The Requirement of β-Dystroglycan, and a Sun1 Protein Interactome, in Myonuclei.

Matthew Cook, BSc (Hons)

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Department of Biomedical Science, The University of Sheffield.

The
University
Of
Sheffield.

Institute of
Medical Biology

A★STAR

January 2021

Abstract

The nuclear lamina (NL) is composed of A- and B-type lamins, which assemble into a complex, insoluble meshwork beneath the inner nuclear membrane (INM). The lamina is anchored to the cytoskeleton by the LINC (Linker of the nucleoskeleton and cytoskeleton) complex which spans the NE and comprises INM SUN-domain proteins which bind their outer nuclear membrane (ONM) counterparts, KASH-domain proteins, in the perinuclear space. The NL and LINC complex, together with other NE proteins, have diverse functions in genome tethering, transcriptional regulation and nuclear positioning. β-dystroglycan is an established cell adhesion protein which is thought to also function in the nucleus to stabilise a number of structures including the NL, but mechanisms behind this are elusive. It is hypothesised that β-dystroglycan exerts its influence on the nucleus through the LINC complex, since previous mass spectrometry analyses found that Sun1 co-precipitated with β-dystroglycan. A Sun1-β-dystroglycan interaction was interrogated using several techniques, with little evidence being found to support its existence in human myoblasts. Moreover, CRISPR-mediated ablation of dystroglycan in human myoblasts appears to not disturb nuclei as previously reported. β-dystroglycan is also extensively modified post-translationally and proteolytic fragments are persistent within cells. Whether distinct β-dystroglycan species have differential functions is unknown. Epitope-tag detection of fragments *in situ* revealed that proteolytic fragments become separately organised, furthering the notion that β-dystroglycan is intricately controlled within the cell.

Sun1 is crucial during development as well as being implicated in muscle-wasting disorders both as a causative and disease modifying component. *Lmna*, encoding lamins A and C, is widely expressed in somatic cells, yet mutations give rise to a range of tissue-specific diseases, including muscular dystrophies, for reasons not fully understood. Ablation of Sun1 in *Lmna* mutant mice suppresses many of the *Lmna*-associated disease phenotypes revealing a potential functional interaction between the LINC complex and NL. It is postulated that nucleoplasmic protein interactions of Sun1 are altered in a tissue-specific manner, in response to the specific *Lmna* mutation present, that then elicit disease phenotypes. To identify

potential nucleoplasmic Sun1 interacting proteins, the proximity dependent biotinylation 2C-BioID technique was used in myoblasts. 2C-BioID identified Lamin A and 6 novel Sun1 interactors which were verified using bimolecular fluorescence complementation. One of these novel Sun1 interactors was Ppm1a, so further implicate the NE and the LINC complex in the regulation of TGFβ signalling.

Acknowledgments

Contents

viii

List of Figures

List of Tables

Abbreviations Used

µl	Microlitre
µm	Micrometre
µM	Micromolar
(p)Y890/892	(phosphorylated) tyrosine 890/892
α-SMA	α-smooth muscle actin
2C-BioID	Two component biotin identification
AD	Autosomal dominant
AKT	Protein kinase B
APEX	Ascorbate peroxidase
APS	Ammonium persulfate
ARM	Armadillo
BiFC	Bimolecular fluorescence complementation
BiFC-FC	Bimolecular fluorescence complementation-flow cytometry
BioID	Biotin identification
bp	base pair
BSA	Bovine serum albumin
CMD	Congenital muscular dystrophy
CMT	Charcot Marie Tooth
CRISPR	Clustered regularly interspace short palindromic sequences
DABCO	1,4-Diazabicyclo[2.2.2]octane
DCM	Dilated cardiomyopathy
DDR	DNA damage response
DG	Dystroglycan
DGC	Dystrophin glycoprotein complex
DMD	Duchenne's muscular dystrophy
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dioxymethyl sulfoxide
DNA	Deoxyribonucleic acid
DNAPK	DNA-dependent protein kinase
Dox	Doxycycline

DSB	Double stranded break
DTT	Dithiothreitol
ECM	extracellular matrix
EDMD	Emery-Dreifuss muscular dystrophy
EDTA	Ethylenediaminetetraacetic acid
EEA1	Early endosome antigen 1
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
ER	Endoplasmic reticulum
ERK	Extracellular signal-regulated kinase
FA	Focal adhesion
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FCMD	Fukuyama congenital muscular dystrophy
FGF	Fibroblast growth factor
FKBP	FK506 binding protein
Flx	Floxed
FPLD	Familial partial lipodystrophy
FRB	FKBP-rapamycin binding
FTI	Farnesyltransferase inhibitor
GAP	GTPase activating protein
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GEF	Guanine exchange factor
GFP	Green fluorescent protein
GO	Gene ontology
gRNA	Guide ribonucleic acid
GST	Glutathione-S-transferase
GTP	Guanine triphosphate
H3K27me3	Histone 3 tri-methylated on lysine 27
H3K9me2/3	Histone 3 di-/tri-methylated on lysine 9
HA	Hemagluttinin
HBSS	Hank's balanced salt solution

HDR	Homology-directed repair
HEAT	Huntingtin, EF3, PP2A, TOR1 (domain)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HGPS	Hutchinson-Gilford progeria syndrome
IBB	Importin-β binding
ICD	Intracellular domain
Ig	Immunoglobulin
IKNM	Interkinetic nuclear movement
Indel	Insertion/deletion
INM	Inner nuclear membrane
IVT	*In vitro* translation
KASH	Klarsicht, ANC-1, Syne Homology
kDa	Kilo Dalton
KI	Knock in
KO	Knock out
LAD	Lamina associated domain
LAP2β	Lamin associated protein 2β
LARGE	Like-acetylglucosaminyltransferase enzyme
LBR	Lamin B receptor
LC-MS	Liquid chromatography mass spectrometry
LEM	Lap2, Emerin, Man1 (domain)
LGMD	Limb girdle muscular dystrophy
LINC	Linker of the nucleoskeleton and cytoskeleton
LPA	Lysophosphatidic acid
MAF	Mouse adult fibroblast
MAPK	Mitogen activated protein kinase
MCK	Muscle creatine kinase
MDa	Mega Dalton
MEB	Muscle-eye-brain
MH	Muscle hybrid
MKL1	Megakaryoblastic leukaemia 1
ml	Millilitre

MMP	Matrix metalloprotease
MOI	Multiplicity of infection
mRNA	Messenger ribonucleic acid
MTJ	Myotendinous junction
MyHC	Myosin heavy chain
NE	Nuclear envelope
Neo	Neomycin
NES	Nuclear export sequence
Nesprin	Nuclear envelope spectrin-repeat
NET	Nuclear envelope transmembrane
NHEJ	Non-homologous end joining
NL	Nuclear lamina
NLS	Nuclear localisation sequence
nm	Nanometre
nM	Nanomolar
NMIIA	Non-muscle myosin IIA
NMJ	Neuromuscular junction
nNOS	Neuronal nitric oxide synthase
NO	Nitric oxide
NPC	Nuclear pore complex
Nup	Nucleoporin
OHC	Outer hair cell
ONM	Outer nuclear membrane
PAM	Protospacer adjacent motif
PBS	Phosphate buffer saline
PCA	Protein complementation assay
PCR	Polymerase chain reaction
PD	Pull down
PDE-5	Phosphodiesterase-5
PDZ	PSD95, Dlg1, Zo-1 (domain)
PGK	Phosphoglycerate kinase
PH	Pleckstrin homology

PIPES	Piperazine-N,N'-bis(2-ethanesulfonic acid)
piRNA	Piwi-interacting ribonucleic acid
PLA	Proximity ligation assay
PNS	Perinuclear space
PoI	Protein of interest
POMT	Protein O-mannosyl-transferase
pRB	Phosphorylated retinoblastoma (protein)
rDNA	Ribosomal deoxyribonucleic acid
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate poly-acrylamide gel electrophoresis
SH	Src homology
SREBP	Sterol regulatory element-binding protein
SRF	Serum response factor
SU	Syntrophin unique
SUN	Sad1p, Unc-84
TAE	Tris-acetate-EDTA
TAN-line	Transmembrane actin-associated nuclear line
TEAB	Triethylammonium bicarbonate
TFA	Trifluoro acetic acid
TGF	Transforming growth factor
TIMP	Tissue inhibitor of metalloproteases
UB	Unbound
WT	Wild type
WS	Werner's syndrome
WW	Tryptophan-Tryptophan (domain)
WWS	Walker-Warburg syndrome

Chapter 1: Introduction

1.1. Dystroglycan and the dystrophin glycoprotein complex

Dystroglycan is the central protein within the dystrophin glycoprotein complex (DGC); a cell membrane associated scaffold for structural and signalling components. The DGC is a large multi-protein complex that is situated at regularly spaced intervals at the sarcolemma of myofibres, in alignment with costameres (Ervasti *et al.*, 1990; Minetti *et al.*, 1992). Costameres are structural, protein-containing components of a myofibre which denote points of attachment of contractile apparatus to the sarcolemma. In muscle, the DGC is composed of several securely bound proteins, including integral membrane components sarcospan, the sarcoglycans and dystroglycan which provide anchorage for intracellular proteins dystrobrevins, syntrophins, nNOS and dystrophin (Gao and McNally, 2015). Perhaps the most well-defined function for the DGC is its ability to form a molecular bridge between extracellular matrix (ECM) components and the cytoskeleton of the cell, thus contributing to membrane integrity (Ervasti and Campbell, 1991, 1993). However, it is clear that the DGC also has numerous signalling functions (Moore and Winder, 2010), and the identification of non-DGC complexed dystroglycan (Johnson *et al.*, 2013) across diverse subcellular localisations, including various nuclear compartments, indicates as yet unidentified roles for dystroglycan.

1.1.1. *Dystroglycan*

In humans, the dystroglycan protein is encoded by the Dag1 gene, located to chromosome 3 p21.31. It is comprised of two exons, separated by a large intron (Ibraghimov-beskrovnaya *et al.*, 1993). The transcript gives rise to a predicted 97 kDa type II transmembrane pre-peptide, which becomes autoproteolytically cleaved by an integral SEA (sea urchin, enterokinase, agrin) domain. This cleavage generates two dystroglycan subunits; the transmembrane β-subunit non-covalently tethers the α-subunit, which resides extracellularly (Ibraghimov-Beskrovnaya *et al.*, 1992; Holt *et al.*, 2000; Akhavan *et al.*, 2008). Within the DGC, dystroglycan functions as an

adhesion receptor. The amino-terminal α-dystroglycan is extensively post-translationally modified by both N-linked and O-linked glycosylation (Endo, 2007, 2015), which permits a well-characterised binding to the extracellular matrix (ECM) protein laminin, but also to numerous other LG-domain containing ECM components (Ervasti and Campbell, 1993) (Figure 1-1). On the intracellular side, β-dystroglycan ultimately binds the actin cytoskeleton via cytoskeletal linker proteins such as dystrophin and utrophin (Way *et al.*, 1992; Jung *et al.*, 1995; Rentschler *et al.*, 1999; Huang *et al.*, 2000), thus connecting extra- and intracellular environments, bridging the plasma membrane (Figure 1-1).

1.1.2. *Dystrophin and Utrophin*

Dystrophin is a large 427 kDa cytoskeletal linker protein, which binds β-dystroglycan through an interaction between the WW-domain and PPPY-motif in their respective C-termini (Jung *et al.*, 1995), an interaction stabilised by additional ZZ-domains and EF1-hands in the C-terminus of dystrophin (Ishikawa-Sakurai *et al.*, 2004). In this way, dystrophin completes a mechanical connection from the ECM, through the DGC, to the γ-actin cell cytoskeleton at its N-terminal (Way *et al.*, 1992). Utrophin is a protein structurally analogous to dystrophin (Tinsley *et al.*, 1992), and interacts with dystroglycan in the same manner as dystrophin (James *et al.*, 2000; Ishikawa-Sakurai *et al.*, 2004). Interestingly, utrophin can functionally replace dystrophin, which it does in many non-muscle tissues where dystrophin is not expressed in a utrophin glycoprotein complex (UGC). In diseases where dystrophin is lost, such as DMD, the UGC is upregulated, possibly as a compensatory mechanism (Tinsley *et al.*, 1998; Cullen *et al.*, 2001). However, utrophin can also diversify the roles of the complexes in which it is incorporated. In healthy muscle, utrophin expression is persistent upon myofibre differentiation, where it restricted to the myotendinous (MTJ) and neuromuscular junctions (NMJ) (Khurana *et al.*, 1991), where in the latter, utrophin is important to form postsynaptic folds for efficient acetylcholine receptor retention (Ohlendieck *et al.*, 1991; Deconinck *et al.*, 1997).

1.1.3. *The sarcoglycan complex*

The DGC also incorporates the sarcoglycan complex, a group of tightly associated transmembrane proteins which are resistant to separation by detergent treatment (Yoshida *et al.*, 1994). In skeletal muscle, the major sarcoglycan complex comprises the α-, β-, γ-, and δ-isoforms (Jung *et al.*, 1996; Chan *et al.*, 1998; Holt and Campbell, 1998) (Figure 1-1). Through interactions with central proteins of the DGC, the sarcoglycan complex is proposed to stabilise the binding between α- and β-dystroglycan, and that between β-dystroglycan and dystrophin (Yoshida *et al.*, 2000; Sandonà and Betto, 2009). Mutations in the sarcoglycans result in a variety of limb-girdle muscular dystrophies (Sandonà and Betto, 2009). Interestingly, the interaction between dystrophin and dystroglycan is retained upon the biochemical dissociation of the sarcoglycans from the DGC (Yoshida *et al.*, 1994), which may provide some explanation to why the sarcoglycanopathies are milder (Xie *et al.*, 2019) than other diseases where components central to the DGC are lost.

1.1.4. *Sarcospan*

Another function of the sarcoglycan complex is to target to, and stabilise, sarcospan within the DGC at the plasma membrane (Crosbie *et al.*, 1999). Sarcospan shares topological homology and a number of characteristics with the tetraspanin family of proteins, which have four transmembrane domains (Crosbie *et al.*, 1997) (Figure 1-1). Accordingly, like tetraspanins, sarcospan can laterally associate to form potential membrane micro-domains, scaffolding and enriching certain proteins with signalling or adhesion characteristics to make their functions more energetically favourable (Marshall and Crosbie-Watson, 2013; Termini and Gillette, 2017). Unlike with other DGC components, sarcospan deficiency does not result in catastrophic muscle wastage, (Lebakken *et al.*, 2000) indicating that mechanical linkage is not the sole function for the DGC; rather, the DGC could act as a platform for chemical transduction of mechanical or non-mechanical signals. Consistently with this theme, sarcospan overexpression in dystrophin-deficient mice reveals a role for sarcospan as a positive regulator of the Akt signalling pathway, which upregulates and stabilises

utrophin and integrin $\alpha7\beta1$ at the sarcolemma, thus contributing to an amelioration of disease phenotypes (Peter, Miller and Crosbie, 2007; Marshall, Holmberg, *et al.*, 2012).

1.1.5. *Non-membrane bound DGC components*

Along the same lines, the sarcoplasmic components of the skeletal muscle DGC, α-dystrobrevin and the syntrophins, α and β, seem to be principally involved in signalling. The C-terminus of dystrophin binds the syntrophins through their PH2 and SU (syntrophin-unique)-domains (Suzuki, Yoshida and Ozawa, 1995). The syntrophins are postulated to function as scaffold proteins able to integrate signalling cascades through their extensive array of binding domains. An integral PDZ-domain permits interaction with several other signalling receptors (Bhat, Adams and Khanday, 2013). A noteworthy syntrophin PDZ-domain interactor is neuronal nitric oxide synthase (nNOS) which syntheses nitric oxide (NO). The sarcolemmal localisation of nNOS, and localisation of the short-lived NO is critical for the modulation of muscular blood flow during exercise, since the signalling molecule powerfully stimulates vasodilation (Percival, 2011).

α- and β-dystrobrevins are expressed in muscle and non-muscle tissues, respectively. The C-terminus of dystrobrevin resembles that of dystrophin (Blake *et al.*, 1995). Through their respective coiled-coil domains, α-dystrobrevin binds to the distal C-terminus of dystrophin (Sadoulet-Puccio, Rajala and Kunkel, 1997) (Figure 1-1). Interestingly, α-dystrobrevin can also bind syntrophins (Butler *et al.*, 1992), potentially expanding their signalling capabilities. Fittingly, α-dystrobrevin appears to be essential for signalling from the DGC, and works in synergy with the syntrophins to recruit nNOS to the sarcolemma (Grady *et al.*, 1999) (Figure 1-1), but also to maintain structural characteristics of the muscle including the integrity of mature NMJs and MTJs, binding sarcoplasmic intermediate filaments and stabilising the interaction between dystrophin and β-dystroglycan (Grady *et al.*, 2000; Bunnell *et al.*, 2008; Nakamori and Takahashi, 2011). NMJ maturation requires the recruitment of signalling adaptor Grb2 by phosphorylated α-dystrobrevin (Gingras *et al.*, 2016), indicating the importance of α-dystrobrevin as a signalling platform.

1.2. Diseases of the DGC

Components of the DGC have diverse roles in both the maintenance of structural integrity and signalling. It is therefore unsurprising to find that a range of disorders arise from mutations in the DGC. Figure 1-1 shows the DGC and conditions which arise from its various components.

1.2.1. *Duchenne's and Becker's muscular dystrophy*

Perhaps the most well-known, and common, disorder pertaining to DGC function is Duchenne's muscular dystrophy (DMD). DMD results from mutations that cause frameshift and premature stop codons, leading to the loss of the cytoskeletal linker protein, dystrophin, which is encoded by the *Dmd* gene (Hoffman, Brown and Kunkel, 1987; Koenig *et al.*, 1987; Wilson *et al.*, 2017) (Figure 1-1). *Dmd* is located on the X-chromosome, and so the disease is almost exclusive to boys, with a prevalence of around 1 in 5000 male births (Mendell and Lloyd-Puryear, 2013). Becker's muscular dystrophy, also results from mutations in the *Dmd* gene, although dystrophin expression is maintained and retains some function, leading to milder muscular dystrophy symptoms (Blake *et al.*, 2002). DMD, is characterised by severe muscle wasting, with a large increase in serum creatine kinase levels; a marker for muscle damage. Muscle is extensively replaced with fibrotic tissues, which results in muscle weakness and loss of ambulation by 12 years of age. Patients usually succumb to the disease within their 20s as the diaphragm and cardiac muscles become overly weakened (Hoffman *et al.*, 1988).

1.2.2. *Limb-girdle muscular dystrophy can result from sarcoglycanopathies*

Some of the recessive forms of Limb-girdle muscular dystrophy (LGMD2C-F) result from mutations in any of the sarcoglycans (Sandonà and Betto, 2009) (Figure 1-1). The sarcoglycans are interdependent on each other for sarcolemmal localisation; if one is lost, the entire complex is absent (Lim *et al.*, 1995; Noguchi *et al.*, 1995; Bonneman *et al.*, 1996; Nigro *et al.*, 1996), along with sarcospan (Duclos *et al.*, 1998; Araishi *et al.*, 1999), leading to a muscular dystrophy phenotype in mice. Interestingly, the inverse is not true; loss of sarcospan alone does not destabilise the

sarcoglycan complex, or have an associated dystrophic phenotype (Lebakken *et al.*, 2000). In contrast to other integral components of the DGC, sarcospan appears to take on a compensatory disease modifying function through signalling. In the absence of dystrophin, sarcospan expression activates Akt signalling to induce utrophin upregulation (Marshall, Holmberg, *et al.*, 2012) and cross-talk with both the UGC and α7β1 integrin to stabilise muscle integrity (Peter, Marshall and Crosbie, 2008; Marshall, Chou, *et al.*, 2012).

1.2.3. *The dystroglycanopathies*

The dystroglycanopathies are a rare group of relatively heterogeneous conditions, and consensus seems to segregate them into three groups. Tertiary and secondary dystroglycanopathies arise from mutations which affect the metabolic synthesis and handling of glycosylation precursors, and the enzymes required to adduct them to the nascent dystroglycan peptide, respectively. Primary dystroglycanopathies are those where the *Dag1* gene itself is mutated leading to disease.

1.2.3.1. *Secondary and tertiary dystroglycanopathies*

The α-subunit of dystroglycan is broadly composed of three domains; an N- and C-terminal IgG-like globular domain either side of the mucin-like domain (De Rosa *et al.*, 2011). The mucin-like domain is responsible for ligand engagement, and is heavily modified by O-linked glycosylation (Ervasti and Campbell, 1993). Dystroglycan also contains four N-linked glycosylation sites; two on the amino- and one at the carboxy-terminus of the α-subunit, and one at the amino-terminus of the β-subunit (Figure 1-1). The functions of N-linked glycosylation modifications are unclear, as N-linked glycosylation appears to be dispensable for ligand binding and nuclear translocation of the β-subunit (Ervasti and Campbell, 1993; Oppizzi *et al.*, 2008). O-mannosyl glycans decorate serine and threonine residues in the mucin-like domain of α-dystroglycan and have been assigned to three groups depending on their core; M1-M3 (Yoshida-Moriguchi *et al.*, 2013). All O-mannosyl glycans are initiated by protein O-mannosyl transferases 1 and 2 (POMT1/2) (Manya *et al.*, 2004) and finally extended by the like-acetylglucosaminyltransferases enzyme (LARGE), which adds disaccharide moieties, conferring laminin binding ability to α-

dystroglycan (Yoshida-Moriguchi *et al.*, 2010; Inamori *et al.*, 2012). In total, there are at least 17 genes necessary for proper maturation of O-linked glycosylation on α-dystroglycan, to allow it to function as an extracellular receptor (Taniguchi-Ikeda *et al.*, 2016; Kanagawa and Toda, 2017), and each of these have been implicated in patients with congenital muscular dystrophies (CMDs) such as Walker-Warburg syndrome (WWS), muscle-eye-brain (MEB), Fukuyama congenital muscular dystrophy (FCMD) and several types of LGMD (Kanagawa and Toda, 2006; Mercuri, Bönnemann and Muntoni, 2019) (Figure 1-1). These secondary dystroglycanopathies are heterogeneous in their presentation and typically characterised by muscle wasting along with a variable brain and eye abnormalities (Martin, 2005), which may be partially explained by differential tissue requirements for α-dystroglycan glycosylation. Differences in glycosylation can be coarsely demonstrated by SDS-PAGE and western blotting displaying different apparent molecular weights for α-dystroglycan (Endo, 2015), and probably results from the tissue-specific distribution of various glycosyltransferases (Beltrán *et al.*, 2019), reflecting diverse functions of dystroglycan in various tissue types. For example, dystroglycan appears to be a prerequisite for brain basement membrane formation (Moore *et al.*, 2002), but not in muscle (Cohn *et al.*, 2002). Though, in patients, a reduction in glycosylation to dystroglycan does not increase disease severity or likelihood of brain involvement (Jimenez-Mallebrera *et al.*, 2009). The antibody used, IIH6, recognises the most distal glycosyl moieties made by LARGE, so judging dystroglycan glycosylation on this alone neglects the contribution of the more proximal glycosyl modifications. Consistently, it has also been found that levels of the core α-dystroglycan peptide in various dystroglycanopathies is unchanged, where IIH6 detects decreased levels (Humphrey *et al.*, 2015).

Where secondary dystroglycanopathies result from mutations in glycosyltransferase enzymes, tertiary dystroglycanopathies arise due to defective biosynthesis of the glycan monomers (Sheikh, Halmo and Wells, 2017), which reveals the possibility that many more as yet undefined dystroglycanopathies exist (Brancaccio, 2019).

1.2.3.2. *Primary dystroglycanopathies*

Information gleaned from observations of improper glycosylation in CMDs indicate that ligand engagement is essential for its role as a membrane receptor. It therefore seems that dystroglycan is an essential protein, and this might explain why very few true primary dystroglycanopathies have been reported in the literature. The complete absence of both dystroglycan subunits has only been described in one family. The loss of dystroglycan was caused by frameshift mutation 743C>del which resulted in A248G amino acid substitution, 19 incorrect amino acids followed by a premature stop codon. The patient investigated displayed severe WWS, and died after 3.5 months. Other affected individuals also displayed muscle and neurological defects, consistent with WWS, and died within hours or days of birth often due to respiratory failure (Riemersma *et al.*, 2015; Leibovitz *et al.*, 2018). Since this is the only known example of complete loss of dystroglycan, it is possible that human mutations are most often embryonic lethal. Indeed, Dag1 ablated mouse embryos die at day E6.5 due to the requirement of dystroglycan for the formation of Reichert's membrane which is a murine specific basement membrane of the embryo (Williamson *et al.*, 1997). However, it is plausible that dystroglycan has paralogous fundamental roles in early human development, as well.

Point mutations have also been reported throughout the dystroglycan protein; the T192M mutation within α-dystroglycan was the first to be described (Dinçer *et al.*, 2003). The patient displayed LGMD with cognitive impairment, but without structural abnormalities in the brain (Dinçer *et al.*, 2003). Epiblast and brain-specific Dag1 knockout mice have revealed that dystroglycan has critical roles in brain development including neuronal migration, cortical layering and basement membrane formation (Moore *et al.*, 2002; Satz *et al.*, 2008). These defects arise as a specific result of dystroglycan loss from the end feet of radial glia where the α-subunit is required for glia limitans perception and reinforcement. Radial glia perform a scaffold-like function and their perturbation, together with that of the pial basement membrane results in the misguided migration of other cortical neurone types neuronal migrations (Myshrall *et al.*, 2012). Interestingly, dystroglycan deficiency in the brain also causes over proliferation of neuronal progenitors (Myshrall *et al.*, 2012). Dystroglycan has been previously localised to the cytokinetic

34

cleavage furrow in dividing cells, and its depletion appears to reduce the proliferative rate and increase apoptosis of HeLa and REF52 cells (Higginson, Thompson and Winder, 2008). Yet, how dystroglycan is involved either directly, or indirectly in arrest of proliferation and terminal differentiation of neuronal progenitors is not clear. Given T192M is responsible for LARGE binding, and the resulting protein is hypoglycosylated (Hara *et al.*, 2011), the relatively mild muscular and brain symptoms displayed by the patient compared with other mutations and mouse models where Dag1 is ablated, similarly supports the notion that heterogeneity in dystroglycanopathies results from subtle differences in α-dystroglycan glycosylation, and each tissues dependence on specific modifications.

Three α-dystroglycan mutations have been identified in the N-terminal Ig-like region. The first, L86F, is present in a patient with arrhythmogenic cardiomyopathy in conjunction with a common PKP2 mutation (König *et al.*, 2017). However it is not known whether the L86F mutation is contributory to the disease, or even results in a compromise to dystroglycan function (Brancaccio, 2019). Expression of the L86F mutant and analysis in cells would provide insight into any detrimental features.

Mild muscular dystrophy has been reported in another patient with the Dag1 compound heterozygous mutations Val77Ile and Asp111Asn. Similarly to the T192M mutation, the compound heterozygous mutations affects the laminin binding capability of α-dystroglycan and there is no difference in expression levels (Dong *et al.*, 2015).

There are also primary dystroglycanopathies which originate in β-dystroglycan; to date, there have been two mutations described. Residue 669 resides in the extracellular portion of β-dystroglycan, and a Cys669Phe was discovered in a patient with MEB-like disease with multicystic leukodystrophy (Geis *et al.*, 2013), which results from defective myelination of the central nervous system. It is notable that a disulphide bond between Cys669 and Cys713 has also been identified as essential for cleavage of dystroglycan into the α- and β-subunits (Watanabe *et al.*, 2007). Although Cys669Phe was not tested, Cys669Arg/Ala displays robust plasma membrane localisation and α-dystroglycan immunoreactivity, indicating that glycosylation is not affected (Watanabe *et al.*,

2007). Deletion of dystroglycan has previously been reported to affect Schwann cell function (Saito *et al.*, 2003), and white matter dysfunction has been reported in WWS patients (Leibovitz *et al.*, 2018). However, the implication of multicystic leukodystrophy disease seen is exclusively in the presence of a Cys669Phe mutation insinuates disease mechanisms which are specific to β-dystroglycan dysfunction, perhaps arising from its inability to become liberated from its extracellular counterpart.

The only identified mutation in the cytoplasmic domain of β-dystroglycan is homozygous Arg776Cys, in a patient who showed late onset muscular dystrophy. Arg776Cys reportedly results in reduced levels of α-dystroglycan, though β-dystroglycan was not assessed (Dai *et al.*, 2018). The arginine residue 776 is situated juxtamembranally, and has been reported several times to constitute a residue of β-dystroglycan's NLS (Oppizzi *et al.*, 2008; Lara-Chacón *et al.*, 2010), although its contribution has never been directly tested and studies have instead been concentration on the nearby lysine residues. Given the comparatively minor disease compared with other dystroglycanopathies, it seems unlikely that Arg776Cys affects the membrane orientation of dystroglycan (Brancaccio, 2019), unless this happens in a small fraction of molecules.

In summary, the few primary dystroglycanopathies that have been discovered can perhaps be allocated to two groups. The first concerns mutations that affect glycosylation, and thus ECM binding, which often give rise to heterogeneous diseases aligned with those seen as a result of secondary dystroglycanopathies. Group two comprises those where α-dystroglycan is unaffected, and presumably appropriately engaged with extracellular ligands, while mutations in the β-subunit affect downstream functions in a mutation-site specific manner, dependent on binding partners for affected regions of the molecule. For these reasons, it is important to understand the multifaceted roles of β-dystroglycan throughout the cell.

36
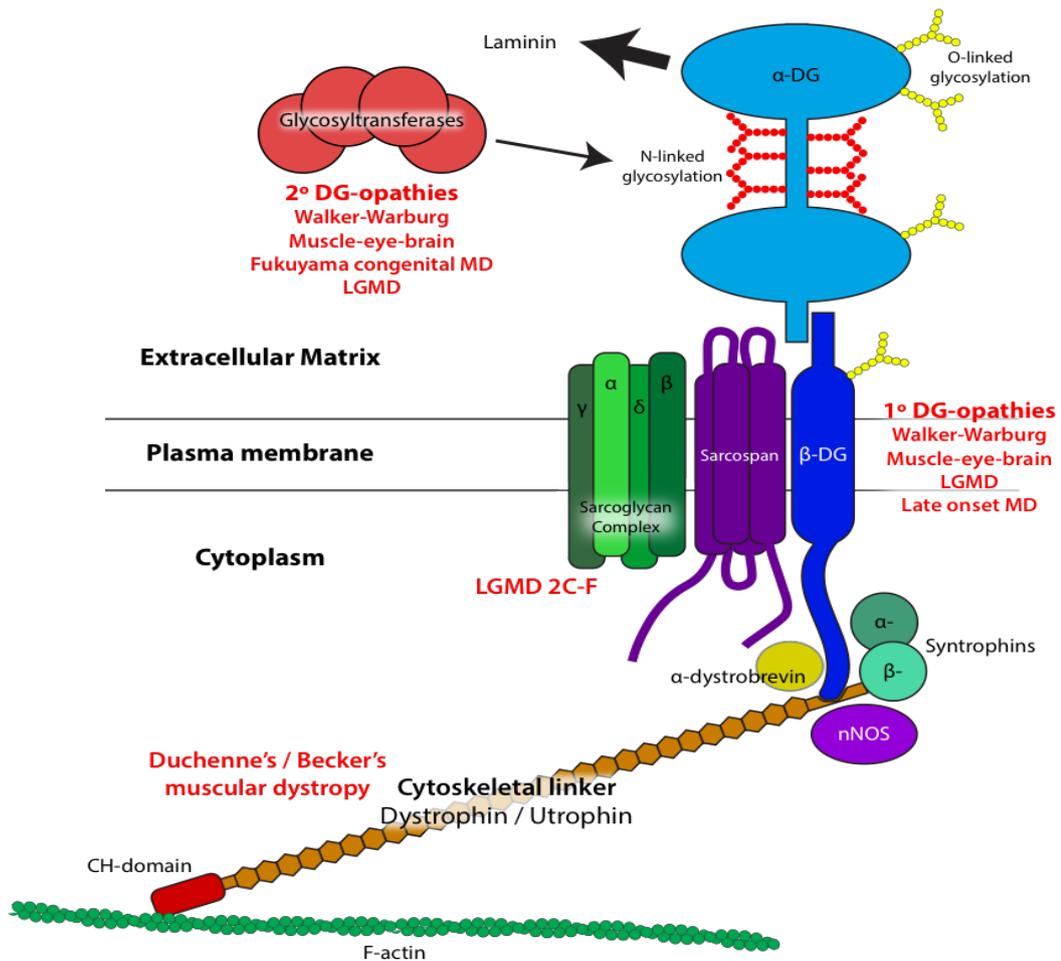
Figure 1-1 – The dystrophin-glycoprotein complex (DGC) and associated diseases. The DGC resides at the plasma membrane providing a connection from the extracellular matrix (ECM) to the cytoskeleton. The heterodimer, dystroglycan, is extensively glycosylated to confer ECM binding, and the β-subunit connects to the cytoskeletal linker proteins to attach the assembly to filamentous actin. Other members of the DGC are the sarcoglycans, sarcospan, dystrobrevin syntrophins and nNOS. Diseases linked to components are coloured red.

1.3. Functions of dystroglycan and the DGC

Research into the components of the DGC has challenged the initial view that its functions simply to connect the ECM to underlying cytoskeleton. New knowledge, and identification of diseases which arise from mutations within DGC proteins, indicate that the DGC has both structural and signalling functions, which, in some cases appear to overlap. In particular, sarcospan, whose loss does not elicit any muscular defects, appears only be involved in signalling, however, these functions emanate from the ability of sarcospan to structurally arrange membrane microdomains.

1.3.1. *The structural role of the DGC*

The notion that the DGC functions to provide the cell membrane with structural integrity has been established in view of the severe muscle necrosis which ensues in DMD, where a functional link between the ECM and the cytoskeleton is absent. In a tissue subject to mechanical strain such as muscle, the DGC is thought to dissipate forces away from the membrane of myofibre, thus acting in a protective capacity (Petrof *et al.*, 1993). In DMD, microtears in the sarcolemma lead to increased permeability and the exchange of ions and proteins between the intracellular and extracellular environments. In particular, a toxic influx of Ca^{2+} to the cell, either through sarcolemmal tears, or stretch activated ion channels is hypothesised to be a significant factor in myofibre degeneration (Van Westering, Betts and Wood, 2015). In addition to providing a structural function at the sarcolemma, dystrophin and the wider DGC, which depends upon dystroglycan, also appears to organise other cytoskeletal elements including microtubules and intermediate filaments (Carlson, 1998; Newey *et al.*, 2001; Stone *et al.*, 2005; Prins *et al.*, 2009), which may alter the structural properties of the myofibre thus increasing susceptibility to damage. Moreover, dystroglycan, within the DGC, together with the UGC, appears to have an important role in aspects of the formation of the NMJ, including acetylcholine receptor clustering and morphology of the folds, which is again defective in DMD (Nagel, Lehmann-Horn and Engel, 1990; Carlson and Roshek, 2001).

1.3.2. *The DGC and dystroglycan as a signalling platform*

In addition to its established structural roles, research over recent years has increasingly associated the DGC with signalling. Illustratively, the depletion of cytoplasmic DGC component α-dystrobrevin, which is involved in nNOS signalling, disrupts binding between dystrophin and β-dystroglycan (Bunnell *et al.*, 2008), consequently breaking the ECM-cytoskeletal axis. However, where this discontinuity leads to sarcolemmal fragility in DMD patients and mouse models, manifesting in severe muscle weakness and degeneration, mice harbouring deletions of α-dystrobrevin display no functional deficits, with normal costameric organisation (Bunnell *et al.*, 2008).

Indeed, recent research into DMD indicates several alternative pathological mechanisms rooted in the signalling functions of the DGC. Damaged myofibres are replaced by the activation, proliferation and differentiation of muscle stem cells, known as satellite cells, that reside in niches which pepper the periphery of mature myofibres. It transpires that the apical polarisation of dystrophin in satellite cells is required for asymmetric cell division giving rise to a daughter cells differentially capable of self-replication or commitment to the differentiation pathway (Dumont *et al.*, 2015). Interestingly, dystrophin is specifically complexed with the DGC as universal muscle depletion of DAG1 results in the same dysfunction of satellite cells (Dumont *et al.*, 2015). These observations substantiate the existence of a pathological mechanism whereby stem cell reserves are exhausted prematurely; initially able to maintain normal muscle regeneration (Cohn *et al.*, 2002), but eventually becoming overpopulated where many cells are unable to progress through the myogenic differentiation programme (Chang, Chevalier and Rudnicki, 2016). The DGC determines the self-replicative daughter cell by sequestration of Carm1, where it becomes phosphorylated by p38γ/MAPK12, thus precluding nuclear translocation. On the contrary, Carm1 remains active in the myogenic daughter cell, and methylates myogenic master regulator Pax7 in the nucleus, which subsequently binds and activates Myf5 by recruiting ASH2L:MLL1/2:WDR5:RBBP5 histone methyltransferase complex, committing to a myogenic programme. In the absence

of a functional DGC, p38γ is not sequestered resulting in exclusion of Carm1 from the nucleus in both cells reducing the differentiation potential (Chang *et al.*, 2018).

Components of the DGC bind and sequester nNOS to the sarcolemma, such that paracrine signalling of NO to adjacent vasculature is efficient. Accordingly, the loss of nNOS in DGC diseases can result in muscular ischaemia, contributing to metabolic deficiencies in exercised muscle and myofibre degeneration (Thomas, 2013).

Finally, fibrosis, the aberrant deposition of tough ECM material, is a common pathomechanism in diseases of central DGC components and not seen upon the loss of α-dystrobrevin (Bunnell *et al.*, 2008). Interestingly, a screen in *Drosophila* revealed that dystrophin and dystroglycan functionally cooperate with the TGF-β signalling pathway (Kucherenko *et al.*, 2008), and TGF-β dysregulation is the main driver of fibrotic deposition (Kharraz *et al.*, 2014). It is possible that the DGC has functions in modulating the TGF-β signalling pathway, through mechanisms currently unclear.

Numerous therapeutic strategies have been attempted to treat DMD, however the sole current universal treatment is glucocorticoids, which improves muscle strength by reducing inflammation and associated fibrosis (Manzur *et al.*, 2008). Other therapies which provide modest improvements in the cardiac, respiratory and skeletal muscle function in mice include phosphodiesterase-5 (PDE-5) inhibitors (Asai *et al.*, 2007; Khairallah *et al.*, 2008; Adamo *et al.*, 2010; Percival *et al.*, 2012). These work by inhibiting the breakdown of second messenger molecule cGMP, thus potentiating the reduced NO signal in DMD tissues to improve vascular function by vasodilation (Lau *et al.*, 1998). However, these observations did not translate well into human DMD patients (Leung *et al.*, 2014; Witting *et al.*, 2014). Since the DGC seems to be a critical nexus for structural integrity and signalling, recent studies are aimed at preserving the sarcolemmal positioning of dystroglycan to potentially alleviate many of the pathological mechanisms identified so far (Lipscomb *et al.*, 2011, 2016; Miller *et al.*, 2012).

Outside of the DGC, dystroglycan is also directly associated with signalling functions which seem to be focussed on two regions; the extreme carboxy-terminus, and the juxtamembrane region. The proline-rich C-terminus of dystroglycan

indicates great flexibility, with no distinct domains, lending it to multiple binding partners. It is also highly conserved among vertebrates, which indicates its importance to some conserved functions. Adhesion receptors, such as dystroglycan, not only anchor cells to their environment, but also act as sensory platforms, from which signalling can be initiated. This concept is well-established in integrins, where signals from ligand binding can be transduced 'outside-in' (Ginsberg, Partridge and Shattil, 2005; Legate, Wickström and Fässler, 2009). Similarly, binding of ECM by α-dystroglycan may result in the phosphorylation of tyrosine residue 890 at the extreme C-terminus of the mouse β-subunit, which corresponds to Y892 residue in the human sequence (James *et al.*, 2000; Sotgia *et al.*, 2001). A complete biological understanding and relevance of this modification remains unclear, but phosphorylation here apparently prevents the binding of dystrophin and utrophin (James *et al.*, 2000), while SH3-domain containing proteins including Grb2 can be recruited (Yang *et al.*, 1995; Russo *et al.*, 2000). In addition, β-dystroglycan is detected within the focal adhesions and podosomes of myoblasts, where it is indirectly responsible for modulating size through its interactions with Tks5, Src and vinexin (Thompson *et al.*, 2008, 2010). Alternatively, dystroglycan phosphorylation, which can result in internalisation (Miller *et al.*, 2012; Lipscomb *et al.*, 2016) maybe an initial step, and may resemble the 'inside-out' signalling phenomenon found in integrins where intracellular chemical changes to a receptor modulates ligand affinity (Calderwood, 2004), although it is unclear if α-dystroglycan is internalised, or if the β-subunit becomes liberated from its extracellular anchor in this context (Moore and Winder, 2010). Moreover, it has emerged that β-dystroglycan can be fragmented by metalloproteases MMP2 and MMP9 (Michaluk *et al.*, 2007; Bozzi *et al.*, 2009; Sbardella *et al.*, 2012), as well as γ-secretase (Leocadio, Mitchell and Winder, 2016). The conformation in which dystroglycan becomes internalised, and indeed which fragments are present, is likely to provide further answers to its signalling and/or structural roles downstream.

The intracellular juxtamembranal region of β-dystroglycan is also a focal point for signalling complex assembly. Dystroglycan overexpression has long been observed alongside exaggerated filopodial structures. β-dystroglycan's binding to

ezrin in the juxtamembranal region, and their co-recruitment of Dbl, locally activates Cdc42 the plasma membrane to promote filopodia construction (Y.-J. Chen *et al.*, 2003; Spence, Chen, *et al.*, 2004; Batchelor *et al.*, 2007). Moreover, dystroglycan is a scaffold for ERK signalling (Spence, Dhillon, *et al.*, 2004). Along with the presence of the DGC at the plasma membrane for anchoring other membrane components, dystroglycan-mediated ERK signalling is, at least in the brain, indispensable for the regulation of AQP4 (Zhang *et al.*, 2019).

1.4. Dystroglycan undergoes extensive fragmentation

Adding further complexity to its functions, it has emerged that β-dystroglycan is extensively processed by extracellular and juxta-membranous proteases, liberating numerous other β-dystroglycan fragments. Cleavage by extracellular matrix metalloproteases (MMPs)-2 and -9 releases a ~31 kDa, β-dystroglycan species which retains its transmembrane domain (Michaluk *et al.*, 2007; Bozzi *et al.*, 2009; Sbardella *et al.*, 2012), and appears to be cancer-associated (Losasso *et al.*, 2000; Singh *et al.*, 2004; Mathew *et al.*, 2013). A further cleavage event in the cytoplasmic juxtamembrane region of β-dystroglycan is mediated by γ-secretase, which releases the C-terminal soluble ~26kDa cytoplasmic fragment (Leocadio, Mitchell and Winder, 2016).

A common property of many DGC components is a transmembrane domain. As such, it might be expected that the membrane-associated nature of these proteins is maintained regardless of their sub-cellular context. Indeed, Oppizzi *et al.*, (2008) found that β-dystroglycan is present in the nucleus in its full length form retaining its transmembrane domain (Oppizzi *et al.*, 2008). Consistent with this, studies have assigned structural roles for β-dystroglycan at nuclear envelope (NE) (Martínez-Vieyra *et al.*, 2013), although immunofluorescence analysis indicates the presence of β-dystroglycan and other DGC components throughout the nucleoplasm of HeLa (Fuentes-Mera *et al.*, 2006), C2C12 (González-Ramírez *et al.*, 2008), and prostate cancer LNCaP cells (Mathew *et al.*, 2013; Mitchell *et al.*, 2013). These observations are supported firstly by the existence of membrane-liberated dystroglycan fragments, and it was also recently suggested that plasma membrane-

derived β-dystroglycan associates with the Sec61 translocon to extract the full-length form from its membranous environment (Gracida-Jiménez *et al.*, 2017). Once cytoplasmic, β-dystroglycan would presumably adopt an alternative soluble conformation where it may function differently and associate with nuclear import machinery.

1.5. <u>Dystroglycan localises to the nucleus</u>

More recently, roles for dystroglycan have emerged within the nucleus. Members of the DGC within the nucleus were first identified when the dystrophin splice isoform Dp71 was found to be two species; the Dp71f and Dp71d isoforms which localise to the cytoplasm and nucleus respectively in HeLa cells (Fuentes-Mera *et al.*, 2006). Interestingly, these observations were recapitulated in C2C12 myoblasts (González-Ramírez *et al.*, 2008), and in both cases immunoprecipitation experiments indicated the existence of an alternative DGC within the nucleus, although its precise functions have remained elusive.

The nuclear localisation of DGC components would appear to be mediated by β-dystroglycan, which contains a nuclear localisation sequence (NLS) in its intracellular juxtamembrane region between amino acids 776-782, thus facilitating nuclear import through recognition by α- and β-importins (Lara-Chacón *et al.*, 2010); a conventional pathway transiting the nuclear pore complex. Significantly, the non-covalent heterodimeric organisation of dystroglycan is required for autonomous localisation of the β-subunit to the nucleus, as is initial transit through the secretory pathways to the plasma membrane (Oppizzi *et al.*, 2008; Gracida-Jiménez *et al.*, 2017). However, it seems that the plasma membrane is not the only source of nuclear β-dystroglycan. At least to some extent, ezrin binding of β-dystroglycan enhances its nuclear localisation specifically from a cytoplasmic pool (Vásquez-Limeta *et al.*, 2014).

The internalisation of β-dystroglycan from the PM is regulated by its phosphorylation at tyrosine 890 ([p]Y890), however, evidence indicates that β-dystroglycan exists within the nucleus in both phosphorylated and unphosphorylated forms (Leocadio, Mitchell and Winder, 2016). Further

complicating the understanding of the origin of nuclear β-dystroglycan is the observations that at least some pY890 β-dystroglycan is ubiquitinated and degraded by the proteasome after internalisation. Indeed, preventing this phosphorylation event or proteasomal degradation both appear to maintain β-dystroglycan and the DGC at the sarcolemma, where it can maintain membrane integrity in the absence of dystrophin, as in animal models of DMD (Lipscomb *et al.*, 2011, 2016; Miller *et al.*, 2012). The mechanisms behind the sorting of β-dystroglycan pools and their designation for degradation or further trafficking are unknown. Indeed, whilst the role of phosphorylation for internalisation is well defined (Sotgia *et al.*, 2003; Miller *et al.*, 2012), whether it has a role in nuclear localisation is equivocal. It is notable that numerous studies investigating the translocation of β-dystroglycan utilised artificial constructs containing just the β-subunit of dystroglycan. Early studies indicated that a non-phosphorylatable (Y890F) point mutation in β-dystroglycan causes increased nuclear translocation (Lara-Chacón *et al.*, 2010), whereas phosphomimetic Y890E mutations cause preferential relocalisation to endocytic compartments, thought to be late endosomes (Sotgia *et al.*, 2003). Without α-dystroglycan and its N-terminal signal sequence, it is unclear whether these synthetic proteins are processed physiologically, which may lead to the misinterpretation of localisation data. A more recent study directly comparing the effects of Y890F and Y890E within the context of a full length dystroglycan construct suggests that phosphorylation is a prerequisite for nuclear import (Gracida-Jiménez *et al.*, 2017). However, this study does not address how both phosphorylated and non-phosphorylated Y890 β-dystroglycan of endogenous origin can be found in the nucleus (Leocadio, Mitchell and Winder, 2016).

Notably, the known C-terminal 31 and 26 kDa β-dystroglycan fragments each retain a functional bipartite NLS, so the existence of a soluble β-dystroglycan fragment also provides another mechanism by which β-dystroglycan can be localised to the nucleoplasm. Accordingly, expression of isolated β-dystroglycan fragments exhibit robust nucleoplasmic accumulation (Oppizzi *et al.*, 2008; Lara-Chacón *et al.*, 2010; Mathew *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). Similar to studies on the significance of Y890 on β-dystroglycan internalisation and nuclear localisation, it is

questionable whether a construct encoding β-dystroglycan or its fragments alone is appropriate to study its biology. Such β-dystroglycan constructs appear to lack plasma membrane localisation, typically signified by enhanced filopodial structures. Ideally, a full-length construct would be used to study the nuclear localisation of β-dystroglycan, however, current studies have utilised C-terminal GFP-fused β-dystroglycan constructs, while all antibodies raised against the protein are directed towards the extreme C-terminus of the protein. Using these tools, each of the β-dystroglycan fragments cannot be distinguished *in situ*, obscuring any important and specific functions.

Recent studies have also identified an evolutionarily conserved a nuclear export sequence (NES) buried within the transmembrane domain (TMD) of β-dystroglycan, which appears to maintain equilibrium of the protein within the nucleus (Vélez-Aguilera *et al.*, 2018).

1.6. The organisation and function of the nuclear periphery

1.6.1. *The nuclear envelope*

Most eukaryotic cells contain a nucleus which spatially and functionally partitions the cells' DNA from the cytoplasm. Demarcating the interface between the nucleus and cytoplasm is the nuclear envelope (NE), which is a double lipid bilayer formed of the inner nuclear membrane (INM) and outer nuclear membrane (ONM). The membrane systems of the INM, ONM are topologically related to each other, and to the endoplasmic reticulum (ER). As such, the INM and ONM are connected where the NE is perforated by nuclear pore complexes (NPCs), and the ONM is continuous with the membrane of the ER. Consequently, the compartment between the INM and ONM, known as the perinuclear space (PNS), is also continuous with the lumen of the ER. Beneath the INM lies the nuclear lamina (NL) which is composed of a complex meshwork of insoluble type V filamentous proteins called the lamins. This functions to provide structural integrity to the nucleus and anchor a plethora of NE associated proteins (Figure 1-2A). The predominant lamins in mammalian somatic cells are A-type (lamins A and C) which are alternatively spliced from the same gene, and B-type lamins (lamins B1 and B2), which arise from separate genes.

Together, elements of the NE function to maintain genomic integrity by providing a barrier to components of the cytosol, but also to organise many aspects of the nucleus (Gruenbaum *et al.*, 2005). The NL and some associated NE proteins are responsible for diverse roles within the nucleus in genome tethering, including transcriptional regulation and DNA damage repair, while other NE proteins are responsible for the active positioning of the nucleus within the cell.

1.6.2. *The nuclear lamina*

Components of NL were first described as an 'amorphous layer' which biochemically purified in association with NPCs. This 'fibrous lamina' contained three protein species which were noted to be resistant to detergent and salt extraction (Aaronson and Blobel, 1975). Further work indicated that these localised to the nuclear periphery (Gerace, Blum and Blobel, 1978), and the three bands of molecular weights 70 kDa, 67 kDa, and 60 kDa became known as lamins A, B and C, respectively. The predominant B-type lamins were discovered to arise from two separate genes, so becoming known as lamin B1 and B2 (Höger *et al.*, 1990). Sequence homology analysis identified the nuclear lamins as members of the intermediate filament (IF) family (Krohne *et al.*, 1987), and like other IFs contain an N-terminal head domain, four α-helical domains in their central rod interspersed by linker peptides (Fisher, Chaudhary and Blobel, 1986; Mckeon, Kirschner and Caput, 1986). The carboxy terminus of the lamin molecules is globular and forms from an immunoglobulin fold with a terminal characteristic lamin motif CaaX (C, cysteine; a, an aliphatic amino acid; X, any amino acid) (Dhe-Paganon *et al.*, 2002). All B-type lamins and immature lamin A, known as pre-lamin A, contain the CaaX motif which provides the site for the successive C-terminal post-translational modifications of farnesylation and methylation (Sinensky *et al.*, 1994). It is thought that farnesylation permits anchorage to the nuclear membrane (Kitten and Nigg, 1991), although it is handled differently between lamin isoforms. Mature B-type lamins are permanently farnesylated, and the final three amino acids (aaX) are cleaved (Rusiñol and Sinensky, 2006). In contrast, pre-lamin A contains an extended C-terminus which is recognised for proteolytic cleavage by the ZMPSTE24 enzyme, thus removing the most C-

46

terminal 18 amino acids, including its modified CaaX motif (Weber, Plessmann and Traub, 1989; Corrigan *et al.*, 2005).

Besides the main lamins A, C, B1 and B2, there are additional isoforms. The lamins AΔ10 and C2 arise from alternative splicing of the *Lmna* gene (Furukawa, Inagaki and Hotta, 1994; Machiels *et al.*, 1996), and the former is found at low levels in many cell types. Lamin C2, on the other hand, was conventionally exclusively found in meiotic cells involved in spermatogenesis (Smith and Benavente, 1992; Furukawa, Inagaki and Hotta, 1994), however more recent reports indicate its presence in both embryonic and mature oocytes, as well (Link *et al.*, 2013; Koncicka *et al.*, 2020), indicating lamin C2 expression is meiosis specific. Similarly, lamin B3, which is splice isoform of the LMNB2 gene, is also expressed in germ cells of the testis (Furukawa and Hotta, 1993), although it is restricted to post-meitoic stages (Schütz *et al.*, 2005). Unlike lamin C2, there is no evidence for lamin B3 expression in female germ cells.

1.6.2.1. *Assembly of the nuclear lamina*

As mentioned, lamins are intermediate filaments, and *in vitro* studies have revealed that lamins self-assemble in a tiered manner. A monomeric lamin molecule measures ~52nm (Heitlinger *et al.*, 1991), and these dimerise through interactions between their rod domains which contain coiled-coil motifs (Dittmer and Misteli, 2011). Lamin dimers can associate longitudinally with each other in a staggered antiparallel fashion to form higher order structures (Stuurman, Heins and Aebi, 1998), and it is these filaments which display a regular pattern of axial beading under electron microscopy, formed of the globular N and C-terminal regions typical of lamin filaments (Heitlinger *et al.*, 1991). A further level of assembly sees the lamin filaments associate laterally (Figure 1-2B), which appears dependent on their globular head and tail domains (Heitlinger *et al.*, 1992). Under some *in vitro* conditions, lamin filaments form vast paracrystalline arrays (Goldman *et al.*, 1986).

1.6.2.2. *Lamin organisation* in vivo

The endogenous NL was first visualised as a filamentous meshwork by electron microscopy analysis of the nucleus of *Xenopus laevis* oocytes (Aebi *et al.*, 1986). For many years, the lamin assembly was assumed to be similar in mammals, despite that

in *X. laevis* being entirely composed of B-type lamins (Krohne, Dabauvalle and Franke, 1981), however, it was also found to be filamentous in nature in mammalian cells (Mahamid *et al.*, 2016). Unlike *X. laevis* nuclei, mammalian cells possess considerable peripheral heterochromatin which contributes to the already crowded molecular environment of the NL to obscure the molecular details lamina organisation *in vivo* (Burke and Stewart, 2013). The mammalian NL was revealed by removing chromatin through DNase treatment of mouse embryonic fibroblast nuclei (Turgay *et al.*, 2017), and measurements in this model corroborated those previously determined for lamina meshwork thickness (10-30nm) (Fawcett, 1966) while the distance between axial beads was equivalent to those found in filaments formed from *C. elegans* lamin monomers (Foeger *et al.*, 2006; Ben-Harush *et al.*, 2009), together indicating that that NL structure is conserved throughout evolution. However, mammalian filament width *in vivo* measure ~4 nm (Mahamid *et al.*, 2016; Turgay *et al.*, 2017), while those in *X. laevis* form filaments of 10nm (Aebi *et al.*, 1986). Notably, *in vitro* polymerisation of lamin filaments from *C. elegans* and rat liver also form stable filaments measuring 10 nm in diameter (Aebi *et al.*, 1986; Wei *et al.*, 1996; Karabinos *et al.*, 2003), while *Drosophila* lamins aggregate and form paracrystalline arrays rather than distinct filaments (Sasse *et al.*, 1997). Lateral lamin assembly appears to be determined by the head domains of lamin monomers (Heitlinger *et al.*, 1992; Moir *et al.*, 2000), although there must also be local determinants, since mammalian filament width in vivo has been reported to measure ~4 nm (Mahamid *et al.*, 2016; Turgay *et al.*, 2017; Ahn *et al.*, 2019). Despite their structural similarities, and ability to hetero-oligomerise *in vitro* (Schirmer and Gerace, 2004), different lamins form separate filamentous networks *in vivo* and seem to be differentially involved in chromatin organisation (Shimi *et al.*, 2008) and NPC anchorage (Xie *et al.*, 2016). The existence of separate lamin networks was confirmed at the ultrastructural level using co-immunogold labelling combined with cryo-EM tomography (Turgay *et al.*, 2017).

