and anti-Flag antibodies.

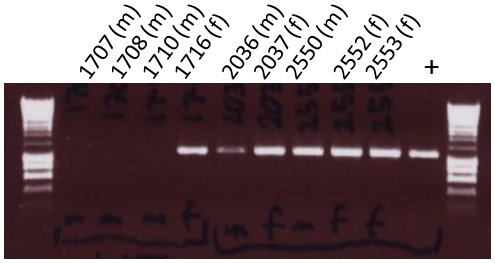


Figure 44. Genotyping of mice received from the Transgenic Facility, University of Oulu, Finland.

Mice born following pronuclear injection were screened for the presence of the transgene and positive ones were selected for shipment to local animal facility. The mice were genotyped again after shipment using primers specific to the transgene. 3 of the selected mice proved not to be transgenic. + = *mUfsp2*^{WT}-Flag plasmid (positive control); M – Hyperladder I DNA marker; (f) – female; (m) - male

5.2.2 Breeding of the transgenic lines

The *mUfsp2*^{BHD}-Flag and *mUfsp2*^{WT}-Flag transgenic mice were crossed with a wild type strain and the offspring sacrificed. The litters were genotyped as before using genomic DNA extracted from tail or ear samples. The outcome of these crosses and subsequent crosses that were performed are summarised in Figure 45. Mouse #1716(f) (the only mouse positive for *mUfsp2*^{WT}-Flag) did not produce any litters. This mouse was sacrificed at 16 weeks on the advice of the vet as it was exhibiting signs of stress. Hence no transgenic lines were generated for the *mUfsp2*^{WT}-Flag. Mice #2036(m) and #2550(m) did not produce any transgenic offspring and so no further crosses were carried out with these mice. Transgenic offspring were generated from crosses with *mUfsp2*^{BHD}-Flag mice #2037(f), #2552(f) and #2553(f). RNA was extracted from brain tissue and the litters were genotyped and transgenic pups were tested for the expression

of the transgene using RT PCR and primers specific for *mUfsp2*^{BHD}-Flag (Figure 46 A; see Appendix 3 for primer sequences). Additionally mRNA was extracted from the hip region of one of the lines and expression of the transgene in the tissue of interest was confirmed by RT PCR (Figure 46 B) As shown in Figure 46 only lines #2552 and #2553 showed positive expression of the *mUfsp2*^{BHD}-Flag transgene and hence subsequent breeding was only carried out using these lines (Figure 45). The next litters from the two lines were weaned and genotyped and the transgenic offspring were again bred with wild type mice. The transgenic offspring (transgenic heterozygotes) of these were subsequently weaned and crossed between themselves to generate litters possibly containing pups with a double dose of the transgene (transgenic homozygotes).

Initial crosses between transgenic heterozygotes produced a lower than expected ratio of transgenic to wild type mice for both lines suggesting possible embryonic lethality. Hence, litters were harvested at day E13.5 and E16.5, newborn and at 6 weeks of age. Specimens were weighed, genotyped and X-ray images were taken of the 6 week old mice. The lower body of the pups was dissected, fixed in 4% formaldehyde and embedded in paraffin blocks (2.2.25).

The observed versus expected ratios of transgenic to wild type offspring were analysed using the *Chi* squared test and the results are summarised in Table 9. The expected litter ratio for a cross between two transgenic heterozygotes is 1:2:1 for the wild type: transgenic heterozygote: transgenic homozygote. As the samples were not tested for copy number but only for the presence or absence of the transgene, the expected ratio from a transgenic heterozygote cross was 1:3 of wild type: transgenic offspring. A significant difference between the observed and expected number of new born and transgenic offspring was observed from the 2552x2552 cross. However, at 6 weeks old and at E13.5 and E16.5 this deviation was not found. The 6 week old data was only obtained from two litters however and so it was possible that transgenic homozygotes were being lost between E16.5 and birth. This was not observed for the 2553x2553 cross or for the 2552x2553 cross and so from these crosses homozygote transgenic offspring should be present.

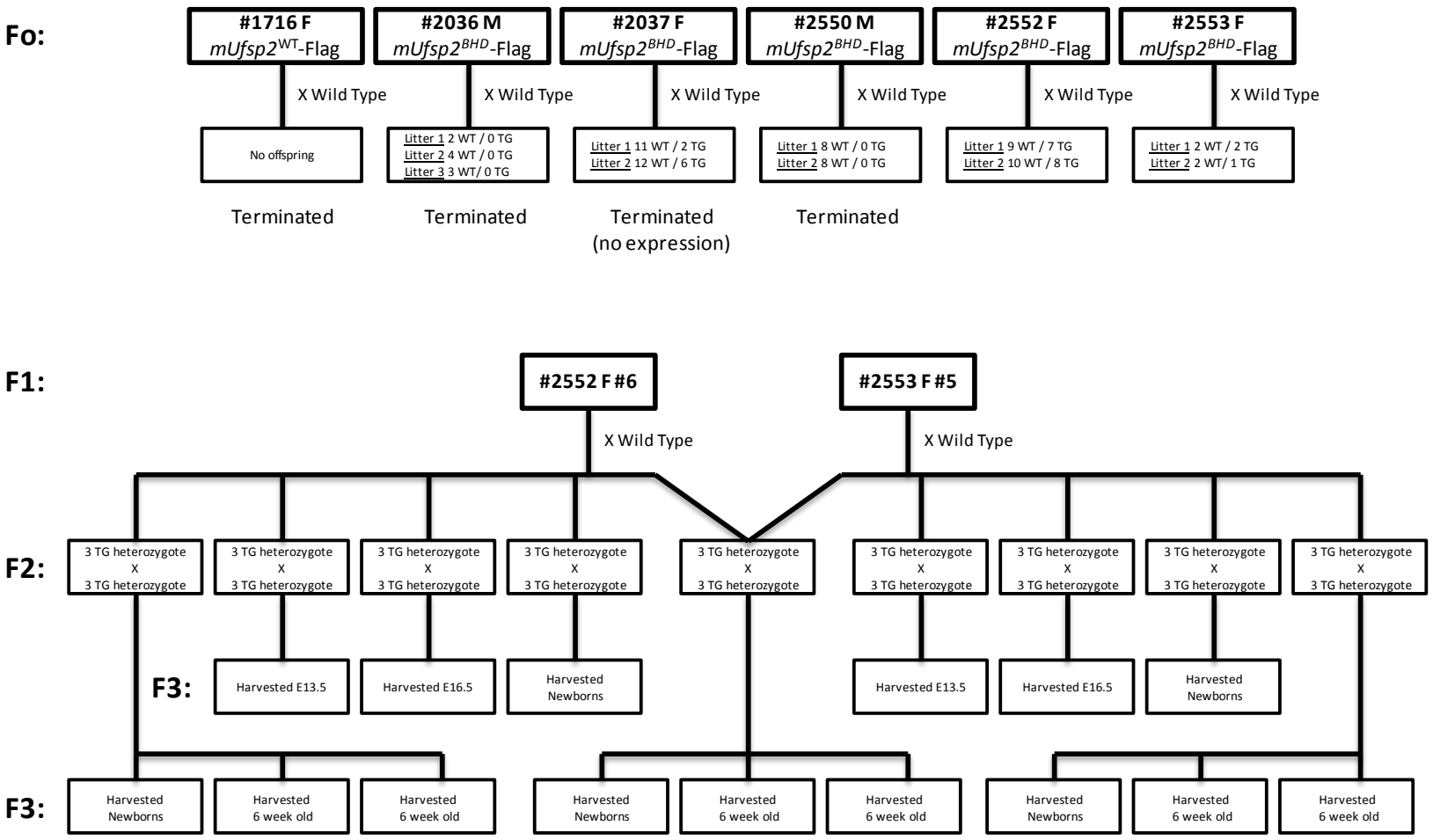


Figure 45. Schematic representation of mouse breeding.

The flow diagram shows crosses performed to obtain specimens for genetic analysis and histology of the transgenic mice. Two of the founders were chosen to establish transgenic lines and the other founders were terminated due to their lack of transgene expression or their inability to produce offspring. TG - transgenic offspring M – male; F – female.

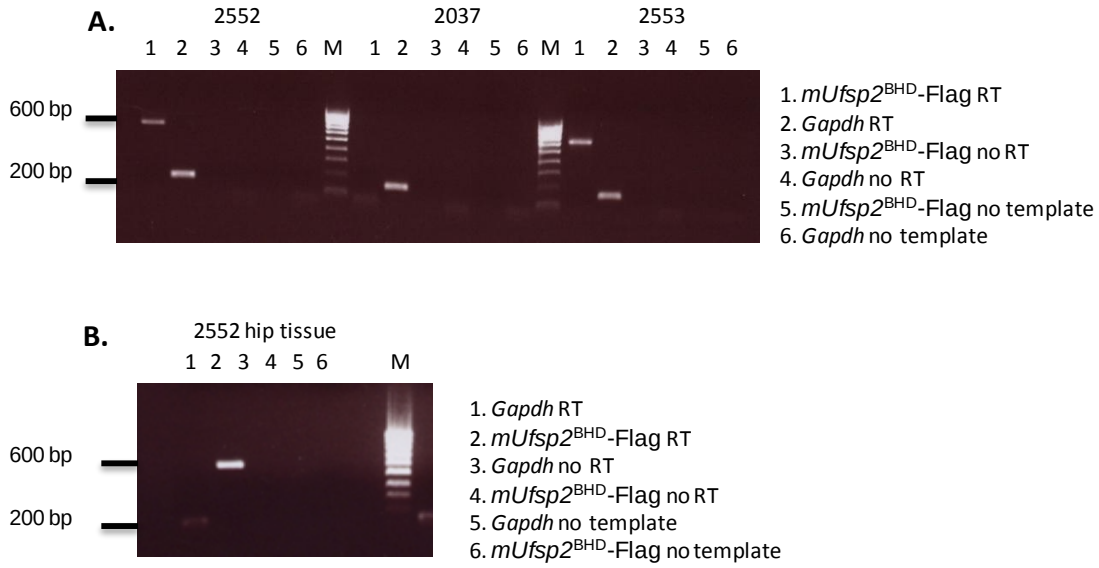


Figure 46. RT PCR showing expression of the *mUfsp2*^{BHD}-Flag transgene in mice.

A. Newborns from crosses of transgenic mice with wild type mice were sacrificed and genotyped. Reverse transcription PCR was performed on mRNA extracted from brain tissue of genotyped pups using primers specific for the *mUfsp2*^{BHD}-Flag transcript. *Gapdh* primers were used as a positive control for the PCR. **B.** mRNA was extracted from the hip region of a transgenic mouse and the expression of the *mUfsp2*^{BHD}-Flag construct was confirmed as described in A. M – Hyperladder IV DNA marker.

Age at genotyping	Litter	Wild type	transgenic	Size of litter
6week	2552x2552 Total	6	13	19 (+3 dead)
	Expected values	6	18	24
	χ^2			***p*=0.3165**
New born	2552x2552 #1	3	5	8
	2552x2552 #2	7	3	10
	2552x2552 #3	4	4	8
	Total	14	12	26
	Expected values	14	42	56
	χ^2			***p*= 0.0052**
E16.5	2552x2552 #1	2	5	7
	2552x2552 #2	3	11	14
	2552x2552 #3	3	9	12
	Total	8	25	33
	Expected values	8	24	32
	χ^2			***p*= 0.4717**
E13.5	2552x2552 #1	1	8	9
	2552x2552 #2	3	7	10
	Total	4	15	19
	Expected values	4	12	
	χ^2			***p*= 0.3909**

* **Table 9** continued on the next page.

Age at genotyping	Litter	Wild type	transgenic	Size of litter
6week	2553x2552 Total	2	14	16
	Expected values	2	6	8
	χ^2			***p*=0.2193**
New born	2553x2552 #1	3	7	10
	Expected values	3	9	9
	χ^2			***p*= 0.7932**

Age at genotyping	Litter	Wild type	transgenic	Size of litter
6week	2553x2553 Total	1	5	6 (+1 dead)
	Expected values	1	3	4
	χ^2			***p*=0.3734**
New born	2553x2553 #1	2	3	5
	Expected values	2	6	8
	χ^2			***p*= 0.5686**
E16.5	2553x2553 #1	0	2	2
	2553x2553 #2	3	3	6
	Total	3	5	8
	Expected values	3	9	12
	χ^2			***p*= 0.2750**
E13.5	2553x2553 Total	8	15	23
	Expected values	8	24	32
	χ^2			***p*= 0.2154**

Table 9. Analysis of the ratios of wild type to transgenic mice in sacrificed litters.

Mice were genotyped after sacrificing and the data was analysed using the Chi squared test in the GraphPad Prism 5 software. The dead mice indicated in brackets were removed by the animal facility before they could be genotyped and were therefore not included in the calculations.

5.2.3 Phenotype analysis

No gross abnormalities were observed when comparing wild type and transgenic mice from all crosses at all ages (see Figure 47 as exemplars). The only exception to this was one E13.5 transgenic embryo (2552x2552) which was underdeveloped. The average weights of offspring from all crosses were compared (Figure 48). No significant differences in weight were found between transgenic and wild type offspring from all crosses other than for the new born mice from the 2552x2552 and 2553x2553 crosses. For the 2552x2552 cross the transgenic mice were slightly heavier (nominal significance) whereas for the 2552x2553 cross the wild type mice were slightly heavier (nominal significance).

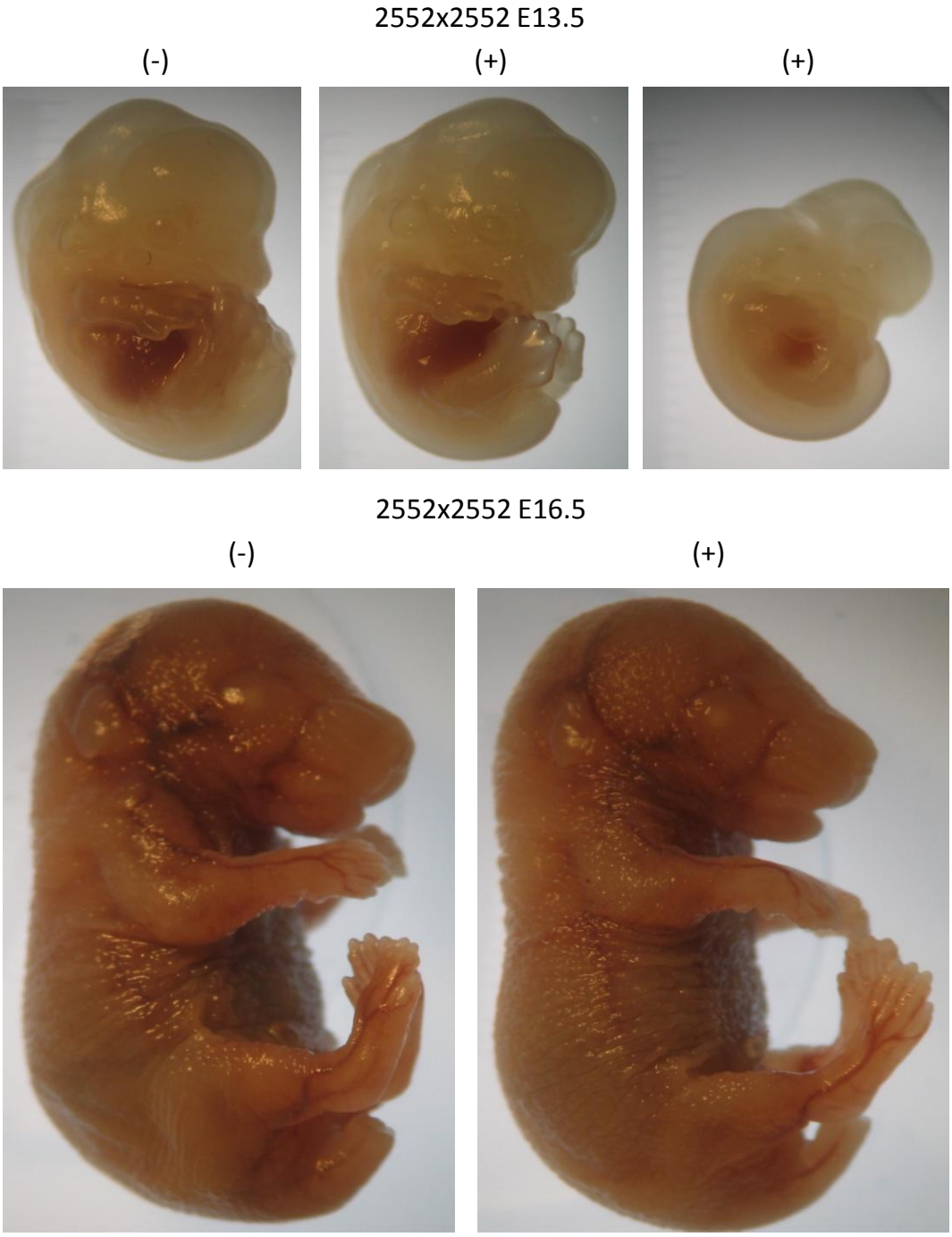


Figure 47. Analysis of the phenotype at E13.5 and E16.5 of embryonic development.

Embryos were harvested from pregnant females. Tail tips were removed to be used as source of gDNA for genotyping. Weights were recorded and images were taken.

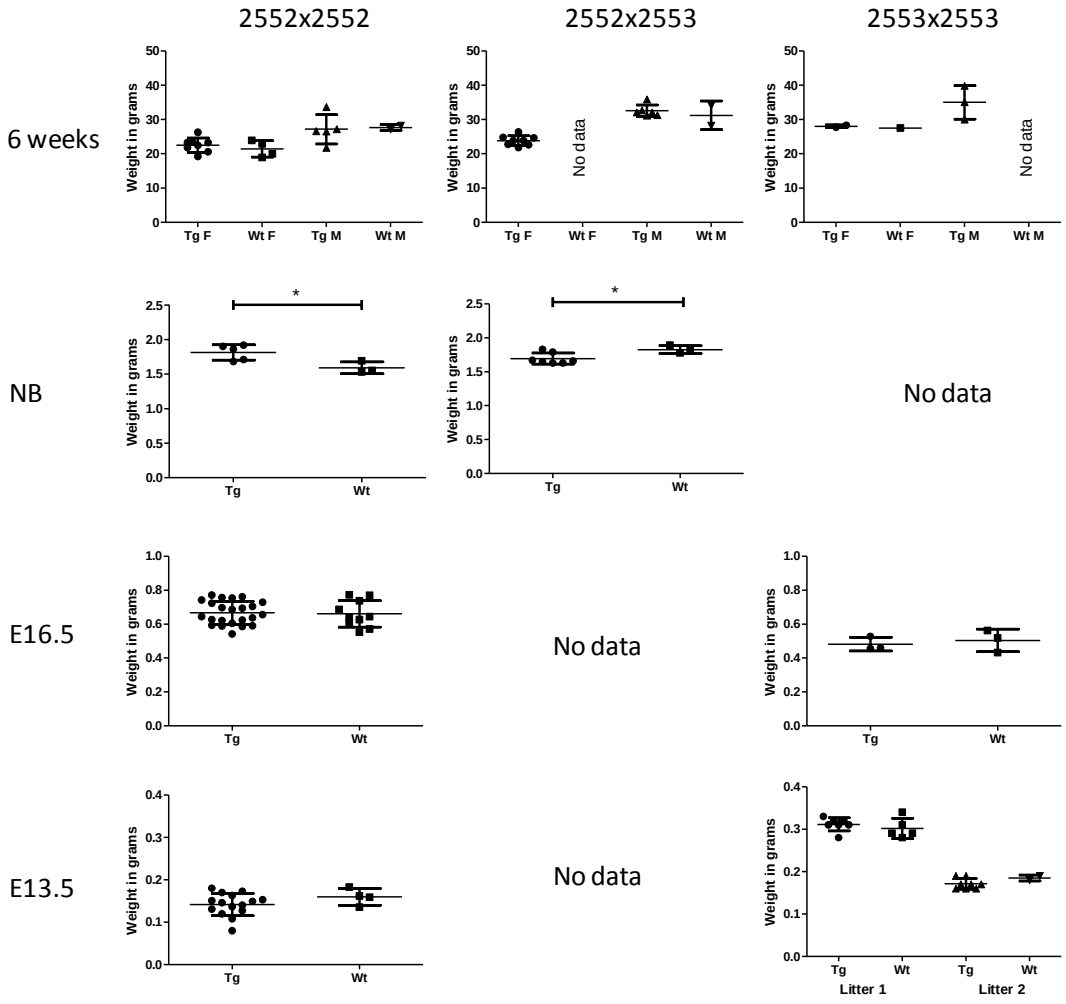


Figure 48. Weight average of transgenic and wild type mice.