1.6.3. *The function of the nuclear lamina*

Given the rigidity and relative stability of the intermediate filament family of proteins, it was initially thought that lamins, as IFs, impart similar physical properties

on the nucleus (Aebi *et al.*, 1986). Indeed, once the lamina is constructed, it remains relatively stable throughout interphase (Broers *et al.*, 1999). However, this structural function for the NL appears more nuanced than first imagined, owing to the existence of multiple lamins. Nonetheless, it seems that physical properties, such as nuclear size, shape and stiffness, are regulated by lamins, but the lamin meshwork also provides a framework on which numerous other peripheral proteins can assemble, diversifying its functional repertoire; incorporating chromatin attachments which facilitate genome organisation, DNA damage repair and regulation of gene expression (Ho and Lammerding, 2012).

1.6.3.1. *Lamins are required for nuclear size and shape*

Lamin B3 is the sole lamin isoform expressed in the oocyte of X. laevis, and its loss, either by genetic ablation or inhibition of nuclear import results in decreased nuclear volume (Newport, Wilson and Dunphy, 1990; Levy and Heald, 2010). However, the lamins are not the only mediator of nuclear size, as disruption of NE proteins such as LAP2β, are also associated with altered nuclear volumes (Yang, Guan and Gerace, 1997; Gant, Harris and Wilson, 1999). Nuclear shape is also regulated by the composition of the NL. Strikingly, the expression of testis specific lamin B3 in fibroblasts induces a hook-shaped nucleus, characteristic of spermatocytes (Furukawa and Hotta, 1993).

1.6.3.2. *The nuclear lamina anchors integral proteins of the nuclear envelope, the nuclear envelope transmembrane (NET) proteins*

In a similar manner to the lamina, NE proteins were first identified biochemically. The presence of lamina associated proteins (LAPs), as they became known, was shown by their relative resistance to extraction under high salt and detergent conditions (Senior and Gerace, 1988; Foisner and Gerace, 1993). Subtractive proteomic analysis of the NE identified 80 nuclear envelope transmembrane (NET), and associated proteins which resided within this region (Schirmer *et al.*, 2003), which represented a great expansion to those previously identified. Some of the well-known NE proteins are widely expressed, however, the cell-type specificity of the NE proteome (Schirmer and Gerace, 2005; Talamas and Capelson, 2015) is becoming increasingly clear, as comparative proteomic analyses

have now identified over 1000 NET proteins within the NE of leukocytes, muscle and liver tissues (Korfali *et al.*, 2010, 2012; Wilkie *et al.*, 2011). The prior discovery that NETs are differentially regulated throughout developmental processes, in this case myogenesis, provides some insight into the plasticity of the NE region (Chen *et al.*, 2006). In fact, the complement of NETs present within a cell also influences global genome organisation, transcription and cell cycle progression (Srsen, Nadia and Schirmer, 2011; Wong, Luperchio and Reddy, 2014).

1.7. <u>Biochemical communication across the nuclear envelope</u>

By far the largest single structures within the NE are the nuclear pore complexes (NPCs). NPCs are composed of multiple copies of 34 different proteins, known as nucleoporins (shortened to Nups), with the resulting pore displaying 8-fold molecular symmetry (Lin *et al.*, 2016). The assembled ~125 MDa NPC protein assembly (Reichelt *et al.*, 1990) is anchored to the NL, and its embedding within the NE forms a channel between the nucleus and the cytoplasm. Similarly to the NL, the NPC is a stable structure in the NE (Daigle *et al.*, 2001). The NPC is constructed from proteins which can be categorised into four groups. These are the pore membrane (POM) domain proteins, which are integral membrane proteins, and the nucleoporins of the coat, adaptor, and channel regions. Filaments extends into the cytoplasm, while the attached nucleoplasmic proteins of the NPC form a nuclear basket (Hoelz, Debler and Blobel, 2011) (Figure 1-2A). The main function of NPCs is to facilitate nucleo-cytoplasmic transport of biomolecules, and properties of Nups within NPCs govern this process. The channel Nups incorporate intrinsically disordered regions rich in phenylalanine and glycine residues, known as FG-repeats, which form the NPC permeability barrier (Ribbeck and Görlich, 2002). The precise mechanisms behind the selectivity are not completely clear, but it seems that FG-repeats are extensively modified by glycosylation, and has been recently demonstrated to partition into a distinct membrane-less phase (Celetti *et al.*, 2020). The permeability barrier permits free diffusion of molecules below 40 kDa, while larger molecules are excluded, unless they are accompanied by chaperone proteins, known as the karyopherins (Yuh and Blobel, 2001).

The karyopherin family of proteins are sub-divided into importins and exportins dependent on the directionality of their cargo movement. The importin complex has two subtypes, α and β (Hoelz and Blobel, 2004). Nuclear import cargoes have a nuclear localisation sequence (NLS) integrated within the peptide; classical sequences consist of a clustering of basic, positively charged residues, either in a monopartite (Kalderon, Richardson, *et al.*, 1984; Kalderon, Roberts, *et al.*, 1984), or bipartite configuration (Robbins *et al.*, 1991). Non-classical sequences tend to be hydrophobic, can be three-dimensional, and are recognisable by conventional import machinery, or be specific to certain karyopherins (Kim, Han and Oh, 2017). NLSs are recognised by the Armadillo (ARM) domains of α- and β-importins (Conti and Kuriyan, 2000). The assortment and diversity of import cargoes is accounted for by the existence of numerous importins (Kim, Han and Oh, 2017), and each contain up to 20 HEAT-domain which form a flexible superhelical structure with concave morphology to presumably sequester a range of transported proteins (Conti *et al.*, 1998; Conti, Müller and Stewart, 2006). Once a cargo is bound, HEAT domains of β-importin can dock with FG-repeat containing Nup98 to facilitate nuclear import (Radu, Moore and Blobel, 1995; Cook *et al.*, 2007), however, some cargo proteins are specifically recognised by α-importin, which lacks the capability to bind to Nups, so cargo laden α-importin complexes with β-importin via an importin-β binding (IBB) domain for import (Görlich *et al.*, 1996; Weis, Ryder and Lamond, 1996; Lott and Cingolani, 2011). Nuclear transport is an active process where cargo release is dependent on the maintenance of a RanGTP gradient across the nuclear membrane. To sustain high RanGTP levels within the nucleus, guanine exchange factors (RanGEF) and GTPase activating proteins (RanGAP) are localised to the nucleus and cytoplasm, respectively (Ohtsubo, Okazaki and Nishimoto, 1989; Matunis, Wu and Blobel, 1998). High levels of RanGTP within the nucleus induce dissociation of the importin-cargo complexes, and promotes the association of α-importin to export karyopherins for cytoplasmic recycling (Moroianu and Blobel, 1995; Richards, Carey and Macara, 1997).

Once inside the nucleus, NE proteins are thought to remain in position by selective retention based on interactions with other resident NE or NL proteins or

genomic elements (Powell and Burke, 1990; Soullam and Worman, 1995; Ellenberg *et al.*, 1997; Ohba *et al.*, 2004). Topologically however, the INM is contiguous with the ONM, but the region in between, known as the pore membrane, is occupied by the NPC, thus presenting an obstacle for transit. Several mechanisms have been proposed to permit translocation of transmembrane proteins to the INM. It seems that proteins with small nucleoplasmic domains appear to be able to diffuse across the pore membrane, while larger proteins are energy-dependent (Ohba *et al.*, 2004). Despite the apparent obstruction posed by the NPC, larger NE proteins potentially utilise similar import machinery as soluble cargoes (King, Lusk and Blobel, 2006), and passage to the nuclear interior appears dependent on disordered regions of protein, which are inherent to INM proteins, and a change NPC conformation through Nup170 (Lusk, Blobel and King, 2007; Meinema *et al.*, 2011). More recently, experiments have substantiated these hypotheses, revealing that large NLS-containing nucleoplasmic domains transit the central channel of the NPC, while smaller INM proteins lacking an NLS take the peripheral route; presumably becoming retained by local interactions. Notably, where the NLS is mutated, INM proteins with large nucleoplasmic domains still transit to the INM, albeit more slowly (Mudumbi *et al.*, 2020).

1.8. A mechanical connection: The linker of the nucleoskeleton and cytoskeleton (LINC) complex

In addition to biochemical communication between the cytoplasm and nucleus, facilitated by the NPC (discussed in section 1.7), the nucleoskeleton is further elaborately interconnected with the rest of the cell cytoskeleton, enabling mechanical communication. Mechanical connections between the structural proteins of the nucleus and cytoplasm enables force stimuli to be efficiently and rapidly transmitted to the nucleus, which enables bulk nuclear movement for cell migration, movement of components within the nucleus and renders the nucleus mechanosensitive. The linker of the nucleoskeleton and cytoskeleton (LINC) complex is a hetero-oligomeric protein structure that forms a chain across the NE, linking within the PNS. SUN (Sad1, Unc-84 homology)-domain containing proteins span the

INM and interact with the KASH (Klarishct, Anc-1 and Syne homology)-domain of KASH-domain containing proteins within the PNS (Crisp *et al.*, 2006), which traverse the ONM. Simply, the LINC complex forms an uninterrupted molecular bridge, mechanically associating the NL to components of the cytoskeleton. SUN-domain proteins anchor to lamins of the nucleoskeleton, while KASH-domain proteins interact with the cytoskeleton (Figure 1-2A). Complexity is built into the LINC complex since many organisms, including those in the class Mammalia, possess at least 5 Sun-domain and 6 KASH-domain proteins, some of which are expressed ubiquitously, while others in a tissue-specific manner. Structural differences in these proteins specify binding regions for different interaction partners, thus diversifying their functions.

1.8.1. *KASH-domain proteins*

The KASH-domain proteins are evolutionarily well-conserved; in *C. elegans*, these are Anc-1 and Unc-83 (Starr and Han, 2002; McGee *et al.*, 2006), and in *D. melanogaster* the KASH-containing protein is Klarishct (Mosley-Bishop *et al.*, 1999). In mammals, KASH-containing proteins include four members of the nuclear envelope spectrin-repeat (Nesprin) family, termed Nesprin-1-4 (Zhang *et al.*, 2001; Wilhelmsen *et al.*, 2005; Roux *et al.*, 2009), together with structurally divergent KASH5 and LRMP (Lindeman and Pelegri, 2012; Morimoto *et al.*, 2012; Kozono *et al.*, 2018). Nesprins-1 and -2 are large proteins of 800-1000 kDa, highly expressed in the striated muscles of the heart and skeleton, and are encoded by the SYNE1 and SYNE2 genes respectively (Zhen *et al.*, 2002; Padmakumar *et al.*, 2004). Like other KASH-domain proteins, Nesprins-1 and -2 interact with Sun-proteins by KASH-peptide binding to the groove formed by adjacent Sun1-domain. As determined through studies using Sun2, conserved cysteine residues within the KASH-peptide are required to form disulphide bonds with the Sun-domain (Sosa *et al.*, 2012). Nesprins-1 and -2 are extensively alternatively spliced and the full-length isoforms, termed Nesprin-1-Giant (Nesprin-1G) and Nesprin-2-Giant (Nesprin-2G), retain an amino-terminal calponin-homology (CH) domain which enables cytoskeletal binding to the filamentous actin network (Starr and Han, 2002; Zhen *et al.*, 2002; Padmakumar *et al.*, 2004) (Figure 1-2A). The central section of Nesprins-1 and -2 is filled with up to

74 or 56 spectrin repeats, respectively (Zhang *et al.*, 2001). Depending upon the Nesprin spliceoform, these spectrin repeats can enable further diversity in cytoskeletal interactions, providing binding sites for actin-bundling formin FHOD1 (Kutscheidt *et al.*, 2014; Antoku *et al.*, 2019) and microtubule motor proteins (Wilson and Holzbaur, 2015; Holt *et al.*, 2019; Gonçalves *et al.*, 2020) (Figure 1-2A).

The SYNE3 gene encodes Nesprin-3 which contains a unique spectrin repeat, enabling its binding to the IF network via plectin (Wilhelmsen *et al.*, 2005) (Figure 1-2A). Nesprin-3 is also highly expressed in myoblasts, but becomes downregulated upon differentiation (Randles *et al.*, 2010; Ketema *et al.*, 2013). Studies indicate that Nesprins-1-3 form an interconnected network with each other, traversing the surface of the NE in order to control nuclear size (Lu *et al.*, 2012; Cartwright and Karakesisoglou, 2014). Interestingly, Nesprin-3 appears to be dispensable for embryogenesis (Ketema *et al.*, 2013), and its absence appears to reduce the accumulation of intermediate filaments from the perinuclear region (Postel *et al.*, 2011).

Expression of Nesprin-4 is restricted to the sensory hair cells of the cochlea and secretory tissues such as the mammary and salivary glands (Roux *et al.*, 2009). Outer hair cells (OHCs) of the cochlea are responsible for signal amplification, and genetic ablation of Nesprin-4 in mice releases the normally basolaterally anchored nuclei, leading to progressive deafness as cells expire, though the pathological mechanism is not known (Horn, Brownstein, *et al.*, 2013). Like other nesprins, the ONM localisation of Nesprin-4 is dependent on Sun-domain proteins (Roux *et al.*, 2009), and in OHCs, a Sun1-Nesprin-4 LINC complex specifically seems to mediate basolateral nuclear anchorage as Sun1 deficient mice are also deaf (Horn, Brownstein, *et al.*, 2013). The centrosome is often observed in close proximity to the nucleus, although in OHCs it is maintained apically through the exclusive ability of Nesprin-4 to engage kinesin subunit Kif5b, causing anterograde microtubule movement of the nucleus (Roux *et al.*, 2009).

KASH5 is meiosis-specific, and functions with Sun1 and dynein-dynactin to mediate telomere motility during meiotic-bouquet formation (Horn, Kim, *et al.*, 2013). (Further discussed in section 1.8.3.2.4)

The final known KASH-domain protein is lymph-restricted membrane protein, *Lrmp*. During the fertilisation of zebrafish, the *Lrmp* homologue, *Fue*, is necessary for pronuclear attachment to the centrosome, and localises to the NE, centrosomes and spindles during this process (Lindeman and Pelegri, 2012). *Lrmp* remains poorly characterised in mammalian cells, but seems to have some involvement in regulating nuclear shape. Like other KASH-domain proteins, *Lrmp* binds Sun1 and Sun2, and appears to localise to NE and ER membranes (Behrens *et al.*, 1994; Kozono *et al.*, 2018). It is not known whether *Lrmp* binds microtubules directly or through adaptors.

1.8.2. *Sun-domain proteins*

Sun-domain proteins are characterised by an evolutionarily conserved carboxy-terminal globular Sun-domain which resides in the PNS (Hodzic *et al.*, 2004; Crisp *et al.*, 2006). As type I transmembrane proteins, the amino-terminus is in the nucleoplasmic region, where it can interact with components of the NL, including A-type and B-type lamins (Haque *et al.*, 2010; Nishioka *et al.*, 2016). Sun proteins are conserved throughout evolution, and homologues can be found in plants (Graumann, Runions and Evans, 2010); *Schizosaccharomyces pombe* as Sad1 (Hagan and Yanagida, 1995); *C. elegans*, as Unc-84 (Malone *et al.*, 1999); and *D. melanogaster*, as Klaroid (Kracklauer *et al.*, 2007). In mammals, there are 5 genes which encode Sun-domain proteins, and these are termed Sun1, Sun2, Sun3, Sun4 (SPAG4), and Sun5 (SPAGL4), of which Sun1 and Sun2 are widely expressed (Crisp *et al.*, 2006). Sun3, SPAG4 and Sun5 appear to be mostly limited to spermatids. Sun3 is expressed post-meiotically, and restricted to the posterior of the NE (Göb *et al.*, 2010), where it interacts with SPAG4, together functioning to shape the sperm head by connecting to the microtubular manchette structure (Calvi *et al.*, 2015; Pasch *et al.*, 2015). Unlike Sun1 and Sun2, which under some circumstances are redundant, the stability of Sun3 and SPAG4 are interdependent, and Sun4 levels are significantly decreased upon genetic ablation of Sun3 (Gao *et al.*, 2020). The form of the sperm head is a crucial determinant for fertility, so it is unsurprising that loss of either Sun3 or Sun4 results in male sterility (Pasch *et al.*, 2015; Gao *et al.*, 2020). Sun5, on the other hand, is localised to the apical surface of the NE (Frohnert, Schweizer and

Hoyer-Fender, 2011), although does not appear to be involved in the attachment of the NE to the acrosome in spermatids (Yassine *et al.*, 2015). The contribution of Sun5 to fertility has not been assessed by any mouse model, while Sun3 and Sun4 appear dispensable for female fertility. Indeed, Sun4 appears not to be expressed in the ovary (Pasch *et al.*, 2015).
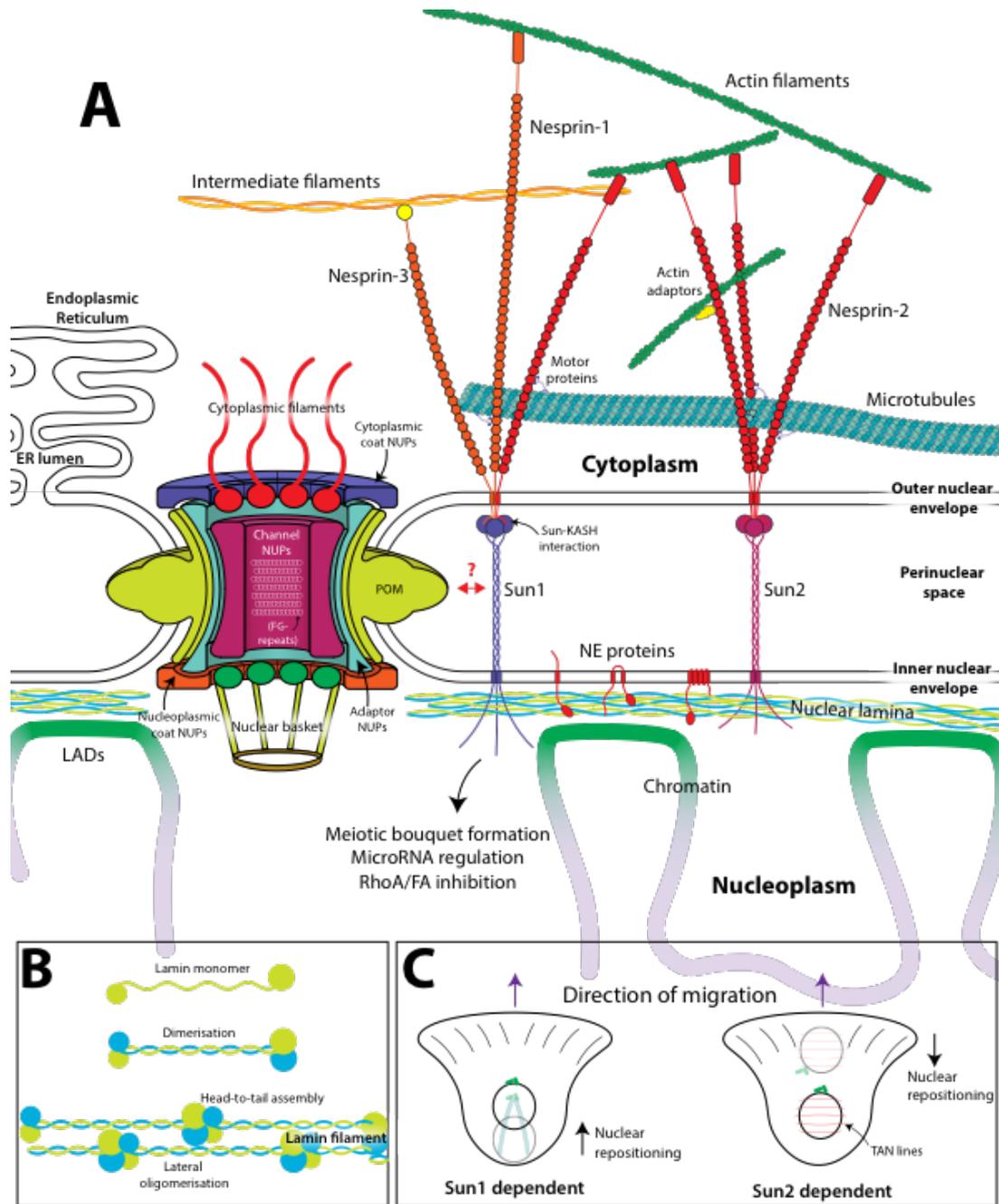
Figure 1-2 – Organisation of the nuclear envelope. (A) The NE is contiguous with the endoplasmic reticulum (ER). Nuclear pore complexes (NPCs) (shown on the left) traverse the NE to allow macromolecule transfer across the NE. LINC complexes (to the right of NPC), composed of the Sun-domain and KASH-domain proteins traverse the NE, providing a connection between the nucleoplasm and cytoskeleton. (B) Assembly of lamin monomers in to the NL. (C) Sun-domain proteins appear to have divergent functions, including in nuclear positioning during cell polarisation. Sun1 and microtubules are required for forward movement of nuclei, whereas Sun2 and TAN lines are crucial for the rearward movement of nuclei.

1.8.3. *The somatic Suns: Sun1 and Sun2*

Sun1 and Sun2 are ubiquitously expressed in somatic cells, and they have numerous structural and topological similarities. In addition to their carboxy terminal Sun domains, Sun1 and Sun2 also share coiled-coil domains in the stalk region, hydrophobic transmembrane regions altogether contributing to 67% identity between the mouse isoforms in total (Haque *et al.*, 2006). Topologically, the Sun-domain of both isoforms localise to within the PNS, while the amino-terminus is nucleoplasmic (Hodzic *et al.*, 2004; Crisp *et al.*, 2006).

1.8.3.1. *Functional redundancy between Sun1 and Sun2*

Given the conservation between Sun1 and Sun2, initial investigations pointed towards functional redundancy between the isoforms (Crisp *et al.*, 2006; Haque *et al.*, 2006). It is possible for the separate gene products to hetero-trimerise through their coiled-coil domains (Minn *et al.*, 2009) and globular SUN-domains (Zhou *et al.*, 2012). Consistent with this, there is a body of evidence which supports partial redundancy in the molecular interactions and physiological functions between Sun1 and Sun2. In addition to binding of lamina components equally (Crisp *et al.*, 2006), Sun1 and Sun2 are capable of interacting with Nesprin-1 and short nesprin isoforms (Lei *et al.*, 2009; Haque *et al.*, 2010) and are both required for the NE tethering of Nesprin-2G (Crisp *et al.*, 2006). NE spacing is only disrupted when both Sun1 and Sun2 are depleted using RNA interference, or when all LINC complexes are disrupted by expression of a dominant negative Sun1 construct (Crisp *et al.*, 2006). In *C. elegans*, deletion of the cc-domain of Sun1 homologue Unc-84 results in nuclear leaflet separation, although the effect was limited to cells under strain like muscle (Cain *et al.*, 2014). During development, neuronal patterning of the cerebral cortex and retina occurs via a process of nucleokinesis and interkinetic nuclear movement (IKNM). Force generation is provided by the microtubule cytoskeleton, to which the nucleus is connected through Sun1 or Sun2-containing LINC complexes. Knockout of both Sun1 and 2 results in aberrant cortical patterning, signifying their redundancy in this process (Yu *et al.*, 2011). Within muscle tissue, Sun1 and Sun2 similarly operate redundantly to anchor nuclei in their clustered arrangement beneath

neuromuscular junctions, though Sun1 deficiency alone does result in some myonuclear anchorage defects (Lei *et al.*, 2009; Meinke *et al.*, 2014).

1.8.3.2. *Sun1 and Sun2 functionally diverge*

However, despite the likenesses between Sun1 and Sun2, there are multiple examples of divergent functions.

1.8.3.2.1. *Binding partners and nuclear stiffness regulation*

The somatic Sun-domain proteins have differing molecular affinities for components of the NE microenvironment. *In vitro* assays have found that emerin preferentially binds to Sun1, and structurally, the binding motif in Sun1 is mostly absent in Sun2 (Haque *et al.*, 2010). Despite this, both Sun1 and Sun2 are independently able to transmit repeated mechanical stimuli through emerin, which, in turn, directs lamina stiffening (Guilluy *et al.*, 2014), indicating that residual emerin-Sun2 binding may be sufficient for force reception by the nucleus. LINC anchorage to the lamin network remains controversial; some studies indicate Sun1 and Sun2 bind lamin A equally (Haque *et al.*, 2010), whereas others suggest that the Sun1-lamin A interaction is of higher affinity (Östlund *et al.*, 2009). Consistent with the latter observation, the mobility of GFP-tagged Sun1 and Sun2 at the NE increases equally in NIH 3T3 fibroblasts lacking of lamin A (Östlund *et al.*, 2009), however, in HeLa cells, the mobility of *C. elegans* Sun1 homologue, UNC84A, appears unaffected by the absence of all lamins (Hasan *et al.*, 2006). Sun1 and Sun2 proteins respond differently to expressed lamins, too. When lamin A is absent, Sun1, but not Sun2, localises properly to the NE (Libotte *et al.*, 2005). However, expression of the sole *C. elegans* lamin, LMN-1, which has homology to human lamin B1, is prerequistite for the NE localisation of the Sun1 homologue, UNC-84 (Lee *et al.*, 2002). To date, the influence of B-type lamins on the localisation of Sun-domain proteins has not been directly assessed in mammalian cells. Recent *in vitro* studies using mammalian proteins suggest that Sun1 also interacts with B-type lamins (Nishioka *et al.*, 2016), despite earlier observations which indicated that the Sun-domain proteins have a strong preference for A-type over B-type lamins (Crisp *et al.*, 2006; Haque *et al.*, 2006). Neuronal cells, which predominantly express B-type lamins, depend upon an intact and functional nucleoskeletal-cytoskeletal axis in order to undergo

nucleokinesis and IKNM to ultimately give rise to the laminated patterning of cortical neuronal cells. Loss of any component which disrupts the mechanical connection to the nucleus, including lamins B1, B2 (Coffinier *et al.*, 2010, 2011), and Sun1 (Wang *et al.*, 2015) results in irregular neuronal layering, thus substantiating the biological relevance of a LINC complex-B-type lamin interaction. Interestingly, however, populations of neuronal cell types are differently affected by the loss of lamins or LINC complexes. In the cerebellum, lamin B2 loss affects both Purkinje and granule cells, whereas Sun1 is most important to maintain nuclear-cytoskeletal connections in Purkinje cells, and granule cells remain unaffected (Coffinier *et al.*, 2010; Wang *et al.*, 2015). Overall, it seems there is an emerging notion that Sun-domain proteins can adapt their lamin binding characteristics differentially in a cell-type specific manner, thus functioning to elicit specific nuclear movements or maintain nuclear position.

1.8.3.2.2. Differential association with NPCs and higher order oligomerisation

In addition to the existence of differential protein interactomes, Sun1 and Sun2 appear to segregate within the NE. Sun1 appears preferentially proximal to NPCs in a variety of cell types (Liu *et al.*, 2007; Lu *et al.*, 2008), complementing the idea of functional diversity between the homologues. Partial Sun2 co-localisation with Sun1 has been observed using antibody staining (Lu *et al.*, 2008), however, Sun1-GFP is appears completely separate from Sun2 (Liu *et al.*, 2007). These studies were both performed in HeLa cells, so a possible explanation for the difference in these observations is that C-terminal GFP on Sun1 impedes its association with other components of the NE, as Sun1 trimers have been reported to laterally oligomerise (Hennen *et al.*, 2018) through their C-terminal Sun-domains (Jahed *et al.*, 2018). Sun1 has been implicated in NPC assembly in interphase nuclei by co-operating with membrane-integral NPC protein POM121 (Talamas and Hetzer, 2011), which may explain why specific silencing of Sun1 results in the clustering of NPCs (Liu *et al.*, 2007). Although an interaction between Sun1 and POM121 could not be demonstrated (Talamas and Hetzer, 2011), this does not exclude the possibility of an interaction which is weak or transient in nature. It is equally possible that the insoluble natures of Sun1 and POM121, as integral proteins of the NE, renders

biochemical binding assays difficult due to the balance of efficient membrane extraction and preservation of endogenous interactions, similarly to the lamins (Burke and Stewart, 2013). BioID analysis of Nups to identify proximal interactors did not identify Sun1. However, a direct Sun1-NPC interaction cannot be discounted, since bait Nups for this study were located in the Y- and pore complexes, and not transmembrane Nups, like POM121 (Kim *et al.*, 2014).

Previous work has revealed that NPCs preferentially populate lamin C, rather than lamin A networks (Xie *et al.*, 2016), which conflicts with reports that Sun1 favours an interaction with lamin A (Haque *et al.*, 2006), based on the clear association of Sun1 with NPCs (Xie *et al.*, 2016), although more recent analyses indicate that Sun1 does in fact bind lamin C (Chojnowski and Xie, unpublished observations). To speculate, it may be that micro-domains exist within the NE, formed of specific lamins and LINC complexes to scaffold structural nuclear characteristics, such as NPC assembly sites. The viability of mice deficient in Sun1 questions the consequences of clustered or depleted NPCs at an organismal level. Interphase NPC insertion may not be essential for every cell type, or nuclear transport capacity generally greatly exceeds normal physiological demand. The contributions of Sun2, in the absence of Sun1, to compensate in these processes has not been explicitly tested.

1.8.3.2.3. Nuclear movement and positioning

A major process in the development and maintenance of mature tissues is cell migration and large organelles, such as the nucleus, must be appropriately handled to enable efficient migration. Cell migration is preceded by polarisation, where the centrosome becomes positioned between the leading edge of the cell and the nucleus. Contrary to early publications reporting centrosome repositioning and passive nuclear movement during polarisation and migration, the current notion indicates that the nucleus is actively positioned in an actin-dependent manner, while the centrosome remains relatively immobile (Gomes, Jani and Gundersen, 2005). Stimulation of a wounded monolayer with serum factor LPA induced the assembly of actin fibres that flowed rearwards, perpendicular to the wound edge. Emerin and myosin IIB are required for the rearward flow of dorsal actin cables (Chang *et al.*,

2013) which can be engaged by NE LINC complexes consisting of Sun2 and Nesprin-2G (Figure 1-2C), transmitting the tractive forces presumably to lamin A/C (Chang *et al.*, 2015). Importantly, rearward nuclear positioning is specific to Sun2 since polarisation is abolished upon Sun2 depletion, and Sun1 expression is not sufficient for restoration (Luxton *et al.*, 2010). Nuclear repositioning is also demonstrably dependent on the actin binding CH-domain of Nesprin-2G because knockdown of nesprin-2G similarly abolishes polarity in mouse C2C12 myoblasts which also express Nesprin-1 (Chang *et al.*, 2015). Depending on the resting state prior to polarisation, cells may also need to move the nucleus forwards. Centrifugation of a wounded monolayer, thus displacing nuclei to the back or front of cells on either side of the wound, provides an elegant system in which to model these possibilities. These experiments revealed that while rear movement of nuclei is Sun2 dependent, forward translocation is facilitated by Sun1-Nesprin-2G LINC complexes with force instead generated by the microtubule network and dynein motors which bind to juxtamembranal spectrin repeats of Nesprin-2G (Figure 1-2C). Indeed, the CH-domain is entirely dispensable for forward nuclear movement (Zhu, Antoku and Gundersen, 2017).

1.8.3.2.4. Meiosis and LINC complexes

Differences in Sun1 and Sun2 have also been observed in meiotic cells. Meiotic prophase I in immature spermatocytes sees meiotic bouquet formation; a pairing of homologous chromosomes, which become gathered at the NE by telomeric attachments to the NE facilitating resolution of DNA damage induced by homologous cross-over, and for the subsequent faithful segregation of sister chromatids. Interestingly, studies disagree whether Sun2 is expressed in spermatocytes, and in mammals, both Sun1 and Sun2 have been observed to link telomeres to the NE (Ding *et al.*, 2007; Schmitt *et al.*, 2007). Indeed, analysis of meiosis in the spermatocytes of mice in the absence of Sun1 reveals that Sun2 alone can also mediate the telomere-NE attachments (Link *et al.*, 2014). Nevertheless, Sun2 expression alone is not sufficient, and Sun1, but not Sun2, deficient mice are sterile (Ding *et al.*, 2007; Chi *et al.*, 2009; Horn, Kim, *et al.*, 2013). More recently, telomere movement at the NE has been identified as dependent on dynein binding to Sun1/KASH5 LINC

complexes in a process which seems to be evolutionarily conserved (Horn, Kim, *et al.*, 2013). In the yeast, *S. pombe*, Sun- and KASH-domain proteins, Sad1p and Kms1 mediates NE linkage to telomeres via meiosis specific adaptors (Chikashige *et al.*, 2006), while in the nematode, *C. elegans*, SUN-1/Matefin-1 binds KASH-protein ZYG-12 (Minn *et al.*, 2009). Defects caused by mutations in dynein, impeding telomere motility, can be partially recovered by the loss of SUN-1 in *C. elegans* (Sato *et al.*, 2009). LINC complexes are also essential for meiotic progression in the plant *A. thaliana*, where the loss of Sun1 and Sun2 results in defective chromosome synapsis (Varas *et al.*, 2015). The loss of KASH5 demonstrated that NE-telomere attachment is partially maintained, potentially through Sun2, indicative of further, non-structural roles for Sun1/KASH5 telomere complexes (Horn, Kim, *et al.*, 2013). Chi *et al.*, (2009) found that, at least in spermatocytes, Sun1 is required for germ cell-specific gene and Piwi-interacting RNA (piRNA) expression. piRNA may function to downregulate retrotransposon DNA elements which could otherwise compromise the integrity of the genome, leading to cell death (Chi *et al.*, 2009).

1.8.3.2.5. LINC complex involvement in DNA damage repair

The LINC-complex dependent pairing of homologous chromosomes in meiotic cells is essential to resolve DNA breaks resulting from homologous recombination. In somatic cells, DNA can be repaired either by homology-directed repair (HDR), or by non-homologous end joining (NHEJ), the latter of which is more error prone. Interestingly, the simultaneous knockout of Sun1 and Sun2 results in deficiencies in the DNA damage response in mouse embryonic fibroblasts (Lei *et al.*, 2012). Notably, sites of DNA damage are often observed to be highly mobile, and at least in yeast, HDR may be mediated by recruitment to the nuclear periphery by a LINC complex consisting of SUN protein Sad1 and KASH protein Kms1 (Swartz, Rodriguez and King, 2014). Interestingly, the loss of *C. elegans* Sun-proteins UNC-84 increases the likelihood of NHEJ, by de-repression of an NHEJ-promoting complex, which, in mammals includes DNA-dependent protein kinase (DNAPK) (Lawrence *et al.*, 2016). Consistently, an interaction between Sun1, Sun2 and DNAPK had been previously observed (Lei *et al.*, 2012), but evidence in HeLa cells indicates that Sun1 KD, or microtubule depolymerisation is sufficient to increase NHEJ, while reducing HDR-

promoting RAD51 complex (Lawrence *et al.*, 2016). Regions of DNA-repeats, such as telomeres are particularly susceptible to double stranded DNA breaks, and similarly have been observed to become repaired via a microtubule and LINC complex dependent mechanism (Lottersberger *et al.*, 2015), though in this study, the contributions of Sun1 were not compared with Sun2. Ribosomal DNA (rDNA) is also relatively unstable and localised within the nucleoli. Indeed, studies have identified an association between Sun1 and the nucleolus, where it is proposed to regulate morphology and rRNA transcription (Matsumoto *et al.*, 2016; Moujaber *et al.*, 2017). Again, DSBs in rDNA appear to be mobilised by LINC complexes, interestingly specifically Sun1 and the actin network in this case (Marnef *et al.*, 2019). While the LINC complex has been robustly implicated in the DNA damage response (DDR), different approaches including the modelled DNA regions and genotoxic agents used means the precise contributions of different LINC complexes remains to be elucidated.

1.8.3.3. *Alternative splicing further refines Sun1 function*

It is clear that various combinations of Sun- and KASH-domain proteins provide a NE platform for multifaceted functions, depending on the subcellular context of the LINC complex. Further diversity can be built into the LINC complex through extensive alternative splicing of the giant nesprins and indeed Sun1. The first indication of a function for specific Sun1 splicing isoforms was the identification of the Sun1η isoform which, in mouse, lacks exons 7-10. Sun1η is specific to the spermatid and spermatocyte where it segregates from Sun3 and Sun4, and together with Nesprin-3, localises to the head of the sperm nucleus, anchoring to local cytoplasmic components (Göb *et al.*, 2010). Analysis of separate tissues by reverse transcription PCR has revealed further diversity in the Sun1 splice isoforms expressed, many in a tissue specific manner (Göb *et al.*, 2011). Splice variation is predicted to alter the juxtamembrane region on the nucleoplasmic domain of Sun1, retaining the extreme amino-terminus, which maintains its ability to bind the NL, and the carboxy-terminal Sun-domain (Nishioka *et al.*, 2016). Notably, both Sun1 and Sun2 are necessary for migration, though depletion of Sun1 more effectively reduces migration. Interestingly, Sun1 isoforms have opposing roles in migration. In breast

cancer cell line MDA-MB-231, knockdown of the full length isoforms and overexpression of the shorter Sun1_785 and Sun1_888 isoforms (exons Δ6-8 and Δ7, respectively) both promote increased migration indicating that full length Sun1 inhibits migration (Nishioka *et al.*, 2016). Interestingly, studies indicate that Sun1 containing LINC complexes suppress migration promoting signals through RhoA inactivation (Thakar *et al.*, 2017). More research is required to identify whether there are specific Sun1 isoforms responsible for RhoA inhibition.

More functions are emerging for Sun1 splice isoforms, not least in muscle, where it has been revealed that Sun1 becomes alternatively spliced upon differentiation from myoblasts to myotubes, altering the expression profile from predominantly full-length Sun1 to various levels of shorter isoforms (Loo *et al.*, 2019). Protein-protein interacting sites are lost and created through alternative splicing, thus changing the nucleoplasmic interactome, and the first glimpses into this property in muscle appear aligned with an identified role of Sun1 in meiocytes where it can regulate the expression of regulatory, short non-coding RNAs (Chi *et al.*, 2009). During muscle regeneration, Rtl1 expression is upregulated, and this is enabled because the Sun1Δ7 splice isoform becomes prevalent in differentiating myoblasts and specifically binds and suppresses the microprocessor complex component Drosha, which is required for the processing of miR-127, a regulatory microRNA from the anti-sense strand of Rtl1. In the absence of Sun1, Drosha is not inhibited, thus increasing miR-127, inhibiting the translation of the Rtl1 protein, resulting in fewer regenerated muscle fibres, which have a smaller diameter, within a region of increased fibrosis (Loo *et al.*, 2019).

1.9. Diseases of the nuclear envelope

Diseases of the NE are collectively termed the nuclear envelopathies, or, where mutations are present in the lamins, the laminopathies (Worman, 2012). Laminopathies linked to the *Lmna* gene give rise to at least 12 diseases, including striated muscular dystrophies such as Emery-Dreifuss muscular dystrophy (EDMD), limb girdle muscular dystrophy 1B, dilated cardiomyopathy (DCM); lipodystrophies and systemic premature aging disorders such as Hutchinson-Gilford Progeria and

Werner syndromes (Worman and Bonne, 2007; Worman, 2012; Burke and Stewart, 2013). A central mystery to the laminopathies is how different disease phenotypes arise, which often affect specific tissues, from mutations in a single gene. In contrast to *Lmna*, there are many fewer diseases caused by mutations in B-type lamins. Disease phenotypes associated with B-type lamins are attributed to their essential role during development, particularly during nucleokinesis required for central nervous system patterning (Coffinier *et al.*, 2010, 2011). Additionally, *Lmna* is unnecessary for early development as it is not expressed until after embryonic implantation into the uterine wall (Wong and Stewart, 2020). In addition to those diseases caused by *Lmna* mutations, at least nine other NE proteins, both membrane and non-membrane bound, have been implicated in laminopathies and envelopathies, including many LINC complex members (Janin *et al.*, 2017).

1.9.1. *Laminopathies of striated muscle*

It is notable that many laminopathies affect striated muscle tissue. Emery-Dreifuss muscular dystrophy (EDMD) was the first identified laminopathy disease as a result of genetic linkage to the X-chromosome (Bione *et al.*, 1994). Mutations in the *EMD* (or *STA*) gene results in disease characterised by striated muscle wasting preceded by contractures of the neck, Achilles tendon and elbow (Worman, 2012). Since the discovery of the *EMD* mutation, an autosomal dominant form of EDMD (AD-EDMD), which is clinically similar to the X-linked form, has been linked to a mutation in the *Lmna* gene (Bonne *et al.*, 1999). Patients with AD-EDMD are often more severely affected than those with the X-linked disorder. Crucially, AD-EDMD patients often display cardiac dysfunctionality including dilated cardiomyopathy (DCM) and heart conduction block, and an affected individual who receives no medical intervention usually succumbs to the disease at 30-40 years of age due to conduction abnormalities (Bione *et al.*, 1994). Another muscle wasting disorder caused by *Lmna* mutations is limb girdle muscular dystrophy type 1B (LGMD-1B), although this has been recently reclassified as EDMD displaying wasting patterns in muscles of the hips and shoulders and without early contractures (Muchir *et al.*, 2000; Straub, Murphy and Udd, 2018). Interestingly, *Lmna* mutations also cause DCM without involvement of skeletal muscle (Fatkin *et al.*, 1999). Up to 10% of the

hereditary cases DCM are caused by *Lmna* mutations, and the disease progresses initially with detectable electrical defects, culminating in structural flaws (Peretto *et al.*, 2018). In DCM, heart ventricles become dilated, while the muscle tissues in the heart become replaced with fibrotic tissue which manifests in heart failure due to reduced ejection fraction or sudden cardiac death from arrhythmias (Cattin, Muchir and Bonne, 2013). Cardiac involvement signifies an overlap between the diseases of the striated muscle, and has led to the understanding that *Lmna*-rooted EDMD, LGMD-1B and DCM are a spectrum rather than distinct disorders (Worman and Bonne, 2007). Indeed, a single mutation in *Lmna* has been reported to produce DCM with a variable amount of skeletal muscle involvement (Brodsky *et al.*, 2001)

1.9.2. *Laminopathies of adipose tissue; the lipodystrophies*

Mutations in *Lmna* also cause diseases of the adipose and skeletal tissues, which again, can overlap to some extent. *Lmna* mutations causing Dunnigan's type familial partial lipodystrophy type 2 (FPLD2) were first uncovered with the single amino acid substitution, R482W (Shackleton *et al.*, 2000). Patients with FPLD2 are overtly normal at birth and throughout childhood, however, during adolescence, subcutaneous fat is lost from the trunk and limbs, while it often accumulates on the face and neck (Jackson *et al.*, 1997). In addition, patients display increased resistance to insulin (Simha and Garg, 2002). FPLD2 mutations, including R482W cluster to the immunoglobulin-like fold of lamins A and C, in contrast to those causing striated muscle diseases which are scattered throughout the protein (Worman and Bonne, 2007). Interestingly, the Ig-like fold may also harbour mutations which cause mandibuloacral dysplasia (MAD), initially R527H (Novelli *et al.*, 2002). MAD shares some features of FPLD2 including fat redistribution and insulin resistance, but patients also display skeletal abnormalities including hypoplasia of the clavicles and mandible (Novelli *et al.*, 2002).

1.9.3. *Laminopathies of the nervous system*

Peripheral nervous tissue may also be affected by *Lmna* mutations. Type 2 Charcot-Marie-Tooth (CMT2) disease can be caused by R298C *Lmna* mutation, which is inherited in an autosomal recessive manner (De Sandre-Giovannoli *et al.*, 2002). In CMT2, the axonal component of the peripheral nerve degenerates, and patients

have a decreased number of myelinated neuronal fibres, which leads to presentation of peripheral sensory and motor neuropathy with onset between 6-27 years of age (Tazir *et al.*, 2004, 2013).

1.9.4. *Systemic laminopathies: the premature aging syndromes*

Other diseases caused by mutations in *Lmna* have a much greater tissue, and system penetrance. These mutations affect multiple organs and result in segmental premature aging disorders. Perhaps the most striking of these is Hutchinson Gilford progeria syndrome (HGPS), which most often arises from a silent mutation in exon 11 at c.1824C->T, creating a cryptic splice site resulting in Progerin expression, a protein product similar to lamin A but lacking the C-terminal 50 amino acids (Eriksson *et al.*, 2003; Sandre-Giovannoli *et al.*, 2003). This removes the cleavage site for ZMPSTE24 and the resulting protein remains permanently farnesylated (Goldman *et al.*, 2004). Infants with HGPS appear normal at birth but signs of accelerated aging surface over the first year of life. HGPS patients fail to grow at normal rate (failure to thrive), have craniofacial abnormalities, delicate skin, osteoporosis, and severe atherosclerosis, which is often the cause of death by early teenage years (Eriksson *et al.*, 2003). Atypical Werner's syndrome (WS) is an adult progeroid disease caused by *Lmna* mutations, and similarly to HGPS, WS patients experience defects in multiple systems which contribute to accelerated aging (Bonne and Levy, 2003; L. Chen *et al.*, 2003).

1.9.5. *Nuclear envelope proteins are increasingly linked to the laminopathies*

As already implied given the existence of the X-linked EDMD caused by mutations in the gene encoding INM protein emerin, laminopathies are not limited those caused by mutations in the *Lmna* gene. Indeed, other proteins of the NE have also been implicated in the aetiology of nuclear envelopathies and anomalies. Notably, mutations in LINC complex components result in a number of striated muscle diseases. Mutations in Nesprins-1 and -2 have been implicated in EDMD, of which many cluster to the C-terminus (Janin and Gache, 2018), some of which are associated with abnormal nuclear morphology and signs of genetic instability

including micronuclei and fragmented nuclei (Zhang *et al.*, 2007). Nesprin-1 has also been implicated in DCM (Zhou *et al.*, 2017) and distal arthrogryposis (Attali *et al.*, 2009; Baumann *et al.*, 2017). In addition, skeletal and cardiac muscle disorders, including an arrhythmia with partial lipodystrophic symptoms may also be caused by mutations in Sun1 and Sun2 (Meinke *et al.*, 2014). Moreover, further mutations in Sun1 and Sun2 have been reported to be disease-modifying in patients with EDMD/CMT and even DMD, offering some explanation as to how variable disease phenotypes arise amongst cases (Taranum *et al.*, 2012; Meinke *et al.*, 2014). A common theme among laminopathies rooted in NE proteins is the reciprocal instability of other NE and NL components, the most commonly analysed being lamins A/C, emerin, and LINC complexes (Zhang *et al.*, 2007; Taranum *et al.*, 2012; Meinke *et al.*, 2014; Zhou *et al.*, 2017). Interestingly, the nucleus-centrosome distance is often increased as a result of LINC complex abnormalities (Taranum *et al.*, 2012; Meinke *et al.*, 2014), and one possibility for the muscle pathomechanism is a failure of efficient myoblast differentiation, which is dependent on centrosome relocalisation to the NE, directed by the short Nesprin-1α isoform (Gimpel *et al.*, 2017).

Mutations in other NE proteins have also been found to cause striated muscle diseases, including LEM domain protein LAP2α which has been implicated in DCM (Taylor *et al.*, 2005). It remains to be elucidated whether LAP2α is functionally connected with the LINC complex in causing striated muscle diseases. It does seem similar to other NE proteins however, in that LAP2α levels adjust in response to depletion of lamin A/C, for example (Osmanagic-Myers, Dechat and Foisner, 2015). Mutations in LAP2α binding partner BAF, causes a progeric disease similar to HGPS, known as Nestor-Guillermo progeria syndrome (Cabanillas *et al.*, 2011; Puente *et al.*, 2011).

1.10. <u>Disease mechanisms and animal models of the laminopathies</u>

Naturally occurring mutations in proteins of the NL and NE which give rise to the nuclear envelopathies and laminopathies have offered insight into how the structure of the nucleus is regulated and unveiled some important functions.

Moreover, the discovery of NE proteins, and their contribution to signalling, both mechanically and biochemically, has provided an expanded understanding of the disease mechanisms manifesting in laminopathy disorders. In a somewhat striking reflection of dystroglycan- and DGC-associated diseases, research into the components of the nuclear periphery, and the diseases which arise from it, indicate that resident proteins have both structural and signalling roles which are similarly intertwined. As such, there are now hypotheses by which the NL functions, and becomes disrupted during disease. The structural hypothesis postulates that cells in the degenerative laminopathies undergo premature death as the nucleus, and underlying genome, becomes damaged as a result of improper nuclear-cytoskeletal connections. On the other hand, there is evidence which suggests that laminopathies affect NL-chromatin connections, which have far reaching consequences for DNA damage repair, chromatin organisation and transcriptional regulation. A number of cellular and genetically modified animals modelling the laminopathies have lent credence to each of these hypotheses and shed light on the essential functions ascribed to the nuclear periphery, however, increasing evidence indicates that the two hypotheses are not as mutually exclusive as previously thought.

1.10.1. *The structural hypothesis: the lamins*

Perhaps the most striking of the laminopathies is the premature aging syndrome, HGPS, and has therefore garnered much research interest. It is also evident that laminopathies disproportionately affect striated muscle, which is indeed an affected tissue in HGPS. Cells derived from patients and mice harbouring mutations of HGPS display extremely deformed nuclei (Eriksson *et al.*, 2003; Goldman *et al.*, 2004). In addition, expression of *Lmna* with mutations associated with EDMD and DCM induces perturbations in nuclear morphology and the Lamin A gene product accrues aberrantly in the nucleoplasm (Raharjo *et al.*, 2001; Barateau *et al.*, 2017). Notably, lamin A R482W associated with FPLD does not display nuclear abnormalities (Raharjo *et al.*, 2001). Similarly, lamin A/C knockout in mice recapitulate the human EDMD condition, and ultrastructural analysis of these nuclei show gross abnormalities (Sullivan *et al.*, 1999). Nuclei in cells lacking lamins A and C were found to deform more easily compared with WT cells (Lammerding *et al.*,

2004). Moreover, deficiency of B-type lamins induces nuclear blebbing, but has no effect on deformability, leading to the understanding that regulating nuclear stiffness is a specific property of A-type lamins (Lammerding *et al.*, 2006). Furthering the concept that nuclear integrity is compromised upon *Lmna* deficiency or mutation, it was found that the nuclei of fibroblasts from a variety of origins, including *Lmna* knockout and progeric patients, exhibited abrupt rupturing events, which is not catastrophic for cells (De vos *et al.*, 2011). Moreover, nuclei in myoblasts derived from *Lmna* deficient mouse muscle are similarly deformable and undergo nuclear rupture (Earle *et al.*, 2020). Nuclear rupture is most likely to occur at the leading edge of the nucleus of a migrating cell, was initially shown to occur in immune cells and cancer cells, *in vitro* and *in vivo*. DNA damage foci can be observed at the site of rupture, which becomes resolved upon NE repair by the ESCRT III complex (Denais *et al.*, 2016; Raab *et al.*, 2016). In addition to migration through small passages, nuclear rupture in myofibres can result from nuclear strain which similarly induces genomic damage (Earle *et al.*, 2020). Blebs precede rupturing events, and these form from nuclear membranes ahead of the lamina which are depleted in many NE components (Sullivan *et al.*, 1999), thus reducing the tension and bending angle the NE can withstand. Presumably in laminopathies where lamins or NE proteins are mutated, optimised nuclear membrane properties are perturbed, leading to an increased likelihood of bleb formation and rupture. Indeed, studies have noted that striated muscle laminopathy-associated *Lmna* mutations result in the abnormal arrangement of lamins (Raharjo *et al.*, 2001) as well as aberrant targeting of NE proteins including emerin, LAP2 and lamin B1 (Sullivan *et al.*, 1999; Raharjo *et al.*, 2001; Holt *et al.*, 2003). In *Lmna*-depleted myotubes, cells towards the ends of the myofibres, which ostensibly experience less contractile force, exhibit fewer nuclear ruptures, while interfering with LINC complex assembly also prevents DNA damage associated with ruptures (Earle *et al.*, 2020).

1.10.2. *The gene regulation hypothesis: the lamins*

An alternative to the mechanical hypothesis is the gene expression hypothesis, which postulates that that changes to the NL can deleteriously affect a cells gene expression programme through disrupted signalling and chromatin organisation.

The Ig-like fold of lamin A is an important region for protein-protein interactions, and disruptions to this region, for instance R482W found in FPLD2 patients, perturbs binding to SREBP1, a transcription factor which may bind to lamin A upon engagement of genomic sterol response targets to induce adipocyte differentiation (Lloyd, Trembath and Shackleton, 2002). Pro-adipogenic SREBP1 target genes are therefore misregulated and cells senescence, which manifests disease symptoms (Vadrot *et al.*, 2014).

Perturbed signalling pathways have also been associated with cardiac striated muscle condition EDMD. The lamin A H222P mutation in lamin A activates MAPK signalling, with increased ERK1/2 nuclear translocation in cardiomyocytes (Muchir *et al.*, 2007). Muscle related genes, including sarcomeric proteins and myosins, become dysregulated, which is temporarily reversed by administering an ERK inhibitor, while preventing the pathological ventricular dilatation (Muchir *et al.*, 2009). Consistently, the NL appears to be an integrator of signalling molecules within the ERK pathway to direct proliferation pathways. In quiescent cells, c-Fos, an AP-1 transcription factor member, is sequestered to the NE by lamin A. pERK translocates to the nucleus, phosphorylating and activating c-Fos which releases it from the NE. c-Fos then heterodimerises with other AP-1 members and binds target genes to direct transcription. Notably, the depletion of lamin A eradicates AP-1 regulation (Ivorra *et al.*, 2006; Gonzàlez *et al.*, 2008).

In addition to signalling-directed gene transcription, alterations to the lamina also affect genomic organisation. Chromatin is precisely arranged within the interphase nucleus and genomic compaction, regulated by histone modifications and NL contacts, serves as a mechanism to differentiate between transcriptional activation and inactivation depending on the circumstantial requirements. These may be developmental, or reactive to biochemical or mechanical signals. In mammalian cells, transcriptionally repressed chromatin is densely associated with the nuclear periphery, in regions known as lamina-associated domains (LADs) (Guelen *et al.*, 2008; Reddy *et al.*, 2008). It seems that there are some standard NL elements required for LAD maintenance, since LADs can be disrupted in *Drosophila melanogaster* by the removal of the only B-type lamin (Kohwi *et al.*, 2013), or in mice by the simultaneous ablation of lamin A and LBR (Solovei *et al.*, 2013). Disruption to

genome organisation is commonly observed as a result of various laminopathies. In HGPS skin fibroblasts, LADs are disrupted as connections between the genome and the NE are weakened, while the genome assumes a differential H3K27me3 pattern leading to alternative transcription (McCord *et al.*, 2013). In addition to signalling disturbances observable in *Lmna* H222P EDMD, peripheral heterochromatin is lost and Sox2, a pluripotency marker, is expressed at higher levels probably due to liberation from a repressed state, thus inhibiting myoblast differentiation (Perovanovic *et al.*, 2016). In contrast, an increased number of LADs have been identified in DCM, which nonetheless causes changes in gene expression (Cheedipudi *et al.*, 2019), although in this study, whether *Lmna* is mutated is not clear. Finally, expression of R388P-mutated Lamin A mutated induces increased histone acetylation, though the implication of this specific mutation is not known (Barateau *et al.*, 2017).

In addition to the lamins, the balance between heterochromatin and euchromatin, and indeed the genes associated with these regions, may be regulated by the array of tissue-specific NE proteins, which control the transcriptional output. The MyoD gene, a master regulator for muscle differentiation, is associated with dense heterochromatin proximal to the NL in myoblasts and becomes repositioned to the nuclear interior during myotube formation, where it can bind transcription factor TAF3 (Yao *et al.*, 2011). Additionally, a number of NETs have been identified which are capable of controlling gene repositioning events, and their differential expression profiles could account for tissue-specific gene regulation (Zuleger *et al.*, 2013).

1.10.3. *The LINC complex can transmit force and regulate genes*

Physical connectivity of chromatin to the cytoskeleton and beyond, though the NL and LINC complex also has considerable influence on force transduction, genome organisation and transcriptional outputs, signifying that the gene expression and structural hypotheses are not mutually exclusive. Mechanical stimuli transmitted to the nucleus through the LINC complex are received by emerin, which can modulate lamina stiffness, but also relocalise to the ONM to nucleate actin polymerisation (Guilluy *et al.*, 2014; Le *et al.*, 2016). G-actin normally binds to signalling molecule

megakaryoblastic leukaemia 1 (MKL1), which is liberated upon actin filament polymerisation. MKL1 is then able to translocate to the nucleus where it can bind serum response factor (SRF) to direct gene expression programmes involved in cell migration (Ho *et al.*, 2013). Given the mechanosensitive function of emerin, MKL1/SRF signalling can also be modulated by the stiffness of substrate that cells are growing on (Willer and Carroll, 2017), thus rendering the LINC complex competent to direct both structural and signalling changes within the nucleus.

Gross alterations in chromatin organisation influenced by external mechanical stimuli can also alter gene transcription. One study revealed the transcription of the exogenous, GFP-tagged bacterial gene DHFT, can be induced by direct force application through integrins (Tajik *et al.*, 2016). Rapid transcriptional changes, combined with the observations that actin cytoskeleton depolymerisation and Sun1/2 knockdown ablates the transcriptional upregulation is consistent with a physical LINC complex-dependent basis for transcriptional initiation (Tajik *et al.*, 2016). Interestingly, this study revealed that the genome also has remarkable spatial sensitivity, with chromatin stretching sensitised to direction of force application (Tajik *et al.*, 2016). However, these mechanical properties have not yet been demonstrated in an endogenous context, so it is unknown whether this pathway would constitute a significant factor in nuclear function, or indeed as a pathological mechanism in the laminopathies.

Mechanical force application to the genome underlying the NE can also modify the epigenetic state of chromatin. LADs detach from the nuclear periphery, as emerin moves to the ONM, where together with non-muscle myosin IIA (NMIIA) it nucleates actin filaments perpendicular to the direction of force application, perhaps to brace the NE (Le *et al.*, 2016). Transcriptionally repressed LADs, which are enriched in di- or tri-methylation on lysine 9 of histone 3 (H3K9me2/3) switch to H3K27me3 histone marks once detached from the NL to retain gene repression (Le *et al.*, 2016). By silencing NMIIA, cells undergo premature differentiation and abnormal epidermis morphology (Le *et al.*, 2016). Although this study does not directly implicate the LINC complex, a reoccurring theme is the implication of emerin in mechanotransduction, which itself is LINC complex dependent (Guilluy *et al.*, 2014). Furthermore, the NL functions to regulate epigenetic markings, via the

polycomb repressor complex (Cesarini *et al.*, 2015), which may be impaired in some laminopathies (Pegoraro *et al.*, 2009; Briand *et al.*, 2018).

As mentioned, HGPS arises from progerin expression. Progerin is a permanently farnesylated version of lamin A, and it has been proposed that progerin contributes to nuclei having increased stiffness and mechanical sensitivity, rendering them more fragile upon encountering mechanical forces. While the use of farnesyltransferase inhibitors (FTI) improves survival in a progerin expressing mouse model (Yang *et al.*, 2006), and returns nuclear stiffness and shape to a more regular pattern, the increased mechanosensitivity and viability of cells is not resolved (Verstraeten *et al.*, 2008), indicating that there are other ramifications of progerin expression. LINC complex component, Sun1, is known to preferentially bind prelamin A or progerin, compared with Sun2 (Crisp *et al.*, 2006; Chen *et al.*, 2014), and given the emergent diverging functions of LINC complexes, an imbalance of Sun-domain proteins is likely to have deleterious consequences. Indeed, mouse models genetically modified for *Lmna* ablation, which models AD-EDMD, or *Lmna*Δ9, which recapitulates HGPS (Mounkes *et al.*, 2003), both succumb to disease symptoms in early life, display an accumulation of Sun1. Simultaneous *Lmna* and Sun1 genetic ablation results in a remarkable suppression of the disease phenotypes and an increase in life span in a gene-dose dependent manner, together implicating the potential for the LINC complex to function as a mutual pathological mechanism in at least some laminopathies (Chen *et al.*, 2012). Many tissues involved in human laminopathies and modelled in the lamin A deficient mice were reported to have more normalised morphology, including bone structure, skeletal muscle fibres with fewer centralised nuclei, and cardiac tissue displaying reduced necrosis along with normalised ejection fraction (Chen *et al.*, 2012). In addition, heterochromatin abnormalities and premature senescence identified in HGPS skin fibroblasts are resolved upon the knockdown of Sun1 (Chen *et al.*, 2012), altogether indicating that the LINC complex might well influence epigenetic regulation discussed above.

Progerin expression and Sun1 accumulation disrupts the assembly of TAN-line Sun2 anchorage, and therefore nuclear migration is disrupted, while the microtubules are increased. The imbalance of nucleocytoskeletal connections may

contribute to cellular dysfunction through disrupted homeostatic nuclear positioning (Chang *et al.*, 2019).

1.11. Summary of LINC-mediated laminopathy pathological mechanisms

In summary, the LINC complex enables mechanotransduction as a mechanism by which the nucleus can be repositioned through application of force to the NL, but also performs its addition functions in chromatin arrangement and epigenetic and transcriptional regulation. Moreover, these functions, together with studies of the laminopathies, have shown that the LINC complex serves to contribute to both structural and gene expression hypotheses of pathology. While it seems that laminopathic nuclei are fragile and rupture more easily, it is unclear whether LINC complexes are entirely complicit in these force-induced pathological mechanisms, since they also contribute to nuclear positioning, DNA damage repair, genome organisation, epigenetic and transcriptional regulation. In particular, it seems unlikely that mechanical damage is solely responsible for the pathology of the laminopathies, since even in the absence of Sun1, Sun2 still connects the nucleus to the cytoskeleton, and evidence shows that, at least to some extent, acts redundantly with Sun1. However, it is unclear how tissue pathologies, including muscle, are alleviated through Sun1 depletion. Given the wealth of nuclear functions in which Sun1 and the LINC complex are integrated, combined with the tissue-specific genome organisation and distribution of NET expression, it is probable that pathomechanisms are nuanced to each cell-type, and cannot be simply explained by the uncoupling of the nucleus to isolate it from the forces of the cytoskeleton. Hence, the discovery of novel nucleoplasmic Sun1 interacting partners would provide further insight into the pathology of the laminopathies.

1.12. The function of dystroglycan and its fragments in the nucleus

There have been no diseases described to date which result from a deficiency in the nuclear function of β-dystroglycan. Similarly, insight into dystroglycan function at an organismal level using murine models have not been powered to divorce the

76

plasma membrane from the nuclear roles. However, more recent *in vitro* investigations indicate that β-dystroglycan localises to numerous sub-compartments of the nucleus, contributing to aspects of nuclear function and NL stability.

Fitting with its identified pan-nuclear localisation, β-dystroglycan has been assigned functions within many nuclear domains. In prostate cancer cells, reports indicate that numerous dystroglycan fragments are translocated to the nucleus in an androgen-dependent manner (Mathew *et al.*, 2013). Expression of the cytoplasmic fragment of β-dystroglycan in these cells induces an alteration to gene expression, which indicates a role in regulating transcription (Mathew *et al.*, 2013), although more research is required to determine how this occurs, since it is not clear whether other β-dystroglycan fragments in the nucleus have a similar function. In addition, the cytoplasmic fragment of β-dystroglycan has been identified within the nucleolus of C2C12 myoblasts, where it interacts with ribosomal DNA (rDNA) regions and resident proteins ribosomal RNA (rRNA) transcription factors UBF and B23 (Azuara-Medina *et al.*, 2019). From overexpression studies, it seems that the cytoplasmic domain of β-dystroglycan mediates the nucleolar stress response and reduces ribosomal component transcription. However, this is not consistent with the observation that nucleolar components are decreased during a dystroglycan knock down, given that the cytoplasmic fragment of β-dystroglycan binds nucleolar proteins (Azuara-Medina *et al.*, 2019).

β-dystroglycan has also been assigned functions within the structural proteins of the nucleus, and immunoprecipitation experiments have identified interactions with NL components lamins A/C and B1 and NE protein emerin (Fuentes-Mera *et al.*, 2006; Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). In addition, perturbation of nuclear levels of β-dystroglycan either by knockdown or accumulation through nuclear export sequence disruption (β-DG mNES) appears to disturb nuclear morphology, perhaps through the simultaneous disturbances in the localisations of NL components lamin B1 and emerin (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). Reinforcing this observation, primary fibroblasts from a dystroglycan-null patient display severe nuclear morphological abnormalities in nuclear morphology (Jacobs, 2017 (PhD thesis)). However, lamin A/C, a well-

established determinant of nuclear shape and integrity is not affected by β-dystroglycan manipulation (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). Interestingly, β-DG mNES, like β-DG, retains its ability to bind NL components, while appearing not to be mislocalised itself (Vélez-Aguilera *et al.*, 2018), which may indicate other binding partners within the NL, thus displacing emerin and lamin B1. Further work is required to identify the molecular mechanisms behind the integration of β-dystroglycan with the NL.

In addition to identified structural roles in the NL, β-dystroglycan has been implicated in nuclear-centrosome tethering. Similarly to the NL, dystroglycan knockdown and nuclear accumulation using the β-DG mNES construct increases the distance between the nucleus and centrosome in myoblasts (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). This is hypothesised to result from an identified interaction between β-dystroglycan and emerin, within their respective transmembrane domains (Gómez-Monsivais *et al.*, 2020), disrupting emerin-dependent centrosome tethering (Salpingidou *et al.*, 2007). These interactions presumably occur on the outer nuclear membrane of the nuclear envelope, though definitive evidence of β-dystroglycan at this position has not yet been reported. To date, transmission electron microscopy only reveals immunogold staining at the INM (Martínez-Vieyra *et al.*, 2013), raising unresolved topological questions how a centrosome tethering-relevant β-dystroglycan-emerin interaction may be possible.

1.13. <u>Summary: the abundant functions of nuclear β-dystroglycan</u>

Dystroglycan possesses critical functions within the DGC to maintain the structural integrity of the sarcolemma, but also crucially regulate signalling pathways which fulfil metabolic demand in healthy muscle, and balance aspects of muscle regeneration, including commitment to myogenic programmes and fibrosis. The identification of β-dystroglycan within the nucleus indicates further diverse functions away from the sarcolemma, which may similarly manifest in both structural and signalling roles. Indeed, β-dystroglycan has been implicated in maintenance of NL integrity, although the molecular basis and downstream consequences of this function are unclear. Interestingly, signalling roles for β-

dystroglycan in the nucleus have not yet been established, although the mechanisms contributing to transcriptional changes in response to nuclear β-dystroglycan are unknown, or indeed if this is a global phenomenon, or specific to androgen receptive cancer lines. The presence of β-dystroglycan within numerous nuclear compartments including the nuclear membrane, nucleoplasm and nucleolus may be explained by various mechanisms to escape its canonical membranous environment, including Sec61 extraction, or proteolytic fragmentation of the cytoplasmic domain. Whether the resulting fragments have propensities to localise to certain compartments or perform specific functions is not clear. To further unravel the functions of β-dystroglycan within the nucleus, it is pertinent to resolve localisation and functional ambiguities relating to various fragments. Improper β-dystroglycan nuclear localisation or function has not to date been implicated in degenerative diseases of the DGC. However, it is notable that β-dystroglycan accumulates in the nucleus in some cancers, and some dystroglycanopathies which cause milder degenerative muscular disease preventing autoproteolysis of the dystroglycan prepeptide, which logically would impair the liberation of β-dystroglycan from the sarcolemma and subsequent nuclear translocation of the full length protein. A comprehensive understanding of the nuclear function of β-dystroglycan is therefore of significant interest.

1.14. Final summary and project hypotheses

Deficiency of dystroglycan appears to affect various aspects of nuclear function as evidenced by gene disruption strategies, and the analysis of tissues and cells. However, the mechanism by which dystroglycan contributes to the maintenance of regular nuclear morphology and NL integrity remains unclear. Preliminary data from the Winder laboratory indicates that β-dystroglycan interacts with Sun1, the INM-spanning member of the LINC complex. In this capacity, Sun1 is partially required for nuclear attachment, transmitting mechanostimuli to the nucleus, and in certain laminopathic conditions, its presence is detrimental to nuclear function and morphology. It is therefore postulated that NE-localised β-dystroglycan exists to stabilise nuclear morphology and integrity through

interactions with the LINC complex, perhaps to regulate forces transmitted to the NL and other components of the nuclear interior.

Secondly, Sun1 has been specifically implicated in the aetiology of the laminopathies; ablation of Sun1 remarkably suppresses simultaneous mutations in, or the absence of the *Lmna* gene (Chen *et al.*, 2012). Despite the seemingly essential role for the LINC complex throughout many developmental processes, it is possible that lamin A disturbances affect the influence of Sun1-containing LINC complexes on as yet unknown effectors critical for disease progression. To further investigate this hypothesis, work presented in this thesis aimed to identify novel nucleoplasmic binding partners for Sun1. Skeletal muscle is a tissue commonly affected in laminopathy diseases, and so novel Sun1 interactors were sought in an appropriate model; cultured primary myoblasts.

1.14.1. *Overview of chapters*

Results presented in chapter 3 describe experiments undertaken to examine a potential association between Sun1 and β-dystroglycan. An interaction is probed using a variety of biochemical strategies, while the functional association is tested using overexpression studies. Functions of dystroglycan within the myonucleus were further probed in chapter 4, which describes the generation and characterisation of dystroglycan-null human myoblasts. Many functions ascribed to β-dystroglycan are ascribed to the full-length version, and so chapter 5 describes experiments aimed at developing new tools to detect β-dystroglycan fragments to permit further dissection of functions.

Chapter 6 concerns the ex vivo validation of an in vi*vo* model for two-component BioID (2C-BioID), a system to perform proximity dependent biotinylation to identify potential interacting partners among vicinal proteins. 2C-BioID performed in myoblasts from the murine transgenic model was used to uncover potential interactors of the nucleoplasmic domain of Sun1 on a wild-type background. The final results chapter centres on the validation of putative Sun1 interactors. Accordingly, experiments presented describe the establishment of Sun1 bimolecular fluorescence complementation for independent verification.

Chapter 2: Materials and Methods

2.1. <u>Materials</u>

Appendix A – Antibodies used

I. Primary antibodies

II. Secondary antibodies

Appendix B – Oligonucleotides

I. Genotyping primers

II. Cloning primers

III. Sequencing primers

IV. Guide RNA sequences for CRISPR

Appendix C – Plasmids used and generated

I. CRISPR plasmids

II. Mammalian expression plasmids

III. Epitope tagged Sun1 candidate interactors

IV. Bimolecular fluorescence complementation constructs

V. IVT plasmids

VI. Targeting plasmids

VII. Bacterial expression plasmids

Appendix D – Recipes for solutions

I. Common buffers

II. SDS-PAGE and western blotting

III. Immunofluorescence microscopy

IV. Bacterial culture

V. Solutions for mammalian cell culture

VI. Solutions for BioID

Appendix E – Antibiotic selection of bacterial and mammalian cells

I. Bacterial selection

II. Eukaryotic selection

2.2. Methods

2.2.1. *Bacterial techniques*

2.2.1.1. *Generation of chemically competent cells*

10ml of antibiotic-free LB was inoculated with a pipette scraping of -80°C frozen *E. coli* stock cells (DH5α, TOP10 or BL21(DE3)). The 10ml culture was incubated at 37°C overnight while shaking at 190 rpm. The next day, 1.5ml of the overnight culture was used to inoculate 100ml antibiotic-free LB which was grown while shaking to log phase. Growth phase was monitored by intermittent culture sampling, measuring the optical density using a spectrophotometer at wavelength 600nm (OD_{600}). Log phase was defined by OD_{600} = 0.4-0.6. All subsequent steps were carried out in a centrifuge pre-chilled to 4°C or on ice. The culture was decanted into tubes and centrifuged at 5000 x *g* for 10 minutes to pellet the cells. The medium was completely decanted by inverting the tubes for one minute. The cell pellet was resuspended in 10ml ice cold 100mM $CaCl_2$ and incubated on ice for 2 hours. Cells were centrifuged as before and the pellet resuspended in 2ml transformation buffer (Appendix D, IV). Cells were aliquoted into 100μl aliquots and immediately snap frozen in liquid nitrogen before storage at -80°C.

2.2.1.2. *Bacterial transformation*

For transformation of miniprepped plasmids, 1μl plasmid DNA (~100ng) was added to 30μl chemically competent cells (prepared as described in section 2.2.1.1). For transformation of a newly cloned plasmid, the 10μl ligation reaction was mixed with 100μl chemically competent cells. Subsequent steps were identical regardless of plasmid origin. Cells with plasmid were mixed gently and incubated on ice for 20 minutes. Cells were then heat shocked in a 42°C water bath for 30 seconds before immediately incubating on ice for a further 2 minutes. 300μl SOC medium was added and cells were recovered while shaking at 37°C for 40 minutes. Bacteria were then spread on prewarmed LB agar plates containing the appropriate selection antibiotic(s). The plates were either incubated overnight at 37°C or for three days at room temperature. For storage, plates with colonies were kept at 4°C for up to one week.

2.2.1.3. *Liquid phase E. coli growth for plasmid propagation*

E. coli were transformed with the desired plasmid and grown on LB agar plates containing the appropriate selection antibiotic(s) as described in section 2.2.1.2. Colonies were picked from the plate using a sterile tip and used to inoculate 4ml LB which was incubated overnight shaking at 190 rpm at either 37°C or 30°C. The next day 4ml cultures were miniprepped or further diluted for maxiprep plasmid preparation.

2.2.2. *Preparation of plasmid DNA*

2.2.2.1. *Miniprep*

4ml liquid cultures were prepared as described in section 2.2.1.3. The next day, bacteria were pelleted by centrifugation at 4200g for 7 minutes at RT. Plasmid DNA was then purified using either a Qiagen or Favorgen miniprep kit according to the manufacturer's instructions. Final DNA elution was carried out using 50μl warmed molecular biology grade water for downstream applications. Purified plasmid DNA was stored at -20°C.

2.2.2.2. *Maxiprep*

To prepare larger quantities of plasmid DNA, 4ml cultures were diluted into 100-150ml LB containing the appropriate selection antibiotic and incubated in a shaking incubator at either 37°C or 30°C overnight, depending on the construct. The next day, bacteria were pelleted by centrifugation at 5000g for 15 minutes at 4°C before proceeding with DNA purification. Plasmid DNA was purified using a Qiagen Hi-Speed Maxiprep kit according to manufacturer's instructions. DNA was resuspended in 1ml warmed molecular biology grade water for downstream applications, and stored at -20°C.

2.2.2.3. *Miraprep* (Pronobis, Deuitch and Peifer, 2016)

Used as alternative method to produce larger quantities of plasmid DNA. A single bacterial colony was used to inoculate a 50ml of LB containing the appropriate selection antibiotic, and incubated overnight at 37°C while shaking at 190 rpm. The next day, bacteria were pelleted by centrifugation at 3220 x *g* for 20 minutes at 4°C. The cell pellet was resuspended in 2ml of resuspension buffer from either Qiagen or

Favorgen miniprep kits supplemented with the fresh addition of 50μg/ml RNase A. Resuspended bacteria were then distributed into 5x 1.5ml Eppendorf tubes. 400μl of lysis buffer was added to each tube and mixed by inversion 6 times before incubating at RT for 4 minutes. The lysis reaction was stopped using 400μl neutralisation and mixed by inversion. Eppendorf tubes were then centrifuged at 18,000g for 10 minutes at RT. Supernatants from each tube were pooled into a 15ml Falcon tube, and 1x volume of 96% ethanol (~5ml) was added and mixed thoroughly for 5 seconds. Plasmid-ethanol mix was loaded onto 5 spin columns placed on the vacuum manifold in sequential 800μl aliquots until all the sample had been run through the spin columns. Spin columns were washed on the vacuum manifold without breaking the vacuum, first using 400μl buffer PB (Qiagen) or W1 (Favorgen) followed by 700μl buffer PE (Qiagen) or Wash (Favorgen). Spin columns were moved to collection tubes and centrifuged at 18,000g for 1.5 minutes at room temperature to dry the resin. The columns were then placed onto new 1.5ml Eppendorf tubes and plasmid DNA was eluted by adding 50μl warmed dH$_2$O, incubating for 2 minutes at RT and centrifuging at 18 000g for 2 minutes at room temperature. Eluates were pooled and plasmid DNA was stored at -20°C.

2.2.3. *Total RNA extraction*

RNA was purified from cultured cells using TRIzol reagent, and all steps were carried out working under a fume hood. 0.3ml of TRIzol reagent was added directly to a culture of 1 x 10^5-10^7 cells and triturated 5 times to homogenise before incubation for 5 minutes at room temperature. The solution was transferred to a 1.5ml tube and 0.2ml chloroform was added per 1ml TRIzol used for lysis. The tube was inverted 3 times to mix and then incubated for 2-3 minutes at room temperature. The sample was then centrifuged at 12,000 x *g* at 4°C, separating the mixture into lower phenol-chloroform, interphase and upper aqueous phases. The RNA-containing aqueous phase was aspirated into a new 1.5ml tube, carefully avoiding contamination from other layers. RNA was precipitated by adding half the TRIzol volume of isopropanol and incubated at room temperature for 10 minutes. Total RNA was then pelleted by centrifugation at 12 000 x *g* for 10 minutes at 4°C. The supernatant was discarded and the pellet resuspended in ice-cold 75% ethanol.

The sample was briefly vortexed and then re-centrifuged at 7500 x g for 5 minutes at 4°C. The supernatant was carefully removed, and the pellet left to air-dry for 10 minutes. Total RNA was resuspended in 40µl Diethyl-pyrocarbonate (DEPC) treated H_2O, and the concentration determined by nanodrop for downstream applications.

2.2.4. *Measurement of DNA/RNA concentration*

DNA and RNA concentration was determined using a nanodrop spectrophotometer reading the UV absorbance at 260nm. DNA purity was determined by the ratio $A_{260/280}$.

2.2.5. *cDNA synthesis*

To isolate BioID candidate genes, cDNA was produced from RNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and reactions were assembled as per the manufacturer's instructions. Briefly, a 2x master mix was prepared using reagents supplied including buffer, random primers, dNTPs, reverse transcriptase and RNase inhibitor, diluted in nuclease-free water. The reverse transcription reaction was set up by mixing 10µl master mix with 10µl total RNA solution (as prepared in section 2.2.3), not exceeding 200ng/µl (maximum input 2µg total RNA per reaction). The reaction was then placed in a thermocycler and programmed as follows:

Step	Temperature (°C)	Time (min)
1	25	10
2	37	120
3	85	5
4	4	Hold

Resulting cDNA was placed at 4°C for short term usage (under 24 hours) or at -20°C for long term storage.

2.2.6. *Agarose gel electrophoresis*

Agarose gels were cast fresh before use. 0.8-1% agarose powder was added to TAE buffer (w/v) and microwaved until the agarose had fully dissolved. For DNA visualisation, ethidium bromide was added to molten agarose to a final concentration of 0.5μg/ml. DNA-containing analytes were mixed with 5x glycerol-containing loading buffer and loaded into gel wells. DNA fragments were separated by applying 120V across the gel for 25 minutes. DNA fragments were analysed by UV transilluminator.

2.2.7. *Molecular cloning*

2.2.7.1. *Oligonucleotide primer design*

Oligonucleotides were designed to have a melting temperature of around 60°C and be at least 18 bases in length for plasmid templates, or 25 bases for cDNA or genomic DNA templates. Oligonucleotides for PCR cloning had an appropriate restriction site incorporated at the 5' end followed by six further bases to ensure efficient restriction enzyme processing of PCR products.

2.2.7.2. *Polymerase chain reaction (PCR) for cloning*

PCR for cloning utilised KOD polymerase for high-fidelity amplification of the appropriate template. A 50μl reaction was set up as follows:

Component	Volume (μl)
10x KOD Buffer	5
dNTP (2mM each)	5
DMSO (100%)	2.5
$MgSO_4$ (25mM)	4
Forward primer (10μM)	1.5
Reverse primer (10μM)	1.5
Template DNA	*x* (5-10ng)
dH_2O	To 50μl

Thermocycling was carried out as follows:

	Step	Temperature (°C)	Time
	Denaturation	95	2 minutes
	Denaturation	95	10 seconds
(x 38)	Annealing	60	20 seconds
	Extension	70	20 seconds / kb
	Complete	4	Hold

Completed PCR reactions were mixed with 5x loading buffer and analysed by agarose gel electrophoresis.

2.2.7.3. *Restriction digest*

For cloning, 3μg purified plasmid or 30μl PCR product (2.2.7.2) were digested using the appropriate restriction enzymes. Restriction digest reactions were set up as follows:

Component	Volume (μl)
10x Enzyme Buffer	5
Enzyme 1 (20 000 U/ml)	1.25
(Enzyme 2 (20 000 U/ml))	(1.25)
Target DNA (plasmid / PCR)	3μg / 30μl
dH$_2$O	To 50μl

For screening ligation colonies, diagnostic restriction digests were performed using 3μl of the miniprepped plasmid DNA. Reactions were set up as follows:

Component	Volume (μl)
10x Enzyme Buffer	2
Enzyme 1 (20 000 U/ml)	0.5
(Enzyme 2 (20 000 U/ml))	(0.5)
Target DNA (plasmid)	3
dH$_2$O	14(.5)

Diagnostic restriction digests and those for subcloning were incubated at 37°C for 2 hours. For PCR-based cloning, restriction digests were incubated at 37°C overnight to ensure complete digestion.

2.2.7.4. *Dephosphorylation of vector backbone*

Dephosphorylation of the backbone was carried out using calf intestinal phosphatase (CIP) to prevent vector recircularisation in the absence of insert. The following components were added directly to the 50μl restriction digest reaction:

Component	Volume (μl)
dH$_2$O	20
10x NEB Buffer 3	7.5
CIP	2.5

The reaction was vortexed to mix, spun briefly to collect liquid and incubated at 37°C for 1 hour.

2.2.7.5. *Agarose gel extraction*

DNA samples were analysed by agarose gel electrophoresis (2.2.6). The gel was then analysed by UV transilluminator, and the appropriate DNA fragment was identified by base pair length. The fragment was excised from the gel using a sharp, clean scalpel and transferred to a 2ml Eppendorf tube. Gel extraction was performed using a spin column kit purchased from either Qiagen or Favorgen and used according to manufacturer's instructions. Extracted DNA fragments were eluted from the resin column in 30μl pre-warmed dH$_2$O.

2.2.7.6. *Ligation*

Ligation reactions were assembled as follows using T4 DNA ligase (New England Biolabs (NEB)):

Component	Volume (μl)
10x T4 ligase buffer	1
T4 DNA Ligase	1
Backbone	1
Insert	7

Ligation reactions were either incubated at room temperate for 2-4hours, or at 16°C overnight.

2.2.7.7. *Cloning by Cold Fusion*

Cold fusion cloning was used to ligate DNA in the absence of appropriate restriction sites. The gene of interest insert was generated using KOD polymerase (2.2.7.2) and oligonucleotides for amplification were designed to include 15 bases homology to vector, adjacent to the site of insertion. Meanwhile, the target plasmid was linearised by one or more restriction enzymes (2.2.7.3), dephosphorylated using CIP (2.2.7.4), analysed by gel electrophoresis (2.2.6) and purified by gel extraction (2.2.7.5). A cold fusion reaction was then set up as follows:

Component	Volume (μl)
5x Cold Fusion MasterMix	1
Linearised vector	0.5
PCR insert	0.5
dH$_2$O	3

The reaction was briefly vortexed to mix the components, spun to collect liquid and incubated for 5 minutes at room temperature followed by 10 minutes on ice. The 5μl reaction was then transformed into TOP10 competent bacteria (2.2.1.1)

2.2.7.8. *Screening cloning colonies by PCR*

Single colonies were picked from cloning plates using a clean pipette tip and used to inoculate 10μl dH$_2$O. 1μl of the colony suspension was used in an EmeraldAmp GT PCR reaction set up as follows:

Component	Volume (μl)
2x EmeraldAmp MasterMix	5
Forward primer (10μM)	0.3
Reverse primer (10μM)	0.3
Template DNA	1
dH$_2$O	To 50μl

Thermocycling was carried out as follows:

	Step	Temperature (°C)	Time
	Denaturation	94	5 minutes
	Denaturation	98	10 seconds
(x 38)	Annealing	60	30 seconds
	Extension	72	60 seconds / kb
	Final extension	72	5 minutes
	Complete	4	Hold

PCR products were then analysed by agarose gel electrophoresis. Positive colonies were used to inoculate liquid cultures for plasmid purification and sequencing analysis.

2.2.8. *PCR from genomic templates*

2.2.8.1. *Genomic DNA (gDNA) extraction from mouse tail biopsy*

DNA for genotyping animals was extracted using HotSHOT alkaline lysis. 75µl HotSHOT alkaline lysis reagent (Appendix D.I) was added to a section of mouse tail no more than 3mm in length. The mixture was incubated at 95°C for 30 minutes, and the reaction stopped by adding an equal volume of HotSHOT neutralising buffer (Appendix D.I). Samples were stored at -20°C, and 1-2µl was used for PCR reactions.

2.2.8.2. *Genomic DNA extraction from cultured cells*

Cells for gDNA extraction were washed twice with PBS, before adding 10µl / cm^2 QuickExtract (Lucigen) lysis buffer. Samples were triturated, aspirated into a PCR strips, and vortexed for 15 seconds to homogenise. PCR strips were transferred to a thermocycler and incubated at 98°C for 6 minutes, followed by 65°C for 2 minutes and then held at 4°C. Samples were stored at -20°C and diluted 1:10 for PCR reactions.

2.2.8.3. *PCR for genotyping reactions*

For genotyping, MangoMix (Bioline) master mix was used to amplify the modified or knocked-in gene. 22µl reactions were set up as follows:

Component	Volume (µl)
2x MangoMix master mix	11
Forward primer (10µM)	0.3
Forward primer (10µM)	0.3
DMSO (100%)	1
gDNA template	1-2
H_2O	To 22µl

Thermocycling was carried out as follows:

	Step	Temperature (°C)	Time
	Denaturation	95	3 minutes
	Denaturation	95	30 seconds
(x 38)	Annealing	60	30 seconds
	Extension	72	30 seconds / kb
	Final extension	72	7 minutes
	Complete	4	Hold

Completed PCR reactions were loaded directly into agarose gels for analysis.

2.2.8.4. *PCR from genomic DNA for sequencing*

PCR for sequencing genomic DNA of modified cells was carried out using the high-fidelity polymerase, KOD. 20µl reactions were set up as follows:

Component	Volume (µl)
10x KOD Buffer	2
dNTP (2mM each)	2
DMSO (100%)	1
MgSO$_4$ (25mM)	1.6
Forward primer (10µM)	0.6
Reverse primer (10µM)	0.6
Template DNA	x (~50ng)
dH$_2$O	To 20µl

Thermocycling was performed as in section 2.2.7.2.

2.2.9. TOPO TA cloning for sequencing

2.2.9.1. *Adenylation of PCR products*

PCR products produced using the high-fidelity polymerase KOD required 3' adenylation for cloning into the TOPO TA vector. This was executed by incubating the gel purified KOD PCR product with *Taq*-containing MangoMix (Bioline) at 72°C for 30 minutes.

2.2.9.2. *TOPO TA cloning for DNA sequencing*

3' adenylated PCR products (produced using *Taq* or as described in section 2.2.9.1) were ligated into the TOPO TA vector as per the manufacturer's instructions. Briefly, a 6µl cloning reaction was set up as follows:

Reagent	Volume (µl)
PCR product	4
Salt solution	1
TOPO vector	1

The reaction was mixed and incubated for 5 minutes at room temperature before transforming into freshly thawed chemically competent bacteria (see section 2.2.1.2)

2.2.10. *DNA Sequencing*

Sanger DNA sequencing was performed by Bio-Basic Asia. Plasmids for sequencing were diluted to 100ng/μl and custom sequencing primers provided at 10μM. 10μl of plasmid or primers were sufficient for 3 sequencing reactions. The read length achieved was ~800-1500 bp, depending on sequence complexity.

2.3. Protein techniques

2.3.1. *Protein sample preparation*

Adherent mammalian cells were washed twice with cold PBS before adding 10μl/cm^{-2} cold RIPA buffer supplemented with 1x protease inhibitor cocktail (Roche) and 1mM sodium orthovanadate. Samples were collected by cell scraping, and DNA was sheared by sonication for 5 seconds, repeated twice. Sonicated samples were then centrifuged at 18,000 x *g* for 10 minutes at 4°C to pellet insoluble material.

2.3.2. *Nuclear extraction*

Adherent cells were washed twice in ice cold PBS, before being scraped from the culture surface using a minimal volume of ice-cold PBS freshly supplemented 1x protease inhibitors. Cells were collected in a 1.5ml tube, and centrifuged at 4°C for 5 mins at 700 x *g*. Pelleted cells were resuspended in ice cold hypotonic lysis buffer (See appendix D.II) at a concentration of 50 x 10^6 cell/ml and incubated on ice for 10 mins. Cell lysis was established by trypan blue staining analysed using a light microscopy. Nuclei were then pelleted by centrifugation at 1500 x *g* for 5 mins at 4°C, Nuclei were washed a total of 5 times in 500μl ice cold PBS with protease inhibitors, centrifuging as before after each wash. Nuclear pellets were then processed as in section 2.3.1.

2.3.3. _Quantification of protein samples by BCA assay (Pierce)_

Protein quantification was carried out in accordance with the manufacturer's instructions for a microplate procedure. Protein standards between 0 and 2000µg/ml were made by serial dilution. 25µl of each standard and sample (10-fold diluted) were mixed with 200µl working reagent (50:1 Reagent A:B). The microplate was thoroughly mixed by plate-shaker for 30 seconds, before being incubated at 37°C for 30 minutes. Upon cooling to room temperature, protein concentration was measured using a spectrophotometer set up for plate-reading absorbance at 562nm. A standard curve was constructed using the measurement for the BSA standards and concentration of samples calculated.

2.3.4. _Protein analysis by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)_

Protein samples were mixed with either 2x, or 4x sample buffer, depending upon concentration, and heated at 95°C for 5 minutes prior to gel electrophoresis. SDS polyacrylamide gels were cast using the BioRad Protean system with an appropriate percentage of acrylamide, depending upon molecular weight of protein of interest (Appendix D, II). First, the variable percentage resolving gel was poured into casting apparatus, and overlain with 1ml isopropanol to yield a level interface between the resolving and stacking gels. The resolving gel was left to polymerise, after which the isopropanol was removed. The 5% stacking gel was poured atop the resolving gel and a comb inserted to mould the wells. The stacking gel was left to polymerise for 10-15 minutes. Gels were assembled into the Bio-Rad gel tank which was filled with tris-glycine running buffer (1st Base, Singapore). Protein samples were then loaded into the wells. Proteins were then electrophoresed first at 100V through the stacking gel (15 mins) and then 150V for 1 hour until the dye-front ran through the gel.

Where proteins with a range of molecular weights were of interest, protein samples were separated using pre-cast NuPAGE Bis-Tris gradient gels (Invitrogen). Gels were assembled into the gel tank which was filled with 1x NuPAGE MOPS-SDS running buffer. Before loading protein samples, a syringe and needle were used to flush wells with running buffer. Protein samples were then loaded into the wells and

initially electrophoresed at 100V for 15 mins followed by 150V for 1 hour until the dye-front ran through the gel.

2.3.5. *Coomassie staining*

Coomassie staining was used to fix and visualise proteins produced by bacterial expression in the gel. Coomassie staining solution was added to the gel and incubated at RT with gentle agitation for 30 minutes. Next, de-staining solution was added and replaced every 20 minutes until background staining was reduced and protein bands were easily visible.

2.3.6. *Western Blotting*

2.3.6.1. *Protein transfer to PVDF membrane*

Proteins were separated by SDS-PAGE gel electrophoresis (as described in section 2.3.4) and then transferred on to PVDF membrane either by wet transfer or by using iBlot-2 apparatus. PVDF membrane was activated by soaking in 100% methanol and then rinsed in either transfer buffer or water. Gels were prepared for blotting by removing the stacking buffer and equilibrating in transfer buffer or water for wet or semi-dry transfer, respectively. For wet transfer, the cassette was assembled as follows: anode-cassette-sponge-filter paper-gel-activated PVDF membrane-filter paper-sponge-cassette-cathode. Each layer was rolled flat to ensure the absence of air bubbles affecting the transfer efficiency. Wet blotting was performed either by applying 100V for 75 minutes, or overnight, applying 10V for 16 hours. Blotting using the iBlot-2 was performed according to the manufacturer's instructions. Briefly, gel was assembled into the pre-packaged kits, and programme P1 (25V) was selected and run for 6 minutes.

2.3.6.2. *Ponceau S staining*

Proteins transferred to PVDF membranes (as described in section 2.3.6.1) were stained using Ponceau S (Sigma Aldrich) following transfer to ensure even blotting. 0.1% Ponceau S was incubated with the membranes for 5 minutes with gentle agitation, followed by washing in distilled water to remove background stain. Ponceau S was removed from the membranes was 3x 5minute washes in TBST (see appendix D.II.).

2.3.6.3. *Antibody probing and western blot development*

PVDF membranes were blocked in blocking solution for 1 hour at RT with gentle agitation. Primary antibodies were diluted appropriately in blocking solution and incubated with the membrane in a rolling Falcon tube either at room temperature for 1 hour or overnight at 4°C. After primary incubation, membranes were washed 3x 10 minutes in TBST at RT with gentle rocking. Secondary HRP- or IRDye-conjugated antibodies were incubated for a further hour, diluted 1:10 000 in blocking buffer. Membranes were washed in TBST as before. Bound antibodies were either detected using ECL (SuperSignal Wester PICO Plus, ThermoFisher Scientific) followed by exposing the membrane to X-ray film (Amersham Hyperfilm) for between 1 second and 40 minutes and development, or imaged using a Li-Cor Odyssey imager.

2.3.7. *In Cell Western assay (LI-COR Biosciences)*

Cells were cultured in a 96 well plate to 80% confluence before analysis by In Cell Western. First, samples were washed using cold PBS twice and fixed using 4% formaldehyde for 10 minutes at room temperature. Formaldehyde was aspirated from the wells, and cells were permeabilised by incubating with 0.1% triton X-100 in PBS for 5 minutes at room temperature with gentle agitation. Permeabilisation buffer was aspirated from the wells and cells were thoroughly washed three times using PBS. Samples were blocked for one hour at room temperature using 5% BSA in PBS. Blocking solution was aspirated and replaced with primary antibodies diluted in blocking buffer. After incubation for one hour at room temperature, primary antibody solution was removed and samples washed three times with PBS. Secondary IRDye-700/800 antibodies (LI-COR Biosciences) were diluted in blocking buffer, added to the cells and incubated with the sample for one hour a room temperature. Samples were then washed three times with PBS, and 150µl PBS was added to each well for imaging by LI-COR Odyssey.

2.3.8. *In vitro transcription-translation*

In vitro transcription-translation (IVT) was performed according to the manufacturer's instructions (Pierce; 1-step human coupled IVT kit). For a 25μl reaction, 12.5μl HeLa lysate, 2.5μl accessory protein, 5μl reaction mix and 1μg plasmid DNA were combined and topped up with molecular biology grade water. Components were mixed by briefly vortexing, collected to the bottom of the tube by pulse centrifugation and incubated in a heating block set to 30°C for 90 minutes. IVT products were stored at -80°C until used. Proteins were subject to no more than 2 freeze-thaw cycles.

2.3.9. *Immunoprecipitation of IVT proteins*

Protein G Dynabeads (Thermo Fisher) were thoroughly vortexed to resuspend, and 15μl per reaction was removed and placed into a Falcon tube. Dynabeads were washed in an excess of water twice as follows. Dynabeads were vortexed to resuspend, placed on the magnetic rack for 20 seconds, and water discarded. After the final wash, protein G Dynabeads were resuspended in a small volume of 1x IP binding buffer (see appendix D.II.). Thereafter, dynabeads, HA (3F10) antibody (1:40 dilution) and 1μl IVT proteins were diluted into 500μl IP binding buffer with 1% BSA (w/v) in a 1.5ml Eppendorf tube. Tubes were incubated overnight with end-over-end agitation at 4°C. The next day, tubes briefly centrifuged to collect liquid, placed on magnetic stands for 20 seconds, and the supernatant was removed. Protein G Dynabeads were then washed three times for 5 minutes in 1x IP buffer, vortexing each time for resuspension. After the final wash, Dynabeads were resuspended in 30μl 1x sample buffer, heated at 95°C for 5 minutes and analysed by SDS-PAGE.

2.3.10. *Immunoprecipitation from cell lysates*

50% protein A slurry (Amintra) were thoroughly vortexed to resuspend, and 30μl of the suspension was removed for each immunoprecipitation condition, and washed three times in cold lysis buffer. Cell lysates were pre-cleared by mixing with washed protein A beads and incubated at 4°C with end-over-end agitation for 1 hour. The pre-cleared sample was separated from the beads by centrifugation at 600 x g for 30 seconds at 4°C, and the sample aspirated into a fresh tube. For

immunoprecipitation, appropriate specific antibodies, or isotype control antibodies were diluted into the pre-cleared lysates overnight at 4°C with end-over-end agitation. The next day, the lysate-antibody mix was combined with washed protein A beads and incubated for a further hour at 4°C with end-over-end agitation. The beads were then pelleted by centrifugation as before, and the supernatant was transferred to a fresh tube. Protein A beads were washed three times in cold lysis buffer, and finally resuspended in 50μl 2x sample buffer and heated at 95°C for 5 minutes.

2.4. GST-binding assays

2.4.1. *Bacterial expression and purification of GST-fusion proteins*

The E. coli strain BL21(DE3) were transformed with pGEX-4T constructs containing the protein of interest, plated on agar plates containing ampicillin and incubated overnight at 37°C. The next day, one colony was picked and used to inoculate a 5ml culture, which was incubated shaking for 8-10 hours at 37°C. The turbid 5ml culture was diluted into an appropriate volume of 2xYT medium, and grown overnight at 37°C while shaking. The next day, IPTG was added to a final concentration of 0.5mM and incubated for a further 3 hours. Bacteria were then centrifuged for 15 minutes at 6000 x g, media poured away, and resuspended in PBS supplemented with protease inhibitor cocktail and 100mM PMSF. To lyse, cells were thoroughly resuspended in PBS, transferred to a Falcon tube and sonicated 5 times for 30 seconds, resting for 30 seconds on ice in between. Samples were then centrifuged at 20,000 x g for 30 minutes at 4°C to pellet insoluble material.

2.4.2. *Preparation of glutathione Sepharose beads.*

Bead slurry was vortexed thoroughly to resuspend, and 100μl was added to a 15ml Falcon tube and 10ml cold PBS added. The mixture was centrifuged at 500 x g for 5 minutes to pellet the beads, and then washed twice more with 10ml PBS.

98

2.4.3. *Purification of GST-fusion proteins*

GST-fusion containing supernatant was applied directly to the washed glutathione Sepharose beads, and inverted to resuspend thoroughly. The mixture was then incubated for 2 hours at room temperature or overnight at 4°C to capture GST-fusion proteins. After binding, the beads were washed 3 times using 10ml cold TBST to reduce non-specific binding, followed by 3 washes using 10ml cold PBS. To elute the fusion protein, the beads were resuspended in 500μl fresh 10mM reduced glutathione in 50mM Tris-HCl, transferred to microcentrifuge tubes and incubated for one hour at room temperature while rotating.

2.4.4. *Quantification of bacterially expressed GST-fusion proteins*

GST-fusion proteins were quantified using a Nanodrop, blanked to reduced glutathione, and measuring absorbance at 280nm (A_{280}). Concentration was determined using the A_{280} and the proteins extinction coefficient.

2.4.5. *Peptide array binding assays*

Peptide arrays (Celluspots, Intavis) containing 15 amino acid peptides corresponding to the cytoplasmic tail of βDG, each with a 14 amino acid overlap (Thompson *et al.*, 2010) were blocked using 5% non-fat milk in TBST for 30 minutes at room temperature. Purified GST-fusion protein was diluted to 10μg/ml in 5% milk in TBST and incubated with the array for 2 hours at room temperature or overnight at 4°C with gentle rocking. The array was then rinsed 3 times with TBST and washed 3 times with TBST, each for 5 minutes with gentle agitation. The primary antibody was diluted into 5% milk in TBST and incubated with the Celluspots array for 1 hour at room temperature, or overnight at 4°C. After primary antibody incubation, the array was rinsed and washed as before. HRP-conjugated secondary antibody in 5% milk in TBST solution was then incubated with the array for 1 hour at room temperature, before rinse and washing the array as before. Excess TBST was drained, ECL substrate was added, and signal from the Celluspots array was then visualised using a BioRad Gel-Doc.

2.4.6. *GST pulldown*

For the probing of cell lysates, GST, or GST-fusion proteins were not eluted from glutathione Sepharose beads. Instead, GST, or GST-fusion proteins bound to glutathione Sepharose beads were mixed with whole cell lysates in RIPA lysis buffer, made from one quarter of a T175 flask at 80% confluence. The lysate and beads were incubated at 4°C for 6 hours with end-over-end agitation. After incubation, the supernatant was removed from the beads which were pelleted by centrifugation at 600 x g for 30 seconds, and washed three times with 1 ml cold lysis buffer. After the final wash, glutathione Sepharose beads were resuspended in 50µl sample buffer, and heated at 95°C for 5 minutes.

2.5. Culturing and genetic manipulation of mammalian cell lines

2.5.1. *Mammalian cell line culture*

HeLa, HEK293T and C2C12 cells were cultured in ZLZ-M media consisting of DMEM medium containing 4500mg/l glucose and supplemented with 10% FBS (v/v). Differentiation of myoblast lines from various origins was induced by changing growth medium to reduced serum differentiation medium at 80% confluence. Cell lines were maintained in a 5% CO_2 incubator at 37°C.

2.5.2. *Passaging cell lines*

Generally, cell lines were passaged upon reaching 90% confluence. Myoblasts were passaged at 60-70%. To passage, culturing medium was removed and the cells were rinsed with DPBS. A small volume of 0.25% trypsin (between 15-30µl/cm^2) was added to the cell culture vessel, distributed evenly, and incubated at 37°C for around 5 minutes while monitoring as cells detached. Once cells had dissociated, trypsin was neutralised using at least the same volume of culturing medium, transferred to an appropriate Falcon tube, and centrifuged at 300 x *g* for 5 minutes. The cell pellet was resuspended and gently triturated to yield a single cell suspension. For continuation, cells were passaged at a ratio of 1:10-1:20, and transferred to a new vessel with fresh media. For experiments, cells were counted using a haemocytometer or automated counter (Bio-Rad) to seed in an appropriate dish / plate.

2.5.3. _Long-term storage of mammalian cells_

Mammalian cells were cultured for at least 2 passages and grown to 80% confluence before re-freezing. Cells were detached from the culturing surface, pelleted and resuspended in freezing medium (10% DMSO in FBS) before being aliquoted into 2ml cryovials. In general, 20μl freezing medium was used per cm^2 culturing area. Cryovials were first stored at -80°C in a freezing container to regulate cooling. After 48 hours, cryovials were transferred to liquid nitrogen or -152°C freezer for long-term storage.

2.5.4. _Plasmid transfer into mammalian cells_

2.5.4.1. **Transfection using Invitrogen Neon Electroporation System**

Transfection of the KM155 human myoblast cell line (Mamchaoui _et al._, 2011) was performed using the Neon Electroporation System. Myoblasts were grown to 70% confluence, trypsinised, and counted. Cells were centrifuged at 300 x g for 5 minutes and the cell pellet was washed twice in PBS. Myoblasts were resuspended in PBS at a concentration of 2×10^6 cells/ml, and maxiprepped plasmid DNA was added to a final concentration of 75μg/ml. Plasmid was mixed by gently inverting the tube 5 times, and incubated at room temperature for 5 minutes. The cell suspension was aspirated into a 100μl Neon electroporation tip, placed in the dock, and electroporated using 1 pulse of 1400V for 30ms. Electroporated cells were immediately dispensed into one well of a 6 well plate containing pre-warmed culture media.

2.5.4.2. **Transfection using Lipofectamine 3000**

Cells were evenly plated 24 hours prior to transfection, such that they would be 70-90% upon transfection. The next day, cells were transfected according to the manufacturer's instructions. Briefly, Lipofectamine reagent was added with OptiMEM (Gibco), and vortexed for 2-3 seconds to mix. Next, up to 680ng plasmid DNA / cm^2 culturing area was diluted into OptiMEM, and P3000 reagent was added at a ratio of 2μl/μg DNA. Diluted Lipofectamine 3000 was then mixed with the diluted plasmid, and incubated at room temperature or 10-15 minutes. Complexed DNA was added to cells and incubated in OptiMEM under standard culture

conditions. Culturing media was changed after 6 hours for most cell types, or left overnight for transfection of HEK293T cells.

2.5.4.3. *Adenoviral transduction*

Adenovirus containing GFP, and GFP-Cre was purchased from SignaGen (SL100708, SL100706, respectively) and diluted with standard culturing media to use at multiplicity of infection (MOI) between 10-100. Cells were incubated as standard overnight and infection efficiency was assessed after 24 hours based on GFP-expressing cells.

2.5.4.4. *Lentiviral transduction*

Primary cells and immortalised cell lines were plated at 1.5×10^5 cells/well in a 6 well plate. The next day, normal culturing media was replaced with 1ml media containing 8μg/ml polybrene (Sigma). Concentrated lentivirus (MOI ~10) was added, the plate gently swirled to mix, and incubated under normal conditions. After 6 hours, media volume was topped up to 3ml with normal culturing media 8μg/ml polybrene. The next day, media was replaced with ordinary growth media and the cells incubated for 2-3 days depending upon confluence.

2.5.4.4.1. Selection of mammalian cells using antibiotics

Primary cells and immortalised cell lines which had been transduced with lentivirus encoding a gene of interest included an antibiotic resistance gene. 4 days after transduction, the antibiotic was added to culture media at a concentration empirically determined for the cell type (Appendix E). Growth media was replaced every 2-3 days, and cells passaged when 70-80% confluent, except myoblasts which were passaged at 50-60% confluence. Cells were maintained in culture under antibiotics for at least one week before further experimentation or manipulation. Antibiotic selection pressure was maintained on cells transduced with lentivirus for prolonged culture.

2.5.4.4.2. Induction of gene expression from Tet-On constructs

The induction of transgene expression in cells selected for stable genomic integration was performed by the addition of doxycycline at a concentration of 1μg/ml into the culture media for three days prior to experimentation.

2.5.5. *Lentiviral production*

2.5.5.1. **Lentivirus production**

Lentivirus was produced in HEK293T using packaging components VSV-G, Δ8.9 and construct-of-interest-containing pTRIPZ-based lentiviral vector. HEK293T cells were plated at a density of 1.52 x 10^5 cells/cm^2 in antibiotic-free medium on day 1 such that they were at 90% confluence the next day. On day 2, plasmids were co-transfected with the packaging, and target plasmids. Transfection medium was replaced with antibiotic-free ZLZ-M after 6 hours. Lentivirus-containing media was then collected in the morning and evening of days 3, 4 and 5.

2.5.5.2. **Lentivirus concentration**

During production, lentivirus in ZLZ-M was stored at 4°C. Once all collections were complete, the media was filtered using a 0.2μm sterile filter unit (Nalgene) to remove cell debris, and centrifuged at 8000 x g for 16 hours overnight to pellet lentivirus. Lentivirus pellet was then resuspended in a small volume of antibiotic-free ZLZ-M contain 8μg/ml polybrene, and stored in aliquots at -80°C.

2.5.6. *CRISPR genomic manipulation*

A CRISPR system using plasmid pSpCas9(BB)-2A-GFP (pX458) was utilised, which encodes the guide RNA, Cas9 enzyme, and EGFP for selection (Ran, Hsu, Wright, *et al.*, 2013).

2.5.6.1. **Guide RNA sequence design**

The gene region to target was queried for potential CRISPR sites using either http://tools.genome-engineering.org (for DAG1; now defunct). Modifications were required at or near the start codon. Guide sequences were then selected based on a balance of parameters such as proximity to intended editing site, off-target prediction, and ability to screen by disruption of endogenously encoded restriction site.

2.5.6.2. *Design and generation of donor plasmids*

Complementary oligonucleotides encoding the gRNA with appropriate cloning overhangs were ordered, and cloned into the linearised pX458 vector digested using BbsI. gRNA incorporation was verified by sequencing, and constructs were propagated and purified by maxi-prep.

2.5.6.3. *Screening for gRNA editing efficiency*

2.5.6.3.1. *The SURVEYOR nuclease assay*

The SURVEYOR nuclease assay was performed as described in (Ran, Hsu, Wright, *et al.*, 2013). Briefly, PCR primers were designed to amplify the region of genomic DNA targeted for CRISPR editing. A single PCR product was verified by running a small amount of product by agarose gel electrophoresis, and remaining was purified by PCR clean-up. DNA heteroduplexes made from edited and WT strands were formed by heating the PCR product to 95°C for 10 minutes and re-annealing by reducing the temperature of the thermocycler by 2°C s^{-1} to 85°C and 0.3°C s^{-1} to 4°C. DNA heteroduplexes were then subjected to digestion by T7 endonuclease I (NEB) to give indication of editing efficiency within the population of cells transfected with pX458.

2.5.6.3.2. *CRISPR efficiency screening by restriction digest*

Similarly to the SURVEYOR assay, a short PCR product was amplified across the targeted region and verified as a single product by agarose gel electrophoresis. PCR product was extracted from the gel using a spin column kit (Favorgen/Qiagen), and subject to restriction digest by the hypothesised disrupted enzyme before analysis by agarose gel electrophoresis.