Weights were recorded at the time points shown. The data was analysed by t test. * - $P < 0.05$; Tg – transgenic; Wt – wild type; F – female; M - male

The hip joints of wild type and transgenic offspring from the 2552x2552, 2552x2553 and 2553x2553 crosses were examined at 6 weeks of age using X-ray and histological analysis. This time point was chosen because according to a previous study [225] this is the age that secondary ossification centre formation in the hip may be observed in mice. Offspring from the 2552x2552 cross were also examined at E16.5 by histological analysis as from the analyses of transgenic offspring numbers (see Table 9) at this age transgenic homozygotes from this cross are most likely viable.

X-ray images of the 6 weeks old mice were analysed for evidence of characteristics like broadening and shortening of the femoral neck, flattened and

irregular capital epiphyses or overgrowth of the greater trochanter, which can be found in BHD patients. None of the BHD characteristics were found and no differences between wild type and transgenic animals were visible on the X-ray images (Figure 49 A). Histological analysis revealed that the secondary ossification centre formation had not commenced in most of the mice. In 6 out of 31 transgenic specimens (3/19 from 2552x2552, 1/13 from 2552x2553, 2/6 from 2553x2553) vascular invasion could be observed. In one of the 9 wild type mice a similar formation was observed at the join margin in contrast to the transgenic mice which had the formation located at the centre (Figure 49 B).

Hip and knee sections of embryos from the 2552x2552 cross were examined at E16.5 for evidence of abnormalities in the shape of the femoral head (Figure 50) or the tibiofemoral components of the knee joint as well as in organisation of the growth plates (Figure 51). Fifteen transgenic specimens and 3 wild type embryos were analysed and representative sections are shown in Figures 50 and 51. No joint shape differences were found between the transgenic and the wild type embryos. The growth plates of the transgenic embryos appeared normal with all zones being comparable to those of the wild type embryos.

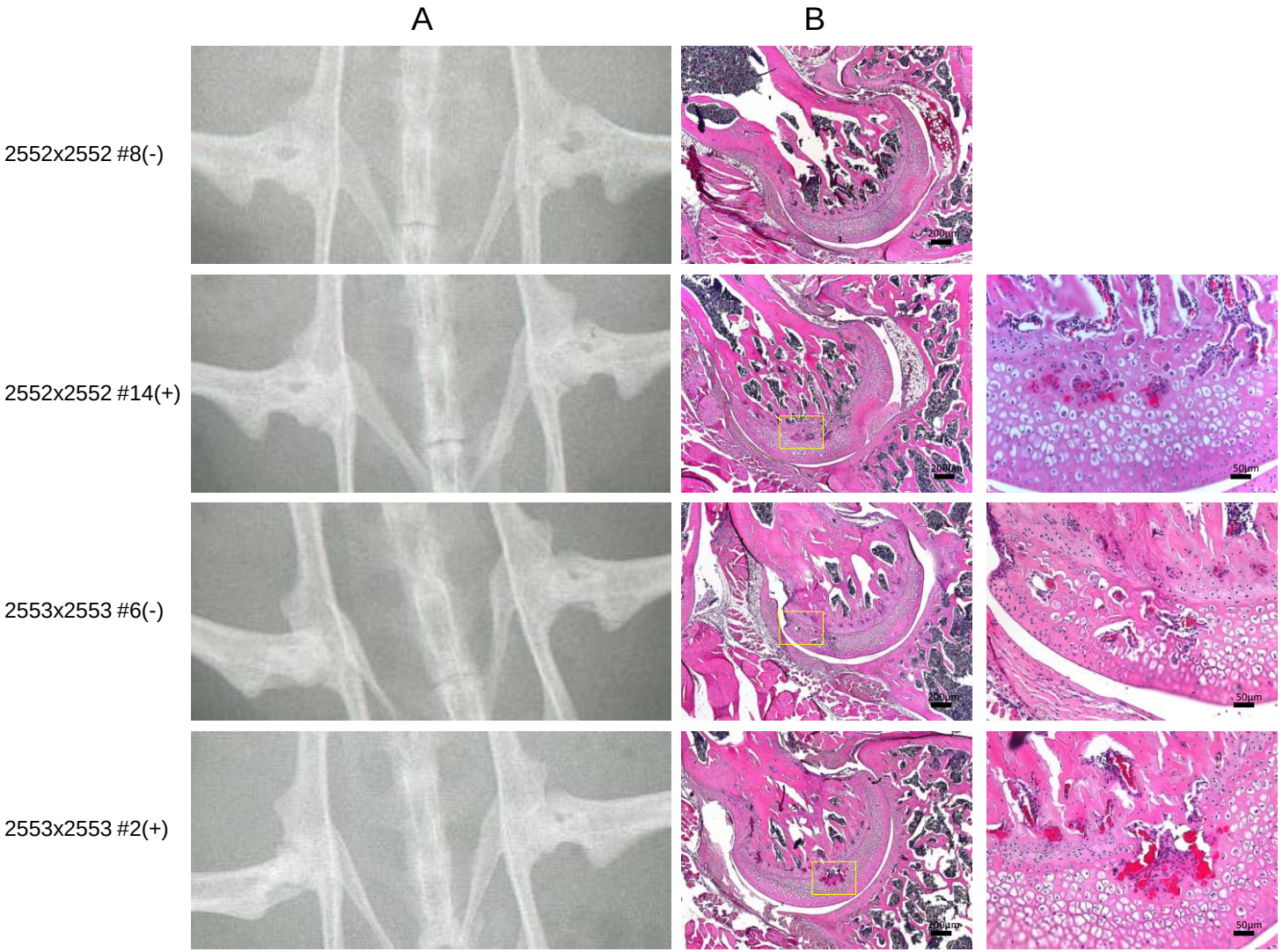


Figure 49. Histological analysis of the hip joints of 6 week old transgenic and wild type mice.

6 weeks old mice were harvested and X-ray images were taken by laying them flat facing down. Hips were dissected and processed for histology. **A.** X-ray images of the hips; **B.** Tissue sections stained with H&E. Scale bar: 200 µm (left, 5x), 50 µm (right 20x).

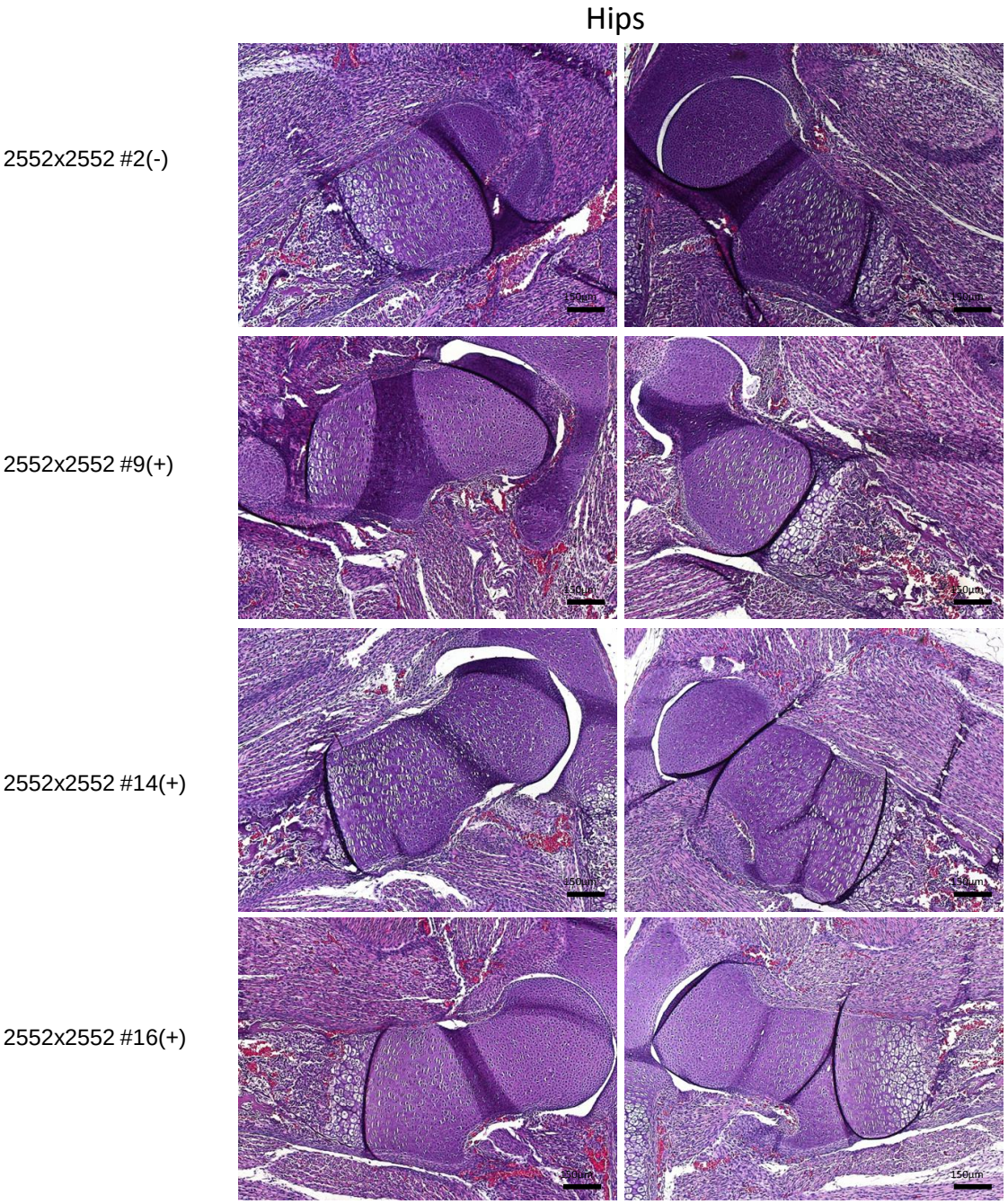


Figure 50. Histological analysis of the hip joints of E16.5 transgenic and wild type mice.

The hips were dissected and processed for histology. Tissue sections were stained with H&E. (-) = wild type; (+) = BHD. Scale bar 150 µm.

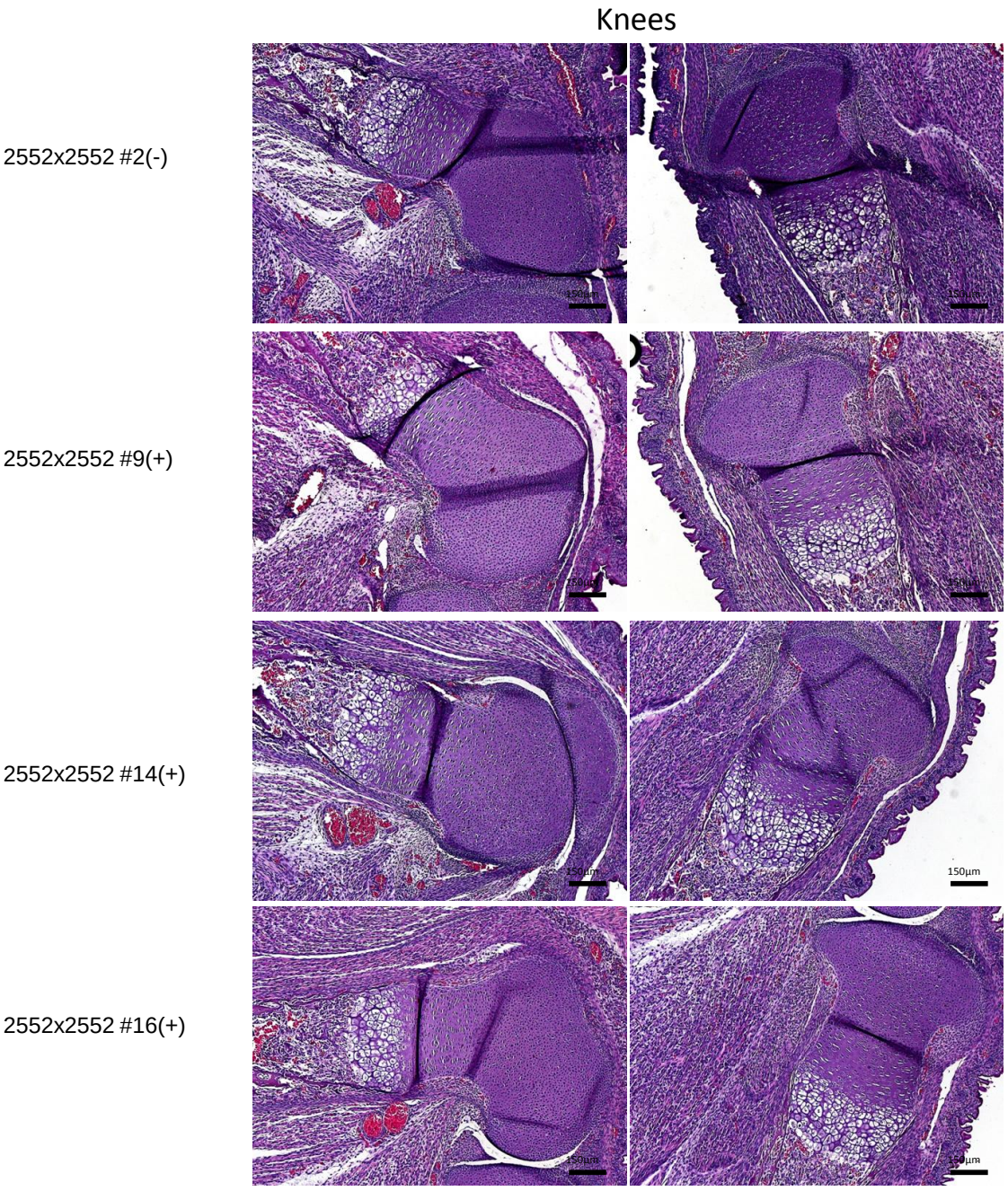


Figure 51. Histological analysis of the knee joints of E16.5 transgenic and wild type mice.

The hips were dissected and processed for histology. Tissue sections were stained with H&E. (-) = wild type; (+) = BHD. Scale bar 150 µm.

5.3 Discussion

To determine whether overexpression of either the WT or mutated (BHD) Ufsp2 in transgenic mice had a phenotypic effect, two expression constructs, *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag under control of the constitutive CMV promoter, were generated. The integrity of these constructs was confirmed by sequence analysis and expression of both WT and BHD Ufsp2 was confirmed by western blotting (Figure 43). Both constructs were then sent to the transgenic facility at the University of Oulu, Finland, for pronuclear injection. Three rounds of pronuclear injections resulted in 5 *mUfsp2*^{BHD}-Flag (2 X male and 3 X female) and 1 *mUfsp2*^{WT}-Flag (1 X female) transgenic mice (Table 8). As summarised in Figure 45, of the 5 *mUfsp2*^{BHD}-Flag mice 2 male mice did not produce transgenic offspring. This is likely to have occurred as the transgene was not incorporated into the germline. The pronuclear injection may sometimes produce mosaic mice when the injected DNA is not immediately incorporated into the genome such that when the fertilised oocyte divides, the transgene is not incorporated into the genome of all daughter cells. The transgene will not be passed onto the offspring if the non-transgenic population of cells gives rise to the germline. Of the 3 remaining *mUfsp2*^{BHD}-Flag mice, 1 mouse produced transgenic offspring but these offspring did not express the transgene (Figure 46). In this case although the transgene was transmitted it may have been silenced due its site of integration in the chromatin. Two of the *mUfsp2*^{BHD}-Flag mice did however produce transgenic offspring and expression of the transgene was confirmed (Figure 46). The only *mUfsp2*^{WT}-Flag transgenic mouse did not produce any transgenic offspring. In this case insertional mutagenesis may be the reason for the infertility but also the overexpression of the active Ufsp2 protein may have prevented production of viable oocytes. In addition, this mouse had to be terminated (at 16 weeks of age) as it failed to thrive and was agitated and stressed. It could not be established as to whether this phenotype was related to overexpression of WT Ufsp2 as no other *mUfsp2*^{WT}-Flag transgenic mice were generated for comparison.

The two *mUfsp2*^{BHD}-Flag lines that expressed the BHD Ufsp2 were bred and the ratio of transgenic offspring was analysed at E13.5, E16.5, new born and 6 weeks old. The ratios of transgenic to WT offspring did not deviate significantly from the expected values for both lines at all ages except for the ratio of newborn mice generated from the 2552x2552 cross where the ratio of transgenic to WT mice was significantly reduced (p=0.0052) (Table 9). This suggests that homozygosity (or a double dose) of expression

from the *mUfsp2*^{BHD}-Flag transgene in this line may be lethal between E16.5 and birth. The ratio of transgenic to WT mice for this line was not however significantly reduced (p=0.3165) for 6 week old mice. This may be due to the small sample size (n=19) as, even though not statistically significant, 1/3 fewer 6 week old transgenic offspring were found than expected. A similar paucity of new born transgenic mice was observed in the 2553x2553 cross but this was not statistically significant which again may be a due to the small sample size. There was however no statistically significant alteration in the numbers of WT to transgenic offspring generated from 2553x2552 cross (either at birth or at 6 weeks) which does indicate that the paucity of newborn mice from the 2552x2552 cross may be result of insertional mutagenesis. However this could also be explained if the level of expression from the transgene was higher in the 2552 than the 2553 line. In addition, as the predictions of the expected number of transgenic mice are predicated on there being a single transgene inserted at a single site in each line, the loss of mice with double the number of transgenes may be masked.

Visual inspection of embryos (E13.5 and E16.5), newborn and 6 week old mice from the 2552x2552 and the 2553x2553 crosses plus newborn and 6 week old mice from the 2552x2553 cross did not reveal any obvious signs of developmental delay or morphological defects. The only exception to this was a single transgenic embryo from the 2552x2552 cross at E13.5 that was obviously abnormal (Figure 47). X-ray images of the hip joints of all 6 week old mice did not reveal any obvious differences between the transgenic and WT mice (Figure 47). The recorded weights from all crosses were not significantly different except for newborns of 2552x2552 and 2553x2552 crosses (Figure 48). In the former the transgenic mice were slightly heavier and in the latter the WT mice were slightly heavier and so this difference in weight was unlikely to be related to the growth rates of these groups. Further, the newborn mice were not separated into groups based on their sex and so these differences might be accounted for by differing numbers of male and female offspring in each group.

Histological analysis was performed on tissue sections of hips and knees of E16.5 offspring from the 2552x2552 cross as at this age it was anticipated from the analysis presented in Table 9 that homozygous (double dose) transgenic mice would be present. This analysis did not reveal any histological differences between the WT and transgenic mice at either of the joint sites at this age. The hip joints of 6 week old offspring from the 2552x2552, 2552x2553 and 2553x2553 cross were examined. Again

there were no obvious histological differences between the WT and transgenic mice in the hip joints and surrounding tissues at this age. The only noteworthy observation was that vascular invasion (which precedes the formation of the secondary ossification centre) had commenced in 6/31 of the transgenic mice but only 1/9 WT mice. Further studies would be required to determine whether the initiation and formation of the secondary centre was significantly different in the transgenic mice than in the WT mice. At this stage however this remains the only potential difference between the hip joints of WT and transgenic mice.

Based on the data presented in this Chapter there was no evidence that overexpression of *mUfsp2*^{BHD}-Flag in mice produces a phenotype in the hip joint. If the mutation in *UFSP2* does indeed cause the BHD phenotype, there are several reasons why this may be the case, including: (i) haploinsufficiency may be the mutation mechanisms and so overexpression of the mutant Ufsp2 may not have an effect; (ii) mouse hip joint formation may not be as affected by the expression of mutated Ufsp2 as is the case for human hip joint development; (iii) that the effects of the expression of the mutated Ufsp2 may only be evident following secondary centre formation or in aged mice; and (iv) the level of expression of the transgene may not be sufficiently high or sustained in the specific tissue of the hip joint to result in a phenotype. To address these issues generation of a knock-in mouse for the *mUfsp2*^{BHD} mutation is required followed by detailed phenotyping at a range of ages.

Chapter 6. Conclusions and future work

As outlined in this thesis, BHD is an autosomal dominant disorder where the abnormal shape of the hip joint leads to secondary OA in early adulthood. At the start of the research reported in this thesis in 2009, a mutation in the *UFSP2* gene had just been found to co-segregate with the phenotype of BHD with a LOD score of 10.4 (see 1.9). There was preliminary evidence at that time that Ufsp2, and the related Ufsp1, were proteases responsible for activation and/or deconjugation of a novel ubiquitin-like modifier,Ufm1 (see 1.8.4). The BHD mutation in *UFSP2* was shown to abolish the catalytic activity of the encoded protein *in vitro* (see 1.9). Further studies of the Ufm1 system published by other research groups during the course of this study (see 1.8.4) included characterisation of the main components of the Ufm1 system, notably the E1, E2 and E3 enzymes (Uba5, Ufc1 and Ufl1, respectively), a single Ufm1 protein target (Ddrgk1) and a Ufm1 associated gene (Lzap). Ufl1, Ufsp2 and Ddrgk1 were found to localise to the ER and there was evidence of either increased or decreased Ufm1 conjugation in response to ER stress. The studies described in this thesis explored the role of the Ufm1 system further in order to identify potential mechanisms as to how alterations to this system might lead to the BHD phenotype. These studies have led to the following conclusions and have indicated areas where further research is required.