2.6. Derivation of primary cells from murine origin

2.6.1. *Mouse adult fibroblast (MAF) derivation from tail biopsy*

A 1cm section of tail was cut from the mouse and rinsed with 1x antibiotic-antifungal solution (Thermo Fisher Scientific) followed by 70% ethanol twice. The sterilised tail biopsy was placed into a 6cm dish and minced using surgical scissors and a scalpel. 0.5ml dissociation solution (see Appendix D, V) was added to the

minced tail and incubated at 37°C for 30 minutes. 6 ml ZLZ-M was added and dish was incubated at 37°C in 5% CO_2 and 3% O_2 undisturbed for 24 hours. The next day, the media was triturated using a large aperture 1ml pipette to aid dissociation of cells from larger pieces of tissue. After 48-96hours, media and larger pieces of tissue were removed from the dish and transferred to a new one. Fresh media was added to the original plate. Fibroblast confluency was monitored over the following days, substituting half the media volume every 2 days until 80% confluence, at which point they were passaged and transferred to a T75.

2.6.2. *Myoblast derivation from muscle*

Plates were pre-coated with 0.1% gelatin solution (Appendix D, V) for at least 30 minutes at 37°C before plating of primary myoblasts. One mouse was sacrificed using CO_2 and sterilised by drenching in 70% ethanol. Skin was removed from the hindlimb from the ankle upwards. Hindlimb was disconnected from the body at the hip joint using scissors, and feet were removed. The entire hindlimb was then placed into 1x antibiotic-antimycotic solution (Thermo Fisher Scientific) on ice. All contaminating hair was rinsed off and hindlimb was placed into a 60mm dish containing 1x HBSS (Gibco) on ice. Fats and nerve tissue were dissected away from the hind muscles as much as possible, before removing the muscle tissue from the bone. Dissected muscle tissue was placed into a new 60mm dish containing 1x HBSS on ice. HBSS was carefully aspirated from the dish, the muscle was minced using a sharp scalpel, transferred to a 50ml Falcon tube and weighed. 4ml sterile filtered enzyme solution (Appendix D, V) per gram of muscle tissue was added to the 50ml falcon tube which was incubated in a 37°C water bath for 30 minutes, triturating 20x using a wide orifice 1ml pipette after the first and second 15 minutes. Tissue digestion was neutralised by adding an equal volume of DMEM containing 10% FBS. Next, the tissue digest was sequentially filtered through 100μm, 70μm, and 40μm sterile filters to remove larger pieces of debris. The resulting suspension was then centrifuged at 300 x g for 5 minutes to pellet cells. Cell pellet was washed with DMEM 10% FBS and centrifuged once more, as before. 5ml DMEM 10% FBS was used to resuspend the cell pellet and plated in a non-gelatin coated T25. Cells were incubated for 60-120 minutes, periodically assessing for attachment of fibroblasts.

When majority of fibroblasts had begun to adhere, and spreading was evident, the culture supernatant was removed and centrifuged at 300 x g for 5 minutes to pellet remaining cells including myoblasts. Fresh DMEM 10% FBS was added to the fibroblasts for growth. The pellet was resuspended in Ham's F10 media supplemented with 20% FBS and 10ng/ml fresh heat stable βFGF (Gibco). Medium was changed on myoblasts every three days and passaged upon reaching 70% confluence or every five days, whichever came first.

2.6.3. *Culturing of primary mouse cells*

All primary mouse cells were cultured in a hypoxic (3% O_2) 5% CO_2 incubator maintained at 37°C. Mouse adult fibroblasts derived from tail biopsies or muscle were grown in ZLZ-M (see appendix D), while primary myoblasts were grown in Ham's F10 medium supplemented with 20% FBS and 10 ng/ml fresh heat-stable βFGF (H20%). Cells were passaged as described in section 2.5.2.

2.6.4. *Differentiation of primary myoblasts*

To differentiate primary myoblasts into myotubes, cells were plated on gelatin-coated surfaces at high density such that they would reach 80% confluence within 24 hours. The next day, H20% growth medium was discarded, cells were gently rinsed with DPBS, and differentiation medium (Ham's F10 medium supplemented with 2% horse serum) was added. Differentiation was monitored over the next 5 days, and characterised by cell fusion, expression of mature muscle markers, and spontaneous contractile events. Differentiation media was changed every 2 days.

2.6.5. *Immortalisation of fibroblasts*

Mouse adult fibroblasts (MAFs) were immortalised by infection by SV40 Large T antigen. Briefly, HEK293T cells stably packaging large T antigen were grown to confluence and media supernatant was collected daily before 0.45μm sterile filtering, aliquoting and storing at -80°C. MAFs below passage 3 were plated at a density of 1.5 x 10^4 cells cm^{-2} and incubated with a minimum volume of viral supernatant supplemented with 8μg/ml polybrene to aid infection efficiency. After

24 hours, media was changed to standard growth medium and MAFs were passaged as they reached confluence.

2.7. Immunofluorescence staining and microscopy

2.7.1. *Immunofluorescence staining for widefield and confocal microscopy*

Cells were propagated on coverslips or in 8 well chamber slides (ibidi). After sufficient culturing time, media was removed and cells were washed twice with 1x PBS before a 10 minute incubation in freshly made or thawed 4% formaldehyde in 1x PBS (Thermo Scientific). PFA solution was removed and cells were incubated in for 5 minutes in 0.1% Triton X-100 for 10 minutes at room temperature to permeabilise. Samples were then washed thoroughly three times with 1x PBS. To stain, cells were blocked using blocking buffer (5% FBS, 3% BSA in 1x PBS) for 1 hour at room temperature to reduce non-specific antibody staining. Primary antibodies were diluted into blocking buffer and incubated with samples for 1 hour at room temperature or overnight at 4°C. Samples were washed three time using PBS, and incubated in solution containing secondary Alexa fluorophore-conjugated secondary antibodies and Hoechst (1μg/ml) diluted in blocking buffer. Following staining, cells were washed three times in PBS followed by once in dH$_2$O. Samples on coverslips or in ibidi chamber slides were then mounted on slides using ProLong Diamond antifade mountant (Invitrogen), or by applying 4 drops of ibidi mounting medium (ibidi) / mounting medium (see appendix D.III.), respectively.

2.7.2. *Widefield microscopy acquisition*

Widefield imaging of samples was typically performed on a Nikon Eclipse Ti inverted microscope equipped with a Nikon Instensilight (C-HGFIE) and CoolLED (pE-100) for fluorescence and brightfield illumination, respectively, and Photometrics CoolSnap HQ2 camera. Samples were analysed at 20x magnification using an air immersion plan apochromat VC objective lens with numerical aperture of 0.75, or at 60x magnification using an oil immersion CFI plan apochromat objective lens with numerical aperture of 1.40. The microscope was controlled using the NIS-Elements

software from Nikon. Alternatively, widefield imaging was performed on a Leica DMIRE2 microscope using a 63x oil immersion lens. The microscope was controlled and images were captured using Qfluoro software using a machine running Windows 2000.

2.7.3. *Confocal microscopy acquisition*

Confocal microscopy was used to acquire high-contrast optical sections through samples. Sample imaging was performed either using a Nikon A1 confocal microscope (Wolfson light microscopy facility in Sheffield, UK) using a 60x (NA 1.40) oil immersion CFI plan apochromat objective lens, or using an Olympus FV3000 confocal microscope (A*STAR microscopy platform (AMP), Singapore) using a 60x (NA 1.35) oil immersion UPlanSApo objective lens. The respective microscopes were controlled using Nikon's NIS-Elements or Olympus's FV31S-SW software.

2.8. Flow cytometry and fluorescence activated cell sorting (FACS)

Cultured cells were trypsinised and washed once in DPBS. Thereafter, cells were resuspended in FACS buffer (see appendix D.V.a) to a maximum concentration of 1×10^6 cells/ml, and stored on ice until all samples were prepared and ready for analysis by flow cytometry. Prior to analysis or sorting, cell suspensions were gently vortexed to declump cells, and mounting on the FACS analyser.

2.9. BioID methods and mass spectrometry preparation

2.9.1. *Culturing of myoblasts for BioID assays*

On day 1, 35×10^5 myoblasts were seeded in a T175 with growth medium supplemented with 1µg/ml doxycycline, depending on the experimental condition. Growth medium was replaced on day 4 to include 1µg/ml doxycycline (Sigma), 50µg/ml biotin (Sigma) and 0.5µM AP21967 (Takara), depending on required conditions. Samples were harvested 24 hours later (Roux, Kim and Burke, 2013; Chojnowski *et al.*, 2018).

2.9.2. *Cell lysis and protein extraction*

After sufficient propagation and culturing with appropriate experimental supplements, cells were washed 3 times in 1x PBS, to remove excess media, and trypsinised to detach from the culture surface. Cells were counted and pelleted by centrifuging at 300 x g for 5 minutes. Pelleted cells were resuspended in PBS and pelleted once more, as before, to remove trypsin. Cells were then lysed using 2ml cold lysis buffer (see Appendix D, VI) / T175. Mixture was vortexed to homogenise the solution, and sonicated 2 times for 8 seconds, placing the sample on ice in between. To remove insoluble components, samples were cleared by centrifugation at 16 000 x g for 10 minutes at 4°C, and the supernatant transferred to a fresh Eppendorf tube. An 'input' sample was removed after this step for analysis by western blotting, and stored at -20°C.

2.9.3. *Preparation of Dynabeads*

Dynabeads MyOne (Thermo Fisher Scientific, Cat# 650.01) streptavidin magnetic beads were thoroughly resuspended by vortexing for 30 seconds, and 100μl slurry was transferred to a 1.5ml Eppendorf tube. To wash, 850μl PBS was added, and the tube transferred to the magnet stand for 1 minute. This process was repeated three times.

2.9.4. *Capture of biotinylated proteins*

PBS was removed from the final wash of the Dynabeads, and protein extracts added. Tubes were then incubated with gentle end-over-end agitation at 4°C overnight. The next morning, samples were incubated for a further 30 minutes at room temperature. Next, the tubes were transferred onto the magnet stand and the unbound sample in the supernatant was removed. 950μl wash buffer A was then added to the beads, thoroughly triturated to resuspend, and transferred to a 2ml Eppendorf tube. Samples were incubated for 10 minutes at room temperature with end-over-end agitation. Supernatant was removed after separation of Dynabeads on the magnetic stand. Wash in buffer A was repeated once more as described. After the send wash, Dynabeads were then resuspended in 950μl buffer B and incubated as before for 10 minutes. Washes in buffer B were repeated twice more.

2.9.5. *On-bead trypsin digestion of captured protein*

Upon the final wash, beads in buffer B were transferred to a new 1.5ml Eppendorf, and the supernatant removed. Working under the fume hood, the Dynabeads were resuspended in 50µl buffer ED1 and TCEP was added to 20mM. Mixture was vortexed and incubated at 55°C for 20 minutes while shaking. CAA was then added to a final concentration of 55mM, mixed and incubated in darkness at room temperature for 30 minutes. Samples were then diluted to 600µl using 100mM TEAB, and 10µg LysC per reaction was added before incubating at 37°C for 4 hours with shaking. 100mM TEAB was added to dilute the samples to 750µl and 10µg trypsin per reaction was added. Samples were incubated at 37°C while shaking overnight. Dynabeads were separated using the magnet stand, and the supernatant transferred to a new Eppendorf tube. 1% TFA (v/v) was added to acidify the samples and stop the reaction, inverting to mix.

2.9.6. *Peptide sample desalting*

C18 desalting cartridges (Waters) were prepared by flowing though 2ml buffer D, followed by equilibrating using 2ml buffer E, both eluted under low vacuum pressure, preceding sample application. Samples were diluted into Falcon tubes containing 4ml buffer E, applied to the C18 cartridge, and a slow through-flow was controlled using low vacuum pressure. Samples were washed by slowly flowing 4ml buffer E through the cartridge, and fully removed by increasing the vacuum pressure. Desalted samples were then eluted by adding 1ml buffer D to the cartridge, and a syringe used to apply pressure to transfer into an Eppendorf tube.

2.9.7. *Mass spectrometry*

Samples were vacuum concentrated, and analysed using an Easy nLC1000 (Thermo) liquid chromatograph equipped with 50cm x 75µm id Easy-Spray columns (C-18, 2µm particles, Thermo) coupled to an Orbitrap Fusion (Thermo) mass spectrometer. Samples were resolved through a 120 minute acetonitrile gradient (0.0%-99.9% (v/v), containing 0.1% formic acid (v/v)). An Orbitrap and iontrap analyser (OT-MS 4xE5 ions, 60K resolution, IT/MS/MS 8E3 ions, normal scan) was

used in data dependent mode at a speed of -3 seconds per cycle (Chojnowski *et al.*, 2018).

2.10. BioID data handling and *in silico* analysis

2.10.1. *BioID scoring of candidate interactors*

Candidate interactors were scored using a mixture of parameters including abundance, CRAPome prevalence (Mellacheruvu *et al.*, 2013), effect size and primary sequence coverage (Chojnowski *et al.*, 2018).

2.10.2. *GO term network assembly*

ClueGo, a plugin for Cytoscape was used for GO term analyses (Bindea *et al.*, 2009). The top ranked 200 interacting candidates were identified by BioID score, and queried through the cellular components GO term database. GO term fusion was enabled to reduce redundancy. Pathways were only shown where the significance of enrichment (pV) was ≤0.05. Analyses were restricted to returning umbrella (global) or mid-level GO terms, depending on aims of each analysis, to render network maps more readable. GO term significance graphs were constructed using Prism, and pV values were corrected for repeated sampling using Bonferroni's step down analysis built-in to ClueGO.

2.11. Data and statistical analysis

Imaging data was gathered using CellProfiler or FIJI. Densitometry of western blots was performed using FIJI. Data was processed using Microsoft Excel, while GraphPad Prism was used for graphing and statistical analyses.

2.11.1. *Densitometry of western blots*

Western blots were quantified using FIJI. Briefly, lanes were defined using the box tool, which were then transformed to display the intensity profile. The background was subtracted using the straight line tool, and the area of the curve measured using the wand tool. Intensity measurements were copied into Microsoft Excel where experimental conditions were normalised to intensities measured from the loading control.

2.11.2. _Statistical analyses_

All statistical analyses were performed within the GraphPad Prism 7/8 software. Generally, datasets were tested for Gaussian distribution using the D'Agostino-Pearson test. Where a Gaussian distribution was confirmed, data was tested using parametric methods. Alternatively, non-parametric tests were used for non-normal distribution. Where p-values are reported, $p \leq 0.05$ is considered to be statistically significant. Statistical strength was considered in the analysis of large datasets using the calculated effect size (d). Effect sizes (d) were categorised as follows (Cohen, 1988; Sawilowsky, 2009):

Effect size	d
Very small	0.01
Small	0.20
Medium	0.50
Large	0.80
Very large	1.20
Huge	2.00

Chapter 3: Investigating a putative interaction between Sun1 and dystroglycan

3.1. Introduction

Mutations in components of the LINC complex and the DGC both give rise to muscle wasting disorders. Mutations in INM LINC component Sun1 are implicated in EDMD and modify disease progression when other known EDMD mutations are present (Meinke *et al.*, 2014). Mutations in dystroglycan, the central protein of the DGC, cause muscle wasting with brain and eye involvement. In striking resemblance to one another, mutations in either the LINC complex or the DGC perturbs their respective association to cytoskeletal elements, rendering their resident membranes, either the NE or sarcolemma, more susceptible to mechanically inflicted damage (Petrof *et al.*, 1993; Lammerding *et al.*, 2004; Earle *et al.*, 2020). The molecular mechanisms behind the muscular dystrophies rooted in the LINC complex and DGC currently are thought to be different, since they arise from fragility or signalling defects emanating from the NE or sarcolemma, respectively, and indeed give rise to distinct clinical conditions (Bertini *et al.*, 2011). A common cellular phenotype observable in muscle wasting diseases associated with NL components, including AD-EDMD, LGMD-1B and HGPS (Muchir *et al.*, 2003, 2004; Bridger and Kill, 2004; Goldman *et al.*, 2004) is an aberrant nuclear morphology. Interestingly, although the nuclear functions of dystroglycan have not yet been linked to the aetiology of muscle wasting disorders, as distinct from muscular dystrophies, there is an emerging body of evidence which implicates it, in a similar manner to the LINC complex, in the maintenance of NL composition, nuclear morphology and centrosome tethering (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018).

3.1.1. *Inter-dependence of NE components*

The stability of NL components appears to be exquisitely inter-dependent on other members, and there are multiple examples of this reciprocal reliance or instability in genetic manipulation experiments or disease states. Lamin B receptor

(LBR) functions redundantly with lamin A/C during embryonic development. Genetic ablation of lamin A/C appears to be compensated for by extended expression of LBR in suprabasal keratinocytes to maintain normal chromatin organisation (Solovei *et al.*, 2013). In a pathological example, the lack of lamin A/C in mice, which results in disease reminiscent of the human AD-EDMD, results in the upregulation of Sun1 expression, an effect consistent with other lamin A/C laminopathies, since human dermal fibroblasts derived from patients with HGPS which express progerin also have increased levels of Sun1 (Chen *et al.*, 2012). Likewise, some EDMD-causing mutations in Sun1 appear to increase its stability at the NE, and concomitantly increase Nesprin-2 levels (Meinke *et al.*, 2014). Interestingly, manipulation of β-dystroglycan leads to alterations in components of the NL. For example, the knockdown or nuclear accumulation of β-dystroglycan leads to disruption in the levels and localisations of emerin and lamin B networks, however, lamin A/C remains relatively undisturbed (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). Emerin is a common target for disruption by both Sun1 mutations and β-dystroglycan manipulation. The expression of EDMD-causing Sun1 mutations, or changing levels of nuclear β-dystroglycan increases nucleus-centrosome distance in C2C12 mouse myoblasts (Martínez-Vieyra *et al.*, 2013; Meinke *et al.*, 2014; Vélez-Aguilera *et al.*, 2018). In patients with Sun1 mutations, centrosomal proteins fail to relocalise to the NE, impairing myotube differentiation (Li *et al.*, 2014; Meinke *et al.*, 2014), however, centrosomal relocalisation during differentiation in myoblasts with β-dystroglycan aberrations has not previously been tested.

3.1.2. *Common phenotypes in the laminopathies and dystroglycanopathies*

Muscle is a tissue affected by many of the laminopathies, and the exogenous expression of lamin A harbouring muscular dystrophy disease causing mutations often perturbs nuclear morphology in otherwise normal cells (Raharjo *et al.*, 2001; Lammerding *et al.*, 2004, 2006). In a lamin A/C KO mouse model, the increase in Sun1 is displayed alongside aberrantly shaped nuclei, and the genetic ablation of Sun1 in conjunction with *Lmna* improves the morphology of fibroblast nuclei (Chen *et al.*, 2012). Expression of EDMD-mutated lamin A with a disease-modifying version of

114

Sun1 also increases the likelihood of aberrantly formed nuclei (Meinke *et al.*, 2014). As Sun2 is unchanged, these studies together support the structural hypothesis for the pathological mechanisms causing the laminopathies, since progeric cells have imbalanced connections to the cytoskeleton (Chang *et al.*, 2019). Disturbances to nuclear-cytoskeletal connections can impair nuclear positioning (Chang *et al.*, 2019), but may also transmit excessive force to an already weakened nucleus, potentially causing the aberrant nuclear morphology observed in laminopathic cells. Depletion of dystroglycan levels has also been demonstrated to disrupt nuclear morphology in C2C12 myoblasts (Martínez-Vieyra *et al.*, 2013). Moreover, nuclei in fibroblasts derived from a patient with a premature stop codon in DAG1 are severely malformed (Jacobs, 2017 (PhD thesis)), as are nuclei of myofibres in mice lacking dystroglycan (Côté, Moukhles and Carbonetto, 2002). Since various fragments of the β-subunit of dystroglycan can be trafficked to the nucleus (Oppizzi *et al.*, 2008; Mathew *et al.*, 2013; Leocadio, Mitchell and Winder, 2016), these studies implicate β-dystroglycan in the maintenance of nuclear morphology. However, nuclei in the skeletal muscle of patient with a mutation in POMT1, one of the enzymes which is responsible for α-dystroglycan glycosylation, are also misshapen (Sabatelli *et al.*, 2003). β-dystroglycan appears to transit to the sarcolemma prior to nuclear localisation, while α-dystroglycan is thought not to be concomitantly trafficked to the nucleus (Gracida-Jiménez *et al.*, 2017). Therefore, it is possible that hypoglycosylation of α-dystroglycan may prevent the liberation and subsequent nuclear localisation of the β-subunit from the sarcolemma, or the reduced engagement of the ECM by α-dystroglycan may perturb mechanosignalling to mechanosensitive proteins of the nucleus through the cytoskeleton. Although the β-subunit of dystroglycan has been revealed to bind NL components in co-immunoprecipitation assays (Fuentes-Mera *et al.*, 2006; González-Ramírez *et al.*, 2008; Martínez-Vieyra *et al.*, 2013), the molecular basis for its function in regulating nuclear morphology is not clear.

Mutations in β-dystroglycan cause severe congenital muscular dystrophies, such as Walker-Walker Syndrome and muscle-eye-brain disease, which also significantly involve tissues of the central nervous system. Typically, dystroglycanopathies display a disruption in the layering of cerebral neurons; a

phenotype which can be attributed to the essential role of β-dystroglycan in forming the pial basal membrane. This structure, also known as the glia limitans, is required for the structural arrangement radial glia which guide neuronal migration during development, and the brain deletion of DAG1 induces poor distribution of granule cells (Moore *et al.*, 2002; Myshrall *et al.*, 2012). Mutations in the LINC complex and B-type lamins can also give rise to disruption of the laminary structure of central nervous system (CNS) tissues, however the pathomechanisms in this case involve intracellular defects in nuclear migration. Despite the apparent different mechanisms responsible for brain patterning, dystroglycan does appear to have other functions which overlap with those of LINC complex and B-type lamins.

Similar to dystroglycan-deficiency, the loss of Sun-domain proteins or B-type lamins results in disrupted layering of the cerebral cortex, cerebellum and the hippocampus (Zhang *et al.*, 2009; Coffinier *et al.*, 2011), but in contrast to dystroglycan-null brains, which are up to 20% larger than wild-type (Moore *et al.*, 2002), Sun1 and B-type lamin deficient cerebra and cerebella exhibit microcephaly (Coffinier *et al.*, 2011; Wang *et al.*, 2015). It may be that the lack of integrity of the glia limitans in dystroglycan-deficient brains does not provide an effective boundary for neuronal migration, which, while not reported, is potentially appropriately formed in mice lacking Sun1 and B-type lamins. As well as proper arrangement of radial glia, generation of the neuronal layering within the cerebral and cerebellar cortices and hippocampus is dependent on interkinetic nuclear movement (IKNM), which is the mitosis synchronised oscillation of nuclei in the apico-basal axis of the developing neuroepithelium, and nucleokinesis, the process by which neurons outwardly migrate along radial glia towards the pial membrane (Tsai and Gleeson, 2005). Both processes require the B-type lamins and LINC complexes. Forces that enable these nuclear movements are generated by the microtubule cytoskeleton (Keays *et al.*, 2007), and these are transmitted to the nucleus through Sun1 and Sun2-domain containing LINC complexes which anchor to lamins B1 and B2 in the NL, since cells of the central nervous system do not express lamins A/C (Zhang *et al.*, 2009; Coffinier *et al.*, 2011). Separately, the B-type lamins and Sun-domain proteins appear to act at least somewhat redundantly with their counterparts in this setting, since depletion of both lamin B1 and B2, or Sun1 and Sun2 displays a more severe

neuronal patterning phenotype than individual knockout (Zhang *et al.*, 2009; Coffinier *et al.*, 2011). However, the observation that Sun1, but not Sun2, deficient mice develop cerebellar ataxia, caused by smaller cerebella containing misplaced Purkinje cells with malformed arborisation (Wang *et al.*, 2015) indicates there are distinct functions between the Sun-domain isoforms. Finally, retinal development also depends on LINC complexes for IKNM and the migration of both rod and cone cells (Razafsky *et al.*, 2016). Similarly, B-type lamins are essential for IKNM and nucleokinesis. Lamin B1 is essential for IKNM, whereas lamin B2 is required for proper migration and lamination of photoreceptors (Yu *et al.*, 2011).

Dystroglycan-deficient CNS tissue displays several similarities to LINC or lamin mutant mice. In addition to the function of dystroglycan in proper formation of the pial basal membrane, defects in brain development in the absence of DAG1 also extend to the cerebellum which is devoid of folia while aberrantly placed neurons are also evident in the hippocampus (Moore *et al.*, 2002). Consistently, dystroglycan is highly expressed in in cerebellar Purkinje cells and hippocampal pyramidal cells (Zaccaria *et al.*, 2001). Moreover, it seems notable that neuronal layering in hippocampi is not as aberrant as that observed in the cerebral cortex (Moore *et al.*, 2002), despite a shared requirement for radial glia to scaffold the migratory routes for neurons (Hatten, 1999), hence indicating potential alternative compensatory or pathological mechanisms. Like Sun1 and B-type lamins, dystroglycan has also been implicated in the development of the retina. Unlike in other CNS tissues where dystroglycan depletion compromises basal laminae thus disrupting neuronal migration (Moore *et al.*, 2002; Myshrall *et al.*, 2012), the structure of the basement membrane in the retina appears normal. However, the expression of a dystroglycan construct mutated such that it cannot undergo autoproteolysis into α- and β-subunits appears to interfere with IKNM causing arrested neuroepithelial cell proliferation and accumulate at the basal, vitreal surface of developing retina (Schröder *et al.*, 2007). Similarities in defective synaptogenesis can also be drawn in the functions of Sun1, B-type lamins and dystroglycan. As mentioned, Sun1-depleted Purkinje cells have reduced arborisation which may contribute to cerebellar ataxia (Wang *et al.*, 2015), and neuronal circuitry of the retina appears to be impaired (Yu *et al.*, 2011). Similarly, depletion of lamin B2 leads to abnormalities in retinal

synaptogenesis (Razafsky *et al.*, 2016), while long-term potentiation in the hippocampus of dystroglycan ablated mice is significantly diminished (Moore *et al.*, 2002).

Taken together, it is plausible that dystroglycan has significant other functions which align with the roles of Sun1 and B-type lamins, aside from its essential role in radial glia endfeet for the sensation and formation of the glia limitans (Myshrall *et al.*, 2012), thus further substantiating functions for β-dystroglycan in the nucleus. Common affected neurodevelopmental processes include the regulation of nuclear movement, through nucleokinesis or IKNM to facilitate neuronal patterning, and also synaptogenesis. On the latter, how proteins of the NE like Sun1 and B-type lamins influence axonal and dendritic processes in the first instance is not clear, and if this process is common to all nervous system tissues, how dystroglycan is integrated, is unknown.

3.1.3. *Preliminary data supporting this project*

Prior to the start of this project, the Winder laboratory had found that dystroglycan-null primary fibroblasts had grossly malformed nuclei, which display a range of defects including deformation, blebbing, twisting and micronucleation (Jacobs, 2017 (PhD thesis)), implicating dystroglycan in the regulation of nuclear morphology. Nuclear functions of β-dystroglycan were explored by immunoprecipitation from a nuclear extraction of fractionated cells followed by proteomic analysis. This method, to identify proteins which co-immunoprecipitated with β-dystroglycan, revealed Sun1 as a potential interactor. In human fibroblasts, proximity ligation assays (PLA) revealed a maximum 30-40nm distance between Sun1 and β-dystroglycan using antibodies detecting endogenous protein levels, and β-dystroglycan appeared to be co-immunoprecipitated with myc-mSun1. In human fibroblasts, proximity ligation assays (PLA) revealed a maximum 30-40nm distance between Sun1 and β-dystroglycan using antibodies detecting endogenous protein levels, and β-dystroglycan appeared to be co-immunoprecipitated with myc-mSun1. Given the association of Sun1 with the NL, and previous evidence that β-dystroglycan co-immunoprecipitates with other NL components including emerin and A- and B-type lamins (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018), these

emerging data may further implicate Sun1 as part of a NE β-dystroglycan complex..
Finally, the overexpression of myc-mSun1 demonstrated that β-dystroglycan levels
could be destabilised in human fibroblasts, indicative of a functional association
similar to that observed with emerin and lamin B1 (Laredo and Winder, unpublished
observations) (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018).

3.1.4. *Research aims and hypotheses*

Preliminary data indicated that Sun1 and β-dystroglycan interact in human
fibroblasts. Various diseases can arise from mutations in components of the NL and
the DGC, which disproportionately affect muscle. Additionally, mutations in Sun1 are
known to affect the progression of muscular dystrophies, while its absence appears
to suppress lamin-associated disorders including AD-EDMD and HGPS-like conditions
in mouse models. Genetic evidence indicates that β-dystroglycan may be linked to
the NL since its mutation, or ablation result in defects to CNS lamination which to
some extent phenocopy those observed for Sun1 and Lamin B1. It was therefore
hypothesised that an interaction between Sun1 and β-dystroglycan may exist in
many tissues, including muscle, which contributes to the normal function of nuclei
in muscle tissue. Research presented in this chapter aimed to further examine an
association between Sun1 and β-dystroglycan in human myoblasts initially using
biochemical strategies. Given the established notion that some NL components,
including emerin and lamin B1, become destabilised from their peripheral nuclear
positioning upon β-dystroglycan manipulation, a second aim of research presented
here was to test whether this principle is extended to Sun1. Constructs
corresponding to nucleoplasmic parts of Sun1 and β-dystroglycan are overexpressed
in myoblasts and peripheral nuclear positioning of other NL components monitored
by immunofluorescence staining. Together, these studies would contribute to
understanding of muscle wasting diseases processes, and further molecular details
of the role of β-dystroglycan in the myonucleus.

3.2. <u>Results</u>

3.2.1. <u>*Assessing a putative interaction between β-dystroglycan and Sun1*</u>

The following series of experiments were designed to test the hypothesis that β-dystroglycan and Sun1 interact in myoblasts using recombinant *in vitro* binding assays, endogenous and overexpression immunoprecipitation assays, and PLA.

3.2.1.1. *In vitro protein purifications*

Sun1, Sun2 and β-dystroglycan proteins used in the following studies were fused to glutathione-S-transferase (GST), and were produced in *E. coli* BL21DE3 cells before purification by affinity chromatography on immobilised glutathione Sepharose. Purified proteins were assessed for purity and fragmentation by SDS-PAGE analysis and Coomassie staining. GST-fusion proteins were efficiently recovered, as evidenced by the clean GST lane (Figure 3-1) Notably, there is considerable fragmentation of the Sun1 and Sun2 GST-fusion proteins (Figure 3-1), which has been previously observed, but nevertheless used successfully to map Sun1 and Sun2 binding sites on lamin A, emerin, and Nesprin 2 (Haque *et al.*, 2006, 2010; Haque, 2011 (PhD thesis); Martina Maric, personal communication).
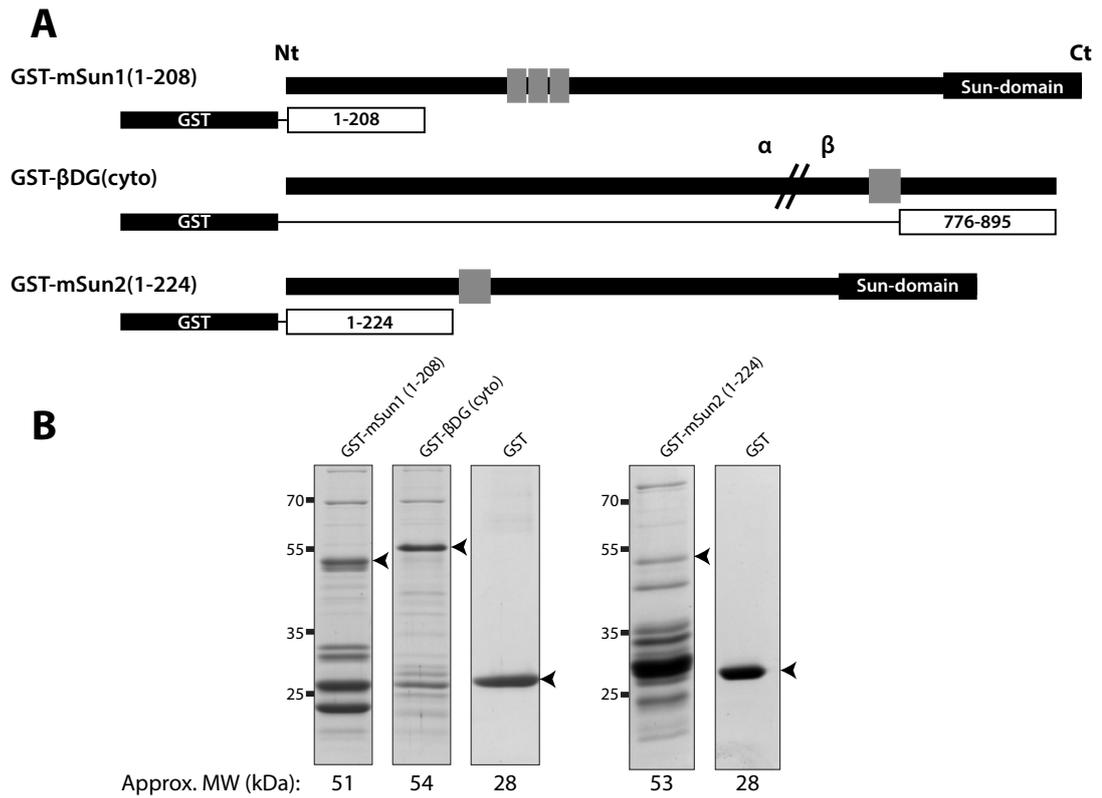
Figure 3-1 - GST-fusion protein purifications. (A) Schematic diagrams depicting GST-fusion constructs used. Full length proteins are shown above GST-fusions of nucleoplasmic regions. Amino- (Nt) and carboxy- termini (Ct) are shown on the left and right of the page, respectively, and grey boxes indicate hydrophobic transmembrane regions. (B) The GST-fused topologically nucleoplasmic domains of Sun1, Sun2 and β-dystroglycan were fused to glutathione-S-transferase and expressed in *E. coli* BL21(DE3) and purified recombinant proteins GST-mSun1 (1-208), GST-mSun2 (1-224), GST-βDG(cyto) and GST were analysed by SDS-PAGE analysis and stained using Coomassie. Predicted molecular weights (MW) are shown and corresponding protein species are indicated (arrowheads).

3.2.1.2. *In vitro interaction studies do not consistently indicate binding between β-dystroglycan and Sun1*

As discussed above (3.1.3), Sun1 was identified within the immunoprecipitated fraction of β-dystroglycan in a mass spectrometry screen. The cytoplasmic carboxyl-terminus of β-dystroglycan is known to be of importance for protein-protein interactions (Moore and Winder, 2010). If Sun1 and β-dystroglycan bind each other, it is therefore conceivable for the interaction to occur between their respective nucleoplasmic amino- and carboxyl-termini which also respects the presumed topology for dystroglycan in the NE. To address this possibility, and identify binding interfaces, a CelluSpots peptide SPOT array containing 15 amino acid peptides

A21:	IKKGVPIIFADELDD
D6:	EYTPLRDEDPNAPPY
D7:	YTPLRDEDPNAPPYQ
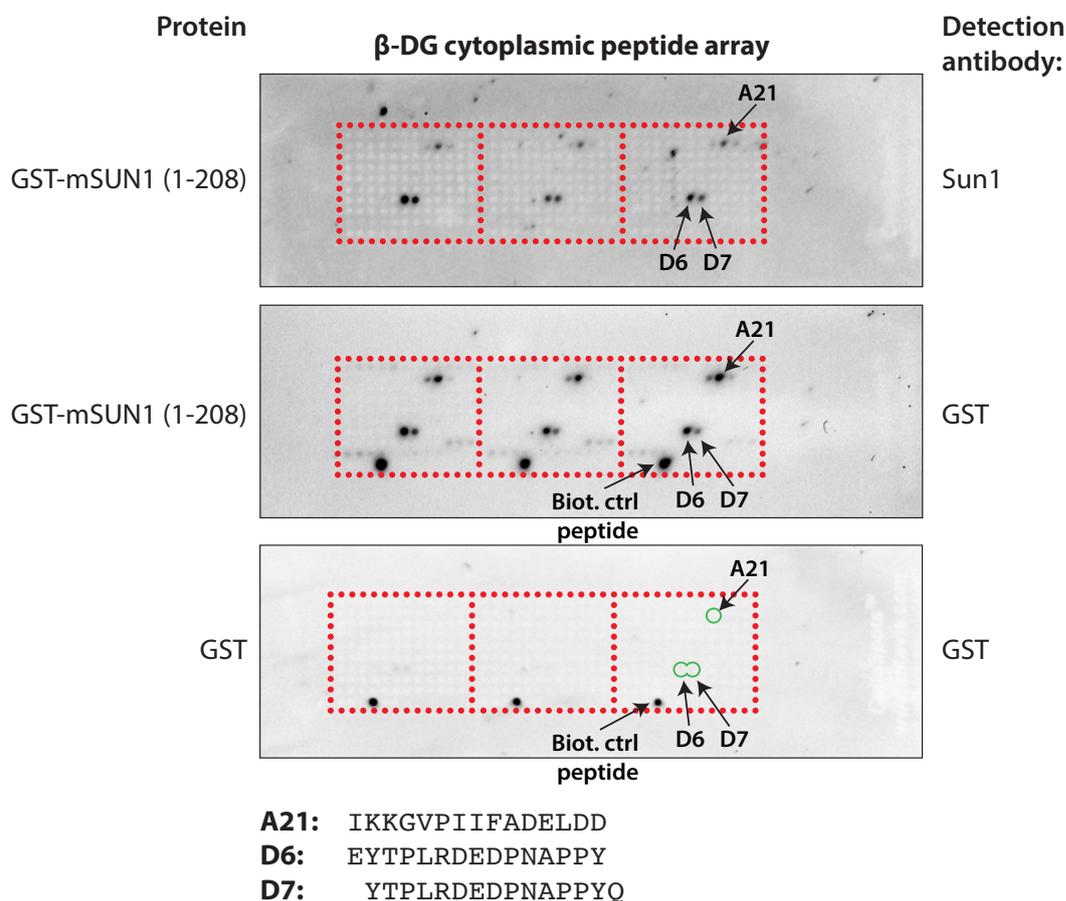
Figure 3-2 – The nucleoplasmic part of Sun1 binds to the C-terminus of β-dystroglycan. A CelluSpots peptide array synthesised with peptide 15 amino acids in length, corresponding to the carboxyl-terminus of β-dystroglycan was incubated with GST-mSun1 (1-208) or GST alone. Bound GST-mSun1 (1-208) is detected using both Sun1 and GST antibodies. Binding peptides are indicated and the sequences are shown below Three biological replicates per slide, each bounded by red dots.

corresponding to the C-terminus of β-dystroglycan, with a 14 amino acid overlap with the adjacent peptide, was probed using an extreme nucleoplasmic portion of Sun1 (aa1-208), fused to GST (GST-mSun1 (1-208)). GST-mSun1 (1-208) was either detected using an antibody raised against nucleoplasmic domain of Sun1 or GST, revealing two potential interaction motifs which were absent upon probing with an unfused GST (Figure 3-2). Probing the peptide with non-fused GST, and detection with the GST antibody shows that the GST-antibody appears to bind a biotinylated control peptide, which is not apparent when the Sun1 antibody is used (Figure 3-2).

To substantiate the finding that a potential interaction between β-dystroglycan and Sun1 occurs in their respective nucleoplasmic domains, it was hypothesised that Sun1 could be recovered from a whole myoblast cell lysate probed using the cytoplasmic domain of β-dystroglycan fused to GST (GST-βDGc). Inconsistent with the peptide SPOT array, Sun1 was not significantly recovered from the precipitated fraction, compared to the GST control (Figure 3-3). In the reverse experiment, the GST-mSun1(1-208) protein, which bound regions of the β-dystroglycan peptide array, was used to probe a myoblast whole cell lysate. Consistent with the GST-βDGc probe of cell lysate, but in opposition to the CelluSpots peptide array, antibodies directed against total β-dystroglycan (JAF1), non-Y892-phosphorylated (MDG2) or Y892-phosphorylated (pY892) (1709) (Figure 3-4) failed to detect β-dystroglycan in the precipitated fraction (Figure 3-5). To positively control this experiment, and validate that a GST-pulldown approach with the binding conditions can be used to detect protein-protein interactions of the NE, GST-mSun2 (1-224) precipitated Sun1 (Figure 3-5) which has been previously indicated, at least *in vitro* (Wang *et al.*, 2006). Taken together, evidence for an interaction between Sun1 and β-dystroglycan gleaned from *in vitro* studies is contradictory and therefore inconclusive, and so alternative methods were sought to further investigate this interaction.
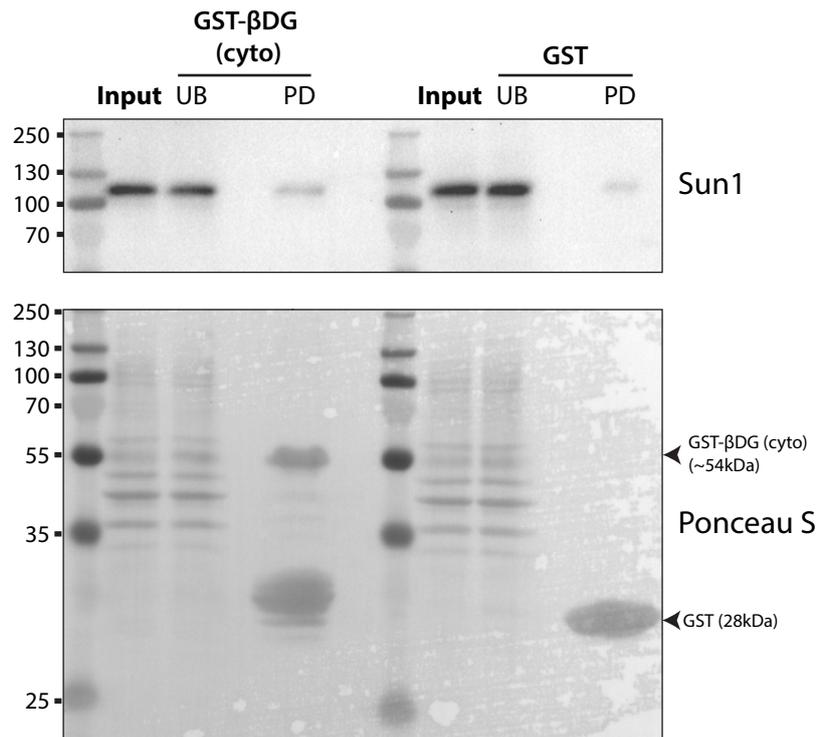
Figure 3-3 - GST-βDG(cyto) does not recover Sun1 from a whole cell lysate. Human myoblast cell lysates were incubated with gluthathione sepharose immobilised GST-βDG(cyto) or GST alone. Unbound (UB) and pull down (PD) samples were analysed by SDS-PAGE and western blotting using antibodies against Sun1. Ponceau staining indicates the bait proteins.

Given the inconsistent results gathered from *in vitro* binding assays, the hypothesis that Sun1 and β-dystroglycan are interacting proteins was further interrogated through alternative techniques which focused on demonstration of binding under endogenous or physiological settings. Therefore, β-dystroglycan from myoblast lysates was instead immobilised using different dystroglycan antibodies 1709 (pY892 β-DG), MDG2 (Y892 β-DG) and JAF1 (total β-DG) (Figure 3-4) bound to protein A Sepharose. Immunoprecipitated β-dystroglycan fractions were analysed by SDS-PAGE and western blotting, where they were probed for the presence of Sun1. Consistent with GST-pulldown experiments, Sun1 did not co-immunoprecipitate with β-dystroglycan (Figure 3-6).

124

Figure 3-4 – antibodies used for β-dystroglycan detection. Upper schematic depicts the protein structure of the dystroglycan protein. α- and β-dystroglycan are non-covalently bound at the plasma membrane. β-dystroglycan antibodies used in this study are shown below. Each antibody recognises regions at the extreme C-terminus of the cytoplasmic domain of β-dystroglycan, though the epitope differs by phosphorylation status of Y892 (position 890 in mouse sequence). Mandag2 and 1709 differentially recognise both non-phosphorylated and phosphorylated β-dystroglycan, whereas LG5 and JAF1 recognise both forms.

Figure 3-5 - GST-mSun1 (1-208) and GST-mSun2 (1-224) probing of whole cell lysates derived from myoblasts. GST-mSun1 (1-208), GST-mSun2 (1-224) and GST were immobilised on glutathione Sepharose, incubated with whole cell lysate followed by analysis by SDS-PAGE and western blotting using antibodies directed against Sun1 and different species of β-DG. MDG2 and 1709 are directed against the non-Y892 phosphorylated and pY892 versions, respectively, while JAF1 is indiscriminate.

One caveat of a co-immunoprecipitation assay is that weak or transient interactions may be perturbed during the lysis of the sample, contributing to a low recovery rate, below the detection limit of a western blot. For this reason, it was hypothesised that the overexpression of Sun1 might increase the likelihood of detecting an interaction with β-dystroglycan. To test this, a myc-tagged Sun1 construct (myc-mSun1) was electroporated into human myoblasts and concentrated using an antibody against the myc-tag. The resulting immunoprecipitated fraction was analysed by SDS-PAGE and western blot, which was probed for the presence of β-dystroglycan. Despite myc-mSun1 becoming efficiently concentrated in the immunoprecipitated fraction, β-dystroglycan was not detected (Figure 3-7), a result consistent with GST-pulldowns and endogenous co-immunoprecipitation assays.
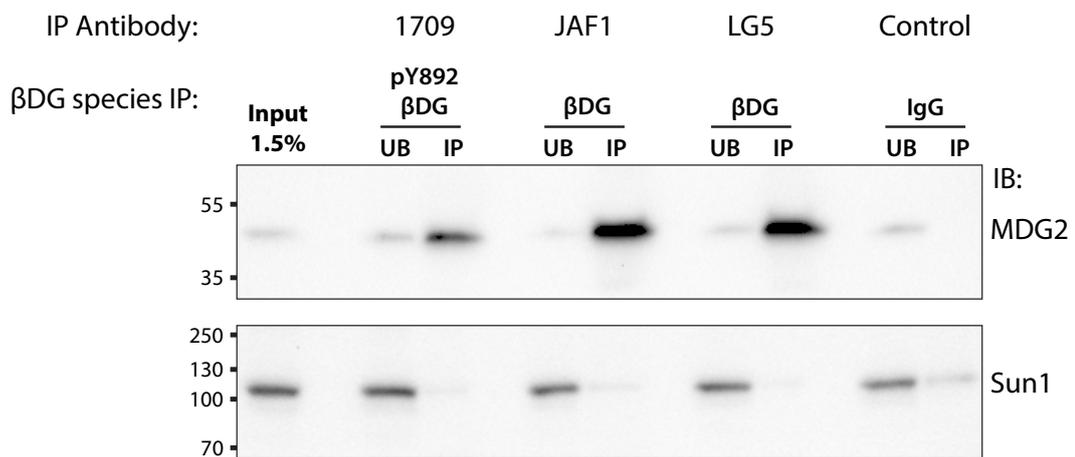
Figure 3-6 - Immunoprecipitation of β-dystroglycan does not recover Sun1. β-dystroglycan was immunoprecipitated using three antibodies directed against both Y892 and pY892 and both forms (1709, JAF1 and LG5). The immunoprecipitated (IP) and unbound (UB) fractions were analysed by SDS-PAGE and western blotting revealed the successful concentration of β-dystroglycan , however probing using a Sun1 antibody indicates its absence in the IP fraction.
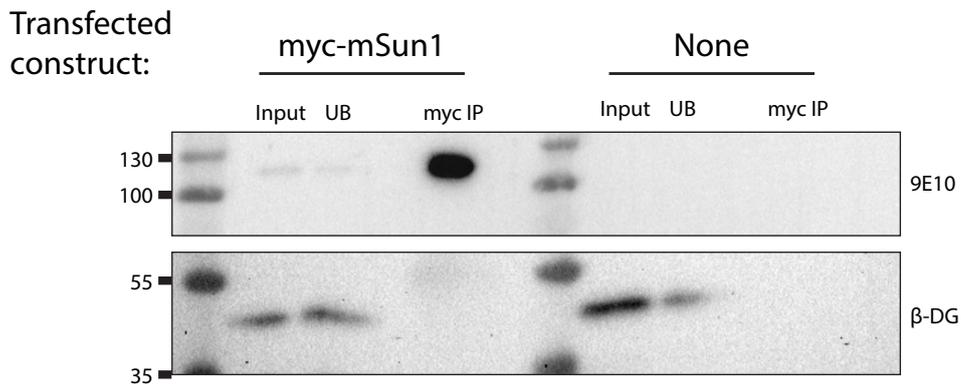
Figure 3-7 - β-dystroglycan does not co-immunoprecipitate with myc-mSun1 in human myoblasts. Human myoblasts were transfected with a myc-mSun1 construct by electroporation, which was immunoprecipitated 48 hours later using anti-myc antibody 9E10. 1.5% of the input material was analysed by SDS-PAGE alongside a 1.5% unbound (UB) and myc-immunoprecipitated (IP) fractions. The resulting western blot was probed using 9E10 revealing successful expression and concentration of myc-mSun1, however probing using MDG2 reveals an absence of β-dystroglycan in the IP fraction.

Biochemical assays are commonly used to reveal protein-protein interactions, however, analyses of the lamins and lamin-binding proteins are complicated by their relative insolubility, resulting in the requirement to balance the stringency of lysis buffers for membrane extraction of proteins while preserving endogenous protein-protein interactions. As such, it is possible that an interaction between Sun1 and β-dystroglycan was obscured during the previous experiments, particularly if it is weak or transient in nature. To negate the limitations of conventional biochemical strategies, a Sun1 and β-dystroglycan interaction was investigated *in situ* by proximity ligation assay (PLA). PLA is performed similarly to immunofluorescence microscopy, however, specialised secondary antibodies are conjugated to complementary oligonucleotides which, when in proximity of less than 40nm (Alam, 2018), can be ligated and amplified. The complex DNA structure which results from the rolling-loop amplification is then intercalated with fluorescent dye and detected by fluorescence microscopy. As a positive control for the assay, an interaction between Sun1 and lamin B1 (Nishioka *et al.*, 2016) could be confirmed (Figure 3-8A).
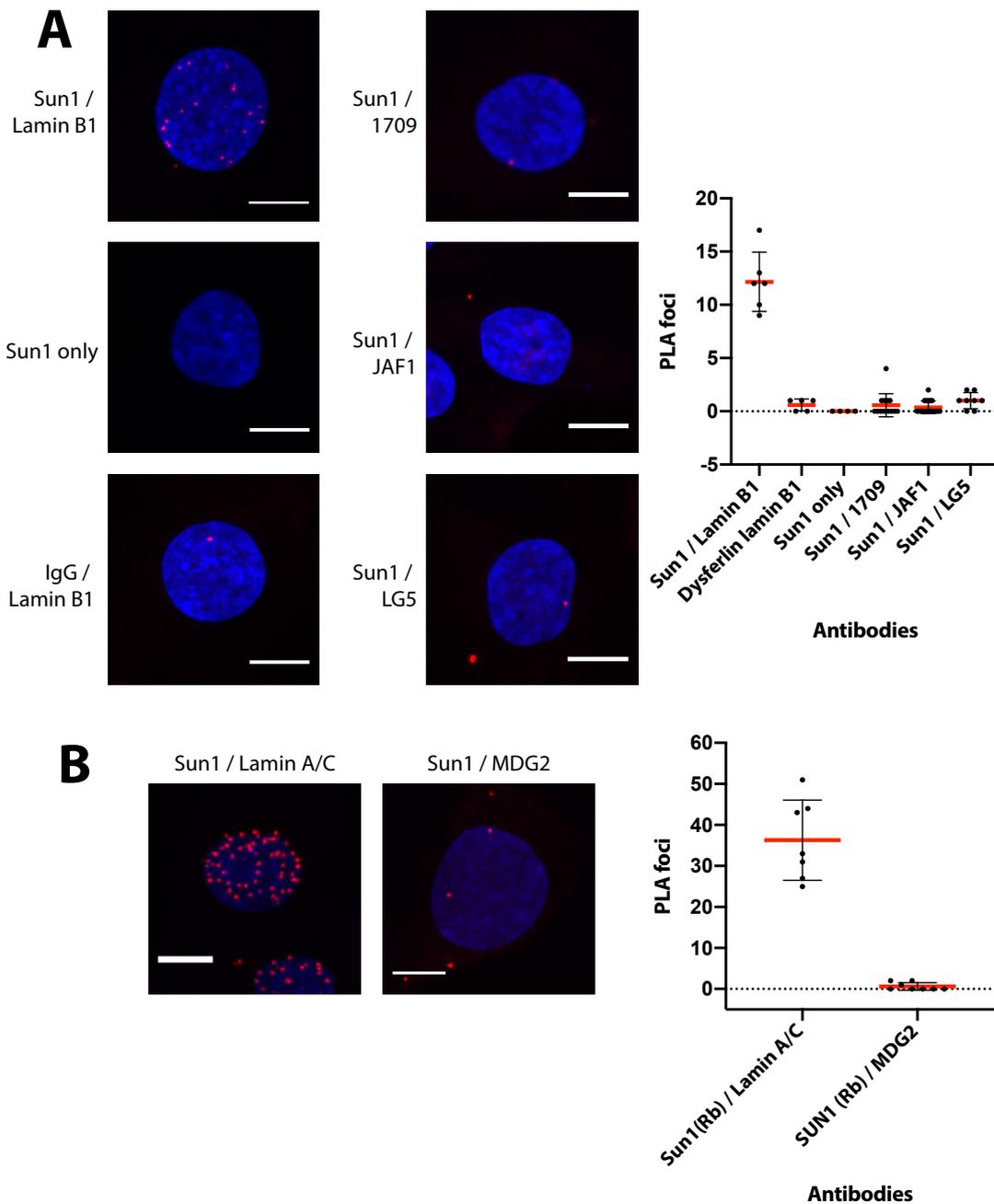
Figure 3-8 - Proximity ligation assay between Sun1 and β-dystroglycan. (A) Antibodies against β-dystroglycan raised in rabbit were used in conjunction with murine Sun1 antibody. As a positive control for the assay, a robust interaction between Sun1 and lamin B could be detected. (B) MDG2, a β-dystroglycan raised in mouse were used in conjunction with a rabbit Sun1 antibody. The interaction between Sun1 and lamin A/C used as a positive control. Number of foci per cell is graphed in each instance. Scale bar = 10μm.

Primary antibodies against β-dystroglycan determined its established cytoplasmic and perinuclear localisation in human myoblasts (Figure 3-16), however in contrast to the positive control, no PLA signal could be detected when using antibodies against β-dystroglycan raised in rabbit, including 1709, LG5 and JAF1 which are pY892-, non-pY892-specific and non-specific for the phosphorylation status, respectively (Figure 3-8A). Consistently, signal from a PLA performed using the anti-β-dystroglycan mouse monoclonal antibody, MDG2, and a Sun1 antibody raised in rabbit could also not be detected, while the established interaction between Sun1 and lamin A/C was strongly identified (Figure 3-8B).

Despite the preliminary *in vitro* confirmation of an interaction between the amino- and carboxy-termini of Sun1 and β-dystroglycan, respectively, subsequent assays using human myoblasts failed to recapitulate and substantiate these findings. At least in myoblasts, an interaction between Sun1 and β-dystroglycan remains elusive, and there is insufficient evidence to indicate that it is present. Therefore, it was alternatively hypothesised that an interaction between Sun1 and β-dystroglycan in muscle is developmentally regulated and may exist upon maturity.