6.1 *Ufsp2* is predominantly expressed in the bone of the hip joint

To determine whether *Ufsp2* has a differential pattern of expression within the multiple tissues of the hip joint and hence potentially indicate the tissue most likely affected by the BHD mutation, radioactive RNA *in situ* hybridisation was performed on mouse tissue sections (see 3.2.1). This very sensitive method was chosen as preliminary real-time PCR experiments performed by Chris Watson described in his PhD thesis [179] showed very low level of *Ufsp2* expression. No *Ufsp2* expression was detected in embryonic and newborn hip joints but distinctive patterns of expression were observed in 10 day old mouse hip and knee joints. In the 10 day old hip joints, the expression was localised mostly in the bone and in the secondary centre of ossification of the knee. There was some evidence of expression in the proliferative zone of the growth plate in the knee and in surrounding ligaments of both the hip and the knee. Although *Ufsp2*

expression was not confined to bone, these studies did identify that osteoblastic cell lines may be a potential cell culture system for studies of the role of the Ufm1 system.

6.2 The Ufm1 system is upregulated during osteogenic differentiation in response to ER stress

The ER is an important cellular organelle, the homeostasis of which is crucial for normal functioning of the cell and its disturbance may lead to cell cycle arrest as well as apoptosis. ER homeostasis is of particular importance in active secretory cells such as chondrocytes and osteoblasts but also in cells that require vast Ca^{2+} storage capability like the muscle cells. Gene expression of the Ufm1 system components was investigated during *in vitro* osteogenic and chondrogenic differentiation assays and components of the system were found to be significantly upregulated (see 3.2.2 and 3.2.3, respectively). In these studies it was found that the Ufm1 system upregulation coincided with induction of ER stress. Chemically induced ER stress led to upregulation of expression of Ufm1 system genes in both osteogenic and chondrogenic cell lines with a lesser effect in a myogenic cell line (see 3.2.4). Silencing of *Ufsp2* using RNAi method in the osteoblast cell line could answer the question whether osteogenic differentiation and/or ER stress response is impaired in consequence of *Ufsp2* deficiency.

At the protein level induction of ER stress in 2T3 cells resulted in the increase in intensity of five bands on a western blot detected with anti-Ufm1 antibody (27, 30, 40, 45 and 60 kDa). Three of these are presumed to be Ufc1 (27 or 30 kDa), Ddrgk1 (45 kDa) and Uba5 (60 kDa) and two remain unidentified. Three bands detected on western blot of proteins purified from 2T3 osteoblasts were not present in the stable HEK293T cell line which suggests that the Ufm1 system may not be fully active in this cell line and that Ufm1 targets may be cell line (or tissue) specific. Purification and mass spectrometry analysis of the Ufm1 conjugated proteins detected in the 2T3 cell line may provide an insight into the role of the Ufm1 system.

6.3 *Ufm1, Uba5, Ufl1* and *Lzap* gene promoters possess Unfolded Protein Response Elements (UPRE)

Analysis of the promoter regions of the Ufm1 system components revealed UPRE sequences upstream of *Ufm1, Uba5, Ufl1* and *Lzap* genes. The ~1.5 kb promoter regions of *Uba5* and *Lzap* were shown to be upregulated by thapsigargin induced ER stress in a luciferase reporter assay and the UPRE sequences were shown to be required for activity of the *Ufm1, Uba5* and *Lzap* promoters. Attempts to amplify the *Ufl1* promoter region failed and a potential downstream UPRE sequence in the Ufm1 gene was not analysed. Thus, further research is required to confirm that the putative UPREs in these genes are indeed ER response elements. Published literature suggests that these UPREs may be regulated by XBP1 (see 1.7). Further research is required to determine whether XBP1, or other transcription factors of the bZIP family known to bind similar sequences, are the mediators of the upregulation of the Ufm1 system in response to ER stress. Specifically, coexpression of the spliced form of XBP1 with the luciferase vectors used in the study could confirm the role of XBP1 as the transcription factor specific for the UPREs described here.

6.4 Development of a Tandem Affinity Purification (TAP) method for isolation of Ufm1 protein targets

The TAP method for isolation of Ufm1 conjugation targets involved denaturing and reducing lysis conditions coupled with Ni-NTA resin purification of His-tagged proteins, buffer change and Strep-Tactin resin purification of proteins possessing StrepTagII. His-tag is known to suffer from nonspecific copurification of proteins possessing high histidine content and StrepTagII, although more specific, is limited by the buffer composition. The use of two step purification utilising two different tags allows combining the strengths of the two systems and circumvention of the limitations. Other alternatives exist on the market like FLAG tag, GST and more recent HaloTag or CaptureSelect C-tag. The method was successfully tested in the HEK293T cell line stably overexpressing Ufm1ΔC2 construct and bands corresponding to the E1 and E2 enzymes of the Ufm1 system were identified by mass spectrometry. No new Ufm1 conjugation targets were however found. In contrast, purification and western blot analysis of Ufm1 coupled proteins from the 2T3 cell line revealed additional putative Ufm1 targets not present in the stably overexpressing HEK293T cell line. Large scale

purification and mass spectrometry analysis of proteins from the 2T3 cell line should yield novel Ufm1 system targets. Additionally, the mass spectrometry data should be analysed using an algorithm detecting an Ufm1 "fingerprint" consisting of peptides with a lysine modified by an isopeptide bond with several C-terminal amino acids of Ufm1 left over after trypsin digestion. The new candidates would then need to be validated by immunoprecipitation of the target and detection of Ufm1 or/and an *in vitro* conjugation assay.

6.5 Overexpression of *Ufsp2* containing the BHD mutation in transgenic mice does not produce an overt hip joint phenotype

Transgenic mice overexpressing the BHD mutated form of *Ufsp2* were generated. Gross morphology, weight, X-rays and histology of hips and knees were analysed. No obvious differences between the wild type and transgenic mice were observed. The number of transgenic mice per litter was slightly lower than expected when two heterozygotes were crossed and was significant lower in one of the crosses. A larger number of crosses would be needed to confirm whether mice (potentially the homozygote mice) die shortly before or at birth when two heterozygotes are crossed. The lack of a phenotype in mice overexpressing mutated *Ufsp2* gene may suggest: (i) intrinsic differences between hip joint morphology and development in mice and humans; (ii) that expression of the mutant *Ufsp2* was not at a sufficient level to cause the phenotype; (iii) and that haploinsufficiency of *UFSP2* is the mutation mechanism underlying the BHD phenotype. A *Ufsp2* knock-out or mutant *Ufsp2* knock-in mouse model should be generated to investigate the above possibilities and to confirm that the *UFSP2* mutation causes the BHD phenotype.

In conclusion, the above investigations of the Ufm1 system suggest its involvement in the homeostasis of the ER and in ER stress responses. That the *UFSP2* mutation may cause disruption of the Ufm1 mediated response to ER stress and thereby lead to BHD is consistent with ER stress being a known pathological consequence of mutations causing forms of chondrodysplasia with associated secondary OA. It is not known however what the effect of Ufm1 conjugation/deconjugation has on its target proteins and how the mutation in *UFSP2* disrupts the Ufm1 system. Further studies of

the Ufm1 system are therefore required to precisely define the BHD genotype to phenotype pathway. Such studies may lead to a better understanding of the regulation of ER stress but also as to how normal regulation of ER stress responses is related to hip joint morphogenesis and development.

7. References

[1] M.B. Goldring, Update on the biology of the chondrocyte and new approaches to treating cartilage diseases, Best Pract Res Clin Rheumatol 20 (2006) 1003-1025.

[2] J. Bondeson, Are we moving in the right direction with osteoarthritis drug discovery?, Expert Opin Ther Targets 15 1355-1368.

[3] N. Arden, M.C. Nevitt, Osteoarthritis: Epidemiology, Best Practice & Research Clinical Rheumatology 20 (2006) 3-25.

[4] L.J. Sandell, Etiology of osteoarthritis: genetics and synovial joint development, Nat Rev Rheumatol 8 (2012) 77-89.

[5] J.J. Ryder, K. Garrison, F. Song, L. Hooper, J. Skinner, Y. Loke, J. Loughlin, J.P.T. Higgins, A.J. MacGregor, Genetic associations in peripheral joint osteoarthritis and spinal degenerative disease: a systematic review, Ann Rheum Dis 67 (2008) 584-591.

[6] C.H. Evans, S.C. Ghivizzani, P.D. Robbins, Arthritis gene therapy and its tortuous path into the clinic, Transl Res 161 (2013) 205-216.

[7] H.J. Cilliers, P. Beighton, Beukes familial hip dysplasia: an autosomal dominant entity, Am J Med Genet 36 (1990) 386-390.

[8] P. Roby, S. Eyre, J. Worthington, R. Ramesar, H. Cilliers, P. Beighton, M. Grant, G. Wallis, Autosomal dominant (Beukes) premature degenerative osteoarthropathy of the hip joint maps to an 11-cM region on chromosome 4q35, Am J Hum Genet 64 (1999) 904-908.

[9] A.G. Cole, B.K. Hall, Cartilage is a metazoan tissue; integrating data from nonvertebrate sources, Acta Zoologica 85 (2004) 69-80.

[10] K. Pelttari, E. Steck, W. Richter, The use of mesenchymal stem cells for chondrogenesis, Injury 39 Suppl 1 (2008) S58-65.

[11] L. Quintana, N.I. zur Nieden, C.E. Semino, Morphogenetic and regulatory mechanisms during developmental chondrogenesis: new paradigms for cartilage tissue engineering, Tissue Eng Part B Rev 15 (2009) 29-41.

[12] A.M. DeLise, R.S. Tuan, Alterations in the spatiotemporal expression pattern and function of N-cadherin inhibit cellular condensation and chondrogenesis of limb mesenchymal cells in vitro, J Cell Biochem 87 (2002) 342-359.

[13] R.A. Oldershaw, T.E. Hardingham, Notch signaling during chondrogenesis of human bone marrow stem cells, Bone (2009).

[14] E. Wright, M.R. Hargrave, J. Christiansen, L. Cooper, J. Kun, T. Evans, U. Gangadharan, A. Greenfield, P. Koopman, The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos, Nat Genet 9 (1995) 15-20.

[15] H. Akiyama, Control of chondrogenesis by the transcription factor Sox9, Mod Rheumatol 18 (2008) 213-219.

[16] G. Zhang, B.F. Eames, M.J. Cohn, Chapter 2. Evolution of vertebrate cartilage development, Curr Top Dev Biol 86 (2009) 15-42.

[17] T.E. Hardingham, R.A. Oldershaw, S.R. Tew, Cartilage, SOX9 and Notch signals in chondrogenesis, J Anat 209 (2006) 469-480.

[18] S. Grassel, N. Ahmed, Influence of cellular microenvironment and paracrine signals on chondrogenic differentiation, Front Biosci 12 (2007) 4946-4956.

[19] E. Koyama, Y. Shibukawa, M. Nagayama, H. Sugito, B. Young, T. Yuasa, T. Okabe, T. Ochiai, N. Kamiya, R.B. Rountree, D.M. Kingsley, M. Iwamoto, M.

Enomoto-Iwamoto, M. Pacifici, A distinct cohort of progenitor cells participates in synovial joint and articular cartilage formation during mouse limb skeletogenesis, Dev Biol 316 (2008) 62-73.
[20] G. Hyde, S. Dover, A. Aszodi, G.A. Wallis, R.P. Boot-Handford, Lineage tracing using matrilin-1 gene expression reveals that articular chondrocytes exist as the joint interzone forms, Dev Biol 304 (2007) 825-833.
[21] G. Hyde, R.P. Boot-Handford, G.A. Wallis, Col2a1 lineage tracing reveals that the meniscus of the knee joint has a complex cellular origin, J Anat 213 (2008) 531-538.
[22] G. Lizarraga, A. Lichtler, W.B. Upholt, R.A. Kosher, Studies on the role of Cux1 in regulation of the onset of joint formation in the developing limb, Dev Biol 243 (2002) 44-54.
[23] C.W. Archer, G.P. Dowthwaite, P. Francis-West, Development of synovial joints, Birth Defects Res C Embryo Today 69 (2003) 144-155.
[24] A.A. Pitsillides, D.E. Ashhurst, A critical evaluation of specific aspects of joint development, Dev Dyn 237 (2008) 2284-2294.
[25] A.J. Hayes, S. MacPherson, H. Morrison, G. Dowthwaite, C.W. Archer, The development of articular cartilage: evidence for an appositional growth mechanism, Anat Embryol (Berl) 203 (2001) 469-479.
[26] J.A. Buckwalter, M.J. Glimcher, R.R. Cooper, R. Recker, Bone biology. I: Structure, blood supply, cells, matrix, and mineralization, Instr Course Lect 45 (1996) 371-386.
[27] V. Abad, J.L. Meyers, M. Weise, R.I. Gafni, K.M. Barnes, O. Nilsson, J.D. Bacher, J. Baron, The role of the resting zone in growth plate chondrogenesis, Endocrinology 143 (2002) 1851-1857.
[28] J.C. Lui, A.C. Andrade, P. Forcinito, A. Hegde, W. Chen, J. Baron, O. Nilsson, Spatial and temporal regulation of gene expression in the mammalian growth plate, Bone 46 (2010) 1380-1390.
[29] F. Las Heras, H.K. Gahunia, K.P. Pritzker, Articular cartilage development: a molecular perspective, Orthop Clin North Am 43 (2012) 155-171, v.
[30] D.A. Garzon-Alvarado, J.M. Garcia-Aznar, M. Doblare, Appearance and location of secondary ossification centres may be explained by a reaction-diffusion mechanism, Comput Biol Med 39 (2009) 554-561.
[31] E.J. Mackie, L. Tatarczuch, M. Mirams, The skeleton: a multi-functional complex organ: the growth plate chondrocyte and endochondral ossification, J Endocrinol 211 (2011) 109-121.
[32] E.J. Mackie, Y.A. Ahmed, L. Tatarczuch, K.S. Chen, M. Mirams, Endochondral ossification: how cartilage is converted into bone in the developing skeleton, The international journal of biochemistry & cell biology 40 (2008) 46-62.
[33] G.A. Wallis, Bone growth: coordinating chondrocyte differentiation, Curr Biol 6 (1996) 1577-1580.
[34] D.R. Eyre, The collagens of articular cartilage, Semin Arthritis Rheum 21 (1991) 2-11.
[35] A.R. Poole, T. Kojima, T. Yasuda, F. Mwale, M. Kobayashi, S. Laverty, Composition and structure of articular cartilage: a template for tissue repair, Clin Orthop Relat Res (2001) S26-33.
[36] D. Heinegard, Proteoglycans and more--from molecules to biology, Int J Exp Pathol 90 (2009) 575-586.
[37] D. Heinegard, T. Saxne, The role of the cartilage matrix in osteoarthritis, Nat Rev Rheumatol 7 (2011) 50-56.

[38] S. Weiner, W. Traub, H.D. Wagner, Lamellar bone: structure-function relations, J Struct Biol 126 (1999) 241-255.
[39] S.C. Marks, Jr., S.N. Popoff, Bone cell biology: the regulation of development, structure, and function in the skeleton, Am J Anat 183 (1988) 1-44.
[40] F. Long, Building strong bones: molecular regulation of the osteoblast lineage, Nat Rev Mol Cell Biol 13 (2012) 27-38.
[41] R. Nishimura, K. Hata, K. Ono, K. Amano, Y. Takigawa, M. Wakabayashi, R. Takashima, T. Yoneda, Regulation of endochondral ossification by transcription factors, Front Biosci 17 (2012) 2657-2666.
[42] A. Cheng, P.G. Genever, SOX9 determines RUNX2 transactivity by directing intracellular degradation, J Bone Miner Res 25 (2010) 2680-2689.
[43] T. Komori, Regulation of bone development and maintenance by Runx2, Front Biosci 13 (2008) 898-903.
[44] B.F. Eames, P.T. Sharpe, J.A. Helms, Hierarchy revealed in the specification of three skeletal fates by Sox9 and Runx2, Dev Biol 274 (2004) 188-200.
[45] A. Yamaguchi, T. Komori, T. Suda, Regulation of osteoblast differentiation mediated by bone morphogenetic proteins, hedgehogs, and Cbfa1, Endocr Rev 21 (2000) 393-411.
[46] G. Xiao, D. Jiang, C. Ge, Z. Zhao, Y. Lai, H. Boules, M. Phimphilai, X. Yang, G. Karsenty, R.T. Franceschi, Cooperative interactions between activating transcription factor 4 and Runx2/Cbfa1 stimulate osteoblast-specific osteocalcin gene expression, J Biol Chem 280 (2005) 30689-30696.
[47] C. Hartmann, Transcriptional networks controlling skeletal development, Curr Opin Genet Dev 19 (2009) 437-443.
[48] T. Tohmonda, Y. Miyauchi, R. Ghosh, M. Yoda, S. Uchikawa, J. Takito, H. Morioka, M. Nakamura, T. Iwawaki, K. Chiba, Y. Toyama, F. Urano, K. Horiuchi, The IRE1alpha-XBP1 pathway is essential for osteoblast differentiation through promoting transcription of Osterix, EMBO Rep 12 (2011) 451-457.
[49] P.A. Downey, M.I. Siegel, Bone biology and the clinical implications for osteoporosis, Phys Ther 86 (2006) 77-91.
[50] D.W. Sommerfeldt, C.T. Rubin, Biology of bone and how it orchestrates the form and function of the skeleton, Eur Spine J 10 Suppl 2 (2001) S86-95.
[51] S.L. Teitelbaum, Bone resorption by osteoclasts, Science (New York, N.Y 289 (2000) 1504-1508.
[52] D.L. Rimoin, D. Cohn, D. Krakow, W. Wilcox, R.S. Lachman, Y. Alanay, The skeletal dysplasias: clinical-molecular correlations, Ann N Y Acad Sci 1117 (2007) 302-309.
[53] P. Kannu, J.F. Bateman, D. Belluoccio, A.J. Fosang, R. Savarirayan, Employing molecular genetics of chondrodysplasias to inform the study of osteoarthritis, Arthritis Rheum 60 (2009) 325-334.
[54] P. Kannu, J. Bateman, R. Savarirayan, Clinical phenotypes associated with type II collagen mutations, J Paediatr Child Health 48 (2011) E38-43.
[55] B.A. Alman, Skeletal dysplasias and the growth plate, Clin Genet 73 (2008) 24-30.
[56] T. Wagner, J. Wirth, J. Meyer, B. Zabel, M. Held, J. Zimmer, J. Pasantes, F.D. Bricarelli, J. Keutel, E. Hustert, U. Wolf, N. Tommerup, W. Schempp, G. Scherer, Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9, Cell 79 (1994) 1111-1120.