3.2.2. *Sun1 becomes alternatively spliced upon myotube differentiation, but an interaction with β-dystroglycan also cannot be detected*

Fragmentation and nuclear localisation of carboxy-terminal β-dystroglycan is known to change with increased cell density in some cell lines (Mathew *et al.*, 2013; Mitchell *et al.*, 2013; Leocadio, Mitchell and Winder, 2016). Moreover, dystroglycan becomes assembled into a mature DGC upon differentiation into myotubes from myoblasts, where the UGC is predominant (Cullen *et al.*, 2001). Clearly, a driver of differentiation is the re-purposing of existing proteins and complexes for new functions specific to the emerging cell type. If β-dystroglycan and Sun1 are functionally associated, changes in Sun1 might also be expected upon increased confluence or myotube differentiation. In precedent for the concept that LINC complexes may change and function alternatively in a developmentally-specific manner, Nesprin-1α has been found to be crucial for nucleating the microtubule

organising centre (MTOC) in myotubes, a requirement for subsequent nuclear positioning in these syncytial cells (Gimpel *et al.*, 2017). To test the hypothesis that Sun1 levels or species change in a developmentally-dependent manner, human myoblasts were seeded at increasing density, and differentiated into myotubes by incubation in low serum medium for 4 days, before maintenance in growth or differentiation medium for a further 24 hours (Figure 3-9A). SDS-PAGE and western blotting analysis showed that while Sun1 levels were unchanged in response to cell density, differentiation into myotubes induced a faster migrating band (Figure 3-9B). Two possibilities were identified to explain the presence of this lower molecular weight Sun1 species; either it arises during proteasomal degradation of Sun1 upon myotube differentiation, or it signifies a genuine, functional isoform of Sun1 specific to mature muscle. To address the first hypothesis, chemical inhibition of the proteasome using MG132 was used, on the premise that the faster migrating Sun1
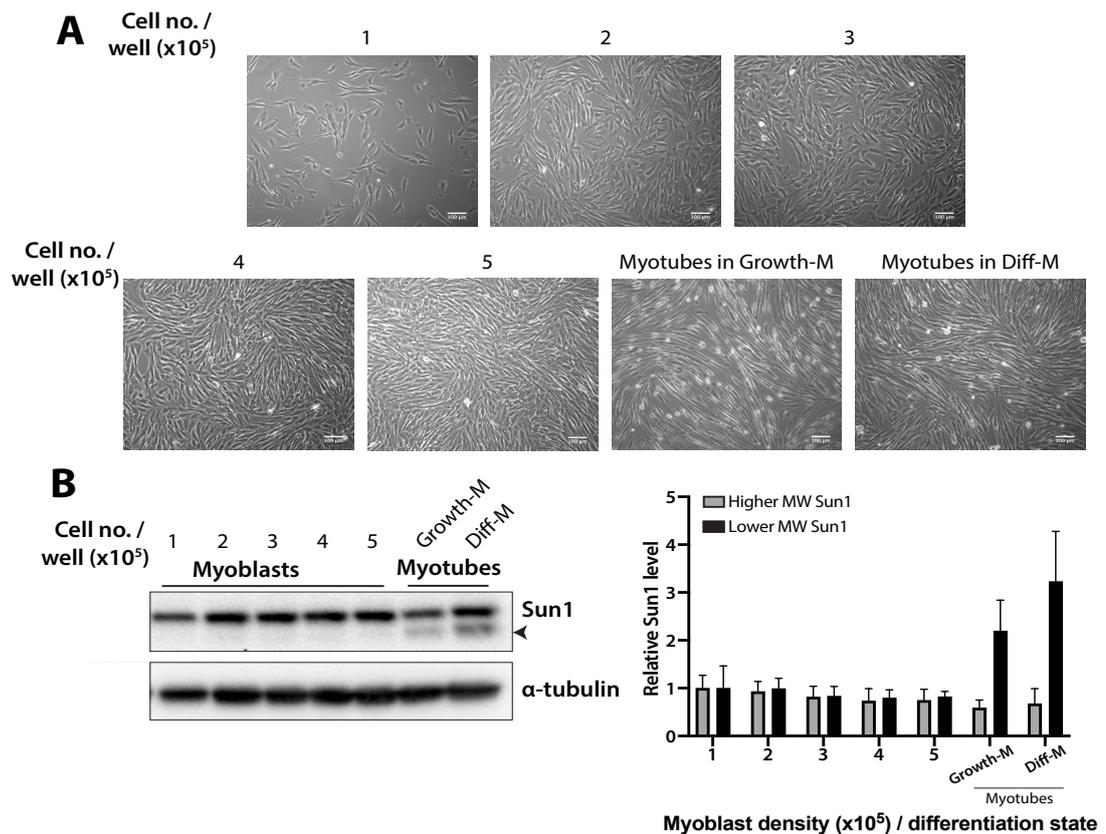
Figure 3-9 - A faster migrating Sun1 species is evident upon differentiation of myoblasts into myotubes. (A) Bright-field images showing increasing cell density of myoblasts and upon myotube differentiation. (B) Sun1 probing of western blot of myoblast and myotube samples separated by SDS-PAGE. The graph on the right hand side shows quantified levels of Sun1 species displayed relative to levels of the indicated Sun1 species in myoblasts at a density of $1x10^5$ cells/well.

band would be depleted, while levels of the larger isoform would increase should proteasomal degradation be responsible for its existence. Evidently, the smaller molecular weight Sun1 species persisted throughout MG132 treatment, while there was also no significant increase in the change in the amount of the larger version (Figure 3-10). Asides from investigating other degradative pathways, these data further substantiate the possibility that the smaller isoform may instead be a genuine Sun1 protein. Interestingly, recent publications have indicated that Sun1 is extensively spliced in a tissue specific manner (Göb *et al.*, 2011), and splice isoforms may have opposing functions within a cell and throughout myogenesis (Nishioka *et al.*, 2016; Loo *et al.*, 2019). To investigate whether alternative splicing occurs during human myoblast differentiation into myotubes, reverse-transcription PCR was performed to amplify the variable region of Sun1 in each condition, using established primers which span the junctions between exons 4/5 and 11/12 (Nishioka *et al.*, 2016). This experiment did not reveal any novel products in either condition (Figure 3-11). Not all cultured myoblasts are competent for differentiation into myotubes, so the resulting RNA extract from a myotube culture is actually a mixture, which probably contributes to the persistence of some products. Nevertheless, this semi-quantitative method does reveal clear differences in the relative abundance of Sun1
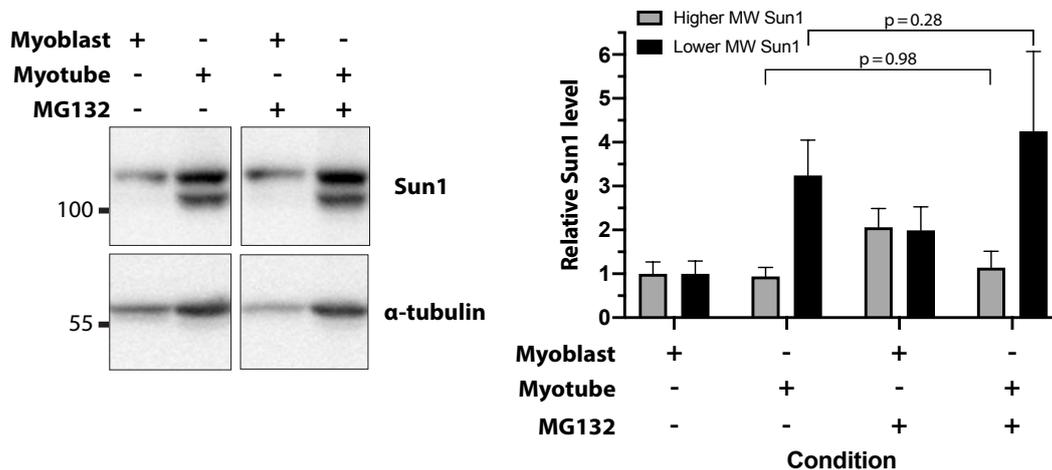
Figure 3-10 - Lower molecular weight Sun1 band persists upon proteasomal inhibition by MG132. Human myoblasts and myotubes were treated with 15 µM MG132 for 24 hours before lysate harvesting. Proteins were separated by SDS-PAGE, western blotted and membranes probed for Sun1. Levels of Sun1 species were quantified and displayed relative to myoblast levels of that Sun1 species in the absence of MG132. Western blots are representative of four biological replicates and statistical significance was determined by two-way ANOVA with Tukey's multiple comparison test.

isoforms between myoblasts and myotubes (Figure 3-11). Although the PCR products were not sequenced, it seems likely that one of these gives rise to the lower MW band, although it is not known which isoform identified by RT-PCR corresponds to the lower molecular weight SUN1 band seen in western blotting (Figure 3-9). In addition to these findings in human myoblasts, the alternative splicing of Sun1 upon differentiation of muscle progenitors was also recently confirmed to occur in mouse muscle and the C2C12 mouse myoblast cell line (Loo *et al.*, 2019), indicating that this process is a conserved phenomenon during mammalian muscle development. Interestingly, Drosha was found to specifically bind to a myotube Sun1 isoform lacking exon 7 (Loo *et al.*, 2019). Presumably, this is just one example of an alternative interactome for Sun1 which can comprise diverse functions (Nishioka *et al.*, 2016; Loo *et al.*, 2019) and can be established by the creation of new binding sites through amalgamation of previously distinct peptide sections. Until now, myoblasts were used since they represent a valuable model of muscle, displaying a number of specific markers in which to contextually place tissue-specific interactions, while being mitotic makes them convenient to handle. However, they do not accurately recapitulate mature muscle since they are mononucleated, migratory and non-contractile. Given that an interaction between Sun1 and β-dystroglycan could not be identified in myoblasts, and that different Sun1 isoforms appear to possess different functions and binding partners which may contribute to
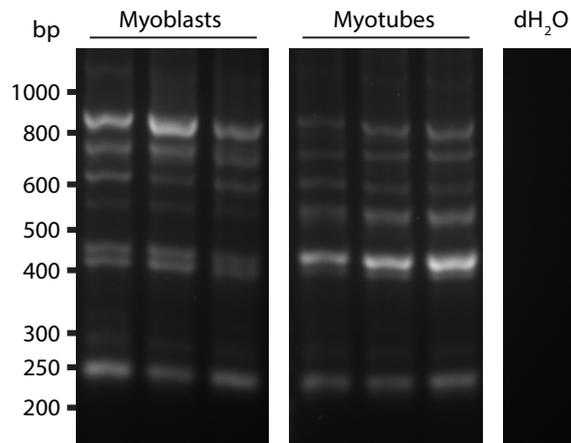
Figure 3-11 - Reverse transcription PCR of the variable region of Sun1 reveals different relative amount of splice isoforms. Myoblasts were cultured in growth medium or differentiated into myotubes before total RNA extraction, reverse transcription and amplification of the variable region of Sun1 mRNA. Products were then run on 0.8% agarose gel and stained with ethidium bromide revealing splice variants. Each lane represents a biological replicate.

muscle differentiation mechanisms (Nishioka *et al.*, 2016; Loo *et al.*, 2019), it was reasoned that an alternative isoform of Sun1, which is only expressed in the myotube, might bind β-dystroglycan. To test this hypothesis, β-dystroglycan was immobilised from a culture of myotubes using the antibody MDG2, and the immunoprecipitated fraction was probed for the presence of Sun1. The co-immunoprecipitation assay revealed that whereas the smaller Sun1 isoform could be detected in the input, it was not present in the immunoprecipitated fraction (Figure 3-12). To ensure that the interaction was not disrupted by the lysis conditions required for the efficient membrane extraction, PLA was again employed as an *in situ* technique to detect binding. Consistent with the results garnered from co-immunoprecipitation and GST-pulldown assays, a suite of antibodies against β-dystroglycan used in conjunction with Sun1 did not produce a PLA signal, despite the interaction between Sun1 and emerin (Haque *et al.*, 2010) being robustly detected used to positively control the assay (Figure 3-13).
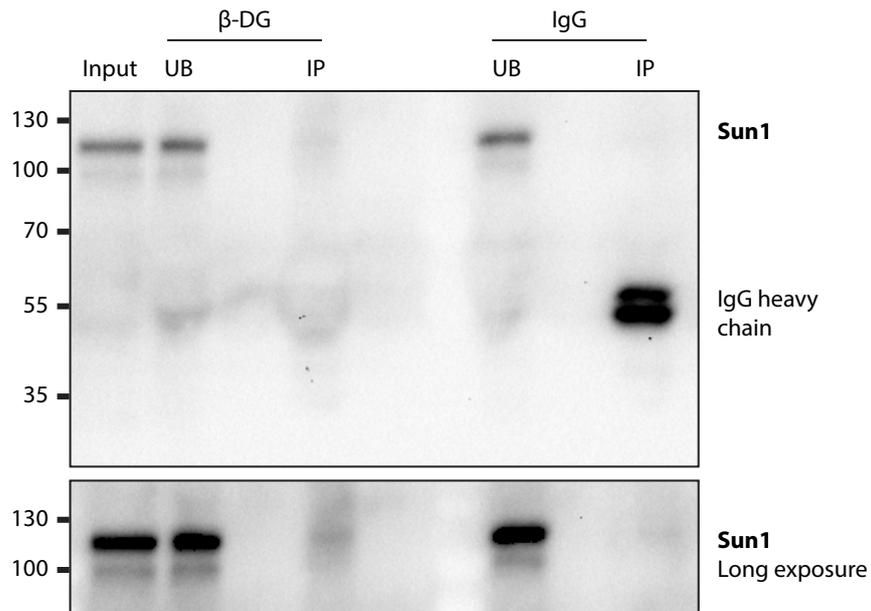
Figure 3-12 - An interaction between β-dystroglycan and Sun1 cannot be detected in myotubes by co-immunoprecipitation. Myotubes were differentiated for 5 days prior to lysis and β-dystroglycan immunoprecipitation. The input, MDG2 immunoprecipitated (IP) β-dystroglycan and unbound (UB) fractions were analysed by SDS-PAGE and western blotting for the presence of Sun1. Increased western blot exposure revealed that Sun1 could be detected only in the input and unbound samples.

3.2.3. _Assessing a functional link between β-dystroglycan and Sun1_

Previous work indicates that manipulation of the levels of β-dystroglycan in the nucleus leads to instability of other nuclear components. In the nucleolus, B23, fibrillarin, and UBF proteins are destabilised by depletion of β-dystroglycan (Azuara-Medina _et al.,_ 2019). In the NE, β-dystroglycan overexpression or forced nuclear accumulation by prevention of its nuclear export affects the expressed levels and localisation of emerin and lamin B1 (Vélez-Aguilera _et al.,_ 2018). Similarly, depletion of β-dystroglycan by RNA interference also perturbs emerin and lamin B1 levels in the NL, among other nuclear components (Martínez-Vieyra _et al.,_ 2013). Given that malformations in the brains of mice deficient in Sun1, lamin B1 and β-dystroglycan appear to phenocopy (Moore _et al.,_ 2002; Coffinier _et al.,_ 2011; Wang _et al.,_ 2015), and that β-dystroglycan seems to be functionally associated with NL components at

a cellular level (Martínez-Vieyra *et al.*, 2013), it was postulated that β-dystroglycan overexpression might similarly affect the levels and localisation of Sun1 at the NL. Topologically, the carboxy-terminus of β-dystroglycan is nucleoplasmic facing, and has been previously demonstrated to localise to the nucleus and mediate transcriptional changes (Mathew *et al.*, 2013). Therefore, a cytoplasmic β-dystroglycan GFP-tagged construct (GFP-βDGc), which retains its NLS, but lacking an NES was used to investigate the hypothesis that β-dystroglycan overexpression disturbs Sun1. Expression of GFP-βDGc in human myoblasts and staining of cells with antibodies against Sun1 revealed that there was no difference in the levels of Sun1, while the localisation of Sun1 in the nucleus was also observed to be similar to that where the GFP-containing vector had been transfected alone (Figure 3-14). To investigate whether the cytoplasmic fragment of β-dystroglycan construct disrupts the levels and localisation of emerin, as has been reported previously (Vélez-Aguilera *et al.*, 2018), either GFP alone or GFP-βDGc were similarly transfected into human myoblasts, and emerin distribution and levels were assessed by confocal
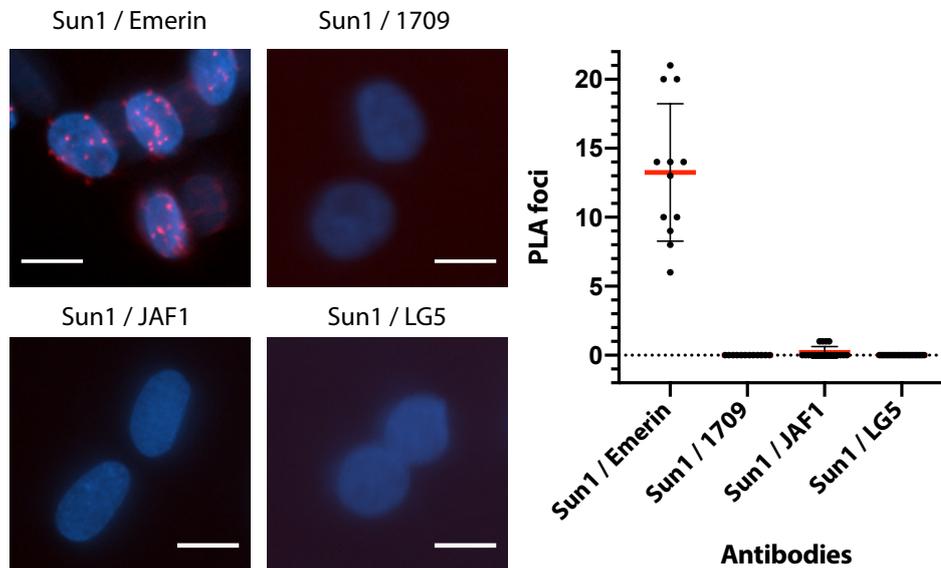
Figure 3-13 - PLA to detect an interaction between Sun1 and β-dystroglycan in myotubes. Myoblasts were differentiated into myotubes over 4 days, and co-stained with antibodies against Sun1 and emerin as a positive control, or β-dystroglycan (1709, JAF1 and LG5). Samples were analysed by fluorescence microscopy, and the number of PLA foci per cell in each condition quantified in the graph. Scale bars = 10 μm.

microscopy. Contrary to published work (Vélez-Aguilera *et al.*, 2018), overexpression of β-dystroglycan revealed no significant difference in the levels of emerin at the NE, and nor was emerin mislocalised in comparison to cells expressing GFP alone (Figure 3-15). Interestingly, previous experiments used a full β-dystroglycan construct, including the transmembrane and ectodomains (Vélez-Aguilera *et al.*, 2018). Results here that show the cytoplasmic fragment of β-dystroglycan has no effect on emerin might therefore indicate that the transmembrane or ectodomain regions may be crucial for the functional association with some NE components.
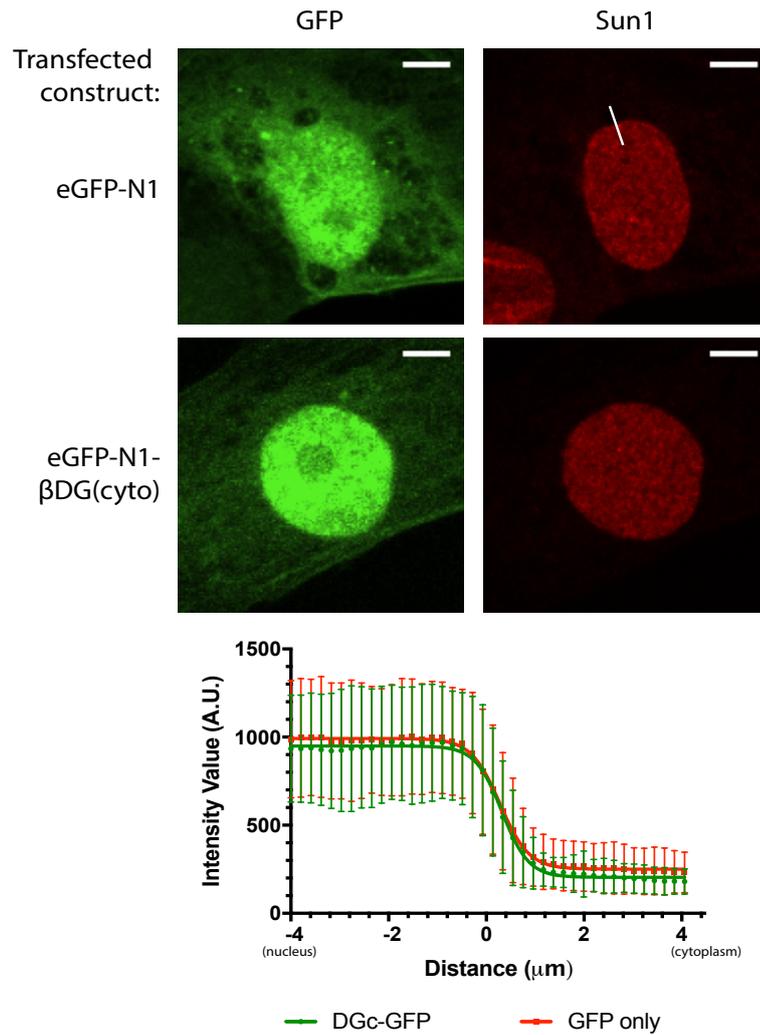
Figure 3-14 - Overexpression of the cytoplasmic fragment of GFP-tagged β-dystroglycan (GFP-βDGc) does not affect Sun1 levels or localisation. Either GFP-βDGc or GFP alone were electroporated into human myoblasts 48 hours prior to fixation and stained with antibodies against Sun1. The graph depicts Sun1 levels across nuclei which were quantified by plotting 6 radial line profiles perpendicular to the nuclear border (one shown as example) for each nucleus and averaged. (GFP-βDGc, n=30; GFP, n=20). Scale bar = 10 μm.
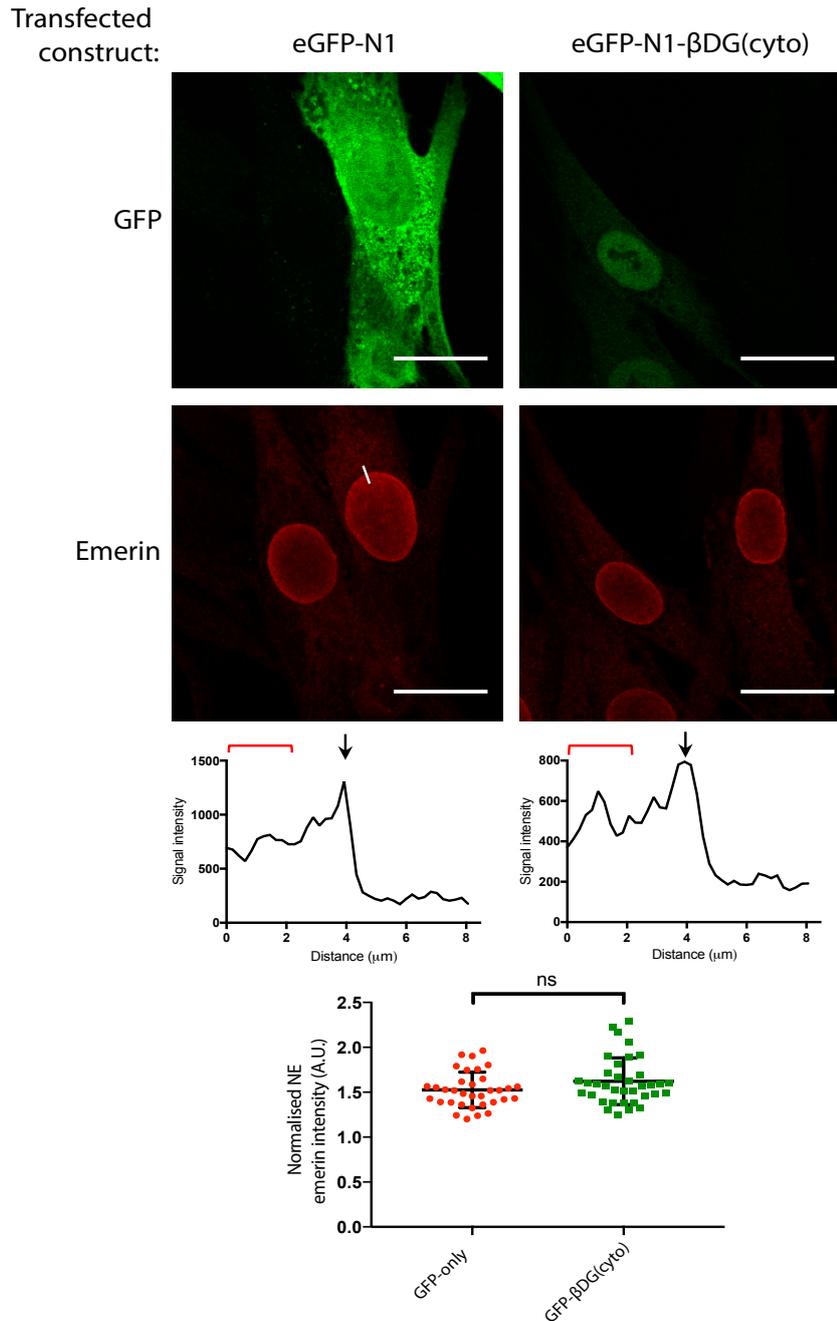
Figure 3-15 - Overexpression of the cytoplasmic fragment of GFP-tagged β-dystroglycan (GFP-βDGc) does not affect emerin levels or localisation at the NE. Either GFP alone or GFP-βDGc were electroporated into human myoblasts 48 hours prior to fixation and staining with antibodies raised against emerin. Samples were imaged by confocal microscopy to obtain optical sections through the nucleus, and emerin levels were quantified by plotting and averaging 6 radial line profiles across the NE (indicated by white line, one shown as example). The NE level of emerin was normalised to the average nucleoplasmic intensity (bracketed in example line profile), and plotted in the graph. Scale bar = 20 μm; n = at least 34 nuclei per condition; unpaired, two-tailed t-test p=0.1155.

3.3. Discussion

The results presented in this chapter aimed to further investigate and confirm the existence of an interaction between Sun1 and β-dystroglycan within the context of muscle. Precedent for an interaction between the proteins is given by various phenocopying in the neuronal patterning defects and abnormalities in synaptic function observed in Sun1 and dystroglycan depleted nervous tissue. An interaction between Sun1 and β-dystroglycan had previously been identified in fibroblasts during experiments performed in the Winder laboratory. However, combined *in vitro* and cell biological approaches aimed at demonstrating an interaction or association between Sun1 and β-dystroglycan in the context of muscle tissue presented here resulted in conflicting evidence. The nucleoplasmic part of Sun1 fused to GST reproducibly binds two regions of the carboxy-tail of β-dystroglycan according to SPOT arrays, in a conformation which respects their expected topology. However, this result could not be recapitulated using alternative strategies in a more physiological setting where a myoblast cell lysate was probed using the same GST-mSun1 protein. Co-immunoprecipitation assays have been extensively used to reveal the interactions of β-dystroglycan with components of the NL (Fuentes-Mera *et al.*, 2006; Martínez-Vieyra *et al.*, 2013), however attempts to identify the proposed binding with Sun1 in this manner were unsuccessful. PLA, an *in situ* approach thought to better preserve interactions between membranous proteins, consistently did not reveal Sun1-β-dystroglycan binding. Investigations also revealed that in human muscle myoblasts, Sun1, like in mouse muscle cells, becomes alternatively spliced upon differentiation in culture. Nevertheless, a Sun1 interaction with β-dystroglycan did not appear to be developmentally regulated, as evidenced by co-immunoprecipitation and PLA experiments. Consequently, a functional association between β-dystroglycan and Sun1 could also not be demonstrated, despite published works indicating that β-dystroglycan interacts with various other NL components, and manipulation of β-dystroglycan levels appears to disrupt lamin B1 and emerin.

3.3.1. *Criticisms of the study design*

One limitation of the present study is the choice of constructs used for β-dystroglycan and Sun1, which were restricted to the cytoplasmic and nucleoplasmic regions of each, respectively. The rationale behind this firstly was topology, and the knowledge that the carboxyl-tail of β-dystroglycan has a disordered structure, which is implicated as a region rich in protein-protein interaction sites (Moore and Winder, 2010). Indeed, many interactors with seemingly disparate functions have been mapped to this region of β-dystroglycan, including numerous SH2-domain, SH3-domain proteins, signalling and structural proteins (Moore and Winder, 2010). This approach, however, cannot address further conceivable interactions within the transmembrane region or ecto-domain of β-dystroglycan. Nevertheless, these possibilities are accounted for through other methods to assess binding between β-dystroglycan and Sun1, since antibodies against the endogenous protein were used. However, functional studies, which aimed to reveal a non-direct association between β-dystroglycan and Sun1, would benefit from a more comprehensive assessment of the entire β-dystroglycan protein, in addition to the promiscuous C-terminal region. It would be predicted that the results of such an experiment, where αβ-dystroglycan, or full β-dystroglycan protein is expressed, would similarly alter emerin levels and localisation, as has been previously observed (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). This could also focus in on important regions of β-dystroglycan necessary for NL integration which have not been previously identified. Interestingly, new data suggests that an identified interaction between β-dystroglycan and emerin might reside within their respective transmembrane domains, although this study does lack some controls to exclude the possibility of non-specific binding between these hydrophobic regions (Gómez-Monsivais *et al.*, 2020).

During this current study, the potential interaction between Sun1 and β-dystroglycan could only be demonstrated through the *in vitro* binding to a CelluSpots peptide array. All other approaches including GST-pulldowns, co-immunoprecipitation and proximity ligation assays failed to substantiate this finding. There are at least two possibilities behind this: first, that the interaction is of low

affinity or transient, meaning it is below the detection limit of other methods; and second, one or more of the breakdown products from the GST-mSun1 purification, which may differ structurally and biochemically from the endogenous nucleoplasmic region, non-specifically binds to the identified β-dystroglycan peptides. Considering the second possibility, it is noteworthy that other Sun1 binding partners have been defined using this construct including emerin, Nesprin-2β and lamin A (Haque *et al.*, 2010). However, it is possible that binding between Sun1 and NL components is of higher affinity, and thus readily detected. Although increasing the amount of detergent or ionic strength would lead to an increased membrane recovery of the relatively insoluble LINC complex and NL components, it is caveated by the increased likelihood of destabilising lower affinity endogenous interactions. Like previous findings in the Winder laboratory, purportedly identifying a Sun1-β-dystroglycan interaction, immunoprecipitation experiments presented here utilised the standard RIPA buffer for cell lysis (Laredo and Winder, unpublished observations). Other reports indicating β-dystroglycan-NL interactions instead use a lysis buffer with physiological ionic strength and 1% Triton X-100 (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). Given the insoluble nature of NL components, it therefore not clear that the NL proteins examined in these studies are sufficiently solubilised in these buffers prior to biochemical assays, which may lead to identification of false positive interactions. Further work in this area might systematically address questions over NL protein solubilisation and their interactions with β-dystroglycan using lysis and immunoprecipitation binding buffers of increasing stringencies. PLA was hypothesised to circumvent solubility problems; however, this assay depends on quality primary antibodies and robust nuclear rim staining using antibodies against β-dystroglycan was notably absent in human myoblasts (Figure 3-16). Previously, β-dystroglycan has been reported to co-localise with components of the NL both at the endogenous level in the C2C12 mouse myoblast cell line (Martínez-Vieyra *et al.*, 2013), however a review of this work does reveal some technical issues compromising conclusions drawn from the published experiments. Double immunofluorescence staining used the anti-β-dystroglycan goat polyclonal antibody C-20 (SCBT) which was detected using chicken anti-goat conjugated to Alexa Fluor

594, however, the secondary antibody used to co-label NE components was also raised in goat. Double staining for β-dystroglycan and the nuclear compartment markers were completed sequentially, however the absence of appropriate staining conditions to control for the likely possibility of cross-recognition significantly impedes meaningful interpretation. Indeed, further panels depicting C2C12 cells co-immunolabelled for β-dystroglycan and nucleolar, rather than NL components, instead do not consistently detect a NE β-dystroglycan stain (Martínez-Vieyra *et al.*, 2013). Mouse cells are commonly used to model human processes, and so it was not anticipated that there would be significant differences in the subcellular distribution of β-dystroglycan in mouse and human myoblasts. However, the prospect that β-dystroglycan localisation changes in a cell-specific manner cannot be completely excluded since β-dystroglycan does appear to accumulate in the nuclei of prostate cancer cells compared with healthy controls (Mathew *et al.*, 2013).

3.3.2. *Re-assessment of preliminary data*

Notwithstanding the limitations identified through the current study, the strength of preliminary data gathered on the potential interaction between β-dystroglycan and Sun1 is also worth considering. Previous experiments which indicate that β-dystroglycan co-immunoprecipitates with an exogenously expressed myc-mSun1 construct would benefit from isotype-matched IgG control immunoprecipitation, to distinguish 43 kDa β-dystroglycan from 50 kDa IgG antibody heavy chain which is present under reducing conditions and sometimes visible depending on detection antibody specificity. Similarly, the rigour of conclusions drawn from initial PLA would be increased if the assay was performed using antibodies against targets thought to be unrelated to Sun1 or β-dystroglycan. Finally, functional studies, where β-dystroglycan is proposed to be decreased upon myc-Sun1 overexpression would benefit from quantification, since some imaged cells do not appear to differ significantly from transfected cells.

3.3.3. *Biological considerations influencing a putative Sun1-β-dystroglycan interaction*

Since this study could not reliably confirm an interaction between Sun1 and β-dystroglycan using multiple approaches, from the results, it is also not possible to disprove the original hypothesis as the experiments performed were not exhaustive. It is therefore worth considering other biological processes which might affect the detection of such an interaction.

It is becoming clear that different cell types possess different constellations of NET proteins; many of which have yet unresolved functions (Schirmer *et al.*, 2003; Korfali *et al.*, 2010, 2012; Wilkie *et al.*, 2011). However, muscle-specific NETs have been identified and some appear to be responsible for nuclear attachment to the specialised cytoskeleton of myofibres and sufficient to influence chromatin organisation and gene expression programmes (Wilkie *et al.*, 2011; de las Heras *et al.*, 2017). Notably, preliminary experiments which initially indicated that Sun1 and β-dystroglycan were interactors were carried out in human fibroblasts, rather than myoblasts. While Sun1 and β-dystroglycan are ubiquitously expressed proteins, it is not known whether their potential interaction is dependent upon the presence or absence of other, tissue- or cell-specific NETs which may be differently regulated in myoblasts (Chen *et al.*, 2006). To further investigate this, it would be worth revisiting the preliminary studies in fibroblasts to confirm the interaction observations, and, should differences between fibroblasts and myoblasts be observed, a proteomic study to identify differentially expressed NETs in myoblasts deficient in either Sun1 or β-dystroglycan might reveal potential adaptors by searching for proteins disrupted in each cell-type.

Recently it was revealed that there are differential functions of Sun1 splice variants both in migrating cells and throughout myogenesis (Nishioka *et al.*, 2016; Loo *et al.*, 2019). Alternative splicing of the variable region located in the nucleoplasmic domain of Sun1 can create binding sites for novel protein interactors, however, investigations using immunoprecipitation and PLA did not detect an endogenous association between Sun1 and β-dystroglycan. Interestingly, while the full-length isoform of Sun1 is predominantly expressed in myoblasts, upon

144

differentiation into myotubes, a wider variety of Sun1 isoforms become expressed (Loo *et al.*, 2019). Noticeably, the antibodies against Sun1 used during this study only identified one other isoform, whereas RT-PCR and others have shown many more. It is therefore not inconceivable that the antibodies do not have pan-Sun1 specificity; indeed, the antibody's recognition epitope may be removed in splice isoforms. It is for this reason that future work might consider controlling for this by sequentially assessing the potential interaction with each Sun1 isoform. This could be performed by *in vitro* binding assays between recombinant proteins, and physiological context could be acquired from expression of tagged Sun1 isoforms in mammalian cells followed by either co-immunoprecipitation or PLA.
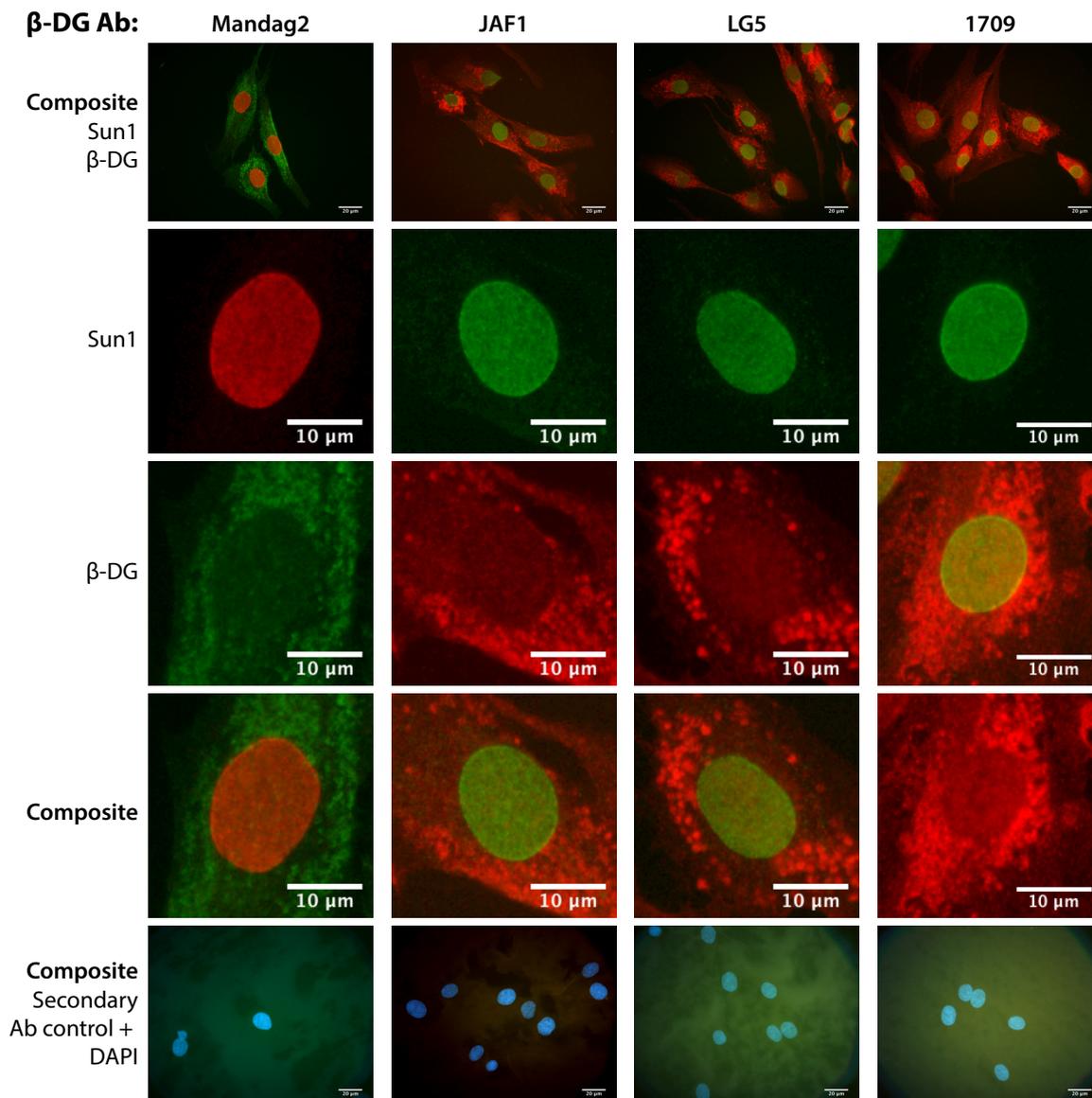
Figure 3-16 - Localisation of β-dystroglycan in human myoblasts. Human myoblasts were stained using antibodies raised against Sun1 and β-dystroglycan to assess subcellular localisation. Mandag2 and LG5 recognise Y892 β-dystroglycan, 1709 recognises phospho-Y892 β-dystroglycan, and JAF1 recognises both forms. Mandag2 is the only β-dystroglycan antibody raised in mouse so a Sun1 antibody raised in rabbit was used for co-staining. Thus, mouse and rabbit primary antibodies were detected using Alexa Fluor 488 (green), Alexa Fluor 568 (red) conjugated secondary antibodies.

Chapter 4: Generation and characterisation of dystroglycan knock-out myoblast cell line.

4.1. Introduction

Results reported in the previous chapter could not confirm a novel protein interaction between LINC complex component Sun1 and β-dystroglycan in human myoblasts. Functional studies also found that the overexpression of the cytoplasmic fragment of β-dystroglycan did not disrupt Sun1 or emerin levels at the nuclear periphery. This chapter therefore reports on the generation of human myoblast cell lines to more generally probe potential functions for dystroglycan in nuclei of human myoblasts.

4.1.1. *Current dystroglycan knockdown strategies*

To date, much of the work on the nuclear roles of β-dystroglycan has been performed in C2C12 myoblasts. Both overexpression and knock-down of dystroglycan by RNAi appears to alter nuclear morphology and the structure of the NL (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). One of the most common and effective methods to begin to interrogate the function of a gene is to remove it and analyse the consequences. For the DAG1 gene, this has been performed through the generation of several conditional mouse models (Williamson *et al.*, 1997; Côté *et al.*, 1999; Cohn *et al.*, 2002; Côté, Moukhles and Carbonetto, 2002; Moore *et al.*, 2002; Satz *et al.*, 2008), while studies in cell culture systems have utilised RNA interference to reduce protein expression (Martínez-Vieyra *et al.*, 2013). Nuclear phenotypes are apparent using RNA interference strategies which typically achieve approximately 50% reduction in protein expression (Martínez-Vieyra *et al.*, 2013). This efficiency of knockdown is consistent with others, who have also performed experiments in H2K myoblasts and Swiss 3T3 cells, which similarly attain ~50% reduction in dystroglycan protein expression (Batchelor *et al.*, 2007; Higginson, Thompson and Winder, 2008; Thompson *et al.*, 2008, 2010). Since knockdown efficiency in these studies was determined by western blotting, actual dystroglycan

knockdown in cells where NL components are disrupted is not clear due to the absence of dystroglycan staining condition (Martínez-Vieyra *et al.*, 2013). Another issue with transient dystroglycan knockdown studies is that it creates a difficult basis on which to perform rescue experiments.

4.1.2. *Dystroglycan in the nucleus*

Dystroglycan is autoproteolytically cleaved into the extracellular, membrane associated α-subunit, and the membrane-spanning β-subunit (Ibraghimov-Beskrovnaya *et al.*, 1992; Holt *et al.*, 2000; Akhavan *et al.*, 2008). At the plasma membrane, these subunits are non-covalently associated, and there appears to be numerous mechanisms by which they can be released (Bozzi *et al.*, 2009; Sbardella *et al.*, 2012; Leocadio, Mitchell and Winder, 2016). Nuclear roles for dystroglycan were first realised through the discovery of dystrophin isoforms in the nucleus. In HeLa cells, the expression of the dystrophin gene is limited to the ~70kDa Dp71 isoform (González *et al.*, 2000). It transpires that there are two distinct Dp71 proteins in HeLa cells which are termed Dp71f and Dp71d, and these differentially localise to the cytoplasm and nucleus, respectively. Alongside Dp71d, a number of other DGC components were also found in the nucleus, including β-dystroglycan (Fuentes-Mera *et al.*, 2006). Further research revealed that DGC components were also present in the nuclei of C2C12 myoblasts. This was further expanded to identify DGC components in an array of cell types, including myoblasts (González-Ramírez *et al.*, 2008). More recent work focussed on pivotal transmembrane DGC component, β-dystroglycan, and concluded the existence of a functional nuclear import (NLS), and nuclear export sequence (NES) (Oppizzi *et al.*, 2008; Lara-Chacón *et al.*, 2010; Vélez-Aguilera *et al.*, 2018). Whilst the functional significance for nuclear localisation of the full DGC assembly remains unresolved, there is emerging evidence implicating β-dystroglycan in the regulation of nuclear morphology and lamina structure in C2C12 myoblasts (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). Additionally, experiments from the Winder laboratory show patient fibroblasts derived from a patient with Walker-Warburg syndrome, caused by a homozygous premature stop codon in dystroglycan (Riemersma *et al.*, 2015) display severely malformed nuclei (Jacobs, 2017 (PhD thesis)). These observations are not unprecedented, since earlier

studies indicate that nuclear aberrations are also present in mouse skeletal muscle lacking dystroglycan (Côté, Moukhles and Carbonetto, 2002), and skeletal muscle from a patient with a mutation in mannosyltransferase POMT1 (Sabatelli *et al.*, 2003). It is notable that the POMT1 mutation results in the apparent total loss of α-dystroglycan in the muscle (Kanagawa and Toda, 2017), however, the antibody used is specific to the LARGE glycan modification which itself is dependent on the O-mannosyltransferase activity of POMT1. Without further analysis using other antibodies to dystroglycan, it is unclear whether dystroglycan is still present in these cells. Throughout these earlier studies, thorough analysis of the nuclear phenotypes has not been performed. Nevertheless, the existence of morphological abnormalities does point towards an authentic biological function for dystroglycan in the nucleus.

4.1.3. *Genome modification using CRISPR*

An efficient, easy to use technology to make genetic modification which has risen to prominence recently is known as CRISPR/Cas9. It joins a group of other programmable site-specific nucleases comprising zinc-finger (ZFNs) and transcription-activator like effector nucleases (TALENs). ZFNs and TALENs are similar in that their site-specificity is programmed through amino-acid sequences of the DNA interacting protein modules, whereas CRISPR-Cas9 site specificity can be manipulated more simply through nucleotide base pairing (Kim and Kim, 2014).

Clustered regularly interspaced short palindromic repeats (CRISPR) was a mechanism originally discovered in bacteria which confers adaptive immunity to bacteriophage infection (Barrangou *et al.*, 2007). The system was later engineered and simplified for use in eukaryotic cells, where it was shown that site-specificity could be easily controlled (Cong *et al.*, 2013; Esvelt *et al.*, 2013).

CRISPR/Cas9 has now been extensively used to create precise genetic modifications in a variety of organisms and cell types. In total, three families of the CRISPR system have been discovered, CRISPR I-III. The most commonly exploited, and that used in this study, is the type II CRISPR system. In a similar manner to the other families, it requires three main elements. The nuclease effector is Cas9, which complexes with the scaffold tracrRNA together with the guide RNA (gRNA). In

engineered systems, the tracrRNA and gRNA are ligated into a single RNA (Jinek *et al.*, 2012). The gRNA measures 20-nucleotides and specifically targets the nuclease machinery to the genomic target site through sequence homology. The choice of target site is only limited to those which are adjacent to a PAM motif (Bolotin *et al.*, 2005; Gasiunas *et al.*, 2012), which, for the type II CRISPR system isolated from *Streptococcus pyogenes*, this sequence is NGG. As with other programmable nuclease methods, the aim of CRISPR/Cas is to induce a double stranded break (DSB) in the DNA of the hosts genome, which can be exploited in various ways. The simplest, is leaving the host to repair the lesion. In eukaryotic cells, the induction of DSB activates the DNA damage response which ultimately results in the repair of the cleavage by either homology directed repair (HDR) or non-homologous end joining (NHEJ). HDR occurs if there is a template with homology to the disrupted region. If the Cas9 cleaves a single allele, this template may be the homologous gene, which would result in the high-fidelity repair of the disrupted region. Alternatively, where there is no template available, NHEJ ligates the ends of the double stranded break. This mechanism is inherently error-prone due to the lack of template, and therefore often introduces unpredictable mutations. However, these mutations, which are often insertions or deletions ('indels'), commonly result in the introduction of premature stop codons, disrupting the targeted gene (Ran, Hsu, Wright, *et al.*, 2013). In addition to deleting genes by relying on NHEJ, harnessing the HDR machinery enables CRISPR/Cas9 to be used to precisely manipulate the genome, to either base edit, or knock-in a sequence of interest. By simultaneously delivering the CRISPR plasmid and a template encoding the modified locus flanked by sequences homologous to the target region, HDR is promoted, thus incorporating the desired sequence.

4.1.4. *DAG1 CRISPR design*

The aim of the research presented in this chapter was to leverage the specific targeting properties of CRISPR/Cas9 derived from *S. pyogenes* (Ran, Hsu, Wright, *et al.*, 2013) to create a genetic lesion in the DAG1 gene in myoblasts, so leading to the absence of the dystroglycan protein. As described, C2C12 cells have been previously established as a model to study the nuclear functions of dystroglycan and, for

consistency, these myoblasts would be a useful in which to create a DAG1 knockout. However, C2C12 myoblasts are inherently difficult to genetically manipulate because they have an irregular chromosome number (Mamchaoui *et al.*, 2011), which would demand an efficient CRISPR/Cas9 process, while requiring many more colonies for screening and sequencing. In light of these known potential complications, a more recently derived and characterised human myoblast cell line was chosen for experimentation. These myoblasts have been robustly characterised, found to express myogenic markers and accurately replicate aspects of myogenesis *in vitro* (Mamchaoui *et al.*, 2011). Moreover, these cells can integrate into the myofibre differentiation pathway upon injection into the muscle of immunocompromised mice, and unlike C2C12 cells, do not form tumours (Morgan *et al.*, 1992; Mamchaoui *et al.*, 2011).

Data from previous studies are suggestive that nuclear phenotypes are persistent throughout a variety of models where β-dystroglycan is reduced or ablated. It was therefore hypothesised that immortalised human myoblasts depleted of dystroglycan would also possess recognisable nuclear aberrations, including alterations to morphology and NL composition, in a similar manner. A knock-out cell line would also provide a valuable model in which to further probe dystroglycan function, by re-expression studies using rescue constructs, for example.

Successfully targeting a gene for disruption requires knowledge of its structure. A common problem with CRISPR-mediated gene disruptions is the presence of downstream alternative start sites, potentially resulting in the restoration of gene expression in the form of a truncated protein. Fortunately, the DAG1 gene has a relatively simple structure, comprising two exons, separated by a large intron, with a single start site (Figure 4-1, 'ATG'). To ensure absence of dystroglycan, it was proposed that two gRNA sequences were delivered simultaneously to the myoblasts with the target sites located either side of the start codon, in order to excise this region so ablating expression (Figure 4-1).

4.1.5. _Aims and hypotheses_

Given the current lack of a dystroglycan-deficient human myoblast cell line, research reported in this chapter aimed to generate multiple clonal myoblast cell lines using CRISPR. The resulting myoblast cell lines would provide a reproducible platform for the analysis of dystroglycan knockout phenotypes, and also for re-expression studies to dissect the functions of dystroglycan at the cellular level. To do this, gRNA sequences required validation and an effective method for screening putative DAG1 ablated myoblast colonies established. Furthermore, validation of DAG1 deleted myoblasts would be validated at the genomic and protein level. Based on previous studies, which indicated that β-dystroglycan has some function in maintaining NL integrity and nuclear morphology, it was hypothesised that genetic ablation of the DAG1 gene would similarly disturb nuclear morphology, levels of NL components such as emerin and lamin B1, and nucleus-centrosome distance.

gRNA_001 (5'-3'):   AGTGCTCTCAGACCTCCACG
gRNA_002 (5'-3'):   TACTTAGCAAGACTATCGAC
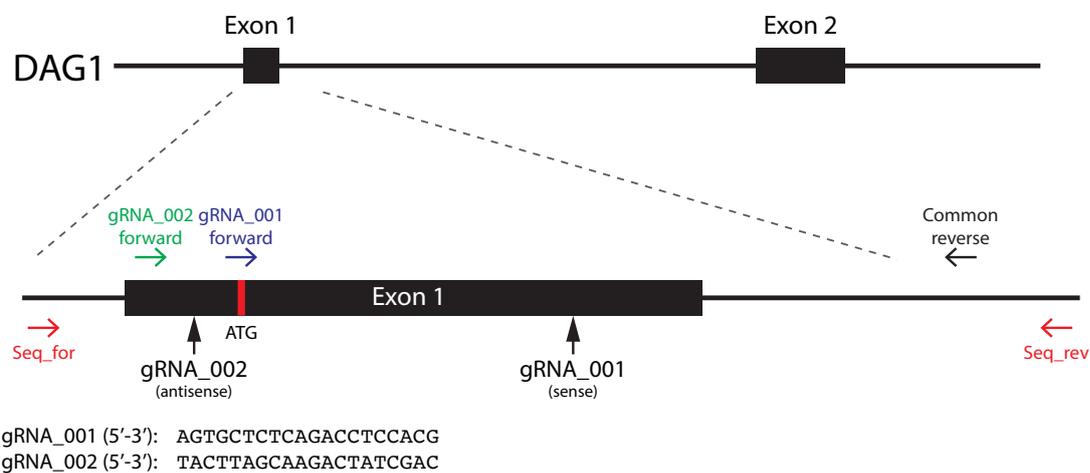
Figure 4-1 - DAG1 CRISPR strategy. The DAG1 gene structure (shown 5'-3'), excluding 5' untranslated region comprises two exons. The first coding exon is expanded to indicate locations of the start codon (ATG) in relation to the guide RNA target site (gRNA_001 and _002). gRNA sequences were designed to flank the start codon. Primer positions to assess the efficiency of each guide site are indicated.

4.2. <u>Results</u>

4.2.1. *<u>Screening gRNA function and DAG1 KO colonies</u>*

The DSB-inducing property of CRISPR followed by error-prone NHEJ DNA repair by the host cell was used to generate genetic lesions in the DAG1 gene. However, to establish myoblast colonies, gRNA activity had to first be assessed for functionality, and a method to screen putative clones needed to be established.

4.2.1.1. ***The SURVEYOR assay***

In the absence of a repair template, Cas9 activity promotes NHEJ DNA repair due the induction of DSB in the genome resulting in the random repair of the modified locus. Within a population of cells, this process incorporates a range of indels, which can then be crudely visualised using the SURVEYOR assay. This assay analyses the efficiency of the gRNA, through the detection of mutated loci in the transfected population. Primers are designed to produce a short PCR product across the targeted region, and this product will contain a mixture of DNA species which are then denatured. Reduction of the temperature causes single DNA strands to randomly reanneal to form heteroduplexes, and the mismatched bases in these oligonucleotides can be recognised and cleaved by the mismatch-specific T7 endonuclease I (Guschin *et al.*, 2010).