[57] R. Nishimura, K. Hata, T. Matsubara, M. Wakabayashi, T. Yoneda, Regulation of bone and cartilage development by network between BMP signalling and transcription factors, J Biochem 151 247-254.
[58] S. Mundlos, Cleidocranial dysplasia: clinical and molecular genetics, J Med Genet 36 (1999) 177-182.
[59] P. Hermanns, B. Lee, Transcriptional dysregulation in skeletal malformation syndromes, Am J Med Genet 106 (2001) 258-271.
[60] M. Czarny-Ratajczak, J. Lohiniva, P. Rogala, K. Kozlowski, M. Perala, L. Carter, T.D. Spector, L. Kolodziej, U. Seppanen, R. Glazar, J. Krolewski, A. Latos-Bielenska, L. Ala-Kokko, A mutation in COL9A1 causes multiple epiphyseal dysplasia: further evidence for locus heterogeneity, Am J Hum Genet 69 (2001) 969-980.
[61] R.P. Boot-Handford, M.D. Briggs, The unfolded protein response and its relevance to connective tissue diseases, Cell Tissue Res 339 (2010) 197-211.
[62] A. Superti-Furga, S. Unger, Nosology and classification of genetic skeletal disorders: 2006 revision, Am J Med Genet A 143 (2007) 1-18.
[63] W.A. Horton, J.G. Hall, J.T. Hecht, Achondroplasia, Lancet 370 (2007) 162-172.
[64] H. Juppner, Jansen's metaphyseal chondrodysplasia A disorder due to a PTH/PTHrP receptor gene mutation, Trends Endocrinol Metab 7 (1996) 157-162.
[65] R.J. Oostra, J.J. van der Harten, W.P. Rijnders, R.J. Scott, M.P. Young, D. Trump, Blomstrand osteochondrodysplasia: three novel cases and histological evidence for heterogeneity, Virchows Arch 436 (2000) 28-35.
[66] J.T. Hecht, D. Montufar-Solis, G. Decker, J. Lawler, K. Daniels, P.J. Duke, Retention of cartilage oligomeric matrix protein (COMP) and cell death in redifferentiated pseudoachondroplasia chondrocytes, Matrix Biol 17 (1998) 625-633.
[67] Y. Hashimoto, T. Tomiyama, Y. Yamano, H. Mori, Mutation (D472Y) in the type 3 repeat domain of cartilage oligomeric matrix protein affects its early vesicle trafficking in endoplasmic reticulum and induces apoptosis, Am J Pathol 163 (2003) 101-110.
[68] K.A. Pirog-Garcia, R.S. Meadows, L. Knowles, D. Heinegard, D.J. Thornton, K.E. Kadler, R.P. Boot-Handford, M.D. Briggs, Reduced cell proliferation and increased apoptosis are significant pathological mechanisms in a murine model of mild pseudoachondroplasia resulting from a mutation in the C-terminal domain of COMP, Hum Mol Genet 16 (2007) 2072-2088.
[69] S. Unger, L. Bonafe, A. Superti-Furga, Multiple epiphyseal dysplasia: clinical and radiographic features, differential diagnosis and molecular basis, Best Pract Res Clin Rheumatol 22 (2008) 19-32.
[70] M.P. Leighton, S. Nundlall, T. Starborg, R.S. Meadows, F. Suleman, L. Knowles, R. Wagener, D.J. Thornton, K.E. Kadler, R.P. Boot-Handford, M.D. Briggs, Decreased chondrocyte proliferation and dysregulated apoptosis in the cartilage growth plate are key features of a murine model of epiphyseal dysplasia caused by a matn3 mutation, Hum Mol Genet 16 (2007) 1728-1741.
[71] E.M. Carter, C.L. Raggio, Genetic and orthopedic aspects of collagen disorders, Curr Opin Pediatr 21 (2009) 46-54.
[72] T. Furuichi, H. Masuya, T. Murakami, K. Nishida, G. Nishimura, T. Suzuki, K. Imaizumi, T. Kudo, K. Ohkawa, S. Wakana, S. Ikegawa, ENU-induced missense mutation in the C-propeptide coding region of Col2a1 creates a mouse model of

platyspondylic lethal skeletal dysplasia, Torrance type, Mamm Genome 22 (2011) 318-328.
[73] M.H. Rajpar, B. McDermott, L. Kung, R. Eardley, L. Knowles, M. Heeran, D.J. Thornton, R. Wilson, J.F. Bateman, R. Poulsom, P. Arvan, K.E. Kadler, M.D. Briggs, R.P. Boot-Handford, Targeted induction of endoplasmic reticulum stress induces cartilage pathology, PLoS Genet 5 (2009) e1000691.
[74] H. Lorenz, W. Richter, Osteoarthritis: cellular and molecular changes in degenerating cartilage, Prog Histochem Cytochem 40 (2006) 135-163.
[75] J.S. Gregory, J.H. Waarsing, J. Day, H.A. Pols, M. Reijman, H. Weinans, R.M. Aspden, Early identification of radiographic osteoarthritis of the hip using an active shape model to quantify changes in bone morphometric features: can hip shape tell us anything about the progression of osteoarthritis?, Arthritis Rheum 56 (2007) 3634-3643.
[76] M.B. Goldring, S.R. Goldring, Osteoarthritis, J Cell Physiol 213 (2007) 626-634.
[77] T. Aigner, A. Sachse, P.M. Gebhard, H.I. Roach, Osteoarthritis: pathobiology-targets and ways for therapeutic intervention, Adv Drug Deliv Rev 58 (2006) 128-149.
[78] M. Hashimoto, T. Nakasa, T. Hikata, H. Asahara, Molecular network of cartilage homeostasis and osteoarthritis, Med Res Rev 28 (2008) 464-481.
[79] J. Martel-Pelletier, C. Boileau, J.P. Pelletier, P.J. Roughley, Cartilage in normal and osteoarthritis conditions, Best Pract Res Clin Rheumatol 22 (2008) 351-384.
[80] D. Studer, C. Millan, E. Ozturk, K. Maniura-Weber, M. Zenobi-Wong, Molecular and biophysical mechanisms regulating hypertrophic differentiation in chondrocytes and mesenchymal stem cells, Eur Cell Mater 24 (2012) 118-135; discussion 135.
[81] M.D. Sewell, K. Rosendahl, D.M. Eastwood, Developmental dysplasia of the hip, BMJ 339 (2009) b4454.
[82] D. Shi, J. Dai, S. Ikegawa, Q. Jiang, Genetic study on developmental dysplasia of the hip, Eur J Clin Invest 42 (2012) 1121-1125.
[83] J. Bracken, T. Tran, M. Ditchfield, Developmental dysplasia of the hip: controversies and current concepts, J Paediatr Child Health 48 (2012) 963-972; quiz 972-963.
[84] A. Mabuchi, S. Nakamura, Y. Takatori, S. Ikegawa, Familial osteoarthritis of the hip joint associated with acetabular dysplasia maps to chromosome 13q, Am J Hum Genet 79 (2006) 163-168.
[85] L.Y. Li, X.K. Sun, Q. Zhao, L.J. Zhang, Q.W. Li, L.L. Wang, H. Gao, [Gene mapping of developmental dysplasia of the hip in chromosome 17q21 region], Zhonghua Yi Xue Yi Chuan Xue Za Zhi 27 (2010) 620-625.
[86] J. Dai, D. Shi, P. Zhu, J. Qin, H. Ni, Y. Xu, C. Yao, L. Zhu, H. Zhu, B. Zhao, J. Wei, B. Liu, S. Ikegawa, Q. Jiang, Y. Ding, Association of a single nucleotide polymorphism in growth differentiate factor 5 with congenital dysplasia of the hip: a case-control study, Arthritis Res Ther 10 (2008) R126.
[87] K. Rouault, V. Scotet, S. Autret, F. Gaucher, F. Dubrana, D. Tanguy, C.Y. El Rassi, B. Fenoll, C. Ferec, Evidence of association between GDF5 polymorphisms and congenital dislocation of the hip in a Caucasian population, Osteoarthritis Cartilage 18 (2010) 1144-1149.
[88] R.B. Vaes, F. Rivadeneira, J.M. Kerkhof, A. Hofman, H.A. Pols, A.G. Uitterlinden, J.B. van Meurs, Genetic variation in the GDF5 region is associated

with osteoarthritis, height, hip axis length and fracture risk: the Rotterdam study, Ann Rheum Dis 68 (2009) 1754-1760.
[89] T. Nakamura, D. Shi, M. Tzetis, J. Rodriguez-Lopez, Y. Miyamoto, A. Tsezou, A. Gonzalez, Q. Jiang, N. Kamatani, J. Loughlin, S. Ikegawa, Meta-analysis of association between the ASPN D-repeat and osteoarthritis, Hum Mol Genet 16 (2007) 1676-1681.
[90] D. Shi, J. Dai, P. Zhu, J. Qin, L. Zhu, H. Zhu, B. Zhao, X. Qiu, Z. Xu, D. Chen, L. Yi, S. Ikegawa, Q. Jiang, Association of the D repeat polymorphism in the ASPN gene with developmental dysplasia of the hip: a case-control study in Han Chinese, Arthritis Res Ther 13 (2011) R27.
[91] J.C. Baker-Lepain, J.A. Lynch, N. Parimi, C.E. McCulloch, M.C. Nevitt, M. Corr, N.E. Lane, Variant alleles of the Wnt antagonist FRZB are determinants of hip shape and modify the relationship between hip shape and osteoarthritis, Arthritis Rheum 64 (2012) 1457-1465.
[92] A.M. Gorman, S.J. Healy, R. Jager, A. Samali, Stress management at the ER: regulators of ER stress-induced apoptosis, Pharmacol Ther 134 (2012) 306-316.
[93] S.B. Stephens, R.D. Dodd, J.W. Brewer, P.J. Lager, J.D. Keene, C.V. Nicchitta, Stable ribosome binding to the endoplasmic reticulum enables compartment-specific regulation of mRNA translation, Mol Biol Cell 16 (2005) 5819-5831.
[94] B.D. Hamman, L.M. Hendershot, A.E. Johnson, BiP maintains the permeability barrier of the ER membrane by sealing the lumenal end of the translocon pore before and early in translocation, Cell 92 (1998) 747-758.
[95] M. Molinari, N-glycan structure dictates extension of protein folding or onset of disposal, Nat Chem Biol 3 (2007) 313-320.
[96] J.D. Oliver, H.L. Roderick, D.H. Llewellyn, S. High, ERp57 functions as a subunit of specific complexes formed with the ER lectins calreticulin and calnexin, Mol Biol Cell 10 (1999) 2573-2582.
[97] L.W. Ruddock, M. Molinari, N-glycan processing in ER quality control, J Cell Sci 119 (2006) 4373-4380.
[98] A.A. McCracken, J.L. Brodsky, Recognition and delivery of ERAD substrates to the proteasome and alternative paths for cell survival, Curr Top Microbiol Immunol 300 (2005) 17-40.
[99] H.P. Harding, Y. Zhang, D. Ron, Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase, Nature 397 (1999) 271-274.
[100] H.P. Harding, Y. Zhang, A. Bertolotti, H. Zeng, D. Ron, Perk is essential for translational regulation and cell survival during the unfolded protein response, Mol Cell 5 (2000) 897-904.
[101] R.J. Kaufman, Regulation of mRNA translation by protein folding in the endoplasmic reticulum, Trends Biochem Sci 29 (2004) 152-158.
[102] P.D. Lu, H.P. Harding, D. Ron, Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response, J Cell Biol 167 (2004) 27-33.
[103] H.P. Harding, Y. Zhang, H. Zeng, I. Novoa, P.D. Lu, M. Calfon, N. Sadri, C. Yun, B. Popko, R. Paules, D.F. Stojdl, J.C. Bell, T. Hettmann, J.M. Leiden, D. Ron, An integrated stress response regulates amino acid metabolism and resistance to oxidative stress, Mol Cell 11 (2003) 619-633.
[104] J. Shen, X. Chen, L. Hendershot, R. Prywes, ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of Golgi localization signals, Developmental cell 3 (2002) 99-111.

[105] J. Ye, R.B. Rawson, R. Komuro, X. Chen, U.P. Dave, R. Prywes, M.S. Brown, J.L. Goldstein, ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs, Mol Cell 6 (2000) 1355-1364.
[106] Y. Adachi, K. Yamamoto, T. Okada, H. Yoshida, A. Harada, K. Mori, ATF6 is a transcription factor specializing in the regulation of quality control proteins in the endoplasmic reticulum, Cell Struct Funct 33 (2008) 75-89.
[107] S. Kondo, T. Murakami, K. Tatsumi, M. Ogata, S. Kanemoto, K. Otori, K. Iseki, A. Wanaka, K. Imaizumi, OASIS, a CREB/ATF-family member, modulates UPR signalling in astrocytes, Nat Cell Biol 7 (2005) 186-194.
[108] S. Kondo, A. Saito, S. Hino, T. Murakami, M. Ogata, S. Kanemoto, S. Nara, A. Yamashita, K. Yoshinaga, H. Hara, K. Imaizumi, BBF2H7, a novel transmembrane bZIP transcription factor, is a new type of endoplasmic reticulum stress transducer, Mol Cell Biol 27 (2007) 1716-1729.
[109] C.Y. Liu, M. Schroder, R.J. Kaufman, Ligand-independent dimerization activates the stress response kinases IRE1 and PERK in the lumen of the endoplasmic reticulum, J Biol Chem 275 (2000) 24881-24885.
[110] J. Zhou, C.Y. Liu, S.H. Back, R.L. Clark, D. Peisach, Z. Xu, R.J. Kaufman, The crystal structure of human IRE1 luminal domain reveals a conserved dimerization interface required for activation of the unfolded protein response, Proc Natl Acad Sci U S A 103 (2006) 14343-14348.
[111] H. Yoshida, T. Matsui, A. Yamamoto, T. Okada, K. Mori, XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor, Cell 107 (2001) 881-891.
[112] D. Acosta-Alvear, Y. Zhou, A. Blais, M. Tsikitis, N.H. Lents, C. Arias, C.J. Lennon, Y. Kluger, B.D. Dynlacht, XBP1 controls diverse cell type- and condition-specific transcriptional regulatory networks, Mol Cell 27 (2007) 53-66.
[113] J. Hollien, J.S. Weissman, Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response, Science (New York, N.Y 313 (2006) 104-107.
[114] M. Ogata, S. Hino, A. Saito, K. Morikawa, S. Kondo, S. Kanemoto, T. Murakami, M. Taniguchi, I. Tanii, K. Yoshinaga, S. Shiosaka, J.A. Hammarback, F. Urano, K. Imaizumi, Autophagy is activated for cell survival after endoplasmic reticulum stress, Mol Cell Biol 26 (2006) 9220-9231.
[115] P. Zhang, B. McGrath, S. Li, A. Frank, F. Zambito, J. Reinert, M. Gannon, K. Ma, K. McNaughton, D.R. Cavener, The PERK eukaryotic initiation factor 2 alpha kinase is required for the development of the skeletal system, postnatal growth, and the function and viability of the pancreas, Mol Cell Biol 22 (2002) 3864-3874.
[116] X. Yang, K. Matsuda, P. Bialek, S. Jacquot, H.C. Masuoka, T. Schinke, L. Li, S. Brancorsini, P. Sassone-Corsi, T.M. Townes, A. Hanauer, G. Karsenty, ATF4 is a substrate of RSK2 and an essential regulator of osteoblast biology; implication for Coffin-Lowry Syndrome, Cell 117 (2004) 387-398.
[117] T.L. Cameron, K.M. Bell, L. Tatarczuch, E.J. Mackie, M.H. Rajpar, B.T. McDermott, R.P. Boot-Handford, J.F. Bateman, Transcriptional profiling of chondrodysplasia growth plate cartilage reveals adaptive ER-stress networks that allow survival but disrupt hypertrophy, PLoS One 6 (2011) e24600.
[118] Y. Honma, K. Kanazawa, T. Mori, Y. Tanno, M. Tojo, H. Kiyosawa, J. Takeda, T. Nikaido, T. Tsukamoto, S. Yokoya, A. Wanaka, Identification of a novel gene, OASIS, which encodes for a putative CREB/ATF family transcription

factor in the long-term cultured astrocytes and gliotic tissue, Brain Res Mol Brain Res 69 (1999) 93-103.
[119] T. Nikaido, S. Yokoya, T. Mori, S. Hagino, K. Iseki, Y. Zhang, M. Takeuchi, H. Takaki, S. Kikuchi, A. Wanaka, Expression of the novel transcription factor OASIS, which belongs to the CREB/ATF family, in mouse embryo with special reference to bone development, Histochem Cell Biol 116 (2001) 141-148.
[120] T. Murakami, S. Kondo, M. Ogata, S. Kanemoto, A. Saito, A. Wanaka, K. Imaizumi, Cleavage of the membrane-bound transcription factor OASIS in response to endoplasmic reticulum stress, J Neurochem 96 (2006) 1090-1100.
[121] T. Murakami, A. Saito, S. Hino, S. Kondo, S. Kanemoto, K. Chihara, H. Sekiya, K. Tsumagari, K. Ochiai, K. Yoshinaga, M. Saitoh, R. Nishimura, T. Yoneda, I. Kou, T. Furuichi, S. Ikegawa, M. Ikawa, M. Okabe, A. Wanaka, K. Imaizumi, Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation, Nat Cell Biol 11 (2009) 1205-1211.
[122] A. Saito, S. Hino, T. Murakami, S. Kanemoto, S. Kondo, M. Saitoh, R. Nishimura, T. Yoneda, T. Furuichi, S. Ikegawa, M. Ikawa, M. Okabe, K. Imaizumi, Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis, Nat Cell Biol 11 (2009) 1197-1204.
[123] S. Izumi, A. Saito, S. Kanemoto, N. Kawasaki, R. Asada, H. Iwamoto, M. Oki, H. Miyagi, M. Ochi, K. Imaizumi, The endoplasmic reticulum stress transducer BBF2H7 suppresses apoptosis by activating the ATF5-MCL1 pathway in growth plate cartilage, J Biol Chem 287 (2012) 36190-36200.
[124] F. Guo, E.A. Lin, P. Liu, J. Lin, C. Liu, XBP1U inhibits the XBP1S-mediated upregulation of the iNOS gene expression in mammalian ER stress response, Cell Signal 22 (2010) 1818-1828.
[125] K. Yanagitani, Y. Kimata, H. Kadokura, K. Kohno, Translational pausing ensures membrane targeting and cytoplasmic splicing of XBP1u mRNA, Science (New York, N.Y 331 (2011) 586-589.
[126] Y. Liu, J. Zhou, W. Zhao, X. Li, R. Jiang, C. Liu, F.J. Guo, XBP1S associates with RUNX2 and regulates chondrocyte hypertrophy, J Biol Chem 287 (2012) 34500-34513.
[127] W.G. Jang, E.J. Kim, D.K. Kim, H.M. Ryoo, K.B. Lee, S.H. Kim, H.S. Choi, J.T. Koh, BMP2 protein regulates osteocalcin expression via Runx2-mediated Atf6 gene transcription, J Biol Chem 287 (2012) 905-915.
[128] J. Wei, X. Sheng, D. Feng, B. McGrath, D.R. Cavener, PERK is essential for neonatal skeletal development to regulate osteoblast proliferation and differentiation, J Cell Physiol 217 (2008) 693-707.
[129] A. Saito, K. Ochiai, S. Kondo, K. Tsumagari, T. Murakami, D.R. Cavener, K. Imaizumi, Endoplasmic reticulum stress response mediated by the PERK-eIF2(alpha)-ATF4 pathway is involved in osteoblast differentiation induced by BMP2, J Biol Chem 286 (2011) 4809-4818.
[130] X. Yang, G. Karsenty, ATF4, the osteoblast accumulation of which is determined post-translationally, can induce osteoblast-specific gene expression in non-osteoblastic cells, J Biol Chem 279 (2004) 47109-47114.
[131] A. Hershko, A. Ciechanover, The ubiquitin system, Annu Rev Biochem 67 (1998) 425-479.
[132] K.H. Ventii, K.D. Wilkinson, Protein partners of deubiquitinating enzymes, Biochem J 414 (2008) 161-175.

[133] C.M. Pickart, Mechanisms underlying ubiquitination, Annu Rev Biochem 70 (2001) 503-533.
[134] A.Y. Amerik, M. Hochstrasser, Mechanism and function of deubiquitinating enzymes, Biochim Biophys Acta 1695 (2004) 189-207.
[135] O. Kerscher, R. Felberbaum, M. Hochstrasser, Modification of proteins by ubiquitin and ubiquitin-like proteins, Annu Rev Cell Dev Biol 22 (2006) 159-180.
[136] C.M. Pickart, M.J. Eddins, Ubiquitin: structures, functions, mechanisms, Biochim Biophys Acta 1695 (2004) 55-72.
[137] R.D. Vierstra, The ubiquitin-26S proteasome system at the nexus of plant biology, Nat Rev Mol Cell Biol 10 (2009) 385-397.
[138] A. Hershko, A. Ciechanover, The ubiquitin system for protein degradation, Annu Rev Biochem 61 (1992) 761-807.
[139] B.T. Dye, B.A. Schulman, Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins, Annu Rev Biophys Biomol Struct 36 (2007) 131-150.
[140] P.K. Jackson, A.G. Eldridge, E. Freed, L. Furstenthal, J.Y. Hsu, B.K. Kaiser, J.D. Reimann, The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases, Trends Cell Biol 10 (2000) 429-439.
[141] D.M. Cyr, J. Hohfeld, C. Patterson, Protein quality control: U-box-containing E3 ubiquitin ligases join the fold, Trends Biochem Sci 27 (2002) 368-375.
[142] Y. Kee, J.M. Huibregtse, Regulation of catalytic activities of HECT ubiquitin ligases, Biochem Biophys Res Commun 354 (2007) 329-333.
[143] L. Shearwin-Whyatt, H.E. Dalton, N. Foot, S. Kumar, Regulation of functional diversity within the Nedd4 family by accessory and adaptor proteins, Bioessays 28 (2006) 617-628.
[144] F.E. Reyes-Turcu, K.H. Ventii, K.D. Wilkinson, Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes, Annu Rev Biochem 78 (2009) 363-397.
[145] E.S. Johnson, Protein modification by SUMO, Annu Rev Biochem 73 (2004) 355-382.
[146] R.T. Hay, SUMO: a history of modification, Mol Cell 18 (2005) 1-12.
[147] G. Gill, SUMO and ubiquitin in the nucleus: different functions, similar mechanisms?, Genes Dev 18 (2004) 2046-2059.
[148] J.S. Seeler, A. Dejean, Nuclear and unclear functions of SUMO, Nat Rev Mol Cell Biol 4 (2003) 690-699.
[149] M. Garcia-Dominguez, J.C. Reyes, SUMO association with repressor complexes, emerging routes for transcriptional control, Biochim Biophys Acta (2009).
[150] T. Kamitani, K. Kito, H.P. Nguyen, E.T. Yeh, Characterization of NEDD8, a developmentally down-regulated ubiquitin-like protein, J Biol Chem 272 (1997) 28557-28562.
[151] G. Rabut, M. Peter, Function and regulation of protein neddylation. 'Protein modifications: beyond the usual suspects' review series, EMBO Rep 9 (2008) 969-976.
[152] D.C. Schwartz, M. Hochstrasser, A superfamily of protein tags: ubiquitin, SUMO and related modifiers, Trends Biochem Sci 28 (2003) 321-328.
[153] Y. Ohki, N. Funatsu, N. Konishi, T. Chiba, The mechanism of poly-NEDD8 chain formation in vitro, Biochem Biophys Res Commun 381 (2009) 443-447.

[154] T. Tanaka, T. Nakatani, T. Kamitani, Inhibition of NEDD8-conjugation pathway by novel molecules: potential approaches to anticancer therapy, Mol Oncol 6 (2012) 267-275.
[155] K.I. Kim, D.E. Zhang, ISG15, not just another ubiquitin-like protein, Biochem Biophys Res Commun 307 (2003) 431-434.
[156] K.J. Ritchie, D.E. Zhang, ISG15: the immunological kin of ubiquitin, Semin Cell Dev Biol 15 (2004) 237-246.
[157] C. Zhao, T.Y. Hsiang, R.L. Kuo, R.M. Krug, ISG15 conjugation system targets the viral NS1 protein in influenza A virus-infected cells, Proc Natl Acad Sci U S A 107 (2010) 2253-2258.
[158] C. Zhao, M.N. Collins, T.Y. Hsiang, R.M. Krug, Interferon-induced ISG15 pathway: an ongoing virus-host battle, Trends Microbiol 21 (2013) 181-186.
[159] L.A. Durfee, N. Lyon, K. Seo, J.M. Huibregtse, The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15, Mol Cell 38 (2010) 722-732.
[160] M. Sakoh-Nakatogawa, K. Matoba, E. Asai, H. Kirisako, J. Ishii, N.N. Noda, F. Inagaki, H. Nakatogawa, Y. Ohsumi, Atg12-Atg5 conjugate enhances E2 activity of Atg3 by rearranging its catalytic site, Nat Struct Mol Biol 20 (2013) 433-439.
[161] P.G. Pedrioli, S. Leidel, K. Hofmann, Urm1 at the crossroad of modifications. 'Protein Modifications: Beyond the Usual Suspects' Review Series, EMBO Rep 9 (2008) 1196-1202.
[162] F. Wang, M. Liu, R. Qiu, C. Ji, The dual role of ubiquitin-like protein Urm1 as a protein modifier and sulfur carrier, Protein Cell 2 (2011) 612-619.
[163] Y.H. Chiu, Q. Sun, Z.J. Chen, E1-L2 activates both ubiquitin and FAT10, Mol Cell 27 (2007) 1014-1023.
[164] G. Schmidtke, B. Kalveram, M. Groettrup, Degradation of FAT10 by the 26S proteasome is independent of ubiquitylation but relies on NUB1L, FEBS Lett 583 (2009) 591-594.
[165] G. Schmidtke, A. Aichem, M. Groettrup, FAT10ylation as a signal for proteasomal degradation, Biochim Biophys Acta (2013).
[166] A. Ciechanover, A.L. Schwartz, The ubiquitin system: pathogenesis of human diseases and drug targeting, Biochim Biophys Acta 1695 (2004) 3-17.
[167] A.L. Schwartz, A. Ciechanover, The ubiquitin-proteasome pathway and pathogenesis of human diseases, Annu Rev Med 50 (1999) 57-74.
[168] J. Golab, T.M. Bauer, V. Daniel, C. Naujokat, Role of the ubiquitin-proteasome pathway in the diagnosis of human diseases, Clin Chim Acta 340 (2004) 27-40.
[169] Y.H. Jiang, A.L. Beaudet, Human disorders of ubiquitination and proteasomal degradation, Curr Opin Pediatr 16 (2004) 419-426.
[170] S. Hussain, Y. Zhang, P.J. Galardy, DUBs and cancer: the role of deubiquitinating enzymes as oncogenes, non-oncogenes and tumor suppressors, Cell Cycle 8 (2009) 1688-1697.
[171] M.J. Edelmann, B.M. Kessler, Ubiquitin and ubiquitin-like specific proteases targeted by infectious pathogens: Emerging patterns and molecular principles, Biochim Biophys Acta 1782 (2008) 809-816.
[172] A. Ciechanover, P. Brundin, The ubiquitin proteasome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg, Neuron 40 (2003) 427-446.
[173] E. Leroy, R. Boyer, G. Auburger, B. Leube, G. Ulm, E. Mezey, G. Harta, M.J. Brownstein, S. Jonnalagada, T. Chernova, A. Dehejia, C. Lavedan, T. Gasser,

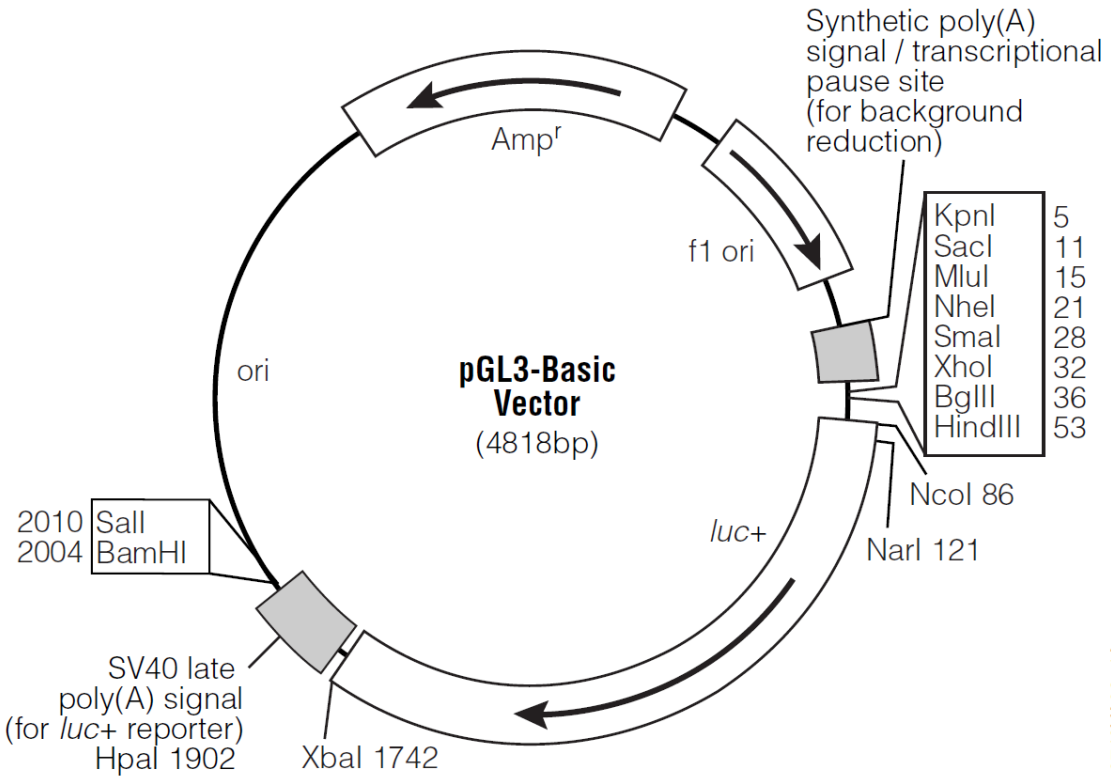
P.J. Steinbach, K.D. Wilkinson, M.H. Polymeropoulos, The ubiquitin pathway in Parkinson's disease, Nature 395 (1998) 451-452.
[174] R. Layfield, J.R. Cavey, J. Lowe, Role of ubiquitin-mediated proteolysis in the pathogenesis of neurodegenerative disorders, Ageing Res Rev 2 (2003) 343-356.
[175] V. Dorval, P.E. Fraser, SUMO on the road to neurodegeneration, Biochim Biophys Acta 1773 (2007) 694-706.
[176] F. Mori, M. Nishie, Y.S. Piao, K. Kito, T. Kamitani, H. Takahashi, K. Wakabayashi, Accumulation of NEDD8 in neuronal and glial inclusions of neurodegenerative disorders, Neuropathol Appl Neurobiol 31 (2005) 53-61.
[177] A. Ferro, A.L. Carvalho, A. Teixeira-Castro, C. Almeida, R.J. Tome, L. Cortes, A.J. Rodrigues, E. Logarinho, J. Sequeiros, S. Macedo-Ribeiro, P. Maciel, NEDD8: a new ataxin-3 interactor, Biochim Biophys Acta 1773 (2007) 1619-1627.
[178] M. Komatsu, T. Chiba, K. Tatsumi, S. Iemura, I. Tanida, N. Okazaki, T. Ueno, E. Kominami, T. Natsume, K. Tanaka, A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier, EMBO J 23 (2004) 1977-1986.
[179] C.M. Watson, Investigating the genetic basis of forms of osteoarthritis., Faculty of Medical and Human Sciences, vol. PhD thesis, University of Manchester, Manchester, 2009.
[180] S.H. Kang, G.R. Kim, M. Seong, S.H. Baek, J.H. Seol, O.S. Bang, H. Ovaa, K. Tatsumi, M. Komatsu, K. Tanaka, C.H. Chung, Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2, J Biol Chem 282 (2007) 5256-5262.
[181] K. Tatsumi, Y.S. Sou, N. Tada, E. Nakamura, S. Iemura, T. Natsume, S.H. Kang, C.H. Chung, M. Kasahara, E. Kominami, M. Yamamoto, K. Tanaka, M. Komatsu, A novel type of E3 ligase for the Ufm1 conjugation system, J Biol Chem 285 (2010) 5417-5427.
[182] H. Shiwaku, N. Yoshimura, T. Tamura, M. Sone, S. Ogishima, K. Watase, K. Tagawa, H. Okazawa, Suppression of the novel ER protein Maxer by mutant ataxin-1 in Bergman glia contributes to non-cell-autonomous toxicity, EMBO J 29 (2010) 2446-2460.
[183] K. Hofmann, P. Bucher, The PCI domain: a common theme in three multiprotein complexes, Trends Biochem Sci 23 (1998) 204-205.
[184] K. Lemaire, R.F. Moura, M. Granvik, M. Igoillo-Esteve, H.E. Hohmeier, N. Hendrickx, C.B. Newgard, E. Waelkens, M. Cnop, F. Schuit, Ubiquitin fold modifier 1 (UFM1) and its target UFBP1 protect pancreatic beta cells from ER stress-induced apoptosis, PLoS One 6 (2011) e18517.
[185] J. Wu, G. Lei, M. Mei, Y. Tang, H. Li, A novel C53/LZAP-interacting protein regulates stability of C53/LZAP and DDRGK domain-containing Protein 1 (DDRGK1) and modulates NF-kappaB signaling, J Biol Chem 285 (2010) 15126-15136.
[186] J. Kwon, H.J. Cho, S.H. Han, J.G. No, J.Y. Kwon, H. Kim, A novel LZAP-binding protein, NLBP, inhibits cell invasion, J Biol Chem 285 (2010) 12232-12240.
[187] A. Azfer, J. Niu, L.M. Rogers, F.M. Adamski, P.E. Kolattukudy, Activation of endoplasmic reticulum stress response during the development of ischemic heart disease, Am J Physiol Heart Circ Physiol 291 (2006) H1411-1420.
[188] H. Lu, Y. Yang, E.M. Allister, N. Wijesekara, M.B. Wheeler, The identification of potential factors associated with the development of type 2 diabetes: a quantitative proteomics approach, Mol Cell Proteomics 7 (2008) 1434-1451.

[189] Y. Zhang, M. Zhang, J. Wu, G. Lei, H. Li, Transcriptional regulation of the Ufm1 conjugation system in response to disturbance of the endoplasmic reticulum homeostasis and inhibition of vesicle trafficking, PLoS One 7 (2012) e48587.
[190] P. Hertel, J. Daniel, D. Stegehake, H. Vaupel, S. Kailayangiri, C. Gruel, C. Woltersdorf, E. Liebau, The ubiquitin-fold modifier 1 (Ufm1)-cascade of Caenorhabditis elegans, J Biol Chem (2013).
[191] K. Tatsumi, H. Yamamoto-Mukai, R. Shimizu, S. Waguri, Y.S. Sou, A. Sakamoto, C. Taya, H. Shitara, T. Hara, C.H. Chung, K. Tanaka, M. Yamamoto, M. Komatsu, The Ufm1-activating enzyme Uba5 is indispensable for erythroid differentiation in mice, Nat Commun 2 (2011) 181.
[192] S. Gannavaram, P. Sharma, R.C. Duncan, P. Salotra, H.L. Nakhasi, Mitochondrial associated ubiquitin fold modifier-1 mediated protein conjugation in Leishmania donovani, PLoS One 6 (2011) e16156.
[193] S. Gannavaram, P.S. Connelly, M.P. Daniels, R. Duncan, P. Salotra, H.L. Nakhasi, Deletion of mitochondrial associated ubiquitin fold modifier protein Ufm1 in Leishmania donovani results in loss of beta-oxidation of fatty acids and blocks cell division in the amastigote stage, Mol Microbiol 86 (2012) 187-198.
[194] B.H. Ha, Y.J. Jeon, S.C. Shin, K. Tatsumi, M. Komatsu, K. Tanaka, C.M. Watson, G. Wallis, C.H. Chung, E.E. Kim, Structure of ubiquitin-fold modifier 1-specific protease UfSP2, J Biol Chem 286 (2011) 10248-10257.
[195] R.K. Smith, L. Zunino, P.M. Webbon, D. Heinegard, The distribution of cartilage oligomeric matrix protein (COMP) in tendon and its variation with tendon site, age and load, Matrix Biol 16 (1997) 255-271.
[196] D. Patra, X. Xing, S. Davies, J. Bryan, C. Franz, E.B. Hunziker, L.J. Sandell, Site-1 protease is essential for endochondral bone formation in mice, J Cell Biol 179 (2007) 687-700.
[197] J. Kahn, Y. Shwartz, E. Blitz, S. Krief, A. Sharir, D.A. Breitel, R. Rattenbach, F. Relaix, P. Maire, R.B. Rountree, D.M. Kingsley, E. Zelzer, Muscle contraction is necessary to maintain joint progenitor cell fate, Developmental cell 16 (2009) 734-743.
[198] N.C. Nowlan, C. Bourdon, G. Dumas, S. Tajbakhsh, P.J. Prendergast, P. Murphy, Developing bones are differentially affected by compromised skeletal muscle formation, Bone 46 (2009) 1275-1285.
[199] N. Ghosh-Choudhury, J.J. Windle, B.A. Koop, M.A. Harris, D.L. Guerrero, J.M. Wozney, G.R. Mundy, S.E. Harris, Immortalized murine osteoblasts derived from BMP 2-T-antigen expressing transgenic mice, Endocrinology 137 (1996) 331-339.
[200] A. Jikko, S.E. Harris, D. Chen, D.L. Mendrick, C.H. Damsky, Collagen integrin receptors regulate early osteoblast differentiation induced by BMP-2, J Bone Miner Res 14 (1999) 1075-1083.
[201] A.R. Haas, R.S. Tuan, Murine C3H10T1/2 multipotential cells as an in vitro model of mesenchymal chondrogenesis, Methods Mol Biol 137 (2000) 383-389.
[202] D. Yaffe, O. Saxel, Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle, Nature 270 (1977) 725-727.
[203] H.T. Harcke, M.S. Lee, L. Sinning, N.M. Clarke, P.F. Borns, G.D. MacEwen, Ossification center of the infant hip: sonographic and radiographic correlation, AJR Am J Roentgenol 147 (1986) 317-321.
[204] A.G.R.a.H.E.R. John J. Callaghan, The Adult Hip 2007.

[205] A.G.R. John J. Callaghan, Harry E. Rubash and Peter D. Simonian, The Adult Knee 2002.
[206] C.K. Li, P. Knopp, H. Moncrieffe, B. Singh, S. Shah, K. Nagaraju, H. Varsani, B. Gao, L.R. Wedderburn, Overexpression of MHC class I heavy chain protein in young skeletal muscle leads to severe myositis: implications for juvenile myositis, Am J Pathol 175 (2009) 1030-1040.
[207] G. Wolf, Function of the bone protein osteocalcin: definitive evidence, Nutr Rev 54 (1996) 332-333.
[208] J.M. Desterro, M.S. Rodriguez, R.T. Hay, SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation, Mol Cell 2 (1998) 233-239.
[209] M.J. Matunis, E. Coutavas, G. Blobel, A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex, J Cell Biol 135 (1996) 1457-1470.
[210] A.A. Fatokun, T.W. Stone, R.A. Smith, Hydrogen peroxide-induced oxidative stress in MC3T3-E1 cells: The effects of glutamate and protection by purines, Bone 39 (2006) 542-551.
[211] M. Kaneko, M. Ishiguro, Y. Niinuma, M. Uesugi, Y. Nomura, Human HRD1 protects against ER stress-induced apoptosis through ER-associated degradation, FEBS Lett 532 (2002) 147-152.
[212] B.G. Hill, C. Reily, J.Y. Oh, M.S. Johnson, A. Landar, Methods for the determination and quantification of the reactive thiol proteome, Free Radic Biol Med 47 (2009) 675-683.
[213] T.G. Schmidt, A. Skerra, The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins, Nature protocols 2 (2007) 1528-1535.
[214] P. Hengen, Purification of His-Tag fusion proteins from Escherichia coli, Trends Biochem Sci 20 (1995) 285-286.
[215] N. Rosenthal, S. Brown, The mouse ascending: perspectives for human-disease models, Nat Cell Biol 9 (2007) 993-999.
[216] S.N. Hardouin, A. Nagy, Mouse models for human disease, Clin Genet 57 (2000) 237-244.
[217] J.A. Fischer-Vize, G.M. Rubin, R. Lehmann, The fat facets gene is required for Drosophila eye and embryo development, Development 116 (1992) 985-1000.
[218] A.L. Cadavid, A. Ginzel, J.A. Fischer, The function of the Drosophila fat facets deubiquitinating enzyme in limiting photoreceptor cell number is intimately associated with endocytosis, Development 127 (2000) 1727-1736.
[219] S. Dupont, A. Mamidi, M. Cordenonsi, M. Montagner, L. Zacchigna, M. Adorno, G. Martello, M.J. Stinchfield, S. Soligo, L. Morsut, M. Inui, S. Moro, N. Modena, F. Argenton, S.J. Newfeld, S. Piccolo, FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls Smad4 monoubiquitination, Cell 136 (2009) 123-135.
[220] J. Cheng, T. Bawa, P. Lee, L. Gong, E.T. Yeh, Role of desumoylation in the development of prostate cancer, Neoplasia (New York, N.Y 8 (2006) 667-676.
[221] J.H. Kim, S.H. Baek, Emerging roles of desumoylating enzymes, Biochimica et biophysica acta 1792 (2009) 155-162.
[222] S. Martin, K.A. Wilkinson, A. Nishimune, J.M. Henley, Emerging extranuclear roles of protein SUMOylation in neuronal function and dysfunction, Nat Rev Neurosci 8 (2007) 948-959.

[223] T. Yamaguchi, P. Sharma, M. Athanasiou, A. Kumar, S. Yamada, M.R. Kuehn, Mutation of SENP1/SuPr-2 reveals an essential role for desumoylation in mouse development, Molecular and cellular biology 25 (2005) 5171-5182.
[224] T. Nishida, Y. Yamada, SMT3IP1, a nucleolar SUMO-specific protease, deconjugates SUMO-2 from nucleolar and cytoplasmic nucleophosmin, Biochemical and biophysical research communications 374 (2008) 382-387.
[225] L.G. Alexopoulos, I. Youn, P. Bonaldo, F. Guilak, Developmental and osteoarthritic changes in Col6a1-knockout mice: biomechanics of type VI collagen in the cartilage pericellular matrix, Arthritis Rheum 60 (2009) 771-779.