To perform the SURVEYOR assay on gRNAs designed for DAG1 gene disruption, KM155 immortalised human myoblasts were transfected with pX458-DAG1-gRNA_001 or pX458-DAG1-gRNA_002 in separate reactions. The CRISPR construct includes a GFP-reporter in addition to the Cas9, guide and scaffold RNAs to identify transfected cells. 24 hours after transfection, cells positive for GFP were bulk sorted by FACS, replated, and genomic DNA was extracted after a further 24-hour incubation. The targeted region was then amplified by PCR using specific primers (Figure 4-1). For gRNA_001, successful CRISPR/Cas-mediated cleavage should result in the 541bp PCR product degraded to ~350bp and ~190bp (Figure 4-2A). T7 endonuclease treatment revealed a faint band at the expected molecular weight of ~350bp, however the shorter fragment was not observed (Figure 4-2B). Similarly, T7 endonuclease treatment of gRNA_002 showed a faint band ~560bp (Figure 4-2B), which correlates with the expected CRISPR cleavage site in this PCR product (Figure

4-2A). Again the smaller expected band could not be observed. It may be that the amount of DNA was below the detection threshold for the shorter bands. Nevertheless, this experiment indicates that the gRNA sequences are functional in
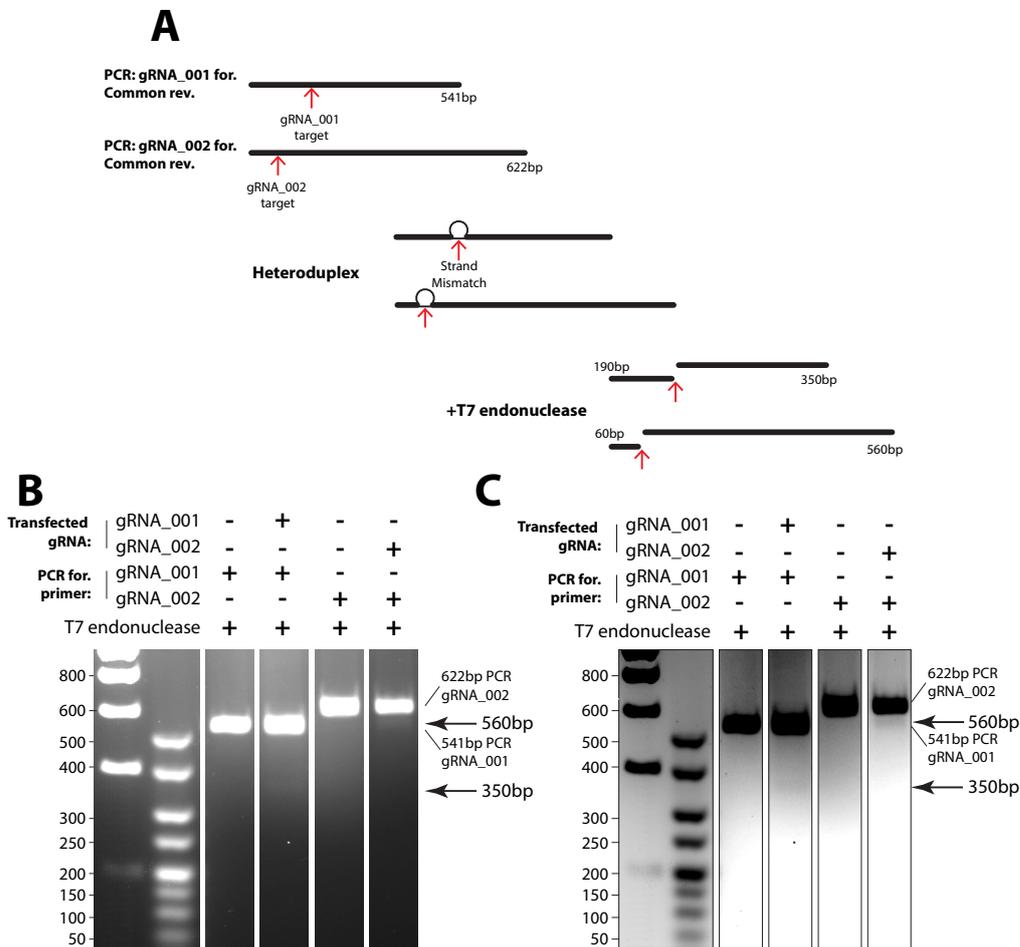
Figure 4-2 - SURVEYOR assay demonstrating activity of DAG1 CRISPR guide RNAs in pX458 vector. (A) Schematic depicting the principle of the SURVEYOR assay for the human DAG1 gene. gRNA_001 and gRNA_002 were separately transfected into human myoblasts. The targeted regions were amplified using PCR, resulting in the products shown. Heat denaturation followed by reannealing forms heteroduplexes which are recognised by the T7 endonuclease and cleaved across both DNA strands. (B) Agarose gel showing SURVEYOR assay results. Transfection of either gRNA_001 or gRNA_002 with exposure to T7 endonuclease results in the detection of restricted DNA products. (C) Inverted and contrast adjusted of gel images shown in (B).

human myoblasts.

4.2.1.2. *Establishing the In-Cell Western assay for clonal screening and selection*

The frequency of gene disruption using CRISPR/Cas9 is not 100% (Ran, Hsu, Wright, *et al.*, 2013)**.** This may be because levels of gRNA and Cas9 expression are insufficient to target the locus and induce the DSB, the DSB occurs on a single allele leading to high-fidelity repair using HDR, the induced indels are in-frame with the rest of the gene, or for some targets, the Cas9 remains bound to the DNA, therefore inhibiting access for repair machinery (Clarke *et al.*, 2018). For these reasons, it is necessary to single cell sort transfected cells followed by clonal expansion to yield cell lines which have the same genetic background. This can be performed using PCR, however, this screening system may disregard DAG1 ablated colonies which have indel-containing repaired DSBs, but not excised the region between the two guide sites. For this reason, it was decided that single cell colonies would assessed based on the absence of dystroglycan protein expression. To do this, the In-Cell Western (ICW) assay was used. Methodologically similar to immunofluorescence, this assay allows the detection of protein epitopes in fixed samples *in situ* in multiwell plates and can be imaged using a LiCor Odyssey.

Antibodies to β-dystroglycan are generally directed against the extreme C-terminus, so another advantage to this method is that it can validate the disruption of the gene at the protein level. Specifically for dystroglycan, the C-terminal antibody epitope also accounts for unknown alternative start sites potentially partially reinstating DAG1 expression. Since single cell clones can grow at vastly different rates, it was important to first establish the sensitivity of the ICW assay using a loading control to indicate the presence of cells in a given well. In the first instance, nuclear DRAQ5 and cytoplasmic Sapphire700 stains were assessed for their sensitivity to cell number. In a 96-well plate, it was found that 1.25×10^4 cells could be observed using DRAQ5 in combination with Sapphire700 (Figure 4-3A, left panels). Interestingly, removing the Sapphire700 stain increased the sensitivity of DRAQ5 staining, identifying wells with 0.63×10^4 cells. MDG2 staining provided a robust signal, revealing wells with 0.16×10^4 cells, validating the use of this antibody for the detection of dystroglycan in this assay (Figure 4-3A, right panels). Compared with MDG2 staining, DRAQ5 staining was not sufficiently sensitive to use as a loading

control for screening so an antibody against globular actin was attempted, instead. While also not as sensitive as MDG2, this approach yielded two-fold increased sensitivity when compared to DRAQ5 staining, showing wells with 0.31 x 10^4 cells (Figure 4-3B). Actin staining was subsequently used as a loading control for this assay. Throughout these experiments, it was noted that there was significant non-specific secondary anti-mouse IRDye800 binding to cells which was taken into account when assessing the MDG2 positive colonies (Figure 4-3B).

4.2.1.3. *In-Cell Western Screening of putative DAG1 KO human myoblasts*

After the functional validation of the guide RNA sequences, and establishing an efficient method to screen potential colonies, a DAG1 KO myoblast cell line was generated. KM155 myoblasts were co-transfected by electroporation with CRISPR/Cas9 plasmids containing gRNA_001 and gRNA_002, and single-cell sorted by FACS 24 hours later into 96-well plates. Cells were cultured until they reached ~80% confluence and passaged for screening and continuation. The In-Cell Western assay was used to initially screen the 32 surviving and expanding colonies for DAG1 KO in the same manner as described (see section: 4.2.1.2). A single clone which displayed reduced MDG2 staining was identified, despite showing robust actin staining (Figure 4-4, clone 11A, green arrowhead). Colonies 11A (MDG2 negative) and 10D (MDG2 positive) were therefore selected for further analysis.
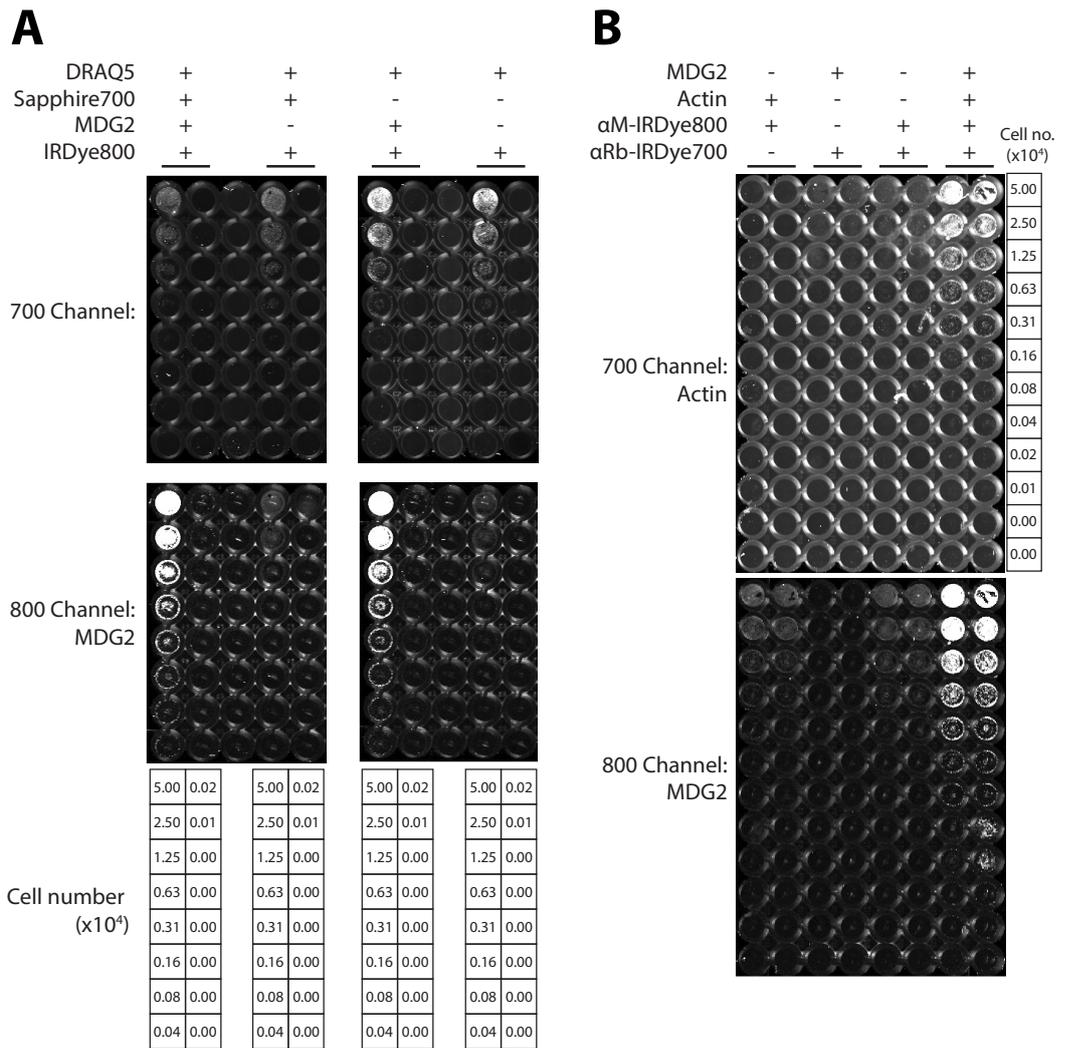
Figure 4-3 - Optimisation of the In-Cell Western assay for screening DAG1 KO myoblast colonies. (A) Determining the sensitivity of the In-Cell Western assay to cell number using a combined DRAQ5 and Sapphire700 stain or with DRAQ5 alone. (B) Determining the sensitivity of antibody staining against actin to the number of myoblasts plated. In both experiments, myoblasts were plated in a 96-well plate at numbers indicated in adjacent tables.
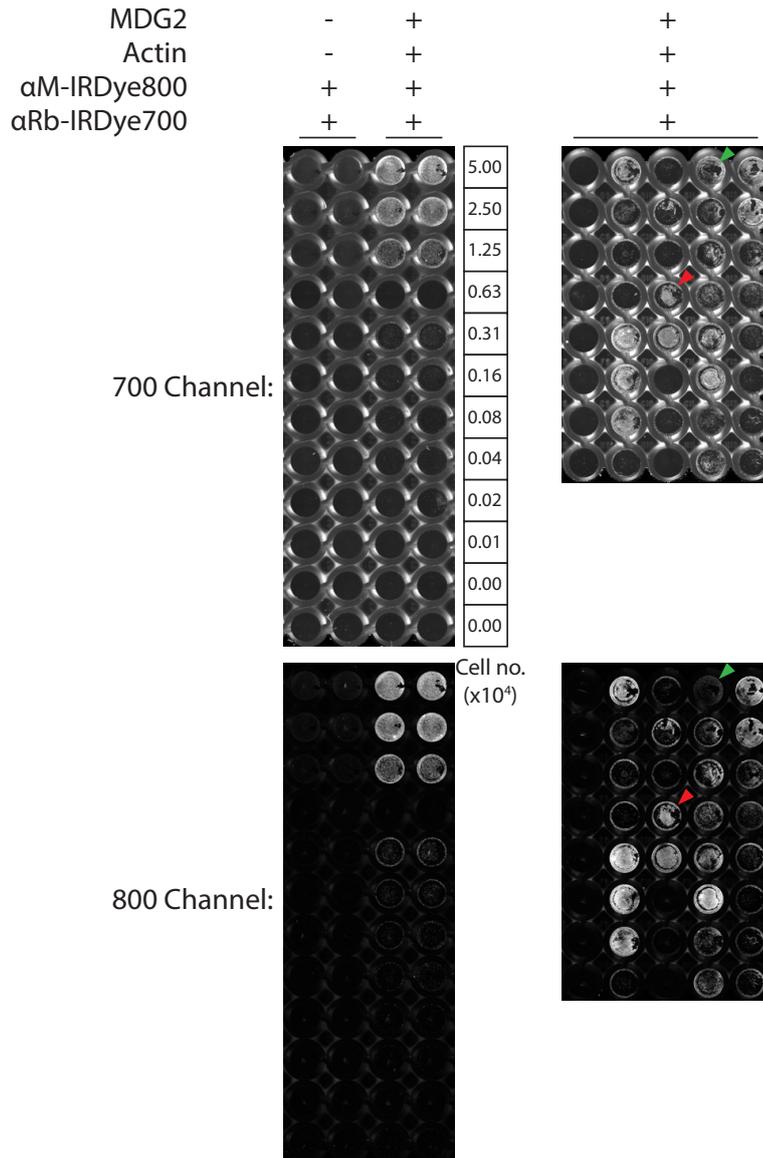
Figure 4-4 – Screening DAG1 CRISPR transfected myoblasts by In-Cell Western. Cells were stained using antibodies raised against actin and β-dystroglycan (MDG2), which were detected using secondary antibodies conjugated to 700nm and 800nm fluorophores, respectively. Control cell dilution series shown in left panels. Well 11A (green arrowhead) was identified as containing DAG1 depleted myoblasts, while the colony in well 10D (red arrowhead) was used as a control.

4.2.2. *Validation of dystroglycan knock-out clone 11A*

The absence of dystroglycan was verified using an alternative method. As such, wild-type, clone 10D and 11A myoblasts were cultured and whole cell lysates were subjected to analysis by western blotting. Blots were probed for β-dystroglycan protein using three different antibodies. JAF1 is a rabbit polyclonal antibody which recognises all forms of β-dystroglycan (François Rivier *et al.*, 1999). MDG2 is a mouse monoclonal specifically raised against non-Y892 phosphorylated β-dystroglycan (Pereboev *et al.*, 2001), and 1709 recognises a similar C-terminal epitope to both JAF1 and MDG2, but is specific for β-dystroglycan phosphorylated at pY892 (Ilsley, Sudol and Winder, 2001). Each of the antibodies detected a band corresponding to 43 kDa β-dystroglycan in both the wild-type and negative control colony 10D. This band was absent in lysates derived from clone 11A, despite positive GAPDH in these lanes indicating protein loading. Together, these data indicate that dystroglycan is not expressed in myoblast clone 11A which is consistent with results from the initial ICW assay.
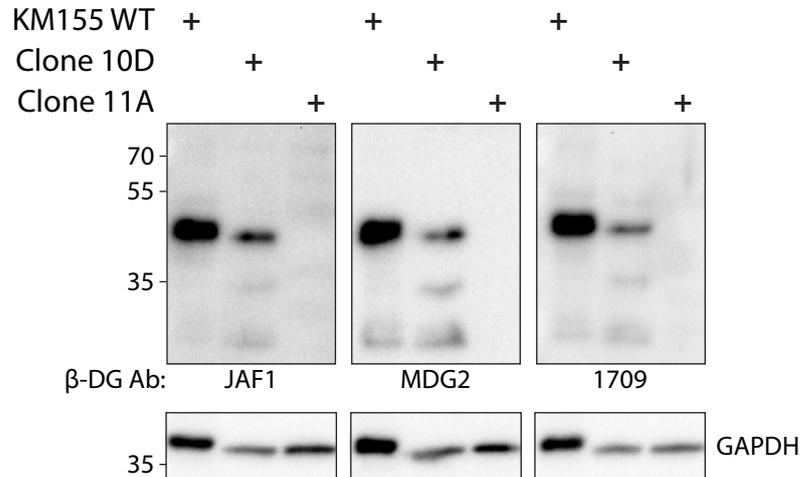
Figure 4-5 - Western blotting DAG1 KO myoblast clone 11A. Indicated myoblast cell lines were lysed and analysed by SDS-PAGE and western blotting. Resulting western blots were probed with the indicated antibodies raised against β-dystroglycan and GAPDH was used as a loading control.

4.2.3. *Determination of the proliferation rate in DAG1 KO myoblast clone 11A*

For an *in vitro*, cell culture DAG1 knock-out model which could be used to stage further expression and phenotype studies to be of maximum use, it is important that it is viable.

Modulations in the expression levels of dystroglycan implicates it in regulating the cell cycle (Hosokawa *et al.*, 2002; Sgambato *et al.*, 2004; Higginson, Thompson and Winder, 2008). Additionally, endogenous levels of pY892 detected within the nucleus have been shown to oscillate depending upon cell cycle stage (Jacobs, 2017 (PhD thesis); Mathew, 2011 (PhD thesis)). The significance of this observation, however, is unknown. If dystroglycan has a critical function in the timing of the cell cycle here, it would likely result in a disruption of myoblast proliferation over time. To test this hypothesis, growth curves were constructed using the WT, negative control clone 10C, DAG1 KO clone 11A, myoblast lines. Growth over the course of one week resulted in no significant differences between the growth rates of the
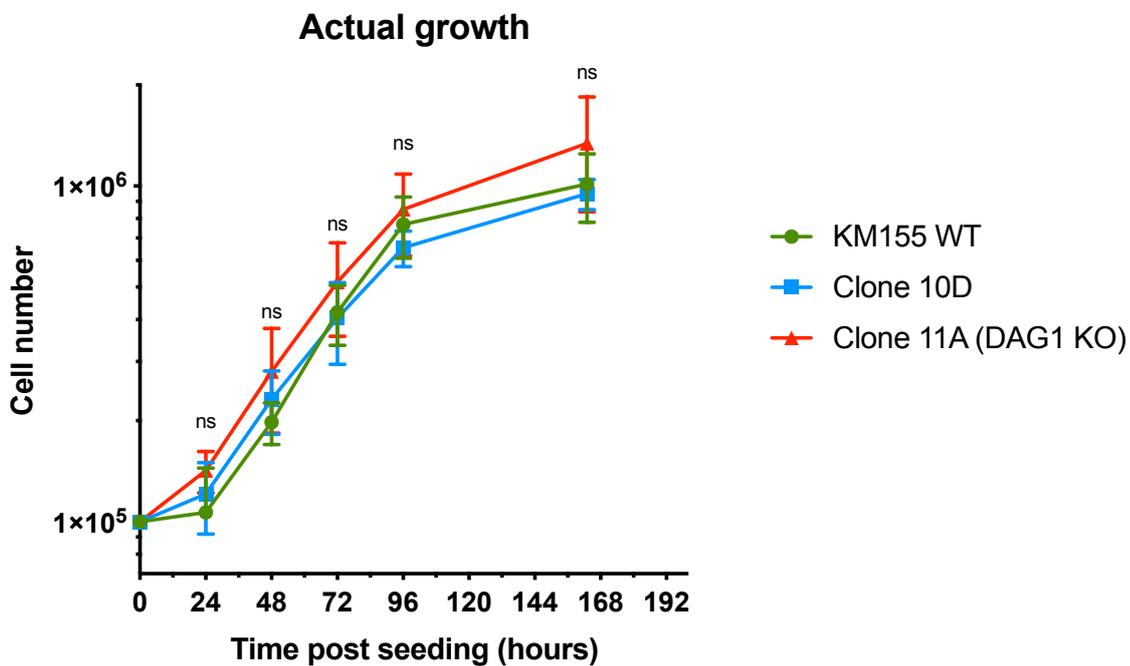
Figure 4-6 - The rate of proliferation of myoblasts lacking DAG1 is unchanged. Indicated human myoblasts were cultured for one week and counted intermittently. Statistical significance was determined using a RM two-way ANOVA with Geisser-Greenhouse correction and Tukey's multiple comparisons test. Error bars signify SD.

different myoblast lines (Figure 4-6). It may be that there are subtler disruptions to the mitotic process, however, this was not pursued.

4.2.4. *The effect of DAG1 KO on NE components lamin B1 and emerin*

β-dystroglycan binds NL components and it is believed that its role in stabilising nuclear morphology occurs through interactions specifically with lamin B1 and emerin since levels of these proteins are reduced when dystroglycan is either overexpressed or modestly depleted by RNA interference (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018; Gómez-Monsivais *et al.*, 2020). Now with access to the dystroglycan-null human myoblast line generated here, it was hypothesised that the disruption to NE components would be more pronounced. Therefore, levels of lamin B1 and emerin were analysed in WT, DAG1 positive clone 10D and DAG1-null clone 11A myoblasts by western blotting. In contrast to previous reports, quantification of immunoreactive bands revealed no difference in lamin B1 or emerin levels in clone 11A compared to control cell lines.
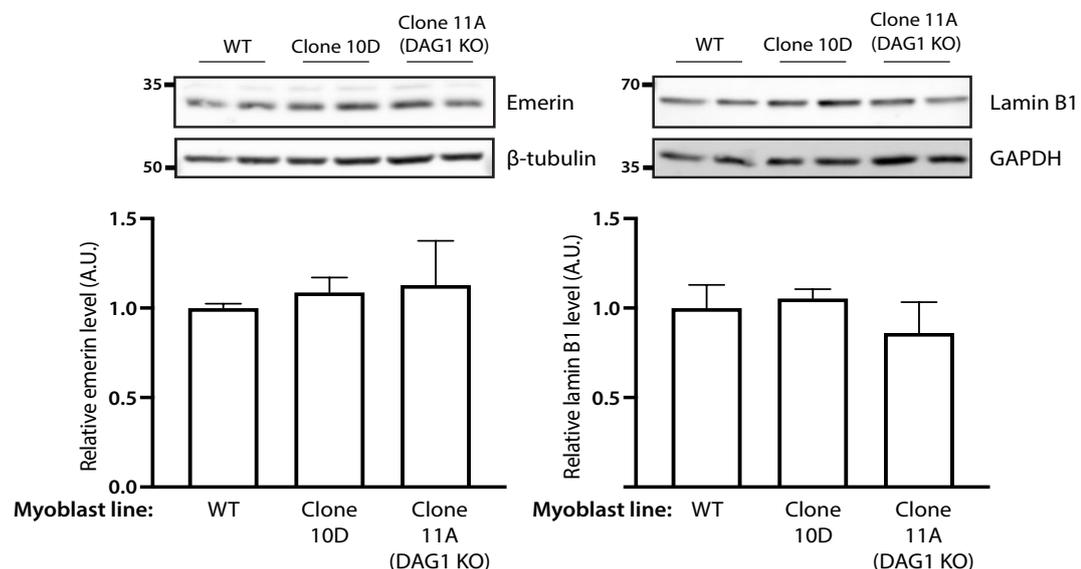
Figure 4-7 - Lamin B1 and emerin levels are unchanged in DAG1 KO myoblast nuclei. Myoblasts originating from the indicated cell lines were lysed and analysed by SDS-PAGE. Nuclear lamina protein levels were revealed by immunoblotting for emerin and lamin B1. Emerin and lamin B1 levels were quantified relative to loading controls β-tubulin and GAPDH, respectively, and normalised to WT cell line (graphed). Error bars signify SD. Statistical differences determined using one-way ANOVA with Tukey's correction for multiple comparisons.

4.2.5. _Generation and sequencing of further DAG1 KO clones_

Using the described CRISPR/Cas9 system, previous experiments verified that DAG1 can be ablated to create a viable immortalised myoblast cell line in which to interrogate dystroglycan function. However, this attempt only yielded one clone lacking DAG1. It is customary that phenotypic analysis should comprise multiple clones, to account for off-target effects of the CRISPR/Cas9 system. It was therefore necessary to repeat the procedure to generate more DAG1 KO myoblast clones. To do this, myoblasts were again transfected with CRISPR/Cas9 constructs, and an increased number of 96 well plates was sorted. ICW screening identified 12 putative DAG1 KO clones (data not shown) which were further analysed by western blotting, indicating the absence of β-dystroglycan as detected using MDG2 in these clones, while negative controls displayed immunoreactivity (Figure 4-8). To understand the genetic aberration behind the DAG1 KO clones, the modified regions were amplified using the Seq_for forward and Seq_rev reverse primers (Figure 4-1). Analysis by gel electrophoresis indicated a range of genetic modifications in this region, however, many clones gave rise to PCR products which corresponded to the excision of the region between the CRISPR target sites (Figure 4-9A). Those clones with more complex genetic aberrations, showing multiple PCR products were excluded from subsequent analyses. PCR products were cloned into TOPO-TA vectors, transformed into competent bacteria, and 10 bacterial colonies from each cell line were sequenced to verify zygosity of the mutation. Sequencing results show that DAG1 expressing negative controls 1.B8 and 1.B9 had not been modified, whereas six of the analysed clones had mutations where the region between the gRNA targets had been excised (1.B6, 1.C2, 1.G7, 2.C7, 2.C9, 2.D8) (Figure 4-9B, Table 4-1). Western blot analysis of clone 1.G4 indicated the absence of DAG1 however yielded a PCR product of normal length (Figure 4-9A). Sequencing found a single base deletion at each of the gRNA target sites, presumably resulting in a premature stop codon at amino acid 79 (Figure 4-9B, Table 4-1).
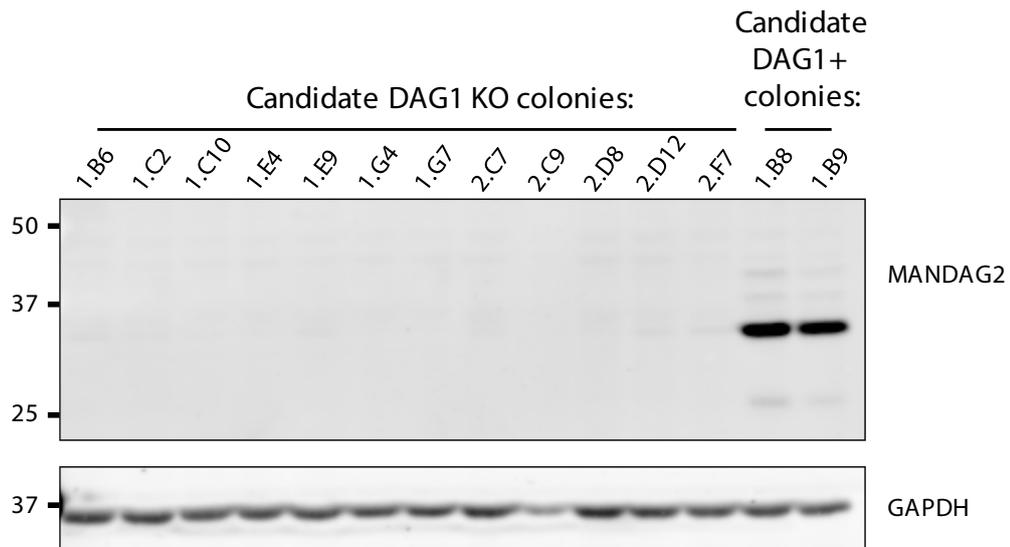
Figure 4-8 – Verification of DAG1 ablation in 12 human myoblast colonies by western blotting. Lysates from putative DAG1 KO myoblast colonies were harvested from cell pellets and analysed by SDS-PAGE. Western blotting for dystroglycan using the MDG2 antibody was used to reveal the level of expression in clones after CRISPR/Cas9 gene deletion. GAPDH was used to show loading in each lane.
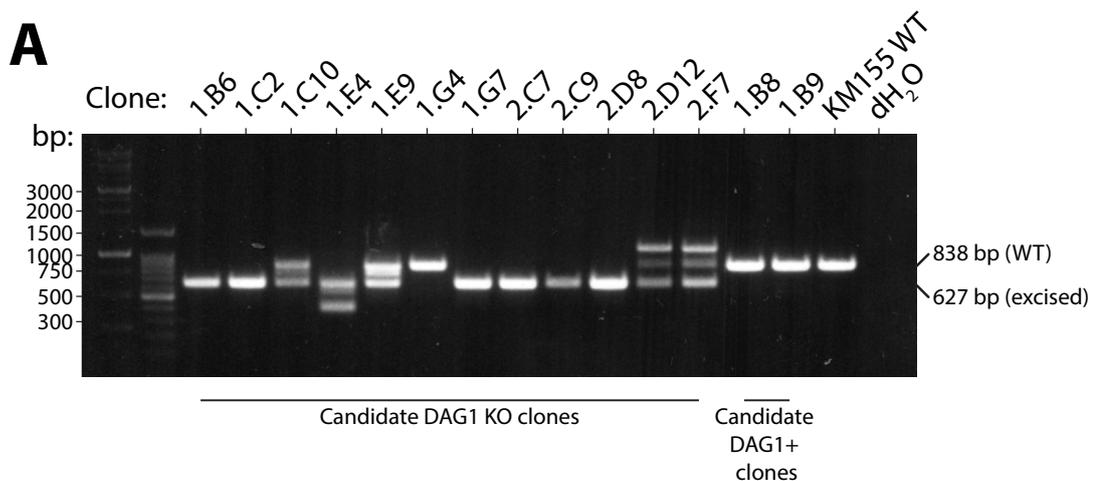
Figure 4-9 - PCR analysis and sequencing of DAG1 KO human myoblast lines. The CRISPR targeted region in the DAG1 gene was amplified using Seq_for and Seq_rev primers and KOD high-fidelity polymerase and analysed by agarose gel electrophoresis (A). Bands were extracted, 5' adenylated, and cloned into TOPO-TA vector, and transformed into competent bacteria. 10 bacterial colonies for each line were sequenced by Sanger sequencing, and a representative trace of the modified region is displayed in (B) (overleaf). Guide RNA sites highlighted in yellow, mutation sites in red and ATG (start) codon in green.

B
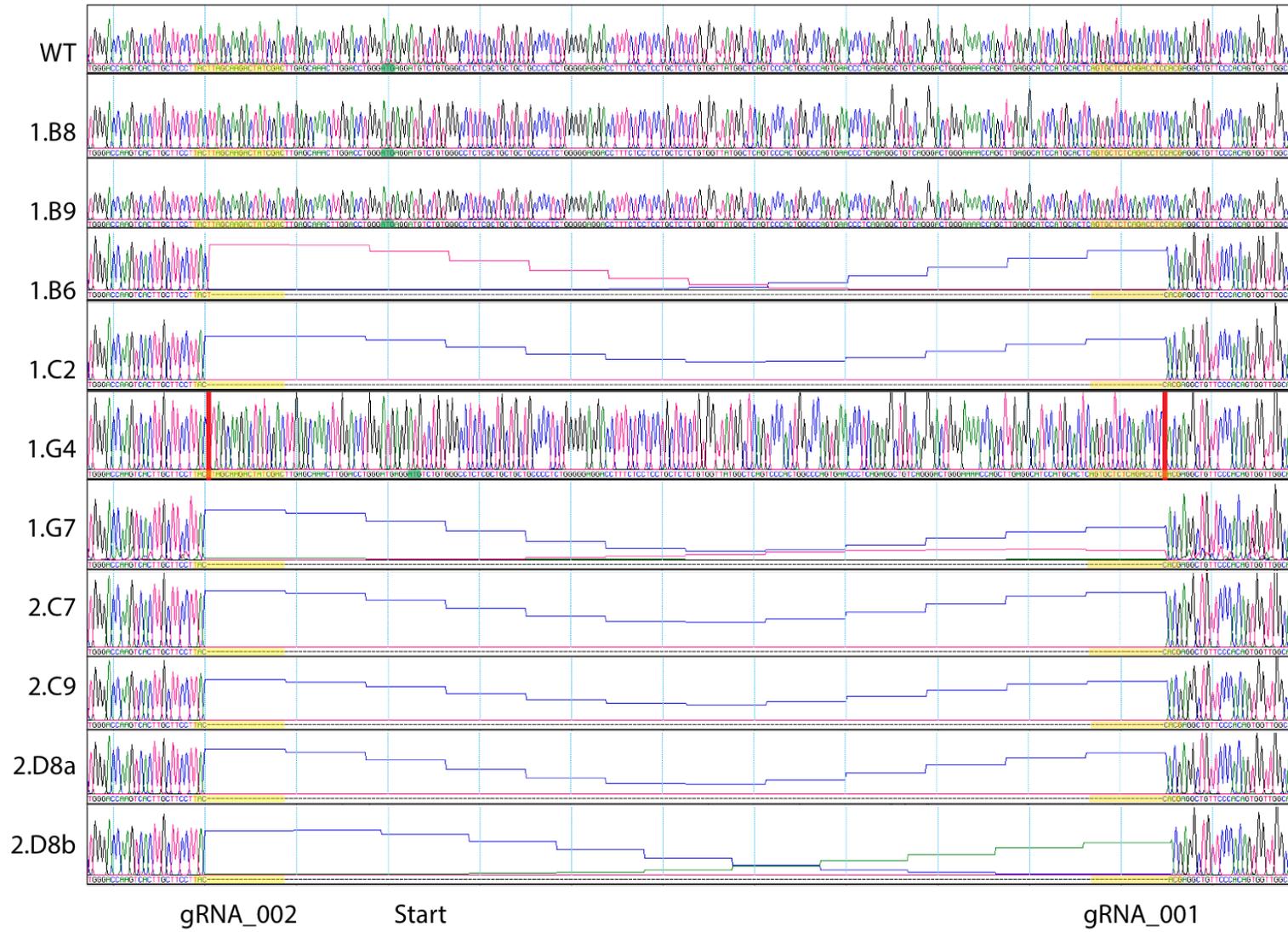
Clone	PCR product (bp)	β-dystroglycan protein	Mutation induced	
			gRNA_001	gRNA_002
WT	~840	Expressed	N/A	
10D	~840	Expressed	N/A	
1.B8	~840	Expressed	N/A	
1.B9	~840	Expressed	N/A	
11A	3 bands	Not detected	1bp deletion	5bp deletion
1.B6	~630	Not detected	Excision: CT insertion	
1.C2	~630	Not detected	Excision: C insertion	
1.G4	~840	Not detected, predicted aa79*	1bp deletion	1bp deletion
1.G7	~630	Not detected	Excision: C insertion	
2.C7	~630	Not detected	Excision: C insertion	
2.C9	~630	Not detected	Excision: C insertion	
2.D8	~630	Not detected	Heterozygous excision: C insertion and end joining.	

Table 4-1 - Summarising CRISPR induced mutations leading to DAG1 ablation in the genome of 10 human myoblast lines.

4.2.6. *Phenotype analysis*

4.2.6.1. **Nuclear shape analysis**

Dystroglycan deficiency has been reported to result in changes to the nucleus, including shape, lamina composition, and nucleolar morphology in a variety of cell types. It was therefore hypothesised that the DAG1 KO human myoblasts would recapitulate these phenotypes. Regular, ellipsoid nuclear morphology in cultured cells appears to be dependent on the NL (Goldman *et al.*, 2004; Lammerding *et al.*, 2006; Dahl, Ribeiro and Lammerding, 2008), and the impaired nuclear shape upon the loss of β-dystroglycan is indicative of a role in maintaining this structure. First, nuclear morphology of the dystroglycan deficient myoblasts was assessed. Myoblasts were seeded on coverslips, stained with Hoechst to delineate the

chromatin and general nuclear shape, and imaged by fluorescence microscopy. Images were analysed using a Cell Profiler pipeline to identify the Hoechst stain and quantify parameters of their shapes. In contrast to previous reports, these experiments revealed there were no significant differences in the nuclear morphologies between DAG1 expressing, and DAG1 KO cells (Figure 4-10). Overall, this examination indicated that all analysed cells had the same surface area and were similarly ellipsoid (shown by form factor and eccentricity analyses). Additionally, measurements of solidity and compactness indicate that nuclei in cells lacking DAG1 are not irregularly shaped, when compared to the negative controls. It should be noted that one DAG1 KO clone, 1.G4, became senescent over time and displayed characteristic features such as enlarged cellular and nuclei areas (Sadaie *et al.*, 2015; Wang and Dreesen, 2018), which is illustrated by its slightly larger surface area measurement and effect size. Notably, and contrary to recent reports (Jimenez-Gutierrez *et al.*, 2020), this phenotype was not reproduced across other DAG1 KO myoblast clones, so clone 1.G4 was excluded from further analyses.
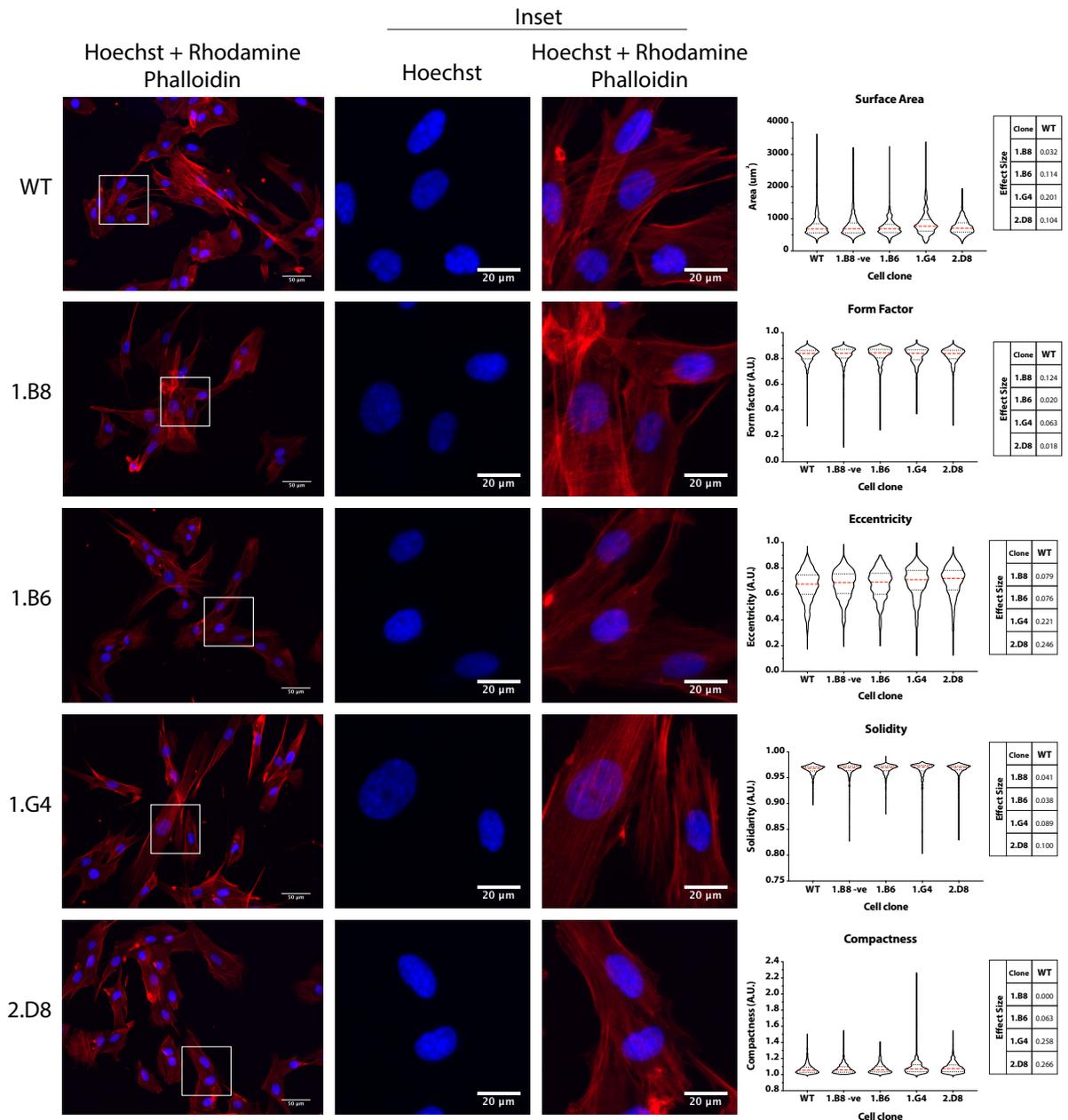
Figure 4-10 - Nuclear shape analysis of DAG1 KO myoblasts. Myoblasts were seeded on coverslips, fixed and stained using Hoechst and rhodamine-phalloidin. Cells were imaged using widefield microscopy and nuclear shape parameters including surface area, form factor, eccentricity, solidity and compactness was analysed using a pipeline built within CellProfiler. N = 3 except clone 1.G4 which became senescent; n = 119-373 cells for each condition. Tested for normality using D'Agostino and Pearson's normality test followed by Kruskal-Wallis Test with Dunn's multiple comparisons. Significance is reported where effect size is >0.2 (small) and DAG1 KO myoblast clones 1.B6 and 2.D8 are consistently different to WT and 1.B8.

4.2.6.2. *Nuclear to centrosomal distance*

The microtubule organising centre is closely associated with the nucleus in many cells types, and is mediated by the microtubule cytoskeleton and NE components (Salpingidou *et al.*, 2007; Burakov and Nadezhdina, 2013). Through its association with components of the NL, recent results indicated that perturbation of dystroglycan levels disrupts the normal close nuclear-centrosomal relationship in myoblasts (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). These results indicate that further dystroglycan depletion by knock-out could also result in an increased nuclear-centrosomal linkage. To test this hypothesis, two DAG1 KO clones, CRISPR negative control, and WT cells lines were stained for γ-tubulin and lamin A/C, to mark the centrosome and NE, respectively. The distance between the centrosome and the nearest point of the nuclear envelope was then measured (indicated by white line in). In agreement with previous reports, the centrosome tends to be more distant from the nucleus in DAG1 KO cells when compared to WT cells. Calculation of the effect size indicates that this increase corresponds to a small difference between WT and DAG1 KO cells. However, a definitive interpretation of these results is compounded by an even smaller trend between the 1.B8 CRISPR negative clone, and the experimental 1.B6 and 2.D8 DAG1 KO clones. This indicates that artefacts introduced as a result of clonal expansion cannot be excluded and the increase in nucleus-centrosome distance cannot be unequivocally attributed to dystroglycan deficiency. While an increase in centrosome distance is broadly in line with previous reports, additional knock out clones would need to be analysed to unequivocally attribute increased nuclear-centrosomal distance to the absence of DAG1. Alternatively, an increase in nuclear-centrosome distance could be confirmed as an effect of DAG1 KO upon re-expression of dystroglycan in deficient myoblasts. However, transduction of DAG1-containing lentivirus and subsequent selection resulted in resistant cells but without dystroglycan expression.
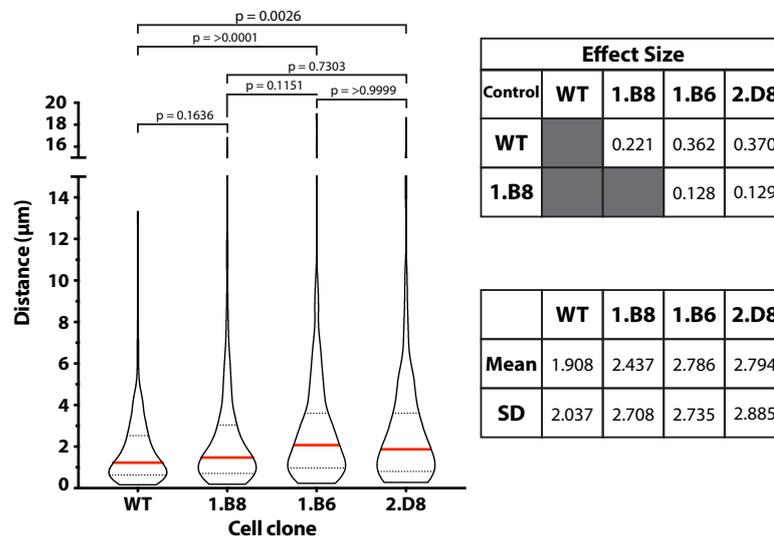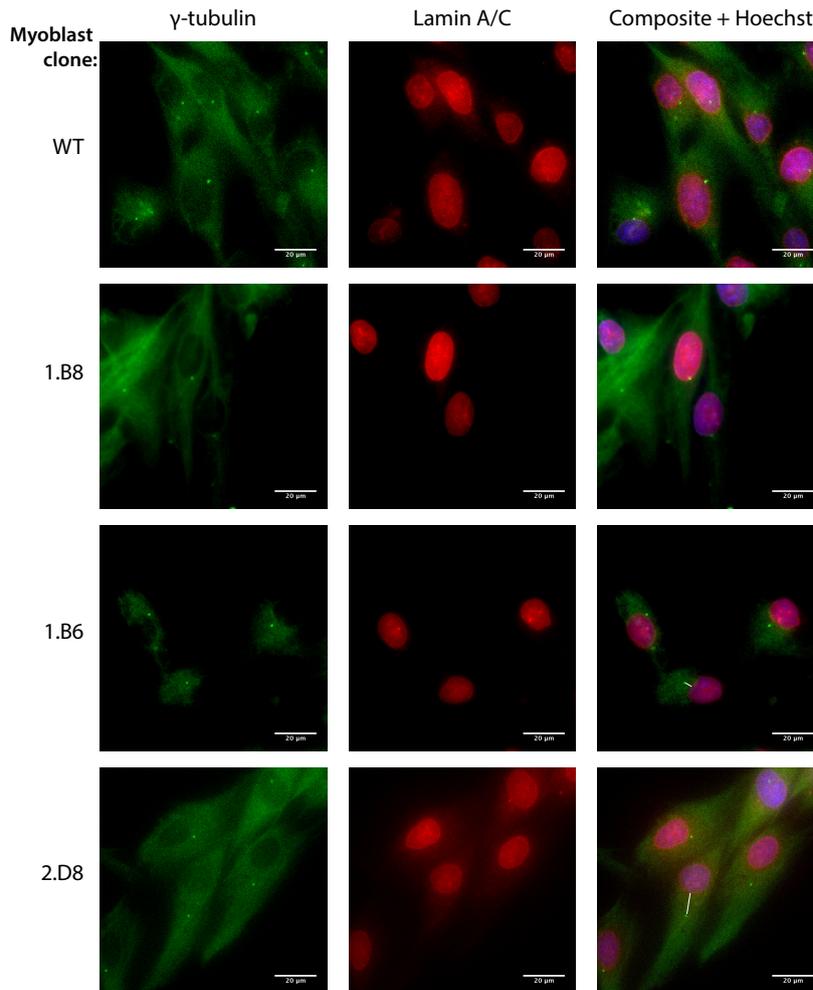
Figure 4-11 - Analysis of the nuclear to centrosomal distance in DAG1 KO myoblasts. Wild-type and DAG1 expressing myoblasts, and two DAG1 negative myoblast lines stained for lamin A/C and centrosomal component, γ-tubulin. The distance between the centre of the γ-tubulin signal and the nearest edge of the nucleus, as delineated by lamin A/C, was measured and plotted. Tested for normality using D'Agostino and Pearson's normality test followed by Kruskal-Wallis Test with Dunn's multiple comparisons. Differences are considered statistically significant where $p \leq 0.05$.

4.3. <u>Discussion</u>

The results presented in this chapter describe the generation of DAG1 knock-out cells using a CRISPR/Cas9 approach. Results show that cells lacking dystroglycan are indeed viable, and appear to proliferate at a normal rate. However, nuclear phenotypes previously attributed to the perturbation of dystroglycan levels were not observed in the analysis of these myoblasts. This includes disruption to NE components lamin B1 and emerin, and nuclear morphology appears not to be disturbed. Using the current experimental system, observations of nuclear-centrosome distance in these could not be explicitly attributed to the depletion of dystroglycan, thus rendering it inconclusive.

4.3.1. *Limitations to CRISPR/Cas9 method*

The CRISPR/Cas9 method used in the DAG1 KO cells described used two gRNA sequences coupled with the Cas9 nuclease from *S. pyogenes*. When directed to the target site, this Cas9 induces a DSB in the DNA. It is possible that the gRNA can anneal to other regions of the genome, through non-specific binding and similarly cause DSBs. The use of two gRNA sequences is understood to increase the risk of off-target effects two-fold. To reduce these, a double-nickase approach could have been used, instead, where the Cas9 protein is mutated, such that it only cleaves a single DNA strand (Ran, Hsu, Lin, *et al.*, 2013). Using this system, specific genetic lesions can be induced using two spaced gRNAs targeting opposite DNA strands, leading to the excision of the intervening region. DNA nicks induced elsewhere in the genome through off-target binding of one of the gRNAs can be repaired with high efficiency. The reasoning to analyse numerous clones generated from CRISPR/Cas gene disruptions is to reduce the impact of off-target cleavage events on any observed phenotype.

4.3.2. *Disparity in phenotypes observed*

The rarity of primary dystroglycanopathies indicates the importance of dystroglycan for organismal development and viability. Myofibres and fibroblasts from the very few patients found with disruptions to the dystroglycan locus have been reported to exhibit nuclear abnormalities (Sabatelli *et al.*, 2003; Jacobs, 2017

170

(PhD thesis)). Perturbed nuclear morphologies have also been observed in mouse models of the dystroglycanopathies, although the function of dystroglycan in this regard has not been thoroughly investigated (Côté, Moukhles and Carbonetto, 2002). More recent cell-based studies suggest that dystroglycan functions to maintain nuclear integrity by binding elements of the NL, maintaining integrity of the NL for normal shape and nuclear-centrosome anchorage (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). Unexpectedly, experiments presented in this chapter, which sought to recapitulate these previously established nuclear phenotypes related to β-dystroglycan, did not show the same effect using a knockout myoblast model generated using CRISPR.

4.3.2.1. *Nuclear lamina components and morphology*

Alterations in the nuclear lamins are known to have profound effects on nuclear mechanics. In mouse embryonic fibroblasts, depletion of lamin A/C results in severely irregular-shaped nuclei and increased deformability, whereas loss of lamin B1 causes a single nuclear bleb and unchanged mechanical stiffness (Lammerding *et al.*, 2006). Similarly, changes to nuclear morphology have been reported to occur in C2C12 cells depleted of dystroglycan by RNA interference, which has been attributed to a functional role for β-dystroglycan in stabilising a complex comprising lamin B1 and emerin (Martínez-Vieyra *et al.*, 2013). Contrary to previous observations, loss of dystroglycan in human myoblasts appeared to have no effect on nuclear morphology, so it is perhaps unsurprising to also observe normal levels of lamin B1 and emerin in human myoblasts lacking DAG1. Reasons behind this disparity may include origin of cells or methods of measurement and quantification of nuclear morphology. Alternatively, the requirement for clonal expansion may mean that phenotypic disparities arise as an artefact within the clonal cell populations. It was thought, however, that the possibility for clonal artefacts was minimised by the analysis of multiple clones.

Unlike the current study, which uses human myoblasts to model dystroglycan deficiency, previous publications use C2C12 mouse myoblasts to probe the cell biology of dystroglycan (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). This may be relevant since there are biomechanical differences between mice and

humans pertaining to whether there are compensatory mechanisms to deal with increased size (Partridge, 2013; Manning and O'Malley, 2015). The striking nuclear abnormalities in primary DAG1-null human fibroblasts (Jacobs, 2017 (PhD thesis)) would indicate that the function of dystroglycan in maintaining nuclear morphology is conserved between mouse and human cells, although whether there are compensatory mechanisms at the cellular level for stressed cell-types in larger organisms is not known. Notably, the current study did not comprehensibly interrogate other NL components, though further experiments in the Winder laboratory have established that the levels of lamins A/C and B2 are also unchanged in multiple clones of DAG1 knockout myoblasts (Ben Stevenson, personal communication), which is consistent with the normal nuclear morphology of these cells. In addition, the stiffness of nuclei as measured by atomic force microscopy is equivalent in control and DAG1 KO myoblasts (Ben Stevenson, personal communication). It is perhaps not unexpected that nuclear stiffness is unchanged, since even in C2C12 cells where depletion of dystroglycan causes decrease of lamin B1, lamin A and C levels remain unchanged (Martínez-Vieyra *et al.*, 2013). Lamins A and C are the major determinant for NL integrity (Lammerding *et al.*, 2006).

The current study sought to identify a quantifiable difference in nuclear morphology, using shape descriptors as calculated using masks of nuclei defined by Hoechst staining. This approach is markedly different from previous studies which quantified morphological differences to nuclei in dystroglycan deficient cells by observation, which carried out unblinded, possibly introduces confirmation bias (Martínez-Vieyra *et al.*, 2013; Jimenez-Gutierrez *et al.*, 2020). It is also possible that the shape descriptors used are not sensitive enough to detect subtle changes to morphology. Cells derived from HGPS patients have extremely deformed nuclei which partially contribute to disease progression (Capell and Collins, 2006). However, unbiased, reproducible quantification of this trait is often lacking (Eriksson *et al.*, 2003; Goldman *et al.*, 2004; Capell *et al.*, 2005; Mallampalli *et al.*, 2005; Driscoll *et al.*, 2012). Appropriately, this problem has been previously addressed by instead extracting the nuclear boundary, delineated by lamin A/C staining, and measuring the proportion of positive and negative curvature in the membrane (Driscoll *et al.*, 2012). To elucidate whether there are more subtle changes in the morphology of

DAG1 KO human myoblasts, it may be worth performing a similar experiment, repeating shape analysis based on a NL stain and calculating membrane curvature instead.

4.3.2.2. *Nucleus-centrosome distance*

In many cell types, the centrosome is located at the convergence of the radial microtubule array. This microtubule organising centre (MTOC) is often located in close proximity to the nucleus. The proper linkage between the nucleus and centrosome is important for nuclear positioning within a cell in preparation for polarised migration (Luxton *et al.*, 2010). Emerin is conventionally known as an INM protein, however it is also thought to reside on the ONM where it can mediate the direct linkage between the nucleus and centrosome (Salpingidou *et al.*, 2007; Chang *et al.*, 2013). In C2C12 cells, loss of β-dystroglycan seems to disrupt this tethering (Martínez-Vieyra *et al.*, 2013). If β-dystroglycan and emerin function synergistically to mediate the attachment of the centrosome to the nucleus, then it might be expected that β-dystroglycan also resides on the ONM. However, ONM-localised β-dystroglycan has not been previously observed. Nevertheless, there is evidence that β-dystroglycan is capable of binding emerin (Martínez-Vieyra *et al.*, 2013), perhaps through their respective transmembrane domains (Gómez-Monsivais *et al.*, 2020). In contrast to previous studies, results indicating increased nuclear-centrosome distance in dystroglycan-deficient human myoblasts generated in this study were equivocal since dystroglycan expressing clone 1.B8 also displayed a slight increase. To investigate if the nucleus-centrosome distance is indeed increased in human myoblasts lacking β-dystroglycan, further KO and control clones should be used to exclude the possibility of clonal artefacts on this parameter. While levels of emerin appear to be unchanged in the DAG1 KO human myoblasts, it could be hypothesised that nucleus-centrosome tethering would not be altered. Nevertheless, processes governing the balance of emerin at the INM or ONM are not fully understood and so it is possible that this could affect nuclear-centrosome distance.