8. Supplementary Figures



Supplementary Figure 1. pGL3-Basic luciferase vector (Promega, UK)

Ufm1

```
rat             ---------------GGAGGGCGACGTGGGCCGCGAAGGACCCCACTG----CCTGCCAT -56
mouse           ---------------GGAGGGCGACGTGGGCCGCGAAGGCCCCCGCTG----CCGGCCAT -52
human           AGGTCCCCAGATTGCAGAGGGAGACGTGGACGTGAGTGGAGCGGGGCGGTCCCCAGCACA -64
cow             AGGACCCCAGCTTGCGGAAGGAGACGTGGACCTGAGTTGGAAGAGACG----CTAGGGTG -58
                                ** ** ******* *       *        *    *  *

rat             CAAGGAAGCT---------AGGCGGGCGGCTG-TTCGCGGTTTGAGTACCAGGCGGTTCC -4
mouse           CAAGGAAGCG---------CGGCGGGCGGCTG-GGCGCGG--TGTGTGTCAGGCGGTTCC -4
human           CTAGAGGAAGTCGTGCTACCCCCGCGGAGTTGTCGTGTGTTCTGGATTCATTCCGGCACC -4
cow             CAAGTGTGAG------GTTCAGCACGGGTTTGCTGTGAGGTTTCGGTCCCGAACGGCGGC -4
                * **                  *  *    **    * *   *   *      ***   *

rat             ACCATGTAAGTGTCAGTCCCCGCGTCCCGGCCTTGACTCCGTGGCCGAA-CTCGACAGTA +56
mouse           ACCATGTAAGTATCCGTCCCTGTGTCCCGGCCTTAACCCCGTGGCCGGT-CCCGACCGTA +56
human           ACCATGTAAGTGTTTG--CTTACCGACTGCCATAATTCCTGGTCCAGCTGCCCGACCCTG +55
cow             ACCATGTAAGTGTTGG--CTTTCCAGCCGCCCAATTCCGCGGTCCCGTTGCCCGACCTTA +55
                ********** *  *  *       * * *         *   * *   * ****  *

rat             ATGACTTCTTGTTCTCTCCCACAGGTCGAAGGTGTCCTTTAAAATCACGCTGACGTCGGA +116
mouse           ACGCCTTCTCGTTCTCTCCGGCAGGTCGAAGGTGTCCTTTAAAATCACGTTGACGTCGGA +116
human           ACTCTCTCCCGCTCTTTTCCTCAGGTCGAAGGTTTCCTTTAAGATCACGCTGACGTCGGA +115
cow             ACT-TCTTCTTCTCTCTTCCTCAGGTCGAAAGTTTCTTTTAAGATCACACTGACGTCGGA +114
                *     *     *** * *  ********* ** ** ***** *****  **********

Primers:
Forward  -1545 - -1524
Reverse    -20 -    +1
UPRE       -85 -   -80
UPRE 2    +108 -  +113
```

Supplementary Figure 2. Fragment of the *Ufm1* promoter region including the UPRE sequence.

Analysis of promoter region of the *Ufm1* gene from four different species. The UPRE sequences are highlighted yellow. ATG codon is highlighted red. The first and second exon only are underlined. Asterisks denote conserved sequences. Positions of primers used to amplify the promoter region and positions of the UPRE sequences in respect to the ATG codon of the mouse gene are given below.

Uba5

```
rat             CGGCGGCGCACCGGAAGTGGCGTGC-GGAAGTCCGGTGGGCATGCCGGAGACGTGTCTTA -148
mouse           CGGCGGCGCACCGGAAGTGGCACGCCGGAAGCCCGGTGGGCGTGCCGGAGACGTGTCTTA -157
human           CTGCGACGCACCGGAAGCGGCTCCGAGGAAGGCCTGTGGGAGTCTCGGAGACGTGTCTGT -156
cow             CAGCAACGCACCGGAAGTGGCCCGCTGGGAGACCCGCAGGAGCGGCGGAGACGTGTCGGG -160
                * **  *********** ***     ** ** ** *  **     ***********

rat             CTGGGA-CCATTGGAGTCGCTTCCCC-GGTGTAGGGATCAGGAAGGCTGTGTC----TAG -94
mouse           CTGGGA-GCCTTGGAGCCGCTTCGCC-GGTGTCTGGACCTGGAAGGCTGTGTC----TGT -103
human           CTGTGAGGCGCTGGGTGCACGTCCCCAGGGCTCTGGGCTAGGAAGGCAGCGGCGAGGTGC -96
cow             CAGGGGCGCGCTGGGAGCGCTTCCCCAGGTTTGTGGGCTGGGAAAGCCGCAGCAAGGTGA -100
                * * *   *  ***   * * ** ** **  *  **    **** ** *   *    *

rat             TCCTCC---TGAACTTCGTGAGCCCAGCA---GAGCAGCGTGGGG---------CGGCGG -49
mouse           TCCTCC---TGCGCTTCGTGGGCCCAGCG---AGGCAGCGTGGGGAGGCGGCGACGACGG -49
human           CTCCCCACGTACCCCTCGCGGGCCCAGC---CGAGCAACGTGGGG---------CGAAGG -48
cow             CTCGCCACATATCCCTCTCGGGGCCAGCAGGCCAGCAGCGTGGGG---------CGGCGA -49
                  * **   *   * **  * * *****      *** *******         **  *

rat             CGACGGCGGCGGTCCGGCCTGGGATCGGGGGCCACCGGGAGCCCAGCCATG +3
mouse           CGACGGCGGCGGTCCGGCCTGGGATCGGAGGCCACCGGGAGCCCAGCCATG +3
human           CGGCGGCGAAGGCCCGGGCTGGGAGCGTTGGCGGCCGGAGTCCCAGC-ATG +3
cow             CGACGGCGACGGCCGGGGCTGGGAGCGCTGGCGGCCGGGGTCCCAGCCATG +3
                ** *****  ** * ** ****** **  ***  ****   ******
```

```
Primers:
Forward   -1348 - -1327
Reverse     -18 -    +2
UPRE       -167 -  -162
```

Supplementary Figure 3. Fragment of the *Uba5* promoter region including the UPRE sequence.

Analysis of promoter region of the *Uba5* gene from four different species. The UPRE sequence is highlighted yellow. ATG codon is highlighted red. Asterisks denote conserved sequences. Positions of primers used to amplify the promoter region and position of the UPRE sequence in respect to the ATG codon of the mouse gene are given below.

Ufl1

```
rat             ---------GGTTTTTGT-----TGAC----------GCAG-GTCTCTTGGCCCTCCGCG -81
mouse           ---------GAATTTCCT-----TGAC----------GCAG-GTCTCTTGGCCCTCTGCG -81
human           GGG-CGCAAGCGTTTGGTG----TGTCTGCGTCGGGGGCGGAGCCTCTTGTCCCTCTGCG -84
cow             AAGTCGCGCACGTTCCGTAGCTGCGTC--------GCGGAGAGTCTCTTGTCCCTCCGCG -81
                            **   *      * *          *  * * ****** ***** ***

rat             CGGCCCGTCTCAGGGTGC--CAGTGGAGGGCTGAAGGCTGACGTGTCTGCCGTTTCCCGG -23
mouse           CGGTCCGTCCCGGGGTGC--CTGCGGCAGGCTGAAGGCTGACGTGTCTGCCGTTTCCCTG -23
human           CGGCCCGTTCCGCCTCTCTTCTCCCACCGCCTGTCGGCTGACGTGTCTGCAGTTCCTCCG -24
cow             CGGCCCGTTCCGCCTCGG--CTCCCGCGACCTTTGGGCTGACGTGTCTGCGGTTTGTCGG -23
                *** ****  *         *         **   ************** ***   * *

rat             TGTC-GCGGAGAGCGGAGCAGTGATG +3
mouse           TGTC-GGGGCGAGCGGAGCAGTGATG +3
human           CGTCTACTGCGAGTCAGGCCGTGATG +3
cow             CGGCTG-GGCGAATAGGGCCGAGATG +3
                 * *    * **     ** * *

UPRE        -43 -   -38
```

Supplementary Figure 4. Fragment of the *Ufl1* promoter region including the UPRE sequence.

Analysis of promoter region of the *Ufl1* gene from four different species. The UPRE sequence is highlighted yellow. ATG codon is highlighted red. Asterisks denote conserved sequences. Position of the UPRE sequence in respect to the ATG codon of the mouse gene is given below.

LZAP

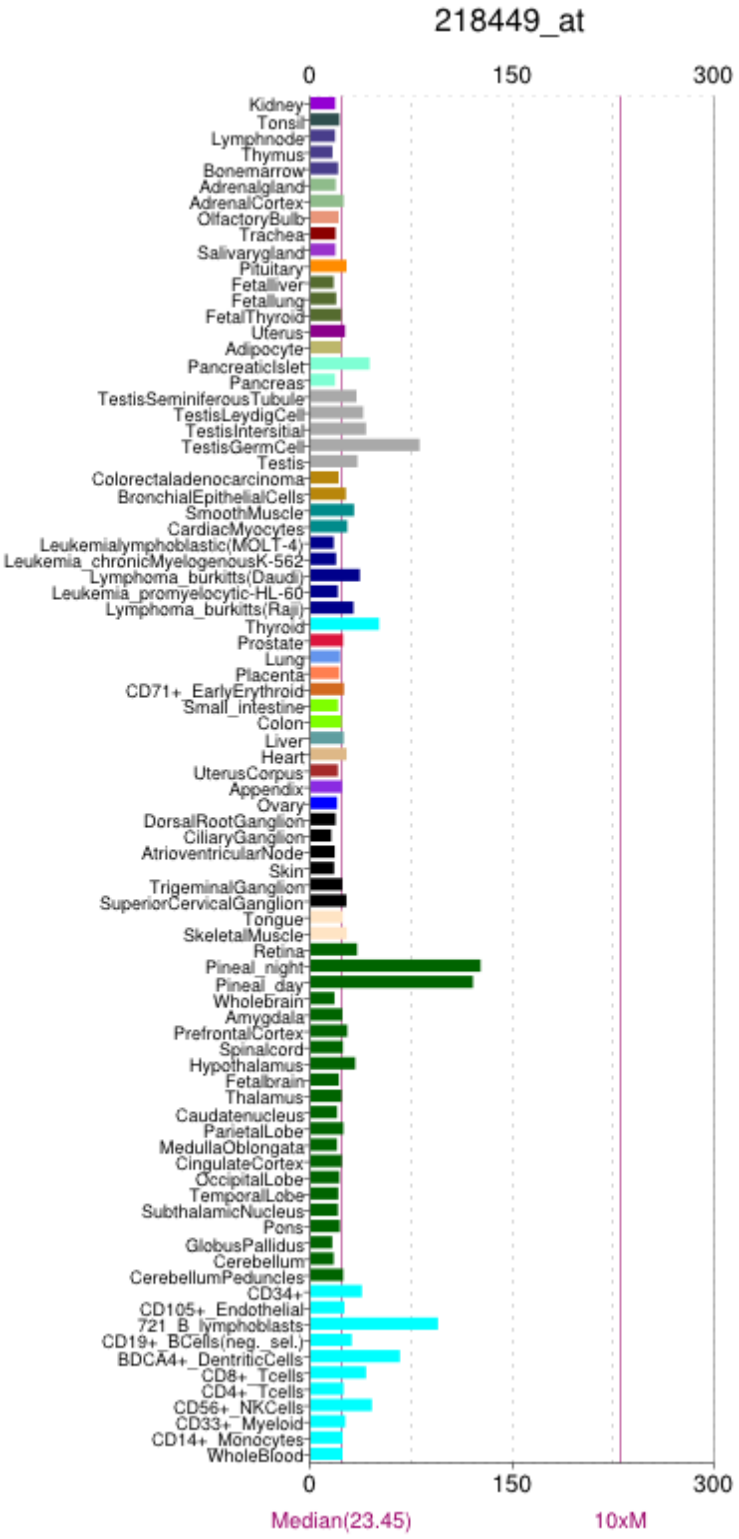
```
rat             TCCGCTGAAGACGGCGACGTGGCTGAAGGAGGCCGGTTTGTTTCTGAATTGAGGCTCAGG -68
mouse           TCCGCTGAAGACGGCGACGTGGCTGAAGGAGGCCGGTTTGTTTCGGAACTGAGGCTCAAG -68
human           TGTACGGAAGGCGGCGACGTGGCCGAAGGATGCCCGTTTGTGTCTAAACGGAGGCTCGGC -73
cow             TCTGAGGAAGGCGGTGACGTGGCCAGAGAAGACCCATCTGTGCCTAACCCGAGGCTCAGC -72
                *     **** *** ********   ** *  **  * ***  *  *   *******

rat             CACGCCGGCTCAGGATTGGTGCTGGGCCGGGCCGGG-----GTCTTTGGCCGGAAGCGGA -13
mouse           CACAGCCGCTCAGGATTGGTGCTAGGCTGGGCCGGG-----GTCTTTGGCCGGAAGCGGA -13
human           CACAACGCCACTGGATTGGTGGTAGGGCGGGGCGGGCCACAGTCTCCAGCCTGAAGCGGA -13
cow             CACAACAGCACTGGATTAATGGTACG-TGGGGCGTGTCCCGGTCAGAGGCCGGAAACGGA -13
                ***  *  * * *****  ** *  *  *** ** *     ***    *** *** ****

rat             AGTGGAAAAAGGATG +3
mouse           AGTGGAGAAAGGATG +3
human           AGTGGAGGAAAGATG +3
cow             AGTGAAGGAAAGATG +3
                **** *  ** *

Primers
Forward   -1356 - -1332
Reverse     -21 -    -1
UPRE        -112 -  -107
```

Supplementary Figure 5. Fragment of the *Lzap* promoter region including the UPRE sequence.

Analysis of promoter region of the *Uba5* gene from four different species. The UPRE sequence is highlighted yellow. ATG codon is highlighted red. Asterisks denote conserved sequences. Positions of primers used to amplify the promoter region and position of the UPRE sequence in respect to the ATG codon of the mouse gene are given below.

Supplementary Figure 6. Expression levels of *Ufsp1* in different human tissues.

The data has been acquired from the BioGPS database and represents expression levels of the *Ufsp1* gene in different human tissues using microarray. The *Ufsp1* exhibits a uniform pattern of expression across all tissues included in the analysis.



Supplementary Figure 7. Expression levels of *Ufsp2* in different human tissues.

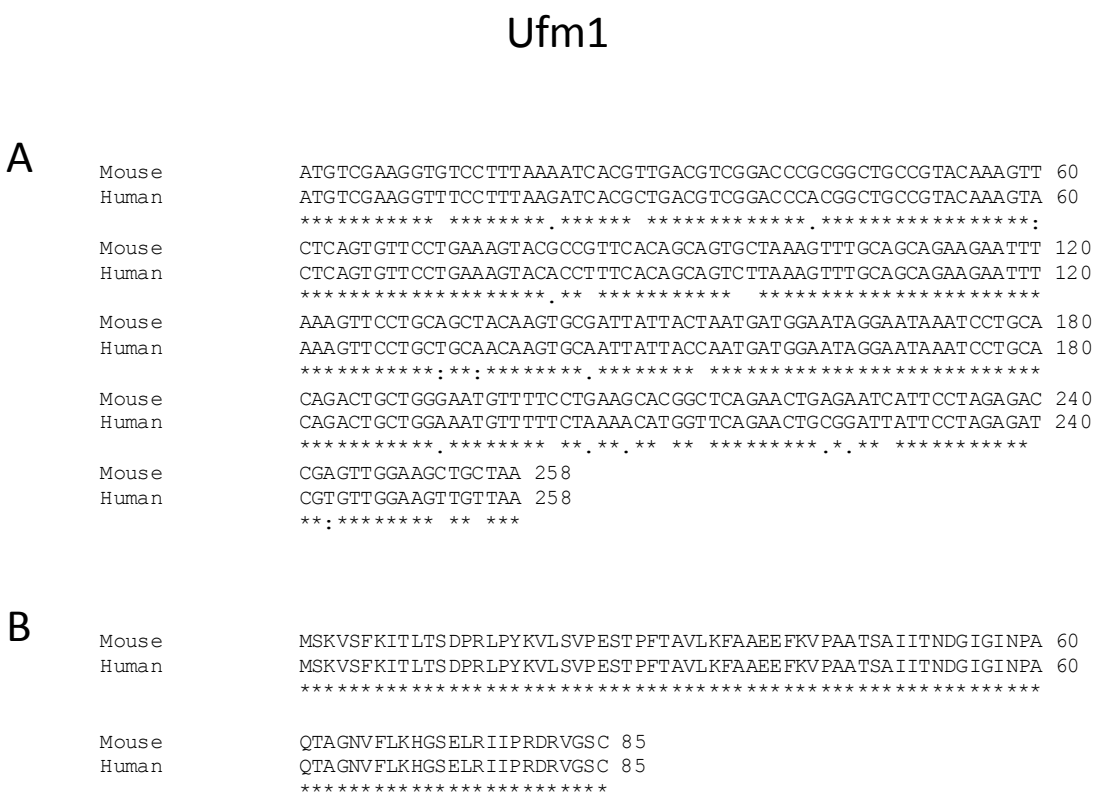
The data has been acquired from the BioGPS database and represents expression levels of the *Ufsp2* gene in different human tissues using microarray. The *Ufsp2* exhibits a more diverse pattern of expression than that of *Ufsp1*.

ATGGCAAGCTGGAGCCACCCGCAGTTCGAAAAGGGTGCATCGAAGGTGTCCTTTAAAATCACGTTGACGTCGGACCCGCGGCTGCCG
TACAAAGTTCTCAGTGTTCCTGAAAGTACGCCGTTCACAGCAGTGCTAAAGTTTGCAGCAGAAGAATTTAAAGTTCCTGCAGCTACA
AGTGCGATTATTACTAATGATGGAATAGGAATAAATCCTGCACAGACTGCTGGGAATGTTTTCCTGAAGCACGGCTCAGAACTGAGA
ATCATTCCTAGAGACCGAGTTGGAAGCTGC**TAA**

MASWSHPQFEKGASKVSFKITLTSDPRLPYKVLSVPESTPFTAVLKFAAEEFKVPAATSAIITNDGIGINP
AQTAGNVFLKHGSELRIIPRDRVGSC**Stop**

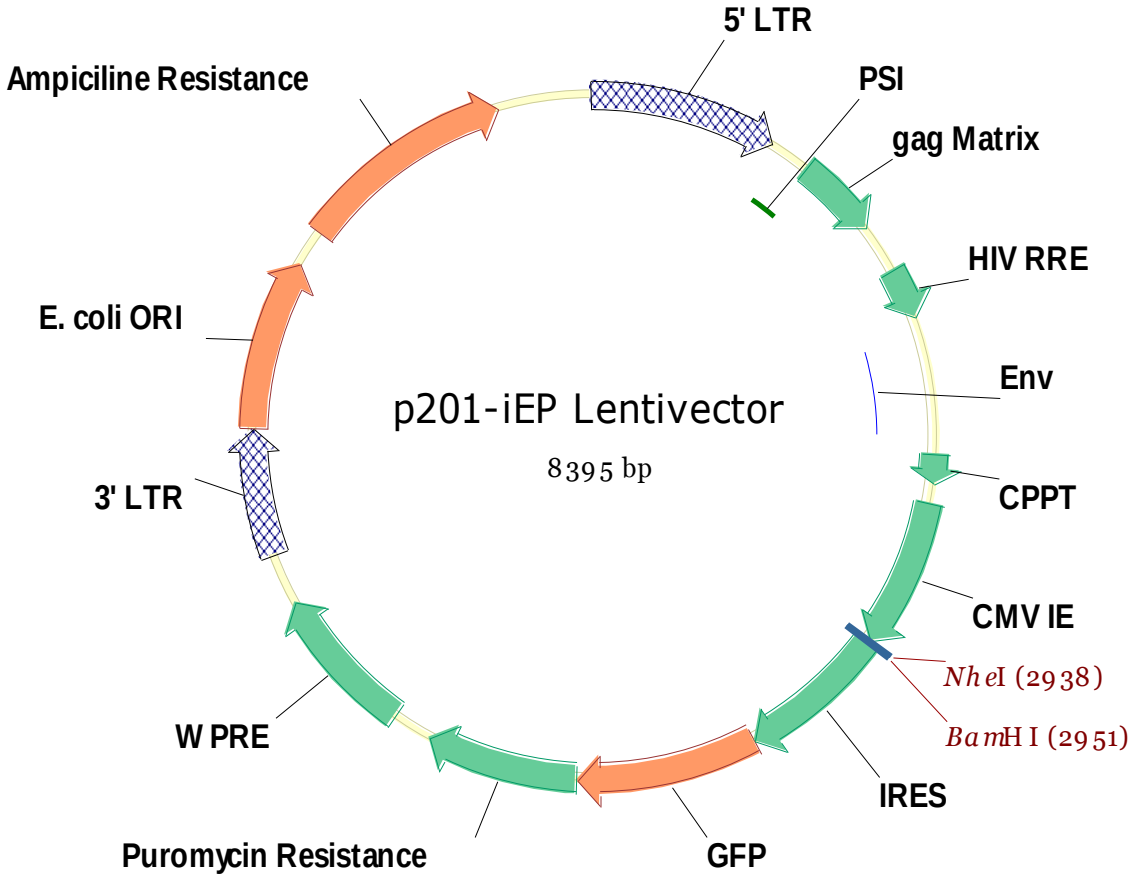
Supplementary Figure 8. Coding sequence and translated amino acid sequence of the Strep-Ufm1 p201.iEP construct.

Sequence of the StrepTagII is highlighted red. (This construct was prepared by former PhD student, Christopher Watson as described in [179])

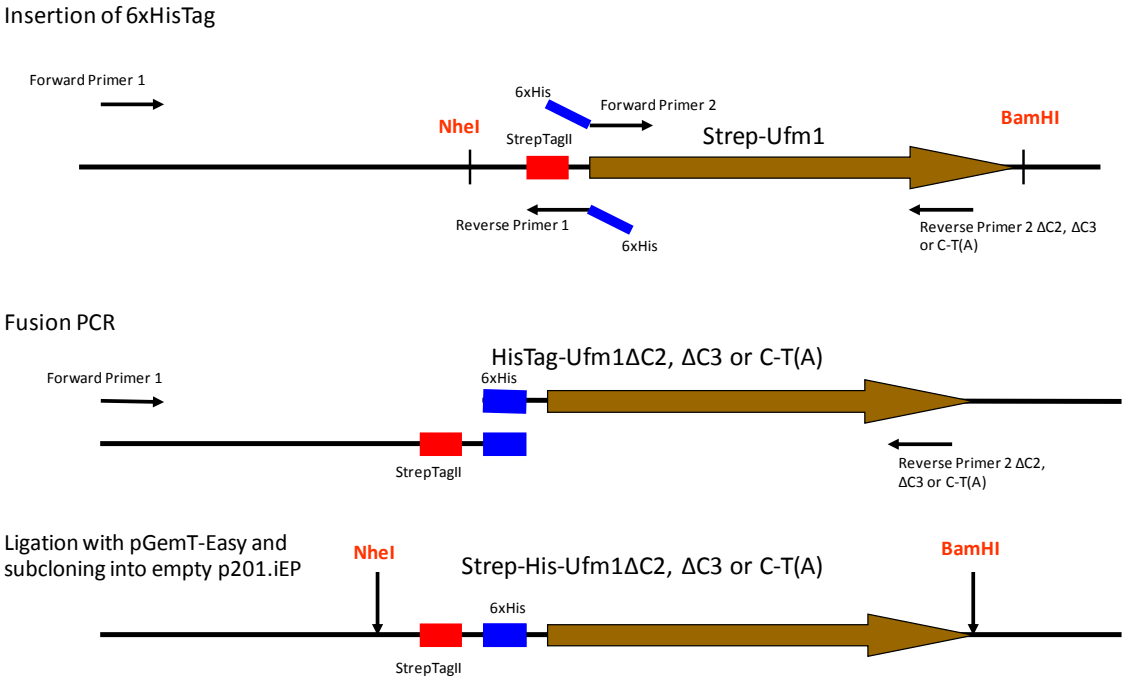


Supplementary Figure 9. Mouse and human Ufm1 sequence alignment.

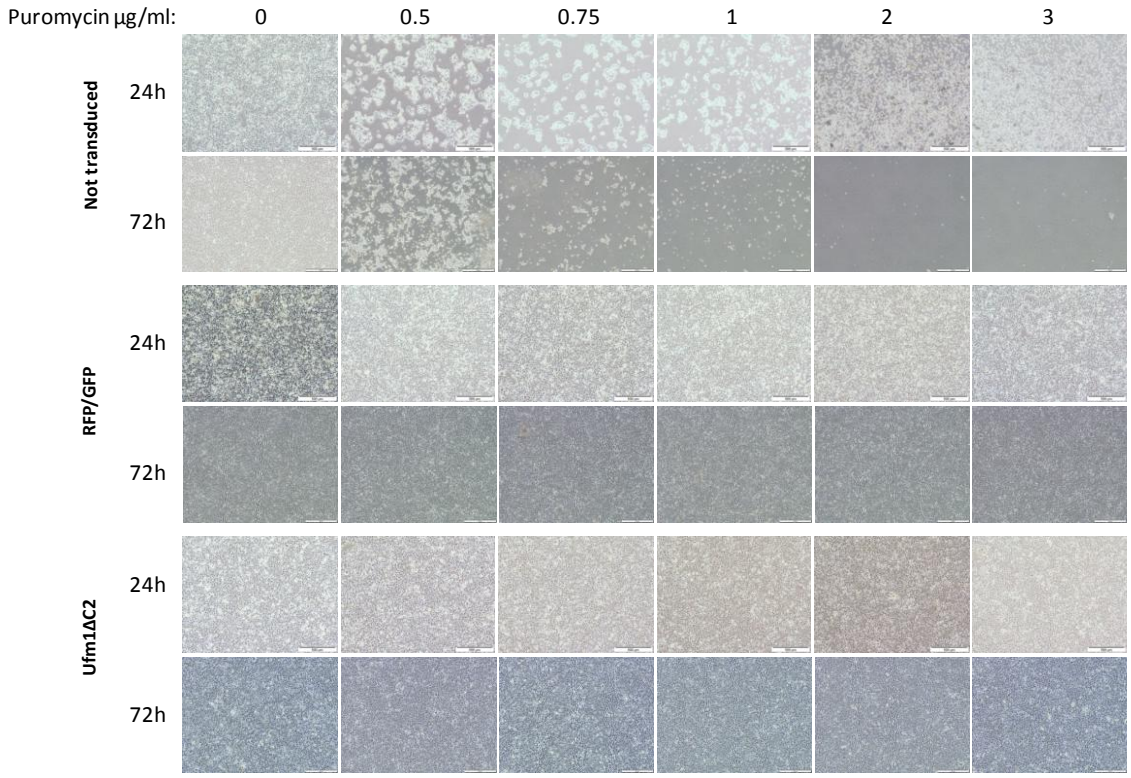
Alignment of the cDNA (A) and amino acid (B) sequences of mouse and human Ufm1.



Supplementary Figure 10. Map of the p201-iEP lentivector

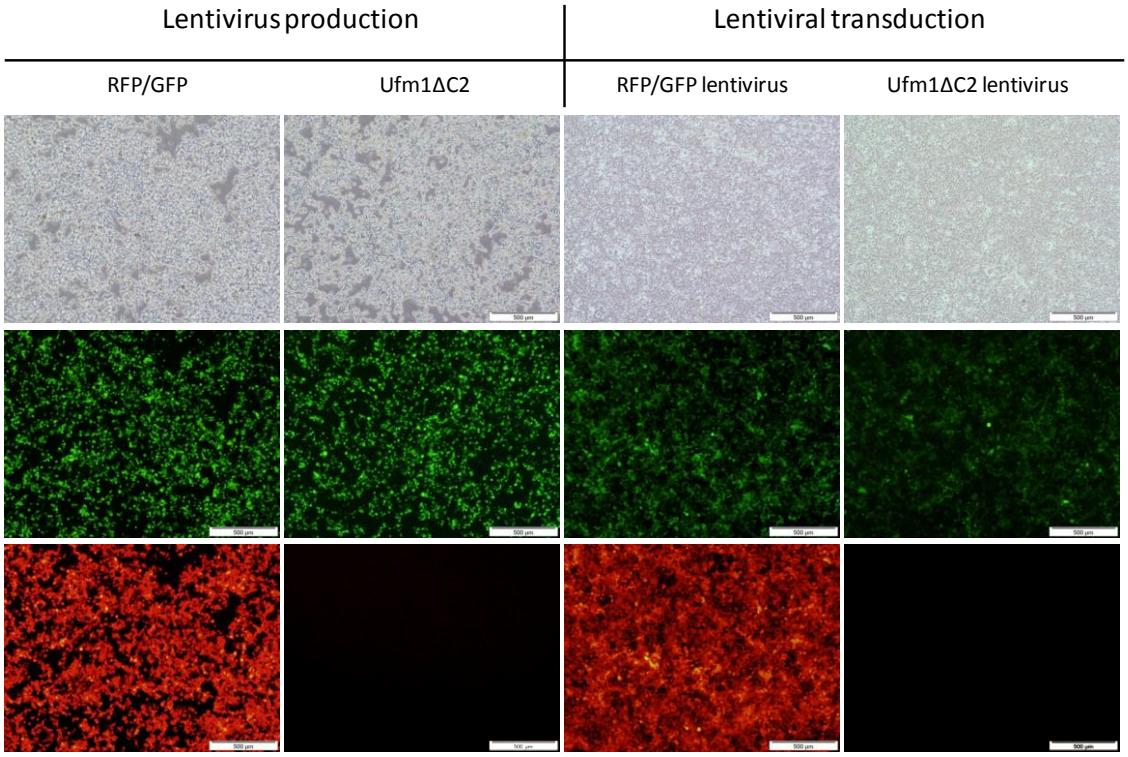


Supplementary Figure 11. Ufm1 expression construct cloning strategy



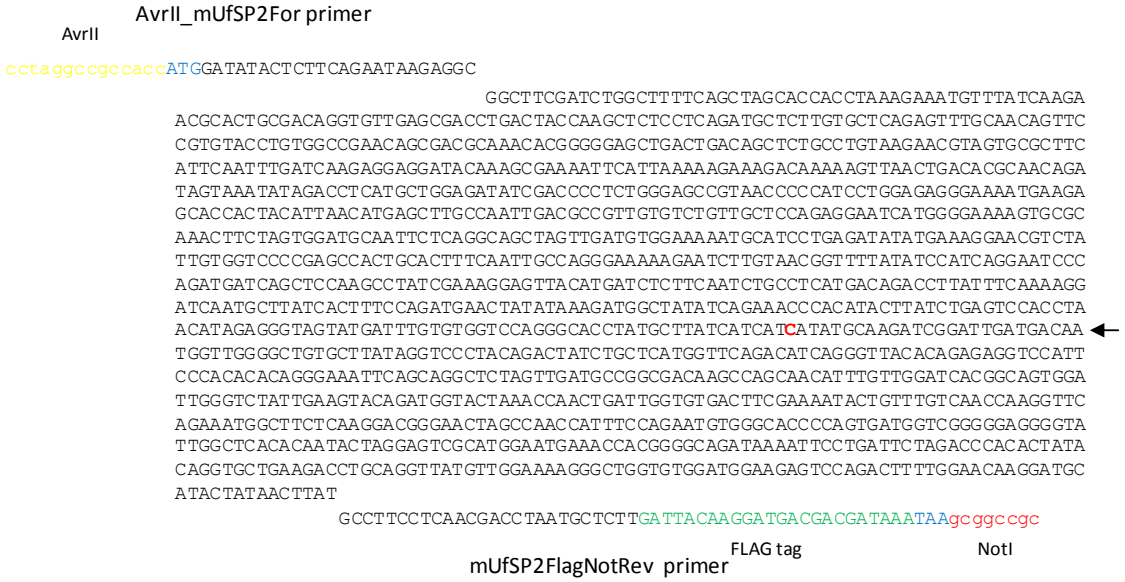
Supplementary Figure 12. HEK293T Puromycin resistance kill curve.

For antibiotic selection of stably transduced cells the optimal concentration was determined as the smallest concentration that killed all cells in a culture of control (not transduced) cells. HEK293T cells were transduced with lentiviruses carrying the RFP/GFP p201.iEP and Ufm1ΔC2 p201.iEP vectors and cultured with increasing concentrations of puromycin. The established stably transduced cell lines were then cultured with the addition of 2 µg/ml of puromicin.



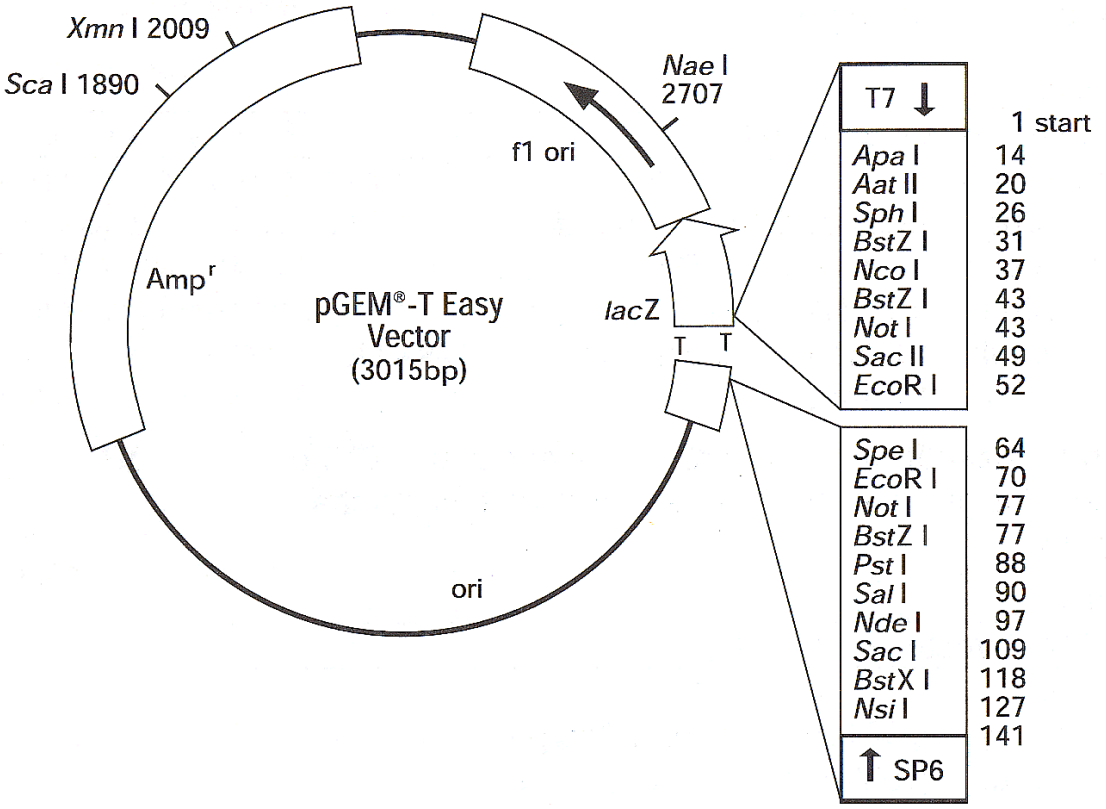
Supplementary Figure 13. Generation of stably transduced HEK293T cell lines.

Left panel: Production of lentiviruses carrying RFP/GFP and Ufm1ΔC2 vectors. Right panel: HEK293T cell cultures transduced with the viruses show overexpression of GFP and RFP. A – normal light; B – GFP; C – RFP;



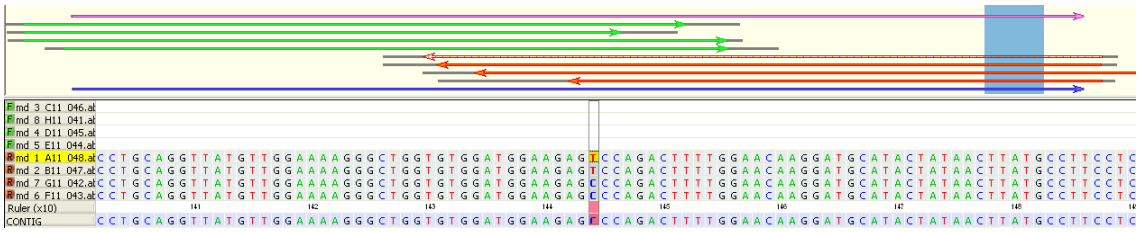
Supplementary Figure 14. *mUfsp2*^{BHD}-Flag cDNA sequence used to generate the transgenic construct.

The primer sequences are shown with the *Avr*II and *Not*I cloning sites in yellow and red respectively. The Start and Stop codons are marked blue and the Flag-tag sequence green. The BHD mutation is marked red and the arrow points to the line it is located in the sequence. Translation of the sequence shows correct in frame incorporation of the Flag-tag.



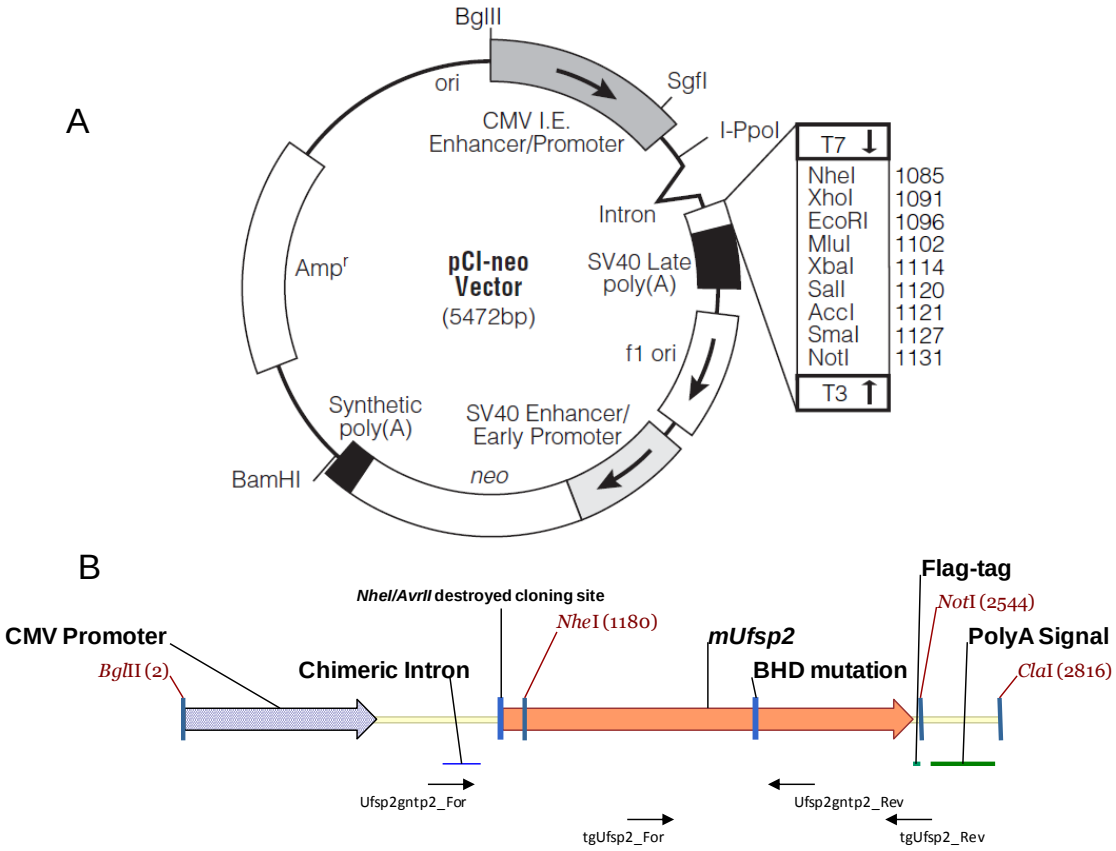
Supplementary Figure 15. pGEM-T easy cloning vector allowing direct ligation of PCR products with A overhangs generated by *Taq* polymerase.

Main features include: ampicilin resistance gene (*Amp*^{r}); origin of replication (*ori*); β-lactamase (*lac*Z); multiple cloning region flanked by T7 and SP6 RNA polymerase transcription initiation sites.



Supplementary Figure 16. Sequencing of the *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag clones.

The pGEM-T Easy *mUfsp2*^{WT}-Flag and *mUfsp2*^{BHD}-Flag clones were sequenced and the result was analysed using DNA baser software (Heracle BioSoft S.R.L.). Clones were confirmed to possess the sequences introduced in both primers and be free of mutations except for the highlighted T to C BHD mutation.