The implication of an increased distance between the nucleus and centrosome in myoblasts remains unclear, though it has been observed in a number of muscle wasting disorders (Taranum *et al.*, 2012; Meinke *et al.*, 2014). The nucleus is thought

to be associated with the centrosome via dynein, where the nucleus acts as a microtubule cargo (Burakov and Nadezhdina, 2013), and it is likely connected to the NL through LINC complexes and NPCs (Agircan, Schiebel and Mardin, 2014). Muscle development requires precisely choreographed nuclear movements to form elongated, syncytial myofibres with evenly spaced nuclei. Nuclear migration in developing myotubes is dependent upon a drastic reorganisation of the cytoskeleton, including the MTOC relocating to surround the NE (Cadot, Gache and Gomes, 2015), and improperly positioned myonuclei can give rise to a number of diseases (Bitoun *et al.*, 2005; Metzger *et al.*, 2012). Once a myoblast has exited the cell cycle, it becomes competent to fuse with other cell, and crucially, centrosomal components become redispersed around the NE, which then acts as the MTOC for subsequent movement (Cadot, Gache and Gomes, 2015). Initially, newly-fused myoblast nuclei are rapidly shuttled towards the centre of the developing or regenerating myofibre in a microtubule and dynein dependent manner (Gomes and Cadot, 2017). Nuclei are subsequently distributed evenly along the length of the myofibre (Metzger *et al.*, 2012; Wilson and Holzbaur, 2012); a process driven by kinesin motors, themselves dependent on microtubules nucleated by Akap450 which is anchored to the NE through Nesprin-1α containing LINC complexes (Gimpel *et al.*, 2017). The precise molecular mechanisms behind the role of emerin, and perhaps dystroglycan, in centrosome tethering and function through muscle development remain to be fully elucidated, however it may be that improper connection inhibits centrosomal relocation to the NE during myogenesis. Indeed, expression of EDMD-related Sun1 mutants in myotubes leads to decreased emerin levels, disrupted emerin binding and a depletion of the NE MTOC (Meinke *et al.*, 2014). In the context of dystroglycan, future work might also focus on the efficiency of MTOC relocalisation to the NE in differentiating DAG1 KO myoblasts. Even though the difference in nuclear-centrosome distance found in DAG1 knockout myoblasts was not consistently significantly greater, there is a visible trend upwards in both mutant clones analysed (1.B6 and 2.D8) compared with the untargeted and CRISPR/Cas9 unaffected controls (WT and 1.B8, respectively). Comparing the data gathered in this experiment to previous studies reveals that the effect on nuclear-centrosome distance produced by a DAG1 ablation in human myoblasts is relatively

mild. Notably, the increase nucleus-centrosome distances reported in C2C12 lines depleted for DAG1 are typically lower than those reported for other cell types (Martínez-Vieyra *et al.*, 2013; Gómez-Monsivais *et al.*, 2020), which includes those derived from patients with muscular dystrophy rooted in NL protein mutations, or in cells specifically disrupted for LINC complex components or microtubule motor proteins (Salpingidou *et al.*, 2007; Morgan *et al.*, 2011; Schneider *et al.*, 2011; Meinke *et al.*, 2014) (TABLE). Combining the data shown here with literary observations may therefore imply that that dystroglycan is a contributory factor to the nucleus-centrosome tethering, downstream of primary determinants including members of the LINC complex, NL and microtubule motor proteins (Salpingidou *et al.*, 2007; Morgan *et al.*, 2011; Schneider *et al.*, 2011; Meinke *et al.*, 2014; Gimpel *et al.*, 2017).

Condition	Gene affected	Cell type	Nuclear-cytoplasmic distance (um)	Reference
Normal		Human dermal fibroblasts	1.53-1.55	(Salpingidou *et al.*, 2007)
X-linked EDMD	EMD	Human dermal fibroblasts	2.94-3.77	(Salpingidou *et al.*, 2007)
Normal		Myoblasts derived from primary patient cells	2.07	(Meinke *et al.*, 2014)
EDMD	Sun1	Myoblasts derived from primary patient cells	4.34	(Meinke *et al.*, 2014)
	Normal	Human aortic endothelial cells	1.42	(Morgan *et al.*, 2011)
	Si Nesprine 3	Human aortic endothelial cells	2.09	Morgan *et al.*, 2011)
	Control siRNA	HaCaT	1.52-1.66	(Schneider *et al.*, 2011)
	siRNA KLC1	HaCaT	4.09	(Schneider *et al.*, 2011)
	siRNA Nesprin 2G	HaCaT	3.63	(Schneider *et al.*, 2011)
Normal		Primary mouse fibroblasts	2.13	(Schneider *et al.*, 2011)
Nesprin 2G KO		Primary mouse fibroblasts	3.66	(Schneider *et al.*, 2011)
	Control RNAi	C2C12	~0.8	(Martínez-Vieyra *et al.*, 2013)
	DAG1 RNAi	C2C12	~1.8	(Martínez-Vieyra *et al.*, 2013)
		C2C12	0.88	(Gómez-Monsivais *et al.*, 2020)
	DAG1 KO	C2C12	2.62	(Gómez-Monsivais *et al.*, 2020)

Table 4-2 - Distances between the nucleus and centrosome previously identified. Studies have also identified increased distance between the nucleus and centrosome in conjunction with DAG1 aberrations and those affecting NL components and microtubule motor proteins.

4.3.2.3. *Dystroglycan and senescence*

A recent report suggested that the CRISPR-mediated ablation of dystroglycan in C2C12 myoblasts leads to senescence (Jimenez-Gutierrez *et al.*, 2020). Interestingly, one DAG1 KO human myoblast line, clone 1.G4, generated and analysed in depth through the current study also displayed features of senescence with a decrease in growth rate, and an enlargement of both cell and nuclear area. Notably, however, other passage-matched DAG1 ablated cells, 1.B6 and 2.D8, continued to proliferate normally, and did not show signs of senescence. Similarly,

the proliferation rate of the initial DAG1 KO myoblast clone 11A was not significantly different from either WT or the passage-matched, clonally expanded DAG1-expressing control line 10D. It therefore seems reasonable to conclude that, in contrast to previous studies, depletion of dystroglycan in human myoblasts does not lead to senescence. In C2C12 cells depleted for dystroglycan either by RNA interference or genetic aberration, lamin B1 levels also decrease. Indeed, the loss of lamin B1 is a marker of senescence (Vergnes *et al.*, 2004; Freund *et al.*, 2012), however whether it is an instigator or a by-product of senescence has been the subject of debate. Both lamin B1 depletion or overexpression have been reported to promote senescence in WI-38 or primary fibroblasts, respectively (Shimi *et al.*, 2011; Barascu *et al.*, 2012). More recent studies revealed that while lamin B1 depletion slows proliferation, additional environmental pressures, such as growth in a low density environment, can induce senescence in cultured cells (Dreesen *et al.*, 2013). This indicates that cells with low lamin B1 levels are not necessarily senescent. Nevertheless, Jimenez-Gutierrez *et al.*, (2020) suggest that senescence caused by the loss of dystroglycan is caused by genomic instability, despite the p53-p21 pathway not being consistently activated in the dystroglycan clones analysed. A significant technical issue with the Jimenez-Gutierrez *et al.*, (2020) study is that a passage-matched, clonal, dystroglycan-expressing cell line is not used as a control; rather the wild-type C2C12 population. Clonal expansion of cell lines is notoriously heterogeneous, thus phenotype analysis should include at least three strains, with appropriate, identically treated controls or, ideally, rescue experiments. This is particularly prominent for the analysis of the susceptibility to DNA damage on a dystroglycan-null background, since lines established from a single cell are likely to have differing sensitivities to the DNA-damage agent; in this case thymidine. Nevertheless, in the context of senescence, re-expression experiments may be difficult to execute due to the irreversible nature of the phenotype, unless early passage cells are transformed with a rescue construct before the onset of senescence. It would be informative to investigate whether dystroglycan depletion has any application to normal physiological aging. This could be achieved by staining muscle tissue samples in organisms of varying age.

4.3.3. *Implications of using a myoblast model*

Myoblasts have been extensively used to model muscle biology *in vitro*, since they are proliferative and retain many hallmarks and markers of muscle lineage. In addition, they retain the capability to differentiate into myotubes which display markers of maturation seen *in vivo*, and some characteristics such as spontaneous contractile events. However, it may be important to recognise that they reflect a distinct, proliferative stage during muscle regeneration: after commitment to the myogenic programme, but before cell cycle exit (Tedesco *et al.*, 2010). Notably, the muscle specific DGC cytoskeletal linker component dystrophin is not expressed in myoblasts, but becomes expressed upon maturity in myofibres (Miranda *et al.*, 1988; Huard *et al.*, 1991). In adult muscle tissues, satellite cells, which are the muscle progenitor stem cells, are niched towards the edge of muscle fibres (Mauro, 1961), and usually maintained in a mitotically quiescent state. Injury to the muscle causes the activation of satellite cells, which then proliferate and differentiate through distinct stages into proliferative myoblasts, then myocytes before cell cycle exit and fusion with existing muscle fibres, regenerating the damaged tissue (Tedesco *et al.*, 2010). Perpetual muscle injury and regeneration is a common pathological mechanism in muscle wasting diseases like DMD. The loss of dystrophin results in the instability of the DGC at the sarcolemma, which usually provides structural integrity to the myofibre by directly connecting the ECM to the cytoskeleton, but also accumulates diverse molecules mediating signalling functions (see also 1.10). Interestingly, studies in mice specifically depleted for DAG1 in mature muscle using an MCK-Cre display a remarkably mild dystrophic phenotype. A global dystroglycan KO reveals a much more severe dystrophy, indicating that satellite cell function is compromised in the absence of DAG1 (Cohn *et al.*, 2002). Dystrophin, and by extension, the DGC, expression in satellite cells was only recently characterised, presumably given the difficulty of working with this cell type (Chang, Chevalier and Rudnicki, 2016). Recent work has identified that dystrophin is actually highly expressed in satellite cells and becomes polarised to the apical surface upon activation (Dumont *et al.*, 2015). Together with dystroglycan, the satellite cell DGC is responsible for recruiting factors which determine the non-myogenic committed

178

replicative daughter cells, maintaining the stem cell population (Chang *et al.*, 2018) (see also 1.3.2). Since myoblasts have reduced dystroglycan expression, it may be that these cells do not overtly depend on dystroglycan for their normal function during myogenesis, as such, nuclear phenotypes analysed in human dystroglycan-null myoblasts may be diminished. For example, cell division multipolarity and amplification of centrosomes has been identified in satellite cells of mdx mice, perhaps due to decrease and failed asymmetry of Mark2, a kinase normally recruited by dystrophin (Kwon *et al.*, 2008; Dumont *et al.*, 2015). Given myoblast proliferation is not an asymmetric division, it is not clear whether Par-1 is required for limiting centrosome number and position. Thus, further research is required to determine why dystroglycan depleted and ablated myoblasts have centrosome amplification and multipolarity (Martínez-Vieyra *et al.*, 2013; Jimenez-Gutierrez *et al.*, 2020), and the downstream implication of this.

It seems worth highlighting that while satellite cells from *mdx* mice do not commit to the myogenic programme, they do continue to proliferate in spite of supposed decreased dystroglycan levels, resulting from the loss of dystrophin (Reimann, Irintchev and Wernig, 2000; Kottlors and Kirschner, 2010). This observation is not consistent with that of dystroglycan depletion driving senescence in myoblasts (Jimenez-Gutierrez *et al.*, 2020). Decreased dystroglycan in tissues from *mdx* and the zebrafish DMD model, *sapje,* was determined by analysis of whole muscle lysates, so further investigations are necessary to identify if there are differences in dystroglycan stability across myogenic progenitor cell-types.

4.3.4. *Conclusions and future directions*

In conclusion, results presented in this chapter described the derivation of a human myoblast cell line genetically ablated for DAG1 which is viable, but does not exhibit nuclear phenotypes previously associated with the depletion of dystroglycan. Further research to establish the relationship between dystroglycan and nuclear-centrosome distance is required, whether dystroglycan is involved in NE MTOC relocalisation during myogenesis, while re-expression studies would identify whether this is a specific effect of DAG1.

Chapter 5: Generation and characterisation of a novel full-length, triple epitope tagged dystroglycan construct.

5.1. Introduction

The β-subunit of dystroglycan is often considered a single entity, but there is a growing body of evidence implicating fragments of β-dystroglycan, thought to arise from proteolysis, in diverse cellular functions. While the full length dystroglycan protein is conventionally understood to reside at the plasma membrane, bridging the ECM to the actin cytoskeleton via its non-covalent association with its α-counterpart, β-dystroglycan can fragment into at least three distinct species. Functions for specific fragments remain somewhat elusive, but it seems that all fragments can be liberated from their established plasma membrane anchorage allowing further roles in other compartments. However, the deficiency of appropriate tools to dissect the multifaceted roles for dystroglycan fragments currently impedes progress. The only reliable antibodies raised against β-dystroglycan are directed at the carboxy-terminus which appears to be preserved in all the most stable β-dystroglycan species, complicating interpretation, while previous strategies for tagging do not always preserve full functionality.

5.1.1. *Fragmentation of β-dystroglycan*

β-dystroglycan can be cleaved within its ectodomain by extracellular matrix metalloproteases (MMP) 2 and 9, which is thought to give rise to a C-terminal fragment of ~31 kDa (Figure 5-1) (Zhong *et al.*, 2006; Bozzi *et al.*, 2009; Sbardella *et al.*, 2012). Cleavage at this site has been implicated in numerous disease processes. In cancer, generation of the ~31 kDa β-dystroglycan species is associated with increased cell migration, perhaps because it reduces cell-matrix contacts (Losasso *et al.*, 2000). If the 31 kDa β-dystroglycan fragment is considered non-functional, these observations are consistent with dystroglycan overexpression which increases cell

attachment and suppresses tumourigenicity, although migration was not explicitly tested in this study (Sgambato *et al.*, 2004). Indeed, defective engagement of the ECM by dystroglycan appears to be common in the function of cancerous cell lines (Yamada *et al.*, 2001; Singh *et al.*, 2004; Mitchell *et al.*, 2013). In addition to cancer, MMP processing of dystroglycan has been implicated in the normal functioning of the nervous system. β-dystroglycan is localised to post-synaptic plates in neuronal cells of the central nervous system. An MMP9-mediated ~31 kDa β-dystroglycan fragment is rapidly produced within 10 minutes of glutamate stimulation of cultured neurones, or within 5 minutes of an induced seizure in mice (Michaluk *et al.*, 2007). A role for β-dystroglycan fragmentation has also been implicated in the maintenance of Schwann cells which form the myelin sheath of the peripheral nervous system. MMP activation is common in autoimmune neuritis, a condition which leads to the loss of myelin sheath, and MMPs 2 and 9 cleave β-dystroglycan (Zhao *et al.*, 2010). Finally, dystroglycan contributes to the integrity of the blood-brain-barrier, and cleavage by MMP2 and 9 comprises this, leading to infiltration by immunogenic cells and inflammation (Agrawal *et al.*, 2006).

However, MMP-mediated reduction in ECM-dystroglycan association may not be an isolated pathomechanism for disease aetiology described. Dystroglycan overexpression in MCF7 mammary tumour cells restores β-, but not α-dystroglycan, and these cells have reduced tumourigenic potential (Sgambato *et al.*, 2004), indicating that β-dystroglycan does have functions beyond that as an anchor for cell-ECM attachment. The carboxyl-terminus of dystroglycan is rich in protein-protein interaction motifs and is able to signal through multiple pathways. It is therefore possible that releasing the C-terminal region from its plasma membrane anchorage may not simply be non-functional, but actually have other effects throughout the cell.
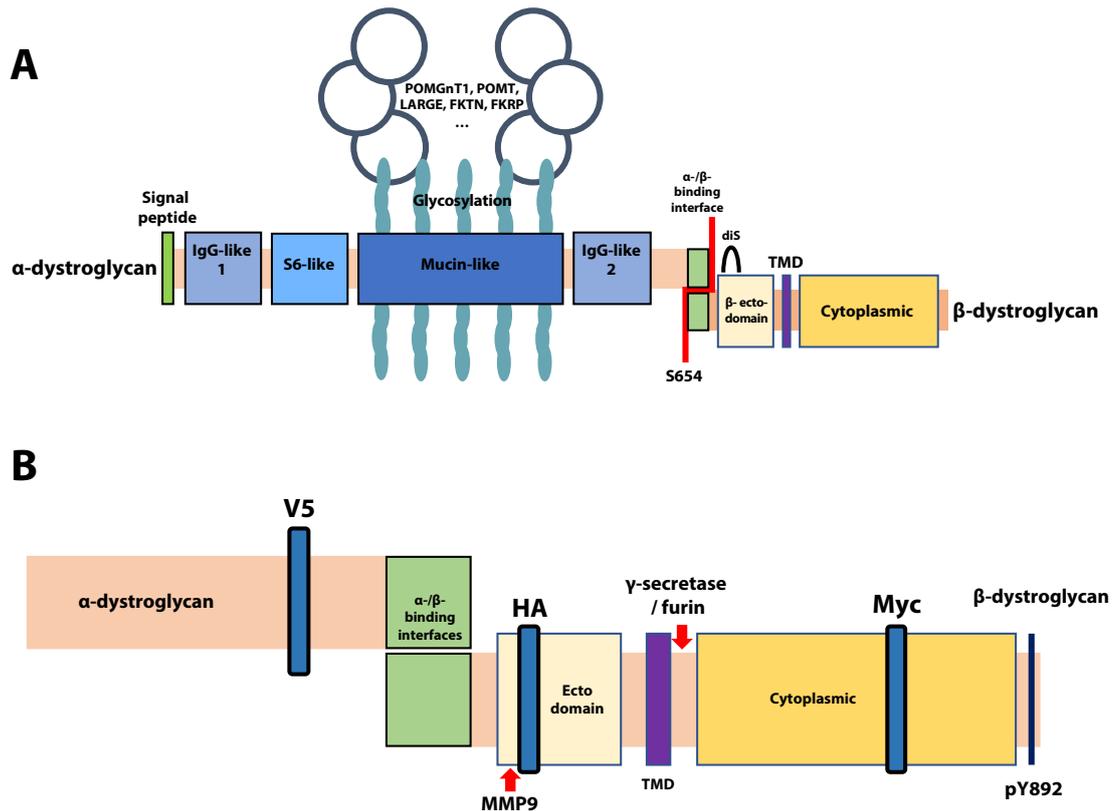
Figure 5-1 - Schematic depicting the dystroglycan pre-peptide showing location of epitope tags in relation to known cleavage sites. (A) Full length dystroglycan. The mucin-like domain is extensively modified by glycosyltransferases. A sea-urchin-enterokinase autoproteolysis (SEA) domain separates the α- and β- subunits which remain non-covalently bound. A disulphide linkage (diS) between Cys669 and Cys713 within β-dystroglycan stabilises SEA domain function and heterodimerisation. (B) Enlargement of β-dystroglycan and C-terminal of α-dystroglycan, indicating points of cleavage by metalloproteases and γ-secretase/furin, together with relative locations of epitope tags V5, HA and Myc. TMD = transmembrane domain.

In accordance, β-dystroglycan can be cleaved in its juxtamembrane region on the intracellular side. Its identity as a C-terminal fragment of β-dystroglycan is confirmed by its detection using antibodies raised against the extreme C-terminal region, and an apparent mass of ~26 kDa as determined by western blotting (Figure 5-1) (Mathew *et al.*, 2013; Mitchell *et al.*, 2013; Leocadio, Mitchell and Winder, 2016). Similarly to the 31 kDa fragment, the 26 kDa β-dystroglycan species has been associated with cancer, and thought to be translocated to the nucleus in a phosphorylation-dependent manner in prostate cancer cells (Mathew *et al.*, 2013). Within the nucleus, 26 kDa β-dystroglycan is thought to direct gene transcription (Mathew *et al.*, 2013). Pharmacological approaches have determined that the 26 kDa β-dystroglycan can be generated by furin and γ-secretase activity (Leocadio, Mitchell and Winder, 2016). Interestingly, and in opposition to previous reports in LNCaP metastatic prostate cancer cells (Mathew *et al.*, 2013) Leocadio et al., (2016) found that 26 kDa β-dystroglycan both Y890 phosphorylated and non-phosphorylated localised robustly to the nuclear fractions (Leocadio, Mitchell and Winder, 2016).

5.1.2. *Dystroglycan trafficking*

It is clear therefore, that dystroglycan has multiple roles beyond those at the plasma membrane, tethering the cell to the ECM via cortical actin filaments and laminin binding. How different pools of dystroglycan are specified for specific functions, however, remains elusive. Recent reports indicate intracellular dystroglycan and its fragments are entirely derived from the plasma membrane, where it is first trafficked, following translation (Gracida-Jiménez *et al.*, 2017). Consistent with these observations, enzymes thought to be involved in proteolytic processing of β-dystroglycan reside extracellularly, in the case of MMPs, and furin accumulates at the plasma membrane (Anderson *et al.*, 2002), thus not interfering with β-dystroglycan structure prior to its cell surface positioning. γ-secretase, on the other hand, is positioned throughout the secretory pathway (Kim *et al.*, 2004), so it is not clear how β-dystroglycan is protected from processing during its anterograde transit. However, inhibition of γ-secretase using DAPT results in a reduced, albeit persistent fragmentation of β-dystroglycan, indicating the existence of other pathways (Leocadio, Mitchell and Winder, 2016). As mentioned, retrograde

transport occurs after plasma membrane localisation, and specific tracing of full length, 43 kDa, β-dystroglycan from the cell surface suggests a transit pathway through dynamin-dependent endocytosis to the ER and then the nucleus (Gracida-Jiménez *et al.*, 2017). How β-dystroglycan is imported into the nucleus is not clear, though β-dystroglycan was found to co-precipitate with the Sec61β translocon subunit, possibly in order to extract the protein from the membrane (Gracida-Jiménez *et al.*, 2017) followed by importin recognition and nuclear import (Oppizzi *et al.*, 2008; Lara-Chacón *et al.*, 2010). However, subsequent handling of membrane associated β-dystroglycan is not clear. Thus far, fragments of β-dystroglycan have not been considered. How these are sorted and specified for various functions is also elusive.

β-dystroglycan can be modified by phosphorylation at tyrosine 890, which has been suggested to have various functional consequences. Y890 phosphorylation is thought to ablate the binding of dystroglycan to dystrophin or utrophin (James *et al.*, 2000), and create an SH2-domain binding site (Sotgia *et al.*, 2001). In the absence of dystrophin or utrophin, it is also thought that the C-terminal region of β-dystroglycan can recruit SH3-domain containing proteins, including Grb2 (Yang *et al.*, 1995; Russo *et al.*, 2000), which may act as an adapter for other signalling cascades. The SH2 and SH3 regions are adjacent and to some extent overlap (Moore and Winder, 2010). One interpretation of the observations that both SH2 and SH3-domain proteins can interact with the carboxyl-terminus of β-dystroglycan is that this region functions as an intersection for signalling (Thompson *et al.*, 2010). How these interactions, and thus signalling outcomes, are regulated, however, remains unclear. More recently, phosphorylation at Y890 was shown to precede internalisation, polyubiquitination and degradation; an important molecular event in the aetiology of DMD since dystrophin no longer protects the residue from tyrosine phosphorylation (Lipscomb *et al.*, 2011; Miller *et al.*, 2012). Indeed, non-phosphomimetic amino acid substitutions and treatment with tyrosine kinase inhibitors aimed at preventing phosphorylation at Y890 retain dystroglycan at the plasma membrane and alleviate muscle wasting in the *mdx* mouse and *sapje* zebrafish models of DMD (Miller *et al.*, 2012; Lipscomb *et al.*, 2016). In addition, both non-phosphorylated and

phosphorylated Y890 β-dystroglycan and its fragments have been observed in nuclear fractions of various cell types (Mathew *et al.*, 2013; Leocadio, Mitchell and Winder, 2016; Azuara-Medina *et al.*, 2019). Provided that dystroglycan first transits to the plasma membrane before internalisation or nuclear localisation (Gracida-Jiménez *et al.*, 2017), it is not clear how tyrosine phosphorylation can be an exclusive signal for internalisation, further signalling functions and degradation.

5.1.3. *Aims and objectives*

Current evidence shows that β-dystroglycan processing beyond the plasma membrane is complex. A number of factors, including proteolytic fragmentation and phosphorylation of Y890 direct dystroglycan to various fates, including degradation and nuclear import. Yet current antibody tools fail to differentiate these β-dystroglycan species given their maintained C-terminal epitope, prohibiting dissection of specific functions. With this mind, the results presented in this chapter aim to begin unravelling the complexities of different β-dystroglycan species by generating reagents to explicitly detect β-dystroglycan fragments. One way to achieve this would be to raise antibodies against the amino-terminus of known fragments. However, this approach may be complicated by the unknown immunogenicity of β-dystroglycan peptides. Instead, internal epitope-tagging was considered to be a simpler approach, such that β-dystroglycan fragments could be detected and differentiated using high quality antibodies raised against the epitope-tags.

5.1.3.1. *Construct design*

The designed construct was to be modified such that various known protein fragments of C-terminal dystroglycan could be identified. Therefore, three epitope tags were selected, allowing the detection of at least three proteins derived from dystroglycan. The α-subunit incorporated a V5-tag, derived from the simian virus 5 (Hanke, Szawlowski and Randall, 1992), while the β-subunit received a HA-, from hemagglutinin protein of influenza (Green *et al.*, 1982; Wilson *et al.*, 1984), and myc-tag, a sequestered sequence from the c-myc oncoprotein (Evan *et al.*, 1985) (Figure 5-1B). The latter HA- and myc-tags are positioned such that their detection corresponds to the full-length fragment and cytoplasmic fragment, respectively, and

avoids disrupting currently known structural domains, interaction motifs and sites of post-translational modification (Figure 5-1B). Dystroglycan exhibits a high degree of evolutionary conservation, so to avoid modifying important regions, dystroglycan amino acid sequences from 13 species were aligned and epitope-tag insertion points were selected adjacent to non-conserved amino acids (Figure 5-2).

5.1.3.2. *Research hypotheses:*

Given the rational design of the epitope-tagged dystroglycan construct, it was first hypothesised that the modifications would not interfere with dystroglycan function or localisation. Secondly, in order for β-dystroglycan fragments to have different subcellular functions, it was postulated that protein species become separately localised as endocytic machinery traffics different β-dystroglycan protein pools. Using current antibodies, it is not possible to test this hypothesis without developing a complicated protocol which deposits β-dystroglycan fragments in distinct subcellular structures, which could be subsequently identifiable by biochemical fractionation. The epitope-modification strategy presented here, therefore, was hypothesised to reveal specific β-dystroglycan protein species detectable by fluorescence microscopy. The ability to distinguish β-dystroglycan species was thought to reveal different fragments populating distinct subcellular structures. For instance, the cytoplasmic domain of dystroglycan has been reported to localise to the nucleus. Thus, using the generated construct, it was expected that myc would be detected in the nuclear region.
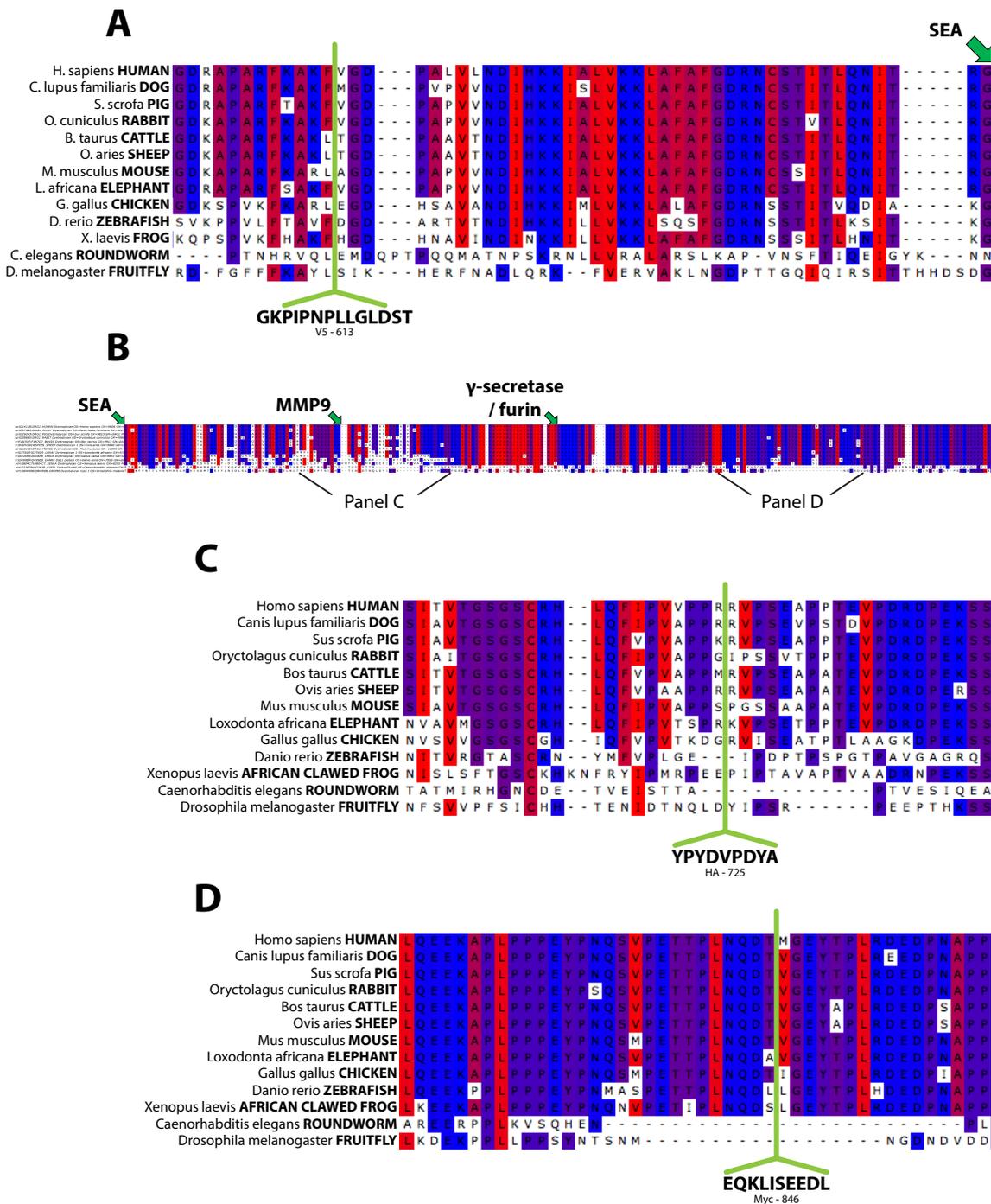
Figure 5-2 – Interspecies peptide sequence alignment of dystroglycan, indicating evolutionary conservation. (A) C-terminus of α-dystroglycan, indicating site of V5 epitope tag insertion. (B) β-dystroglycan, (C) HA insertion site in β-dystroglycan, (D) myc insertion site in β-dystroglycan.

5.2. Results

5.2.1. *Epitope tags within dystroglycan display fragmentation with distinct subcellular localisation.*

Non-physiological epitope and globular peptide tags have been extensively used to study β-dystroglycan, even though modification in this way is well documented to disrupt endogenous protein function. A FLAG-tag engineered into the peptide sequence of β-dystroglycan appears not to localise to the nucleus, whereas carboxy-terminal GFP tagged version is not phosphorylated at Y890 (Leocadio-Victoria, 2015 (PhD thesis)). To determine whether the rational addition of the three epitope tags affects the fragmentation of dystroglycan, HeLa cells were stably transduced with an inducible construct encoding αV5-HAβmyc-DG. Protein expression was induced over three days, and whole cell lysates were analysed by SDS-PAGE. Western blotting for the V5 epitope revealed the α-subunit at 100 kDa compared to its predicted mass of 74 kDa, presumably due to its extensive post-translational modification by glycosylation (Figure 5-3). Probing using antibodies raised against the HA and Myc epitopes revealed the full length β-subunit at ~50 kDa which is increased from the endogenous 43kDa protein probably due to the addition of two epitope tags (Figure 5-3). However, only when probing the western blot with the myc antibody could a band of reduced molecular weight be detected (Figure 5-3). This indicates that the reduced molecular weight β-dystroglycan species corresponds to the cytoplasmic fragment since it was not immunoreactive to HA antibodies. However, the endogenous cytoplasmic fragment is usually detected at around 26 kDa (Mitchell *et al.*, 2013; Leocadio, Mitchell and Winder, 2016), thus confounding this conclusion. It should be noted that the entire endogenous β-dystroglycan protein is predicted to measure 26.6 kDa, but is observed at 43 kDa possibly due to its high proline content. Likewise, the 11 kDa cytoplasmic fragment migrates at 26 kDa on SDS-PAGE. In addition, the HA and myc tags are calculated to add 2.3 kDa to the total molecular weight of β-dystroglycan which is inconsistent with the ~7 kDa mobility shift observed. These inconsistencies might arise due to retained secondary structure after reduction in sample buffer, or the epitope tags

may alter the biochemistry of the protein to reduce SDS binding ability. The addition of epitope tags may further alter protein parameters, or unknown post-translational modifications might be responsible for β-dystroglycan fragment regulation upon overexpression.

Nevertheless, western blot analysis indicates that αV5-HAβmyc-DG undergoes physiological autoproteolytic cleavage separating α- from β-dystroglycan, and indeed that β-dystroglycan becomes further processed. It was therefore appropriate to visualise αV5-HAβmyc-DG *in situ*, to further establish that the construct elicits recognised cellular phenotypes for β-dystroglycan overexpression which would provide additional confidence that the epitope tags do not disrupt normal dystroglycan function. To test this hypothesis, HeLa cells were transfected with either αV5-HAβmyc-DG or co-transfected with dystroglycan constructs containing each of the epitope tags separately. Both transfection conditions gave rise to exaggerated filopodial structures which is characteristic of dystroglycan overexpression and appropriate function at the plasma membrane (Y.-J. Chen *et al.*, 2003) (Figure 5-4A, B). This phenotype arises as a result of the ability of β-dystroglycan to bind ezrin at its cytoplasmic juxtamembrane region (Spence, Chen, *et al.*, 2004; Batchelor *et al.*, 2007), indicating that flanking this region with HA and Myc epitope tags does not interfere with this function. Expression in cells also allowed the testing of the original hypothesis that fragments of β-dystroglycan, which can be alternatively detected by western blotting, could be visualised separately using epitope tags *in situ*. The singly-tagged constructs were co-transfected as a means to positively control the experiment since they arise from different messages and therefore could become separately organised. Indeed, this condition showed regions of epitope-tag colocalisation as well as the separation throughout the cell, mostly as punctate structures (Figure 5-4A). In a similar manner to the expression of separately-tagged constructs, the three epitope-tags could be separately observed, arising from the same αV5-HAβmyc-DG construct (Figure 5-4B). These data therefore indicate that β-dystroglycan fragments become separately organised within the cell, possibly due to divergent functions, and this can be visualised by microscopy, *in situ*, using the epitope-tagged construct.

190
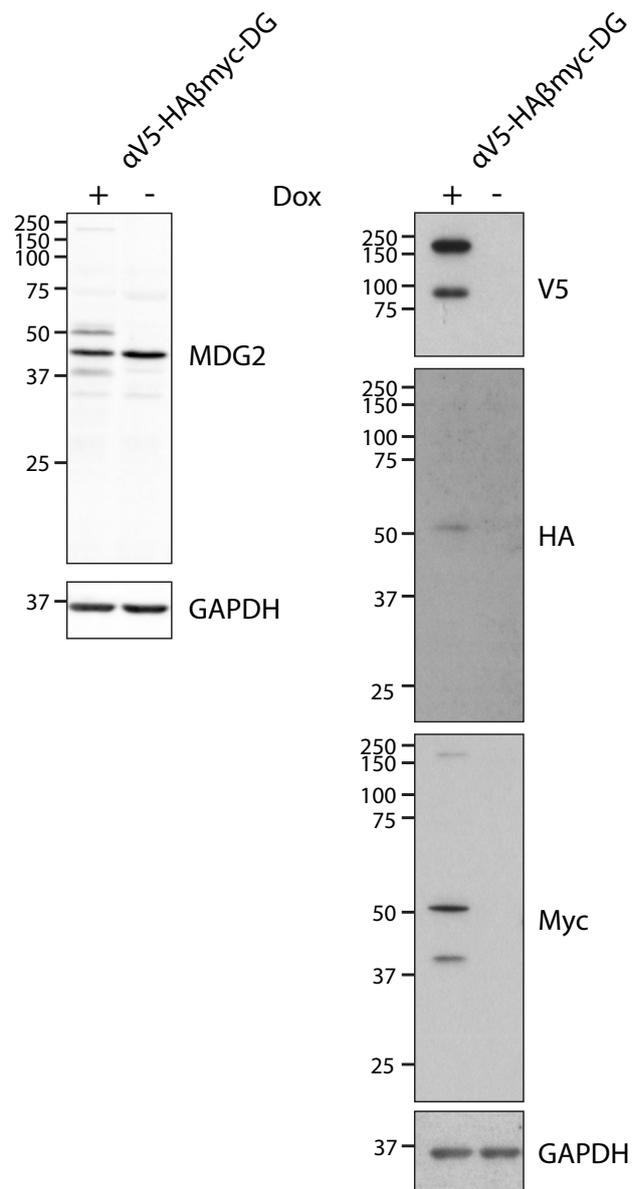
Figure 5-3 - Three internal epitope tags throughout a the single dystroglycan construct genuinely fragment. HeLa cells stably transformed with the inducible αV5-HAβmyc-DG construct were either cultured in the presence or absence of doxycycline for 3 days prior to analysis by western blotting.
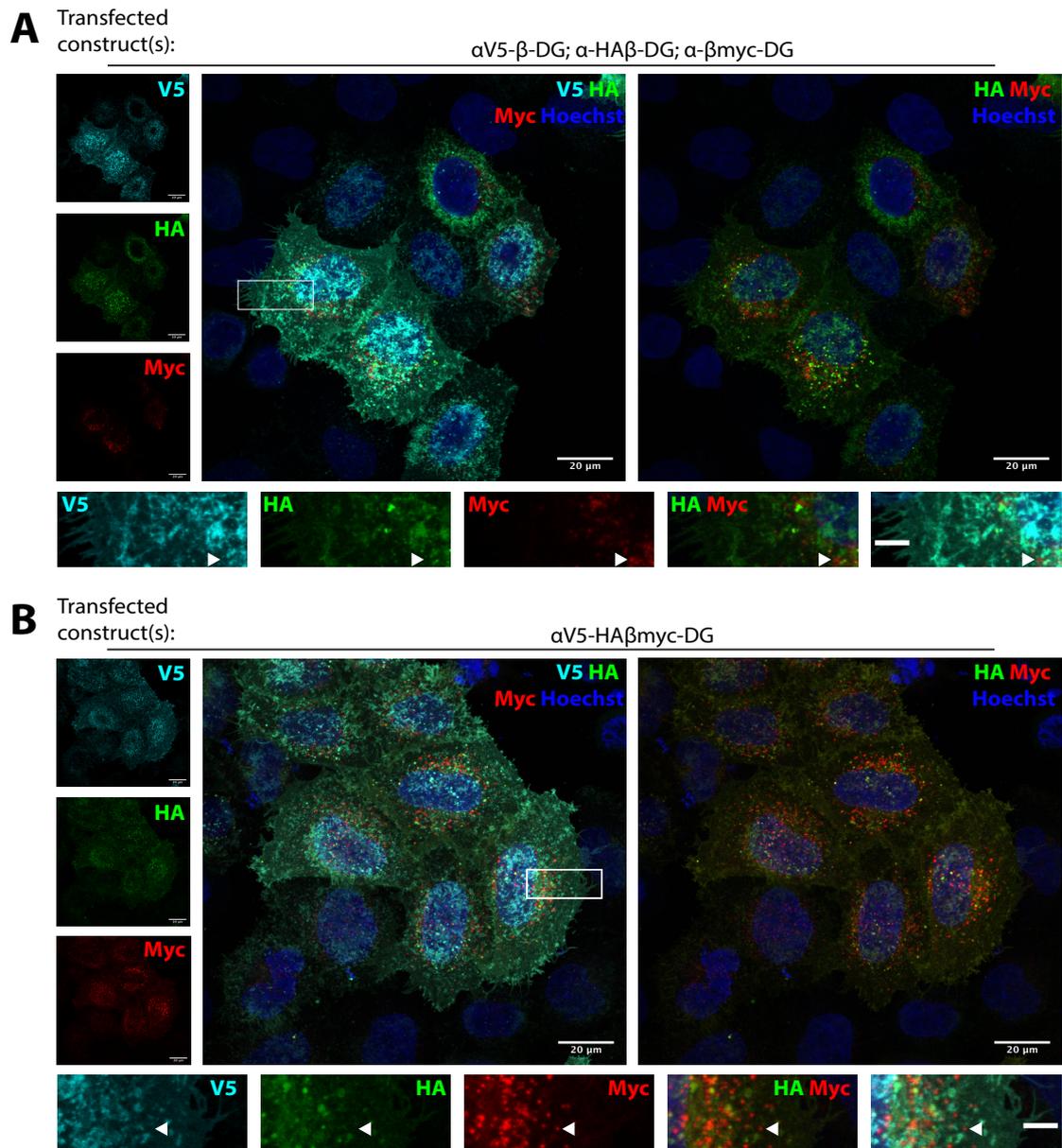
Figure 5-4 - Epitope tags within the αV5-HAβmyc dystroglycan construct become separated upon its expression in HeLa cells. HeLa cells were either triply transfected with a version of DAG1 modified by a single epitope tag (A), or with the αV5-HAβmyc construct (B). Transfected HeLa cells were probed using antibodies specific to the epitope tags through the dystroglycan construct to reveal localisation of fragments relative to one another. Arrowheads denote regions which are myc stained but not HA. Specificity of antibodies was confirmed by co-staining cells individually transfected with a construct with one epitope tag as shown in Supplementary Figure I.

192

5.2.2. *α-dystroglycan V5-tag interferes with SEA-domain mediated autoproteolysis*

Western blotting analysis of β-dystroglycan fragmentation in HeLa cells also revealed a version of dystroglycan with high molecular weight at 150 kDa (Figure 5-3). This protein species was detected using the V5 antibody, located in the N-terminal α-subunit, as well as myc and MDG2 antibodies, both of which have epitopes in the C-terminus of the β-dystroglycan. It therefore seemed possible that this species corresponds to the dystroglycan pre-peptide. Cleavage into α- and β-subunits is mediated autonomously via the autoproteolytic domain of dystroglycan, which has been previously demonstrated using a cell-free system (Akhavan *et al.*, 2008), and immunofluorescence together with western blotting indicates that the triple-epitope tagged αV5-HAβmyc-DG construct does still undergo this process. The presence of the pre-peptide is therefore surprising and implies that the kinetics of the proteolysis had been impeded. However, there are also reports of mutations which impede autoproteolysis (Signorino *et al.*, 2018).

To investigate whether one of the artificial epitope tags is responsible, αV5-HAβmyc-DG was transiently expressed in HeLa and HEK 293T cells alongside constructs encoding dystroglycan modified with each of the tags separately. The high molecular weight band is persistent in the transient αV5-HAβmyc-DG system, and it can be observed when blotting for MDG2, V5, HA and myc in both HeLa and HEK 293T cells. Of the singly-tagged constructs, MDG2 probing revealed that this immunoreactive band only persisted when the V5-tagged construct was expressed and not with either the HA- or myc-only tagged versions (Figure 5-5). These results indicate that the addition of the V5 tag after residue F613 interferes with the kinetics of the SEA autoproteolysis.

Since the V5-tag appeared to be the limiting factor in the efficient autoproteolysis of the epitope-tagged construct, and the consequences of this defect were unknown, it was posited that its removal would result in the reduction in the detection of the pre-peptide. In a similar experiment, HeLa and HEK 293T cells were transiently transfected, this time including a α-HAβmyc-DG construct. Detection of dystroglycan using MDG2 showed that the pre-peptide arising from the

αV5-HAβmyc-DG construct was still somewhat persistent in HeLa, but undetectable in HEK 293T lysates (Figure 5-6). This result indicates that the combined modifications of HA- and myc-tags may also have some effect on the autoproteolytic activity of the SEA domain, however it is reduced. It may be that the proximity of the HA epitope-tag to the autoproteolytic site is responsible.
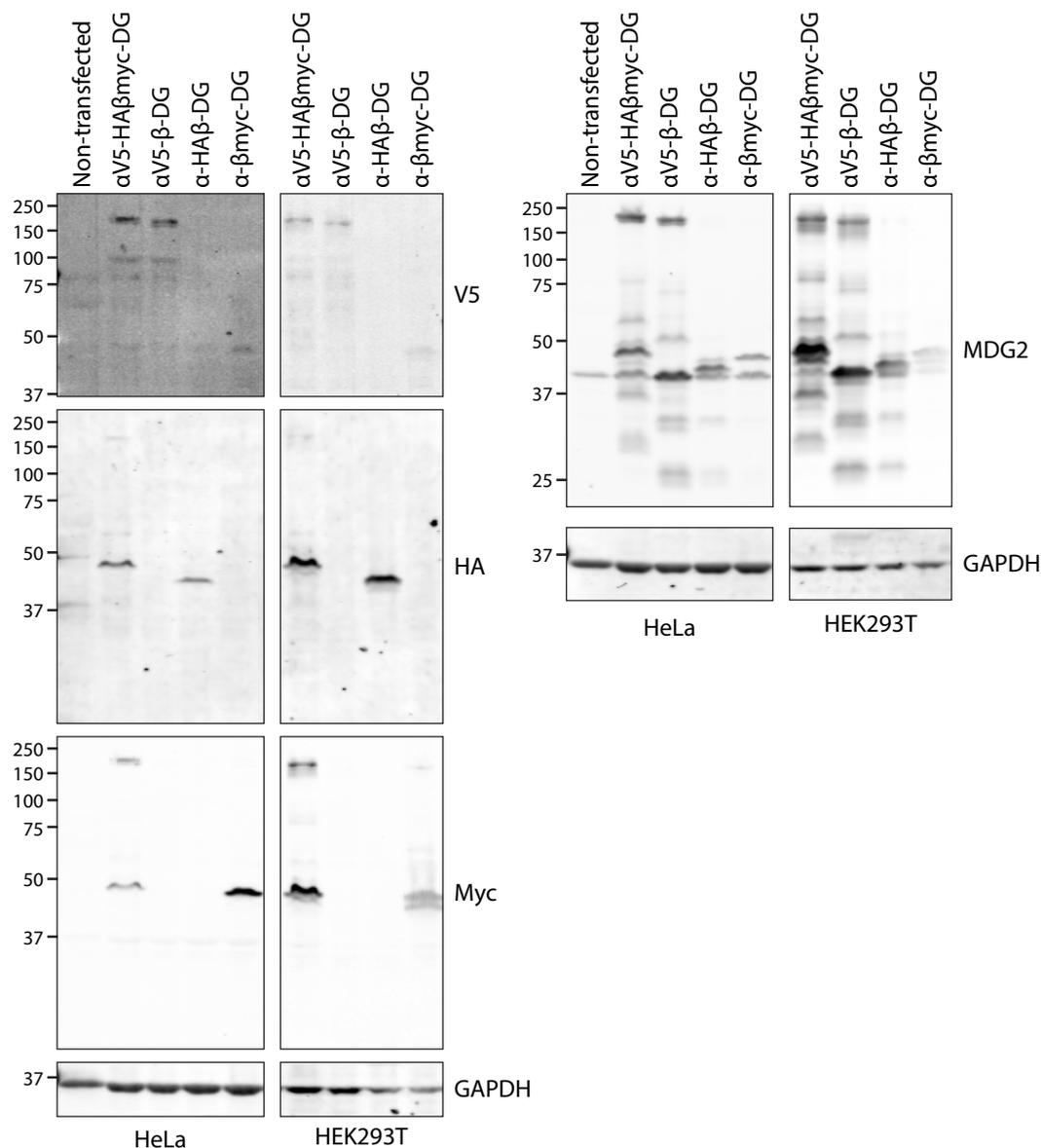
Figure 5-5 – The V5-tag in α-dystroglycan impedes autoproteolysis of dystroglycan into α- and β-subunits. HeLa and 293T cells transiently expressing the indicated constructs were subject to SDS-PAGE. Epitope tags were detected using the indicated antibodies.
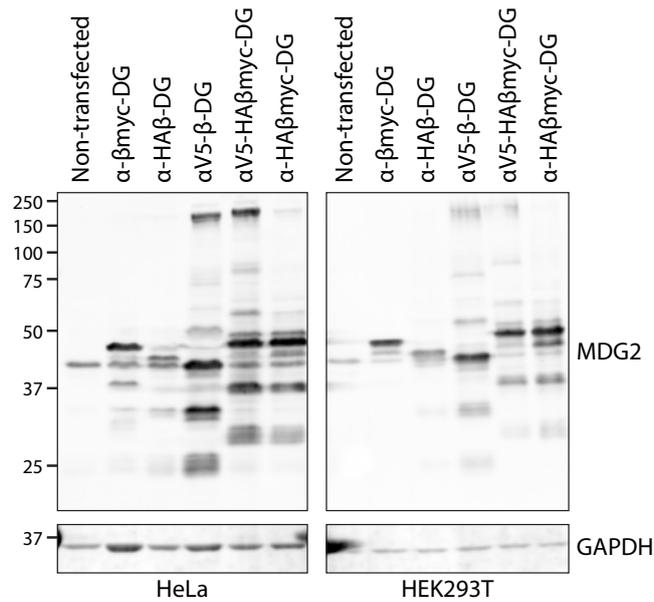
Figure 5-6 – Removal of the V5-tag improves autoproteolysis. HeLa and 293T were transfected with dystroglycan constructs with a variety of epitope tags and subject to SDS-PAGE. High molecular weight dystroglycan bands were detected using the MDG2 antibody.

5.2.3. *The V5-tag in α-dystroglycan does not appear to affect spatial fragmentation of β-dystroglycan.*

Epitope-tagging of dystroglycan has revealed that the C-terminal fragments becomes spatially separated from its N-terminal anchor *in situ*, however, why independent trafficking routes for various fragments exist, and how they become sorted remains elusive. Studies suggest that there are diverse fates for dystroglycan after plasma membrane localisation, ranging from degradation (Lipscomb *et al.*, 2016) to roles in the nucleus (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018). It has also been shown that the cytoplasmic C-terminal fragment can accumulate under experimental circumstances, for example by cell confluence (Mitchell *et al.*, 2013), kinase activation or proteasome inhibition (Leocadio, Mitchell and Winder, 2016). It was therefore considered whether previously identified alterations in dystroglycan fragmentation could be visualised *in situ* and quantified using the α-HAβmyc-DG construct. A CellProfiler pipeline was first established to separately identify regions of interest containing only HA or myc signals using confocal images of HeLa cells transiently transfected either both the αV5-HAβmyc-

DG and α-HAβmyc-DG constructs. Pearson's correlation analysis of the identified regions revealed significantly more myc signal in the absence of HA, indicating the liberation of the carboxy-terminus of β-dystroglycan (Figure 5-7). Despite the reduced autoproteolysis effect of the V5 epitope on the αV5-HAβmyc-DG construct, it appeared to have no difference on separation of β-dystroglycan fragments downstream (Figure 5-7).
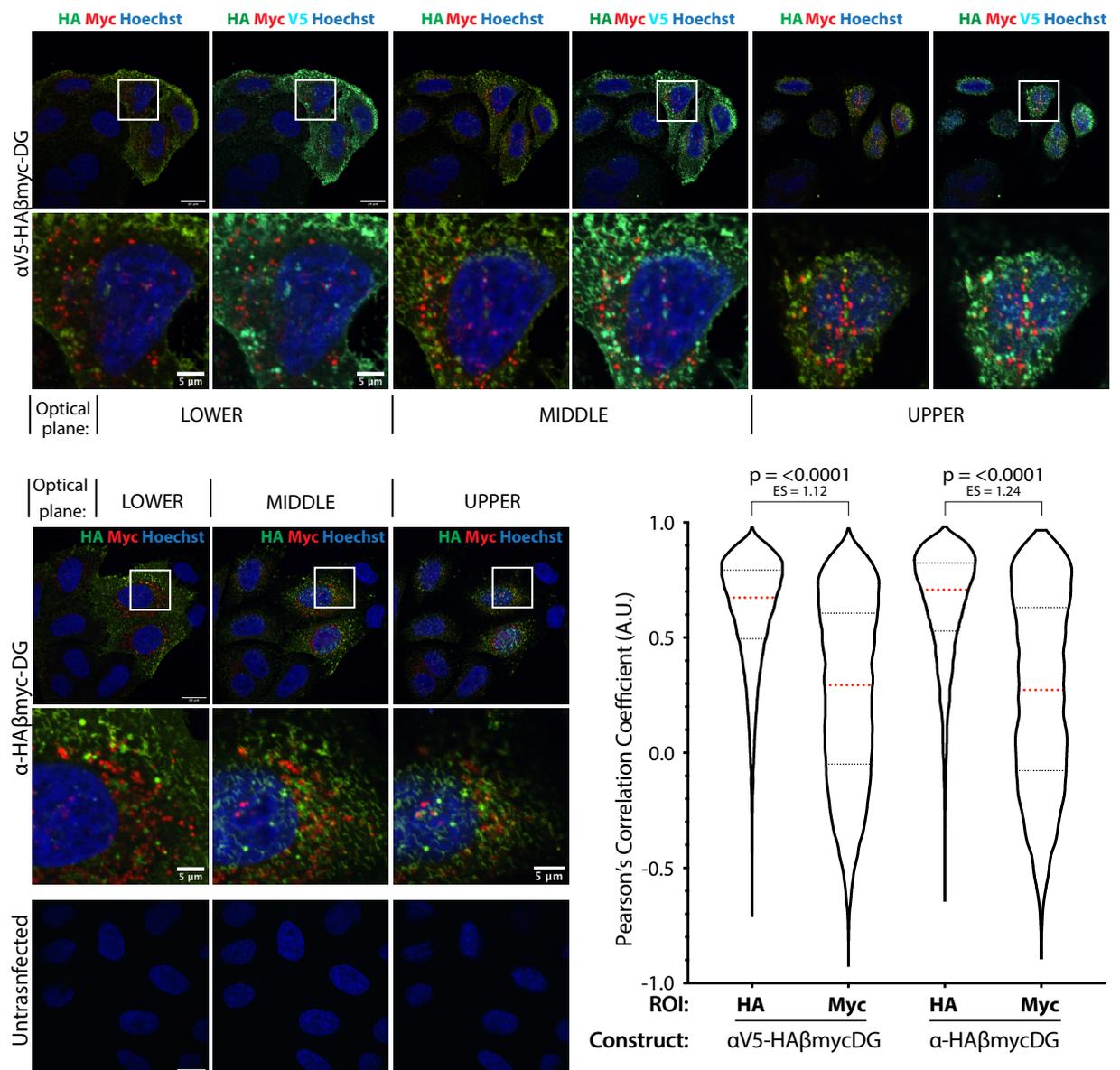
Figure 5-7 – V5 tagged α-dystroglycan does not impede further proteolytic processing of β-dystroglycan. HeLa cells were transiently transfected with either αV5-HAβmyc-DG or α-HAβmyc-DG, stained with antibodies raised against the epitope tags and analysed by confocal microscopy. Images shown are maximum intensity projections of optical sections corresponding to the lower, mid or upper planes of the cell. Violin plot represents correlation of Myc or HA puncta with the opposite signal using CellProfiler. At least 6500 puncta from at least 33 cells were identified in each condition. Statistical significance was calculated using Kruskal-Wallis Test with Dunn's post-hoc analysis for multiple comparisons. P-values shown for differences with effect size above 0.5.