Supplementary Figure 17. Map of the pCI-neo expression vector (A) and the *mUfsp2*$^{WT/BHD}$-Flag transgenic constructs generated by excision from the pCI-neo vector (B).

The CMV promoter drives the expression, a chimeric intron improves expression of the transgene, the *mUfsp2*$^{WT/BHD}$-Flag transgene is followed by a polyadenylation signal. **B.** Black arrows depict annealing sites of primers used for genotyping of mice (Ufsp2gntp2_For/_Rev primers produce a 1177 bp PCR product) and testing of the transgene expression (Ufsp2rt_For/Rev primers produce a 608 bp PCR product).

Appendix 1. Reagents and suppliers

- 35S UTP 1250Ci/mmol (Perkin-Elmer, USA)
- acetic acid glacial (Fisher, UK)
- Acetone (Fisher, UK)
- Agar (Melford, UK)
- Agarose (Lonza, USA)
- Amersham HyperfilmTM ECL (GE Healthcare, UK)
- AmershamTM ECL Plus Western Blotting Detection System (GE Healthcare, UK)
- ammonium acetate (BDH Laboratories, UK)
- antibody rabbit-anti-Ufm1 polyclonal (Boston Biochem, USA)
- antibody goat-anti-mouse HRP conjugated (Dako, Denmark)
- antibody mouse-anti- STREP·Tag® II monoclonal (Merck, UK)
- antibody mouse-anti-Flag-tag (Invitrogen,UK)
- antibody mouse-anti-Ufsp2 (Abcam,UK)
- antibody swine-anti-rabbit HRP conjugated (Dako, Denmark)
- bovine serum albumine - BSA (Sigma, UK)
- bromophenol blue (Sigma, UK)
- Buffer E for Strep-Tactin Sepharose (IBA, Germany)
- Buffer W for Strep-Tactin Sepharose (IBA, Germany)
- chloroquine diphosphate salt (Sigma,UK)
- cOmplete, EDTA-free Protease Inhibitor Cocktail Tablets (Roche, Germany)
- D19 film developer (Kodak)
- dextran sulphate (Sigma, UK)
- Diethylpyrocarbonate DEPC (Sigma, UK)
- disodium carbonate (BDH Laboratories, UK)
- disodium hydrogen orthophosphate (BDH Laboratories, UK)
- dithiothreitol - DTT (ApliChem, Germany)
- DMSO (Sigma,UK)
- dNTP set (Bioline, UK)
- Dulbecco's Modified Eagle's Medium (Sigma, UK)
- EDTA (Fisher, UK)
- Ethanol (Fisher, UK)
- Expand High Fidelity PCR system (Roche, Germany)
- ficoll 400 (Sigma, UK)
- foetal bovine serum - FBS origin: EU Approved (South American) (Invitrogen, UK)
- formamide (Sigma, UK)
- FuGENE® HD Transfection Reagent (Promega, UK)
- glycine (Fisher, UK)
- Guanidinium-HCl (Fluka, UK)
- Ham's F12 medium (Sigma, UK)
- HCl (Fisher, UK)
- hematoxyline

- HEPES (Fisher, UK)
- HyperLadderTM I (Bioline, UK)
- HyperLadderTM IV (Bioline, UK)
- imidazole (Sigma, UK)
- iodoacetamide (Sigma, UK)
- K5 nuclear emulsion (Ilford, UK)
- L-ascorbic acid 2-phosphate (Wako, Japan)
- L-glutamine (Sigma, UK)
- MEM alpha modification (Sigma, UK)
- MEM Non-essential Amino Acid Solution (100X) (Sigma, UK)
- Methanol (Fisher, UK)
- MOPS (BDH Laboratories, UK)
- N-etylmaleimide (Sigma, UK)
- NuPAGE® MOPS SDS Running Buffer (Invitrogen, UK)
- NuPAGE® Novex 10% Bis-Tris Gel (Invitrogen, UK)
- paraformaldehyde - PFA (Sigma, UK)
- PBS (Sigma-Aldrich, UK)
- penicillin/streptomycin (100x) (Sigma, UK)
- PerfectPro Ni-NTA Agarose (5Prime, Germany)
- pGEM®-T Easy vector system (Promega, UK)
- Phenol-Chloroform-Isoamyl alcohol (Sigma, UK)
- Polyvinylpyrrolidone - PVP (Sigma, UK)
- potassium chloride (BDH Laboratories, UK)
- potassium dihydrogen ortophosphate (BDH Laboratories, UK)
- proteinase K (Roche, Germany)
- Puromycin (InvivoGen, USA)
- PVDF transfer membrane Amersham HybondTM - P (GE Healthcare, UK)
- QIAGEN Plasmid Maxi Kit (Qiagen, Germany)
- QIAprep Spin Miniprep Kit (Qiagen, Germany)
- QIAquick Gel Extraction Kit (Qiagen, Germany)
- rhBMP-2 (Invitrogen, UK)
- Riboprobe® System (Promega, UK)
- Rnase (Roche, Germany)
- RNeasy Mini Kit (Qiagen Ltd., UK)
- sodium acetate (BDH Laboratories, UK)
- sodium chloride (Fluka, UK)
- sodium citrate (Sigma, UK)
- Sodium deoxycholate -DOC (Sigma, UK)
- sodium dihydrogen ortophosphate (BDH Laboratories, UK)
- sodium dodecyl sulphate - SDS (Fisher, UK)
- sodium hydrogen carbonate (BDH Laboratories, UK)
- Strep-Tactin Sepharose (IBA, Germany)
- SYBR® green PCR mastermix (Applied Biosystems, UK)
- Taq polymerase (Bioline, UK)

- Thapsigargin (Sigma, UK)
- Trichloroacetic acid - TCA (BDH Laboratories, UK)
- triethanolamine (Sigma, UK)
- Tris Base (Fisher, UK)
- tRNA (Roche, Germany)
- Trypsine-EDTA (Sigma-Aldrich, UK)
- Tryptone (Melford, UK)
- Tunicamycin (Sigma, UK)
- Tween 20 (Sigma, UK)
- Unifix fixing solution (Kodak)
- xylene (Genta Enviremental, UK)
- Yeast extract (Melford, UK)
- β-glycerophosphate (Sigma, UK)
- β-mercaptoethanol (Sigma, UK)

Appendix 2. Buffers components

Hydrolysis Buffer (In Situ Hybridisation)

1 M $NaHCO_3$	40 µl
1 M Na_2CO_3	60 µl
1 M DTT	5 µl
DEPC H_2O	395 µl
Total	500 µl

Neutralising Buffer (In Situ Hybridisation)

3 M NaOAc (pH 5.2)	33 µl
1 M DTT	5 µl
Glacial acetic acid	5 µl
DEPC H2O	457 µl
Total	500 µl

Proteinase K buffer (In Situ Hybridisation)

1M Tris (pH 8.0)	35 ml
0.5M EDTA (pH 8.0)	35 ml
DEPC H2O	280 ml
Proteinase K 10 mg/ml	175 µl
Total	350 ml

Hybridisation Buffer (In Situ Hybridisation)

Formamide	25 ml
Dextran sulphate (50%)	10 ml
1M DTT	2.5 ml
tRNA (10 mg/ml)	2.5 ml
5M NaCl	3 ml
1M Tris (pH 8.0)	1 ml
50X Denhardts	0.5 ml
DEPC H2O	4.5 ml
Total	49 ml

Wash solution (In Situ Hybridisation)

20X SSC	100 ml	
Formamide	500 ml	
dH2O	400 ml	
DTT	1.925 g	Add immediately prior to first wash
Total	1 litre	

RNase buffer (In Situ Hybridisation) r

5M NaCl	160 ml
1M Tris (pH 8.0)	16 ml
0.5M EDTA (pH 8.0)	3.2 ml
dH2O	1420.8 ml
Total	1.6 litres

SSC/DTT buffer (In Situ Hybridisation)

20X SSC	4 ml	
dH2O	796 ml	
DTT	1.54 g	Add immediately prior to first wash
Total	800 ml	

Ammonium acetate / Ethanol (In Situ Hybridisation)

3M Ammonium acetate	35 ml
dH2O	70 ml
EtOH 100%	245 ml
Total	345 ml

Blocking Buffer

PBS, 0.1% Tween 20, 3% BSA

Phosphate Buffered Saline – PBS

NaCL	8g
KCl	0.2g
Na_2HPO_4	1.44g
KH_2PO_4	0.24g

Adjust pH to 7.4
Add H_2O to 1L

20x SSC
NaCl 172.3g
Sodium citrate 88.2g
Adjust pH to 4.5
Add H_2O to 1L

Gu-HCl denaturing lysis buffer pH 8.0
6M Guanidinium-HCl
100 mM Na_2HPO_4/ NaH_2PO_4 buffer pH 8.0
10 mM Tris-HCL pH 8.0
10 mM imidazole
10 mM β-mercaptoethanol
20 mM N-ethylmaleimide
20 mM iodoacetamide
1x cOmplete protease inhibitors

Urea buffer A pH 8.0
8M urea
100 mM Na_2HPO_4/ NaH_2PO_4 buffer pH 8.0
10 mM Tris-HCL pH 8.0
10 mM imidazole

Urea buffer B pH 6.3
8M urea
100 mM Na_2HPO_4/ NaH_2PO_4 buffer pH 8.0
10 mM Tris-HCL pH 8.0
10 mM imidazole

2X HBS
NaCl 16g
KCl 0.74g
Na_2HPO_4 0.27g
Dextrose 2g
HEPES 10g

Adjust pH to 7.0

Add H_2O to 1L

Lysogeny broth (LB)

Tryptone	10g
Yeaset extract	5g
NaCl	10g
+/- Agar	15g for plates

6x Laemmli sample buffer

SDS	1.2g
Bromophenol blue	6mg
Glycerol	4.7ml
0.5M Tris-Cl pH6.8	1.2ml
DTT	0.93g

Add H_2O to 10ml

10x DNA Loading Dye

Bromophenol blue	0.02g
Glycerol	6ml
0.5M EDTA	200µl

Add H_2O to 14ml

10x Transfer Buffer

Tris base	58 g
Glycine	29 g
SDS	3.7g

Dissolve in 800 mL of water and adjust the volume to 1000 mL with water

To prepare 1L 1x buffer take 100 ml of 10x buffer 200 ml of methanol and add distilled water to a total volume of 1L

Appendix 3. List of primers used in this study

NAME	SEQUENCE	Tm °C	COMMENTS
Forward Primer 1	GACAGCAGAGATCCAGTTTGGTTAGTACCG	68.1	Used for generation of the Ufm1ΔC2 expression construct
Reverse Primer 1	GGCAGCGTGATGGTGATGGTGATGTGCACCCTTTTCGAACTGCGGG	>75	Used for generation of the Ufm1ΔC2 expression construct
Forward Primer 2	CATCACCATCACCATCACGCTGCCTCGAAGGTGTCCTTTAAAATCACGTTGACG		Used for generation of the Ufm1ΔC2 expression construct
Reverse Primer 2	CGCGGATCCTTATCCAACTCGGTCTCTAGGAATGATTCTCAG	>75	Used for generation of the Ufm1ΔC2 expression construct
Ufm1dC3_Rev	GGA TCC TCA AAC TCG GTC TCT AGG AAT GAT TCT CAG TTC TG	72.4	Used for generation of the Ufm1ΔC3 expression construct
Ufm1ctermAlaRev	CGG ATC CTT ATG CAA CTC GGT CTC TAG G	68	Used for generation of the Ufm1C-T(A) expression construct
DDRGK1constr_Rev	GGA TCC TTA TTT ATC GTC GTC ATC CTT GTA ATC TGC GGC TGA AGC CTG GGC AGG GAG G	>75	Used for generation of the DDRGK1 expression construct
DDRGK1constr_For	gct agc GAC ACC ATG GTG GGG CCC TGG GTG TAT CTG G	>75	Used for generation of the DDRGK1 expression construct
AvrII_mUfSP2For	cct agg ccg cca ccA TGG ATA TAC TCT TCA GAA TAA GAG GC	74.4	Used for generation of the transgenic Ufsp2 construct
mUfSP2FlagNotRev	GCG GCC GCT TAT TTA TCG TCG TCA TCC TTG TAA TCA AGA GCA TTA GGT CGT TGA GGA AGG C	>75	Used for generation of the transgenic Ufsp2 construct
M13 rev (-29)	CAG GAA ACA GCT ATG ACC	53.7	sequencing primer
T7 for	TAA TAC GAC TCA CTA TAG GG	53.2	sequencing primer
CMV_For	gtg gga ggt cta tat aag cag agc tgg	66.5	sequencing primer
UfSP2gntp2For	GGC TTG TCG AGA CAG AGA AGA CTC TTG C	68	Used for genotyping of the transgenic mice (1176 bp product)
UfSP2gntp2Rev	GAT CCA ACA AAT GTT GCT GGC TTG TCG C	66.6	Used for genotyping of the transgenic mice (1176 bp product)
tgUfsp2_For	AGA GGG TAG TAT GAT TTG TGT GGT CC	63.2	Used for genotyping and RT PCR of the transgenic mice (607 bp product)
tgUfsp2_Rev	CGT CGT CAT CCT TGT AAT CAA GAG C	63	Used for genotyping and RT PCR of the transgenic mice (607 bp product)
50s	GTC AGG AAA CTG TGC TTT ATA G	56.5	collagen alpha-1(XXVII) chain precursor (genotyping positive control)
50as	GGA AAG CAA GGC TTG TAT AC	55.2	collagen alpha-1(XXVII) chain precursor (genotyping positive control)
mUfSP2rt_For	cgg cga caa gcc agc aac	60.5	mouse real-time PCR primers
mUfSP2rt_Rev	CCT CCC CCG ACC ATC ACT G	63.1	mouse real-time PCR primers
mDDRGK1rt_For	CGG GAG CAC GAG GAG TAC C	63.1	mouse real-time PCR primers

mDDRGK1rt_Rev	GGT TTA TGG CGT CCT GAG TCC	61.8	mouse real-time PCR primers
mCdk5rap3rt_For	tgg aga agc agc agg agg cg	63.4	mouse real-time PCR primers
mCdk5rap3rt_Rev	GTG GAG GAC ACA GAG GGA GC	63.4	mouse real-time PCR primers
mUfm1rt_For	ggt att tga ctt gac agc atc tcc	61	mouse real-time PCR primers
mUfm1rt_Rev	CAG TTC CCA CCA TGA CAT ACT GG	62.4	mouse real-time PCR primers
mOCrt_For	ggc aat aag gta gtg aac aga ctc c	63	mouse real-time PCR primers
mOCrt_Rev	GTC CTA AAT AGT GAT ACC ATA GAT GCG	61.9	mouse real-time PCR primers
mCol1a1rt_For	tca aga gaa gtc tca aga tgg tgg	61	mouse real-time PCR primers
mCol1a1rt_Rev	GTA GAC CTT GAT GGC GTC CAG	61.8	mouse real-time PCR primers
mCol2a1rt_For	cca ctg taa gaa cag cat cgc c	62.1	mouse real-time PCR primers
mCol2a1rt_Rev	GGT ACT CGA TGA CGG TCT TGC	61.8	mouse real-time PCR primers
mGAPDHrt_For	cgt gcc gcc tgg aga aac c	63.1	mouse real-time PCR primers
mGAPDHrt_Rev	TGG AAG AGT GGG AGT TGC TGT TG	62.4	mouse real-time PCR primers
mUfl1rt_For	ctc aga ttt cct cat ggc agt gg	62.4	mouse real-time PCR primers
mUfl1rt_Rev	GCA GAA TCC AGG CAA GAC AGG	61.8	mouse real-time PCR primers
mBiPrt_For	gct tcc gat aat cag cca act gt	60.6	mouse real-time PCR primers
mBiPrt_Rev	TCT CAG CAT CAT TAA CCA TCC TTT C	59.7	mouse real-time PCR primers
mOsterixRT_F	AGC CCA GAG CAG AGC AAC CT	61.4	mouse real-time PCR primers
mOsterixRT_R	GGT CGG AGC ATA GGA ACT AGG	61.8	mouse real-time PCR primers
mRunx2RT_F	CCT GAA CTC TGC ACC AAG TCC	61.8	mouse real-time PCR primers
mRunx2RT_R	CCA GAG GCA GAA GTC AGA GG	61.4	mouse real-time PCR primers
mUfSP1-rt_For	CTT GAT ACT GGA CCC ACA CTA C	60.3	human real-time PCR primers
mUfSP1-rt_Rev	CCT GCC CGT ATC ATT GTG CC	61.4	human real-time PCR primers
mUba5-rt_For	GAA GAA TAC AAG AAG AGG GCA GCA G	63	human real-time PCR primers
mUba5-rt_Rev	TCA GGT AAG GTG GGA ACA GGA C	62.1	human real-time PCR primers
hsLZAP-rt_For	TGC AGA AGC AGC AGG AGG CA	61.4	human real-time PCR primers
hsLZAP-rt_Rev	TGT AGA GAA CAT GGA GAG TGC C	60.3	human real-time PCR primers
hsUfBP1-rt_For	GCG GGA GCA TGA GGA GTA CC	63.5	human real-time PCR primers
hsUfBP1-rt_Rev	GCG ATT TAT GGT GTC CTG AGT GC	62.4	human real-time PCR primers
hsUfSP2-rt_For	cag cag gct cta gtc gat gc	61.4	human real-time PCR primers
hsUfSP2-rt_Rev	CTC CTC CCC CGA TCA TAA CTG	61.8	human real-time PCR primers
hsUfm1-rt_For	GGC AAT TTG CTT GAT AGC ATC TGA	59.3	human real-time PCR primers
hsUfm1-rt_For	GGA CAC CAA TTT ATA CAC ATA AGG G	59.7	human real-time PCR primers
hsUfl1-rt_For	CGG ATT TAA TGA TGG CAG TAG ACG	61	human real-time PCR primers
hsUfl1-rt_Rev	CCC CTT TTC ACC ATA ATA TCA CAA GC	61.6	human real-time PCR primers

hsBIP-rt_For	CAG CTT CTG ATA ATC AAC CAA CTG T	59.7	human real-time PCR primers
LZAP-luc_F	gct agc tct gcc tgc ctc tgt ctc cc	71.1	Luciferase promoter region analysis
LZAP-luc_R	ctc gag CCT TTC TCC ACT TCC GCT TCC	69.5	Luciferase promoter region analysis
Uba5-luc_R	ctc gag ATG GCT GGG CTC CCG GTG GC	74.3	Luciferase promoter region analysis
Ufl1-luc_F	gct agc TCC AAC AAG GCC ACA CCT CC	69.5	Luciferase promoter region analysis
Ufl1-luc_R	ctc gag ATC ACT GCT CCG CTC GCC CC	72.7	Luciferase promoter region analysis
Ufm1-luc_F	gct agc TGC CAG GAC GAC TTT CCA TCC	69.5	Luciferase promoter region analysis
Ufm1-luc_R	ctc gag TGG TGG AAC CGC CTG ACA CAC	71	Luciferase promoter region analysis
LZAP-Luc_F2	gct agc TGT TTG AGT CCT CCC TCC TCA CCC	72.2	Luciferase promoter region analysis
Ufl1-LucF2	gct agc ATT CTC CTT TGA GAA CTC TTG AGC C	68.2	Luciferase promoter region analysis
Ufl1-Luc_F3	TGC CCT GGT GGG AAG GTA GG	63.5	Luciferase promoter region analysis
Ufm1mut_Rev	CCT TCG CGG CCC ATT ACG CCC TCC AGC GCG GCT ATC C	> 75	Luciferase promoter region analysis
Ufm1mut_For	GCG CTG GAG GGC GTA ATG GGC CGC GAA GGC CCC	> 75	Luciferase promoter region analysis
Uba5mut_For	GTG CCG GAG TAA TGT CTT ACT GGG AGC CTT GGA GCC G	> 75	Luciferase promoter region analysis
Uba5mut_Rev	GCT CCC AGT AAG ACA TTA CTC CGG CAC GCC CAC CGG	> 75	Luciferase promoter region analysis
LZAPmut_For	CTC CGC TGA AGA CGG CGT AAT GGC TGA AGG AGG CCG GTT TGT TTC G	> 75	Luciferase promoter region analysis
LZAPmut_Rev	CCG GCC TCC TTC AGC CAT TAC GCC GTC TTC AGC GGA GCC AAT CC	> 75	Luciferase promoter region analysis
Ufl1Luc_For4	gct agc CTT GAG CCA TGT CTA CAC ACT ATT CC	69.5	Luciferase promoter region analysis
Ufm1Luc2MutRev	CTC GAG GGT CCG ATT ACA ACG TGA TTT TAA AGG ACA CCT TCG ACC	74.9	Luciferase promoter region analysis