5.2.4. _Increasing cell confluence does not increase separation of HA- and Myc-tags._

As previously mentioned, increased cell confluence can provoke increased β-dystroglycan fragmentation and nuclear targeting to increased fragmentation nuclear targeting (Mitchell _et al.,_ 2013; Leocadio, Mitchell and Winder, 2016). Previously, β-dystroglycan fragmentation events could not be assessed using immunofluorescence microscopy due to unavailability of appropriate antibodies. Western blotting has been useful in displaying fragmentation events, though whether fragments become independently trafficked at the subcellular level for divergent functions remains elusive. Since fragmentation appears to be quantifiable, it was therefore hypothesised that increasing cell confluence would lead to decreased co-localisation of the HA and myc tags in the α-HAβmyc-DG construct, signifying carboxy-terminal liberation. HeLa cells were transduced with α-HAβmyc-DG and selected for stable integration with the aim to yield a cell line with more moderate levels of expression. Confocal images were analysed using the same pipeline as before (Figure 5-7). Similar to the transient expression, co-localisation analysis revealed that regions of myc staining less frequently co-localised with HA, when compared with HA-stained areas, though the difference was not as pronounced (Figure 5-8). However, it seems that an increase in fragmentation cannot be detected using this method on cultures of increasing confluence (Figure 5-8). It may be that fragmentation is occurring, but without clear spatial separation. As cell confluence increases, the cytoplasmic volume decreases leading to a reduction of resolution separating individual puncta, which is essential for this experimental approach. It would therefore be more appropriate to analyse β-dystroglycan fragmentation in HeLa cells by western blotting, or by immunofluorescence microscopy in cells which retain a high cytoplasmic-nuclear volume ratio upon contact-inhibition.

Seeding density (Cells / cm²)	HA	Myc	Lamin B1	Composite + Hoechst
2.11 x 10⁴				
4.21 x 10⁴				
8.40 x 10⁴				
2.11 x 10⁴ No Dox				
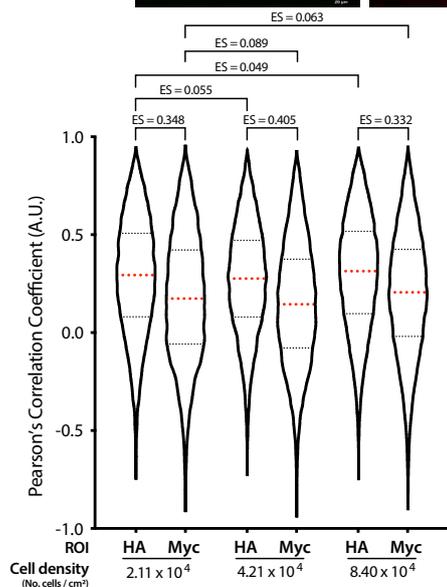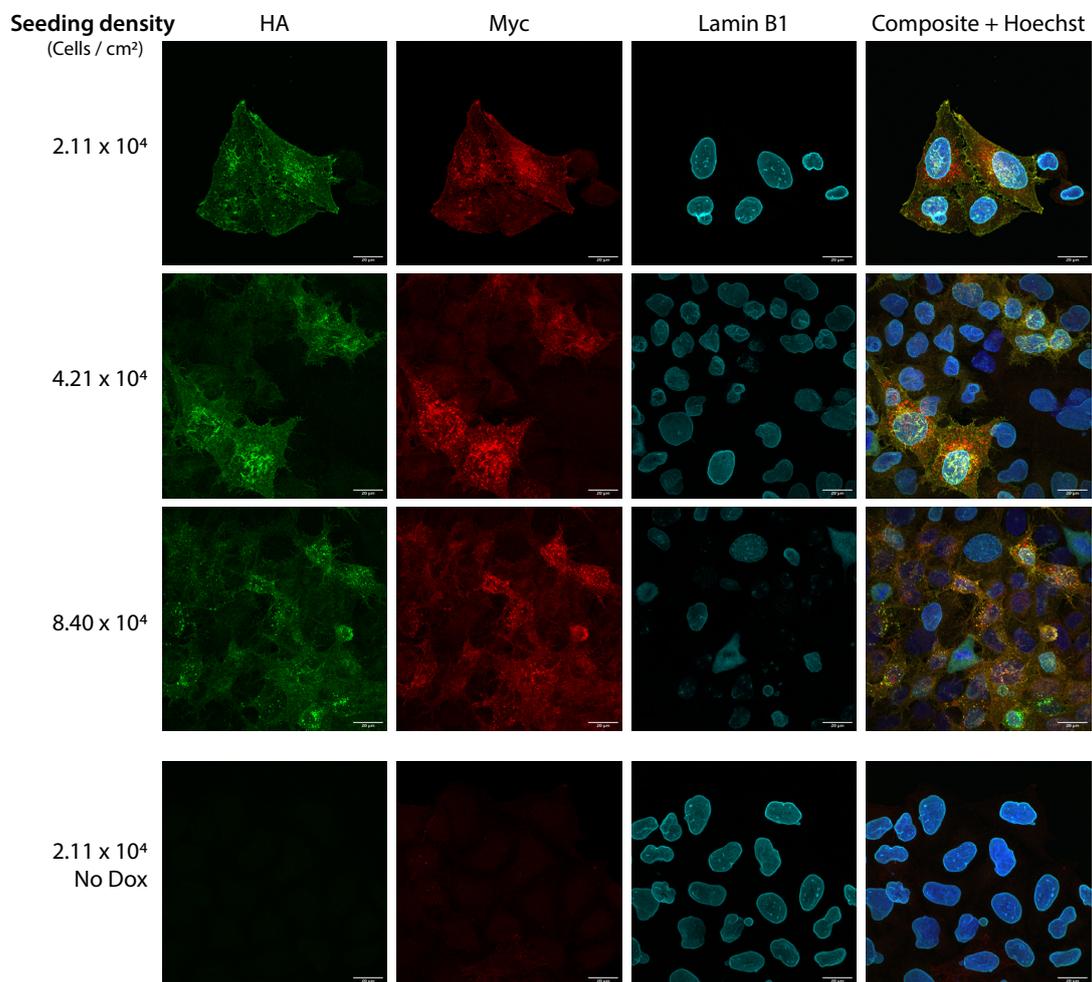
Figure 5-8 – Increasing cell confluence does not affect β-dystroglycan fragmentation *in situ*. HeLa cells stably transduced with α-HAβmyc-DG were seeded and cultured at increasing densities before analysis by confocal microscopy. Sites of HA and Myc staining were identified from each optical section and the correlation with the opposite epitope was calculated and graphed. At least 7800 puncta were identified in each condition, and statistical significance was determined by Kruskal-Wallis test with Dunn's multiple analysis. P-values are shown where effect size is >0.5.

5.2.5. _Nuclear localisation of fragments derived from full-length dystroglycan is not clear._

Throughout previous immunofluorescence experiments, nuclear localisation of any tagged dystroglycan fragments derived from a full-length construct was not observed, despite previous evidence suggesting its presence. However, studies probing nuclear localisation and functions of β-dystroglycan have extensively utilised truncation mutant constructs (Oppizzi _et al._, 2008; Lara-Chacón _et al._, 2010; Martínez-Vieyra _et al._, 2013; Mathew _et al._, 2013; Mitchell _et al._, 2013; Vélez-Aguilera _et al._, 2018). Therefore, it was postulated that expression of truncation mutants of the epitope-tagged dystroglycan construct would recapitulate published observations of the subcellular localisations of β-dystroglycan in HeLa cells, and demonstrate that the behaviour of dystroglycan with epitope tags is consistent with previous studies. The full length β-dystroglycan protein and its cytoplasmic fragment are commonly expressed in isolation to study nuclear β-dystroglycan. These dystroglycan fragments were sub-cloned from the epitope-tagged full length construct and expressed separately in HeLa cells. Immunofluorescence analysis revealed that β-dystroglycan (Δα-HAβmyc-DG) or the cytoplasmic fragment (βmyc-DG(cyto)) displayed vastly different subcellular distributions to each other, but also to the full length α-HAβmyc-DG construct (Figure 5-9A, B). Similar to previous experiments, HA and myc antibody staining did not detect any parts of the full length construct α-HAβmyc-DG in the nuclei of HeLa cells (Figure 5-9A, B, upper panels). Staining of cells transfected with the Δα-HAβmyc-DG construct similarly showed an absence of nuclear signal, though it was possible to discern a nuclear-ring like stain which overlayed the lamin A stain delineating the NE (Figure 5-9A, B, middle panels). However, co-staining with calnexin, a resident ER membrane marker, raised the possibility that staining in the perinuclear region actually arises due to the contiguous nature of the ER and ONM membrane systems (Figure 5-9B, middle panels), and that membrane associated β-dystroglycan freely diffuses throughout these membranes. In addition, the Δα-HAβmyc-DG protein was distributed throughout the cytoplasm in staining that resembled an ER-like pattern, as opposed to the plasma membrane and punctate staining seen with the full-length construct

(Figure 5-9A, B, upper/middle panels). Indeed, Δα-HAβmyc-DG co-localised with staining for calnexin, a resident ER-marker (Figure 5-9B, middle panels). Finally, there was no clear spatial separation of the HA and Myc signals when β-dystroglycan was expressed alone when compared with the full length construct which displays clear isolated regions of myc-staining (Figure 5-9A, B, middle panels). Inconsistent with other dystroglycan constructs, expression of the cytoplasmic fragment of β-dystroglycan (βmyc-DG(cyto)) robustly accumulated in the nuclei of transfected cells, however there was some cytoplasmic staining, which, unlike the Δα-HAβmyc-DG protein, did not mimic calnexin (Figure 5-9A, B, lower panels). This is perhaps unsurprising since the βmyc-DG(cyto) fragment is divided from the rest of the protein at the putative γ-secretase cleavage point, eliminating the transmembrane domain but leaving the NLS intact, presumably releasing the protein from membrane anchorage. Taken together, imaging results show that omitting parts of dystroglycan alters the subcellular distribution of β-dystroglycan which may indicate that the entire protein is necessary for proper regulation. To further clarify observations that dystroglycan derived from the full-length epitope-tagged constructs is not clearly observable in the nucleus, HeLa cells transfected with dystroglycan constructs were fractionated and analysed by western blotting. Inconsistent with immunofluorescence microscopy findings, probing for the myc epitope tag revealed that a minimal quantity of β-dystroglycan could be observed in the nuclear fraction upon long membrane exposure. However, probing for calnexin showed that the nuclear fractions were at least partially contaminated with ER. Calnexin was specifically chosen to control for ER contamination, at it, like dystroglycan, is a membrane protein. Similar to microscopy, it is difficult to truly distinguish between nuclear and residual ER β-dystroglycan given the membranes are contiguous, permitting free diffusion of proteins throughout the ER and ONM.
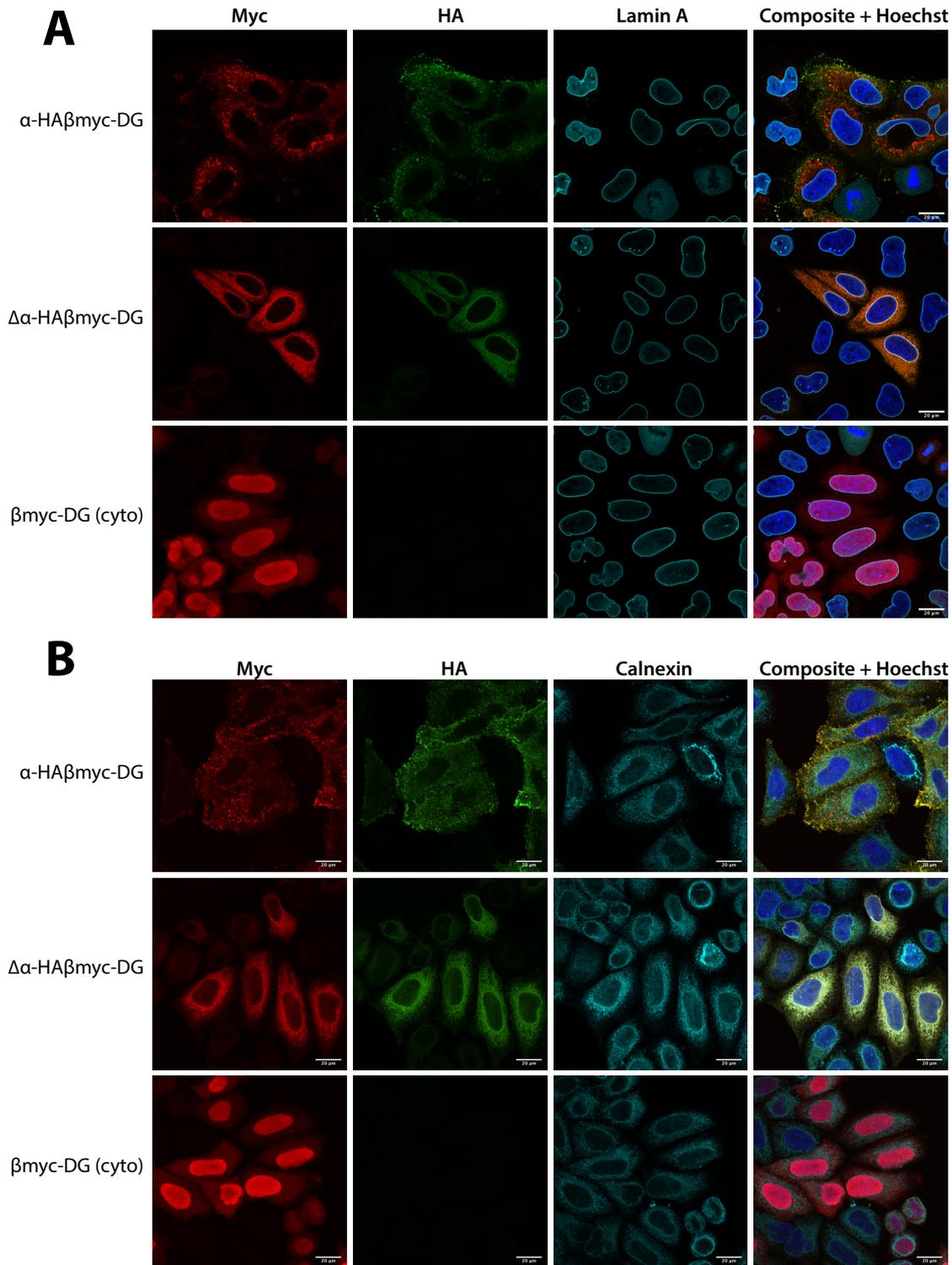
Figure 5-9 – Expression of dystroglycan fragments results in differential localisations. HeLa cells were transfected with either α-HAβmyc-DG, Δα-HAβmyc-DG or βmyc-DG (cyto), and stained with antibodies to detect the epitope tags. Subcellular structures were also stained; the nuclear envelope was delineated with a stain for lamin A (A), while the ER network highlighted with calnexin (B). Specificity of antibodies against HA and Myc was controlled for by staining untransfected cells, shown in Supplementary Figure II.
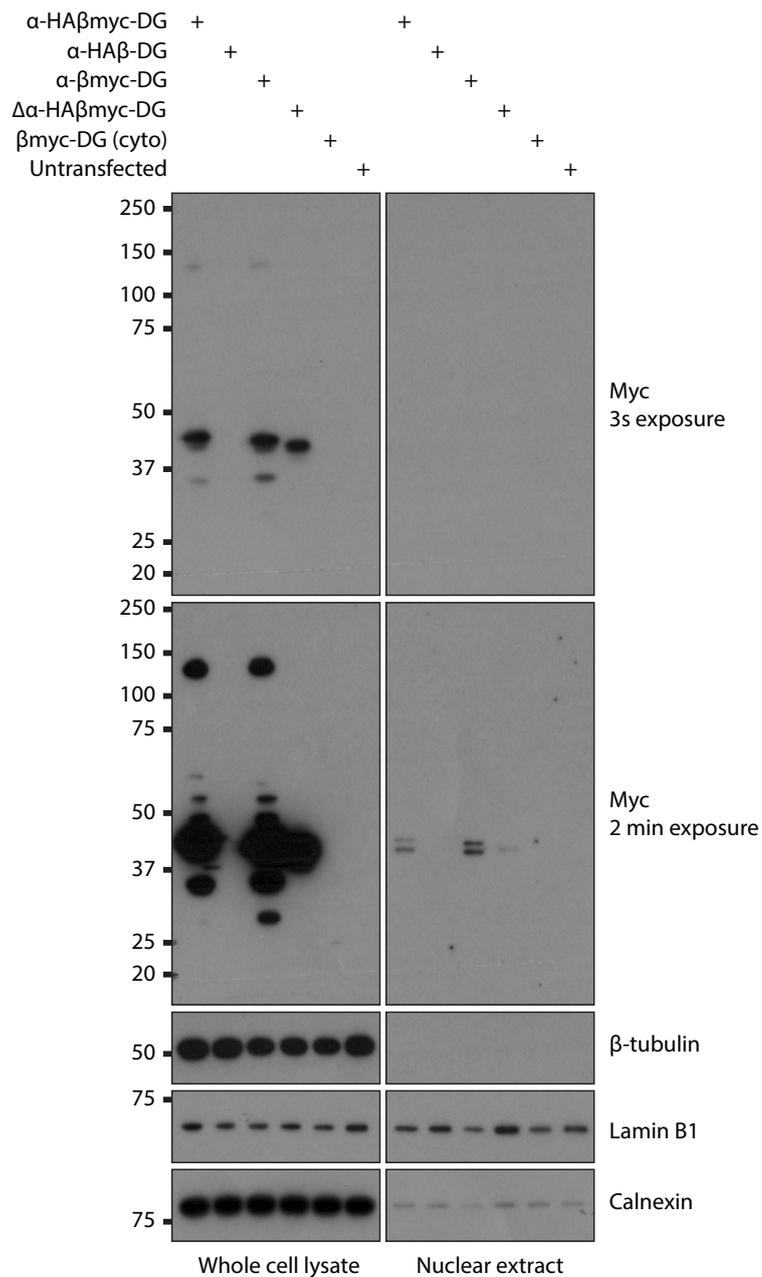
Figure 5-10 – Minimal β-dystroglycan can be detected in a nuclear extract. HeLa cells transfected with the indicated constructs were fractionated to isolate nuclei. Whole cell lysates and nuclear extracts were then analysed by western blotting and probed with antibodies raised against the myc-epitope within the construct. The presence of lamin B1 and absence of β-tubulin was used to verify the enrichment of nuclei, while calnexin was used as a measure of purity.
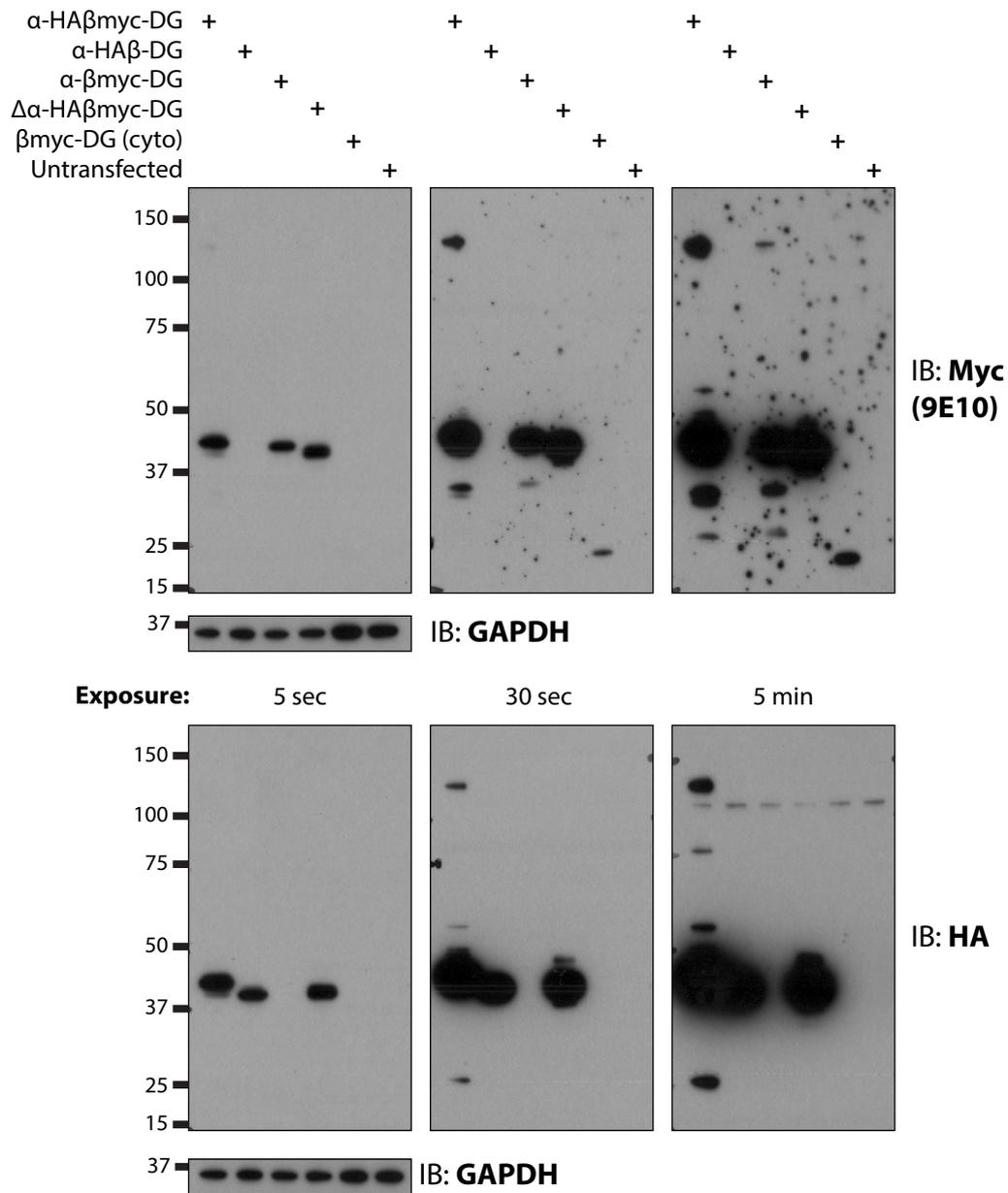
Figure 5-11 – β-dystroglycan fragments differently when expressed the absence of α-dystroglycan. HeLa cells were transfected with the indicated construct, and β-dystroglycan fragments arising from each construct were analysed by western blotting. Upper panels show myc immunoreactive bands, while lower panels show HA. GAPDH was used to show approximately equal loading. Multiple exposure of western blots reveal less abundant fragments.

5.2.6. *Fragmentation of β-dystroglycan is not observed when expressed without the α-subunit.*

The finding that constructs encoding isolated dystroglycan fragments do not consistently localise raised the possibility that the full dystroglycan protein must be expressed for proper regulation. This, together with the observation that Myc and HA tags do not separate upon the expression of β-dystroglycan alone formed the hypothesis that processing of β-dystroglycan is also dependent on the presence of a full length construct. Interestingly, western blotting analysis of HeLa cells expressing Δα-HAβmyc-DG displayed differences between β-dystroglycan species. The full length α-HAβmyc-DG construct gave rise to a ~35kDa Myc immunoreactive band, which is similar to that seen in the whole cell lysate during nuclear fractionation (Figure 5-10, left panels), though here as a doublet (Figure 5-11, upper panels). The myc-tagged protein species pattern is also consistent with that seen in Figure 5-3, though the use of SDS-MOPS gels here may account for slight differences in relative molecular weight of protein species between experiments. In addition to a ~35kDa myc-tagged only species, there was a ~26kDa HA-only immunoreactive band (Figure 5-11, lower panels). Together, the bands described potentially represent the two fragments of β-dystroglycan upon γ-secretase cleavage. This observation would signify the first detection of the transmembrane-ectodomain β-dystroglycan fragment. However, the ~26 kDa HA-tagged protein species is not present when β-dystroglycan is tagged only with HA (Figure 5-11, lower panels, lane 2). It is therefore possible that this species is either non-physiological or somehow protected from degradation when derived from the α-HAβmyc-DG construct. Nevertheless, the Myc-only ~35kDa band is also observed in the α-βmyc-DG construct, albeit as a single species. On the other hand, the Δα-HAβmyc-DG construct displays HA and Myc immunoreactive bands at ~45kDa, which likely corresponds to full length β-dystroglycan, and there is an absence of lower molecular weight bands which is not a reflection of the fragmented protein species displayed by the full length β-dystroglycan proteins, even upon long exposure (Figure 5-11, right-hand panels). The lack of fragmentation using this construct therefore raises the possibility that truncated forms of β-dystroglycan are not processed in a physiological manner, and

this should be considered while ascribing functions to dystroglycan determined using these constructs. Notably, none of the myc-epitope positive bands derived from full length constructs correlated with the β-dystroglycan cytoplasmic fragment, βmyc-DG(cyto) (Figure 5-11, upper panels), undermining the previous hypothesis that the addition of the myc-epitope alters biochemical characteristics of the peptide (see section 5.2.1). Nonetheless, isolated expression of β-dystroglycan does not fragment in the same manner as β-dystroglycan which is derived from a construct also incorporating the α-subunit, further implicating disparate parts of dystroglycan in normal regulation.

5.3. Discussion

Results presented in this chapter outline the characterisation of a dystroglycan construct interspersed with epitope-tags and show that β-dystroglycan fragments arising from the construct can be separately detected by western blotting and by immunofluorescence microscopy. These results therefore indicate firstly that cells have mechanisms to distinguish β-dystroglycan fragments from each other and secondly, that different β-dystroglycan species potentially have specific functions. It was noted that insertion of a V5 tag after F613 was found to reduce autoproteolysis of the dystroglycan pre-peptide, though immunofluorescence evidence indicates that this has no impact on further processing of β-dystroglycan. Previous studies non-specifically detecting β-dystroglycan have assigned a number of functions for non-plasma membrane localised β-dystroglycan including internalisation and targeting for degradation, and stabilisation of nucleoli and NE proteins. It was predicted that these functions cannot be mediated by a single β-dystroglycan species, and results here seem to support this idea, instead implicating the existence of multiple, differentially regulated β-dystroglycan fragments. Isolated expression of epitope-tagged dystroglycan constructs were found to be inconsistently distributed at the subcellular level, and fragment differently, when compared with the full length construct. Almost analogous to the effect of the distally located V5 tag from the SEA domain, these observations together insinuate that full length dystroglycan is required for normal functioning, and that it is sensitive to truncations and

modifications throughout its length. At least one other example of this is the difficulty in correlating fragment sizes to known physiological species, possibly due to unintended and unknown consequences of, or modifications to, the integral epitope tags. Overall, however, results suggest β-dystroglycan is protein of multiple facets which should perhaps not be amalgamated by non-specific detection methods to dissect its subcellular function.

5.3.1. *Western blotting reveals multiple β-dystroglycan protein species with ambiguous identities*

A central aim for experiments focussed on a new dystroglycan probe, was for epitope tags to some extent unravel complexities in visualising fragmentation of the β-subunit. However, western blotting of cells expressing the tagged dystroglycan construct revealed multiple specific protein bands, from which fragment identities are not clear. For example, the myc-only tagged β-dystroglycan species measured ~35 kDa, despite the myc-tag being designed to identify the 26 kDa fragment. It should be noted that previous studies have found varying arrangements of β-dystroglycan protein species by western blotting (Mathew *et al.*, 2013; Mitchell *et al.*, 2013). In addition to fragmentation, β-dystroglycan is extensively modified by a suite of other protein adducts including glycosylation, phosphorylation and ubiquitination which may function to further specify the functions of β-dystroglycan species (Ervasti and Campbell, 1993; James *et al.*, 2000; Endo, 2007, 2015; Lipscomb *et al.*, 2016). Therefore, it is worth considering mechanisms and modifications by which these shifts in apparent molecular weight arise. Full length dystroglycan is subject to an extensive array of post translational glycosylation modifications during its maturation. Appropriate binding to the ECM is ensure by modifications performed by at least 17 enzymes (Taniguchi-Ikeda *et al.*, 2016; Kanagawa and Toda, 2017), raising the possibility that the addition of extra epitope sequences would become aberrantly glycosylated. However, neither HA or Myc epitope tags possess or create the Asn-X-Thr/Ser required for N-glyscoylation or Thr/Ser residues for O-linked glycosylation (Alberts *et al.*, 2008), thus excluding the probability that aberrant glycosylation contributes to the altered relative mass by SDS-PAGE.

Another possible modification leading to increased molecular weight is ubiquitination. Ubiquitin, a small protein with a molecular weight of ~9 kDa can be attached to target proteins by three enzymes, E1-3 (Lecker, Goldberg and Mitch, 2006). The ligation of ubiquitin chains to target proteins is a mechanism used by the cell to mark proteins destined for degradation by the proteasome (Lecker, Goldberg and Mitch, 2006). The degradation of dystroglycan has already been implicated in the pathogenesis of DMD, and proteasome inhibitors, such as MG132 and Velcade, have been shown to restore its expression and ameliorate disease in the *sapje* zebrafish DMD model (Gazzerro *et al.*, 2010; Lipscomb *et al.*, 2011). Moreover, E3 ubiquitin ligase inhibitors have been proposed as a mechanism to retain dystroglycan at the plasma membrane (Lipscomb *et al.*, 2016). The ~9 kDa increase in molecular weight observed between the predicted and observed molecular weights of the cytoplasmic fragment however is not consistent with a polyubiquitin chain thought to signal proteins for destruction, which typically comprises at least 4 ubiquitin molecules (Hicke and Dunn, 2003). Therefore, if the observed ~35 kDa band compromises the cytoplasmic region only, it may be monoubiquitinated, which is consistent with a ~9 kDa addition to a 26 kDa unmodified fragment. The Winder laboratory has previously examined the ubiquitination of β-dystroglycan in the context of the full-length form, and evidence suggests that monoubiquitination is linked to phosphorylation at Y890 (Leocadio-Victoria, 2015 (PhD thesis); Lipscomb et al., 2016; Piggott, 2014 (PhD thesis)). While polyubiquitination can signal for degradation, monoubiquitination is thought to regulate the membrane trafficking, and thus subcellular localisation, of a target protein (Hicke and Dunn, 2003; Erpapazoglou, Walker and Haguenauer-Tsapis, 2014; Sewduth, Baietti and Sablina, 2020). The subcellular separation, together with increased molecular weight of the myc-tagged cytoplasmic fragment indicate that this is a reasonable hypothesis. However, while there is evidence that ubiquitinated residues of β-dystroglycan reside within the cytoplasmic region, there is no current indication that specific fragments of β-dystroglycan can be modified. The cytoplasmic fragment of human dystroglycan is just 121 amino acids in length, and it seems that small peptides of this length can be targeted for degradation using a single ubiquitin modification (Shabek *et al.*, 2012). Transient overexpression experiments, which were frequently

used during this study, can lead to increased protein of interest degradation as the cell compensates (Prelich, 2012; Gibson, Seiler and Veitia, 2013). To establish whether cytoplasmic β-dystroglycan is degraded, a double staining for myc to delineate the cytoplasmic fragment, and a lysosomal marker might be able to distinguish whether a potential monoubiquitination modification signals for further trafficking or degradation. Unequivocal identification of the complicated β-dystroglycan fragments would be ultimately assessed using pulldowns to isolate β-dystroglycan specific modifications followed by mass spectrometry.

5.3.2. *Regulation of dystroglycan fragmentation.*

At the cell membrane, ECM engagement by dystroglycan is thought to transduce extracellular signals to the cells interior by conformational changes in its carboxyl-terminus leading to the uncovering of new docking sites. Full details for the intracellular function of dystroglycan in propagating these signals are still emerging, though it seems that the protein can scaffold a number of signalling proteins through its juxtamembrane region or cytoplasmic domain (Moore and Winder, 2010). However, like other cell surface receptor proteins, namely notch and CD44, it seems some stimuli can also induce the movement of dystroglycan away from the plasma membrane (Kidd, Lieber and Young, 1998; Mumm *et al.*, 2000; Okamoto *et al.*, 2001; Lammich *et al.*, 2002; Mathew *et al.*, 2013; Mitchell *et al.*, 2013). The consequences of β-dystroglycan trafficking are yet to be fully elucidated, but in prostate cancer cells, the cytoplasmic fragment localises to the nucleus where it is thought to direct transcriptional output (Mathew *et al.*, 2013). One stimulus reported to upregulate proteolysis leading to the liberation of the cytoplasmic β-dystroglycan domain, as assessed by western blotting, is increased cell density (Mitchell *et al.*, 2013). However, experiments performed during this study using an α-HAβmyc-DG construct in HeLa cells of increasing density failed to replicate this observation when assessed by immunofluorescence microscopy. Time constraints did not permit further investigation of this, however western blotting could be performed to assess whether proteolysis in response to high cell density had indeed occurred, while fragment separation was not observable *in situ*. Indeed, culturing cells at high confluency does decrease the cytoplasmic volume of each cell, for which the analysis

method chosen was dependent. Pharmacological activation of the notch signalling pathway has previously been shown to increase the amount of 26 kDa β-dystroglycan cleavage, via an unknown mechanism (Leocadio, Mitchell and Winder, 2016). Resveratrol agonism of β-dystroglycan fragmentation in subconfluent cells would make immunofluorescence analysis more amenable. A complementary approach might be to inhibit γ-secretase or furin proteases in order to reduce fragmentation in a subconfluent culture (Leocadio, Mitchell and Winder, 2016).

5.3.3. *Subcellular localisation of full-length dystroglycan does not reflect expression of β-dystroglycan truncation mutants.*

A striking observation made during the course of this study was the clear differences in the subcellular localisations of truncated dystroglycan constructs. Expression of full length dystroglycan was initially thought to localise in such a way that represented a sum of individual dystroglycan fragments. However, expression of truncated dystroglycan proteins showed that this was not the case. The absence of fragmentation upon expression of full length β-dystroglycan indicates the necessity of the α-subunit for proper processing. Interestingly, however, the determined localisations of the truncated β-dystroglycan constructs was not unprecedented, and has been routinely observed throughout previous studies.

The expression of full length β-dystroglycan has been used extensively to study nuclear functions, and is routinely observed in an ER-like distribution (Sotgia *et al.*, 2003; Lara-Chacón *et al.*, 2010; Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018; Gómez-Monsivais *et al.*, 2020). Confocal analysis of C2C12 myoblasts expressing full length β-dystroglycan also indicates presence throughout the nucleus, though how it escapes the membranous environment is unclear (Vélez-Aguilera *et al.*, 2018). Consistent with results from this study, research focussed on the cytoplasmic region of dystroglycan report robust nuclear accumulation (Oppizzi *et al.*, 2008; Mathew *et al.*, 2013; Mitchell *et al.*, 2013; Azuara-Medina *et al.*, 2019), though whether this localisation is relevant in the full length context, is unclear.

5.3.3.1. *α-HAβmyc-DG is not detected in the nucleus of HeLa cells.*

Finally, despite widespread reports of its presence, and the robust localisation of βmyc-DG to the nucleus, this study did not detect a significant accumulation of dystroglycan derived from the full length construct within the nucleus by immunofluorescence analyses. Results from western blotting experiments were ambiguous; the absence of β-tubulin from the nuclear fraction would indicate that β-dystroglycan detected was indeed in the nucleus, however a more rigorous interrogation of the sample revealed ER contamination due to the presence of calnexin. In fact, the absence of detectable β-dystroglycan is somewhat in keeping with previous experiments which use full length dystroglycan constructs. A C-terminal GFP fusion appears to localise to the cytoplasm in REF52 cells and fibroblasts, though epifluorescence microscopy and lack of nuclear counterstain makes this conclusion uncertain (Y.-J. Chen *et al.*, 2003). However, the same construct localises exclusively to the cytoplasm in C2C12 myoblasts (Gracida-Jiménez *et al.*, 2017), unless a Y890E phosphomimetic mutation is introduced leading to an additional diffuse nuclear signal (Gracida-Jiménez *et al.*, 2017). Interestingly, GFP-tagged dystroglycan localises to both the nucleus and cytoplasm in mammary epithelial MEpE cells (Oppizzi *et al.*, 2008). Though epitope modifications used in the current study were rationally placed to avoid important regions of dystroglycan, it seems that protein sequences in proximity to the SEA domain do impact protein function. Indeed, the mutation of residues around the α-/β- cleavage site has various effects on dystroglycan, ranging from normal cell surface positioning to loss of expression (Sciandra *et al.*, 2009). In addition, the Cys667Phe mutation found in an MEB patient results in ER-retention of a non-cleaved dystroglycan pre-peptide, which appears to arise due to the disruption of the Cys669-713 disulphide bridge thought to stabilise the SEA domain (Sciandra *et al.*, 2012; Signorino *et al.*, 2018). Furthermore, N-linked glycosylation sites on α- and β-dystroglycan are crucial for efficient cleavage (Esapa *et al.*, 2003). Overall, there are a range of defects which inhibit physiological characteristics of dystroglycan including autoproteolysis and localisation, but it is not known whether these also result in a disruption to further processing. It is therefore feasible that the multiple epitope modification made to

dystroglycan in the current study affect the protein in ways which are not currently understood. Nevertheless, the existing bank of evidence points towards cell-type specific tendencies for nuclear localisation. Despite initial evidence for the nuclear localisation of dystroglycan and other DGC components being found in HeLa cells (Fuentes-Mera *et al.*, 2006), α-HAβmyc-DG should be expressed in different cell types to elucidate whether it can be detected within the nucleus. Alternatively, the contribution of a phosphomimetic Y892 mutation on nuclear localisation could be assessed. Overall, however, it seems difficult to truly discern whether a protein which is ambiguously localised to the nucleus truly does reside in this compartment. Signal from immunofluorescence microscopy may be below the background measurement in cells, while cellular fractionation suffers from ER-contamination rooted in shared membranes between the ER and ONM.

5.3.4. *Conclusions and future directions.*

Dystroglycan is a cell adhesion protein but has numerous emerging functions throughout other cellular compartments. Moreover, it becomes elaborately post-translationally modified by phosphorylation, ubiquitination and proteolysis. The insertion of epitope tags throughout β-dystroglycan appears not to disturb its normal function and has permitted the tracing of some proteolytic fragments, revealing that carboxyl-terminus becomes spatially separated from the ectodomain. The exact identity of this fragment remains unclear due to its relative molecular weight. Moreover, which of the proteolytic fragments are subjected to other modifications including phosphorylation has not yet been investigated. Finally, evidence for nuclear localisation of dystroglycan is not clear in HeLa cells, and experiments have not yet excluded the possibility that modifying the proteins contributes to disruptions in nuclear transport.

Chapter 6: Validating a murine model for *ex vivo* two-component BioID and generating the Sun1 protein interactome in muscle.

6.1. Introduction

Proximity biotinylation using the BioID technique coupled with mass spectrometry is used to identify potential novel protein-protein interactions (PPIs) *in vitro*. It is unbiased, and since interacting and proximal proteins become covalently labelled by a biotin modification, balancing extraction and the necessity to preserve weaker endogenous PPIs, as is the case for antibody-based affinity purification, is not problematic. As such, proximity biotinylation techniques are now widely regarded as a powerful tool by which to discover novel PPIs (Gingras, Abe and Raught, 2019). Mutations in genes involved in the nuclear envelopathies often affect multiple tissues, and one of the outstanding questions is how multiple diseases arise from a single gene. Use of an *in vivo* BioID system to model these diseases therefore provides several advantages since PPIs are assessed in their physiological context, and be used to uncover how key gene products interact with other proteins in affected cell or tissue types. However, the study of nuclear proteins is complicated by their targeting to the NE. Protein fusions are often blocked at the NPC due to size constraints, preventing access to the nuclear interior, and so a modified BioID system, known as two-component BioID (2C-BioID), was developed to circumvent targeting issues was developed (Chojnowski *et al.*, 2018).

6.1.1. *Unbiased detection of novel protein-protein interactions (PPIs)*

Proteins are required to detect, transduce and integrate diverse signals, originating both within and beyond the boundaries of a cell. Execution of specific functions based on such complex stimuli requires significant cross-talk between individual components. Understanding the molecular function of single proteins

often relies on detecting changes in interacting partners depending on stimuli or disease mutations. A number of biochemical methods exist to detect protein-protein interactions (PPIs) including immunoprecipitation and ELISA, while other strategies, like proximity ligation assays (PLAs), can be used to detect close antigen proximity *in situ*. These have proved powerful tools, however, they are dependent on high quality antibodies raised against the target proteins. Crucially, many conventional PPI detection methods such as co-immunoprecipitation, PLA or ELISA absolutely require prior knowledge of the interacting partners. This is sufficient to test specific hypotheses; however, to discover novel PPIs, this prerequisite introduces significant bias. Unbiased strategies such as two-hybrid assays have been used to address unknown PPIs. This system is commonly used in yeast, where bait and prey proteins are fused to DNA binding domains and activating domains of the Gal4 transcription factor, respectively. Gal4 binds the UAS promoter element, and reporter expression occurs only if the prey binds the bait protein, bringing the Gal4 activating domain into close proximity (Chien *et al.*, 1991). The scalability of the yeast-2-hybrid assay has enabled library scale screening of PPIs (Fields, 2005). However, the system does not accurately reflect the suite of protein modifications found in metazoans (e.g. phosphorylation), and yeast translation machinery may misfold mammalian proteins. In addition, yeast-2-hybrid screens require soluble proteins for nuclear localisation, hence membrane proteins are particularly refractory to this technique. For these reasons, the technique is not ideal to search for interactors of integral membrane proteins of the nuclear envelope.

Proteomics defines a field in which biomolecules are identified by mass spectrometry (MS). The strategy can be ported into many different experimental types, from defining all proteins of a cell or tissue type, or to identify subsets of proteins, enriched using biochemical techniques (Lössl, Waterbeemd and Heck, 2016). MS-proteomics therefore presents a means to identify proteins in an unbiased manner, which is essential to expand our understanding of the multifaceted chemistries of a complex biological system.

Conventional proteomic strategies used to define organelle components or binding partners of a protein of interest (PoI) tend to destroy cellular compartmentalisation; the common initial step being cell or tissue lysis. There are

numerous implications to this approach. Proteins which did not originally populate the same compartment can now associate, resulting in the identification of false components or interactors. Conversely, detergents used to extract proteins from lipids together with huge sample dilution in buffer, protein complexes may dissociate resulting in the loss of weak or transient interactors. Due to their membrane bound, insoluble nature, biochemical analysis of components of the nuclear lamina is often hindered by this balance between efficient membrane extraction and recovering a complete interaction profile intact (Burke and Stewart, 2013).

6.1.2. *Principles of BioID proximity biotinylation assays*

Enzymatic proximity biotinylation has evolved as a method to label interactors of a PoI *in situ,* thus overcoming the issues of conventional proteomics which includes the destruction of compartmentalisation allowing unrelated proteins to mix and spuriously interact, the difficulty in balancing efficient extraction of membrane anchored proteins with preserving weak endogenous interactions. The use of BioID, which can overcome these problems through *in situ*, covalent biotin labelling of proximal proteins, has therefore become increasingly widespread since initially described by Roux, Burke and colleagues in 2012 (Roux *et al.*, 2012). The original BioID1 enzyme is a constitutively active mutant form of the biotin ligase BirA* isolated from *Escherichia coli* and is typically fused to a PoI. BioID1 has a molecular mass of 35.1 kDa (Roux *et al.*, 2012; Roux, Kim and Burke, 2013), and it is long established that large peptide tags on PoIs can affect their physiological function and subcellular localisation. This provided impetus to introduce the smaller 26.6 kDa BirA* enzyme from *Aquifex aeolicus* as BioID2 (Dae In *et al.*, 2016).

The premise behind all BioID assays is the covalent biotin modification of proteins proximal to the BioID enzyme of choice, *in situ*. Biotin is readily recovered from a mixture by immobilised streptavidin, which allows the use of harsh lysis buffers to efficiently extract all proteins from membranous environments. The aforementioned BioID enzymes also utilise labelling periods of 16-24 hours, providing a significant advantage for the capture of transient interactors which might otherwise be lost using conventional immunoprecipitation based methods (Roux *et*

al., 2012). For time-resolved analyses, APEX (Hung *et al.*, 2014) and the recently described TurboID and miniTurbo enzymes (Branon *et al.*, 2018) can provide insight into highly dynamic processes. BioID enzymes have also been adapted to suit various other applications. Split-BioID divides the BirA* enzyme into two non-functional subunits such that it is only active when the engineered subunits are in close proximity; their localisation guided by interacting fusion PoIs. This is particularly useful to define the interactome of PoIs known to have multiple functions or localisations (Schopp *et al.*, 2017).

6.1.3. *The in vivo two-component BioID approach*

Targeting modified integral proteins to the INM is technically challenging. While it is now established that the NPC can accommodate and import huge proteins and complexes many times larger than previously believed (Wang and Brattain, 2007), the comparatively modest BioID or GFP- and associated tags commonly cause localisation issues (Lammerding *et al.*, 2006; Chojnowski *et al.*, 2018). Endogenous nuclear import is probably controlled through intricate biochemical properties between NPCs, transport machinery and their cargoes (Schmidt and Görlich, 2016). The current notion, however, maintains that a smaller modification will, in many cases, more accurately imitate physiological function of a PoI. This concept is exemplified by the development of two-component BioID (2C-BioID), where the conventional BioID enzyme modification to the INM protein LAP2β precluded a proportion of the fusion from positioning at the NE. 2C-BioID employs the well-established rapamycin-directed conditional heterodimerising motifs FRB domain of mTOR (11.2 kDa) and FKBP (11.9 kDa) (Putyrski and Schultz, 2012) fused to the PoI and BioID2 enzyme, respectively. Biotinylation specific to the FRB-tagged PoI therefore only occurs in the presence of the dimerising agent AP21967 (Figure 6-2A), allowing the fusion-PoI to be processed in the absence of the large enzymatic domain. The advantage of using the less intrusive heterodimerising domains is demonstrated in the BioID analysis of INM protein LAP2β. The FRB-tagged version localises akin to endogenous or V5-labelled versions. The two-component system also offers the ability to control for non-specific biotinylation which would otherwise

be detected as the PoI transits through common trafficking pathways (Chojnowski *et al.*, 2018).

The majority of BioID assays have been applied to *in vitro* systems. This has been greatly successful, yielding plentiful potential interacting partners with relative ease, offering investigative routes at the subcellular level. An *in vitro* strategy likely encapsulates the conserved interactors, or perhaps those artefactual to prolonged culture systems. However, this approach alone cannot address the existence of divergent disease conditions arising from a widely-expressed gene, as is the case for *Lmna*, for example. A system which respects the subtleties of protein interactions with regards to tissue-specificities and alternative splicing variants was therefore devised using the 2C-BioID method *in vivo*.

6.1.4. *The Rosa26 locus and DNA recombination technology*

To make the *in vivo* 2C-BioID system portable into other mouse genetic models, it capitalises on the homologous recombination technologies enabling the controlled expression of exogenous transgenes in a mouse model. The myc-tagged FKBP-BioID2 (mycFKBP-BioID2) is an artificial, exogenous gene and must be stably integrated into the genome. To do this, it was knocked into the reverse orientation splice acceptor (ROSA) 26 locus, which is known as a safe harbour for constitutive and ubiquitous tissue expression of transgenes in the mouse genome (Wong and Stewart, unpublished data; Friedrich and Soriano, 1991; Kisseberth *et al.*, 1999). Its serendipitous discovery was the result of a mutagenic gene-trap procedure incorporating a reporter containing a splice acceptor and a promoter-less reporter consisting of a bifunctional fusion between the neomycin-resistance, and *lacZ* genes (β-Geo). This feature enabled the screening of targeted embryonic stem cells before blastocyst injection and demarcated tissue expression in the developing mouse. Prior to this development, 'trapped' genes were often discovered to regulate readily identified phenotypes, while subtle effects went largely unnoticed. In the Rosa(β-Geo)26 mouse line, X-gal staining of mouse embryos throughout development revealed ubiquitous expression without phenotypic abnormalities (Friedrich and Soriano, 1991). Further developments also found that that the expression from the Rosa26 locus persisted into maturity (Kisseberth *et al.*, 1999). Previous technology

relied on the random genome integration of vectors, leading to a mosaic gene expression pattern, known as variegation (Martin and Whitelaw, 1996; Montoliu, Chávez and Vidal, 2000). The precision of a single genetic modification to the Rosa26 locus, combined with a reliable pan-tissue penetrance of the promoter has led to the regions' popularity for the knock-in of exogenous transgenes.

However, ubiquitous expression of transgenes is often not desired, as the exogenous element may contribute to embryonic lethality, or other unintentional deleterious effects on tissues. Precise control of Rosa26 knock-in transgenes can be provided by unison with DNA recombination technology, such as the Cre-lox or FLP-FRT systems (Soriano, 1999). The Cre-lox system originated in the bacteriophage P1 and was shown to be highly efficient at specifically recombining DNA segments flanked by LoxP sequences (floxed) (Sternberg and Hamilton, 1981). The FLP-FRT system was discovered as an analogous system the yeast *Saccharomyces cerevisae* (Broach, Guarascio and Jayaram, 1982), and both Cre-lox and FLP-FRT functioned in same manner when transplanted into mammalian cell lines (Sauer and Henderson, 1988; O'Gorman, Fox and Wahl, 1991). Therefore, mycFKBP-BioID2 was knocked into the Rosa26 locus under the control of a Cre-lox system to mitigate potential negative consequences of mycFKBP-BioID2 expression *in vivo*.

6.1.5. *The importance of the nucleoplasmic Sun1 protein interactome in muscle*

Sun1, as a member of the LINC complex, is responsible for maintaining physical connection between the nucleus and cytoskeleton. This bridge is used to relocate the nucleus within the cell, and transmit mechanical signals to the nucleus, resulting in changes to genome organisation and transcriptional output. Sun-domain proteins bind lamin A of the nuclear lamina (Crisp *et al.*, 2006; Haque *et al.*, 2006). Mutations in *Lmna* give rise to a range of tissue-specific diseases, including muscular dystrophies, cardiomyopathies, lipodystrophies and progeria for reasons not fully understood. Chen *et al.* (2012), found that despite the apparent crucial role of Sun1 in development of auditory and nervous systems and during meiosis I (Lei *et al.*, 2009; Zhang *et al.*, 2009; Horn, Brownstein, *et al.*, 2013; Meinke *et al.*, 2014), loss of Sun1 in *Lmna* mutant mice suppresses many disease phenotypes associated with

Lmna mutations (Chen *et al.*, 2012). The nucleoplasmic Sun1 interactome is of specific interest since the knockout of Sun2 in Lmna mutant mice does not alleviate disease phenotypes (Stewart *et al.*, unpublished data, personal communication).

6.1.6. *Chapter aims*

The work presented in this chapter centred on the validation of the Rosa26-mycFKBP-BioID2$^{+/-}$ mice and the generation of a nucleoplasmic protein interactome for Sun1 (Figure 6-1). Validation experiments were performed *ex vivo* from adult fibroblasts (MAFs) derived from tail biopsies of heterozygous animals which were immortalised using a lentiviral vector containing SV40 large T antigen for experimental convenience. In addition, the study aimed to assess the expression and function of the mycFKBP-BioID2 transgene in MAFs and myoblasts derived from the Rosa26-FKBP-BioID2 mouse, and to generate a Sun1 protein interactome from in myoblasts derived from this model given they are a cell type relevant to the aetiology of the laminopathies.
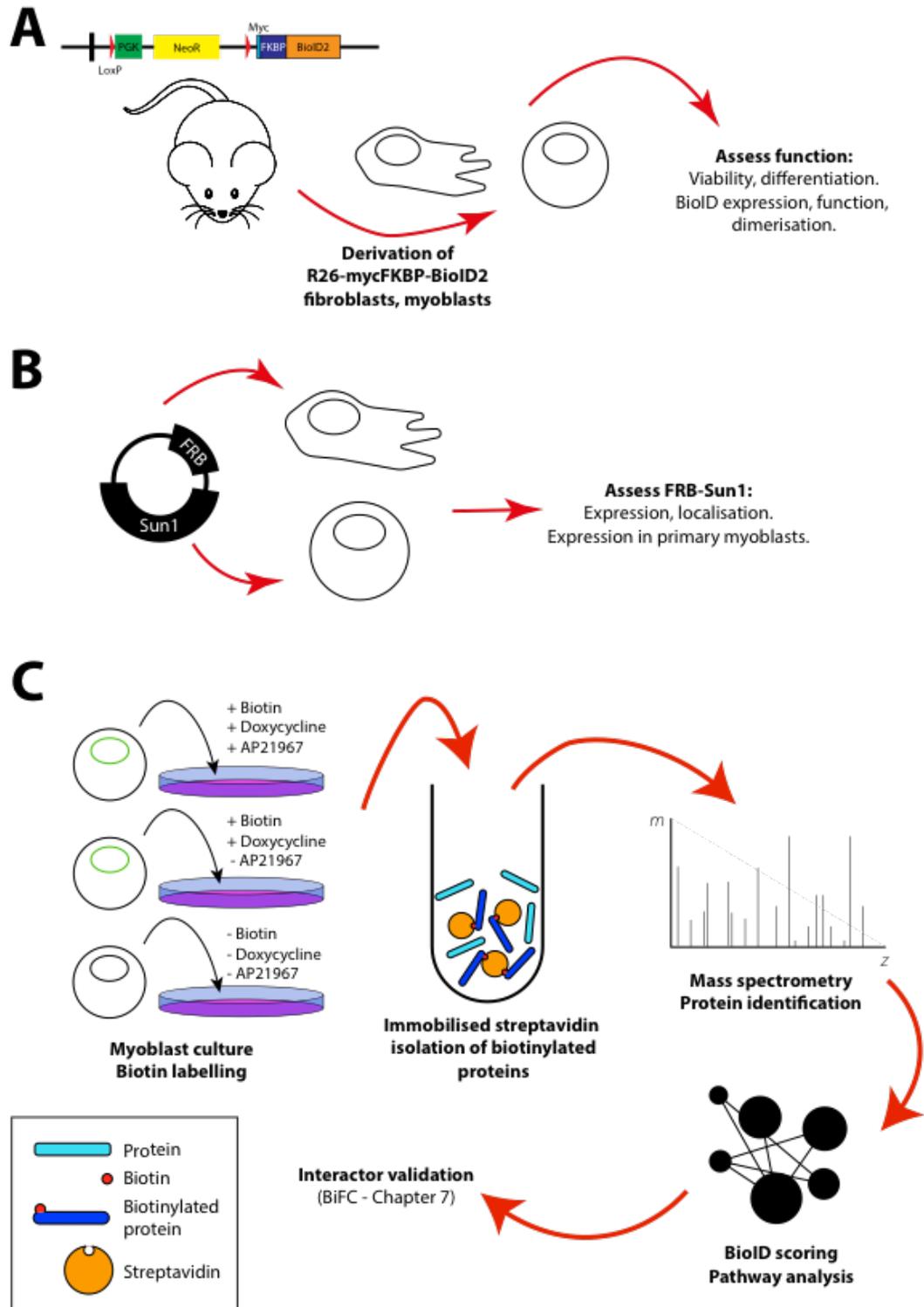
Figure 6-1 – 2C-BioID strategy for defining Sun1 protein interactome. (A) Functional assessment of 2C-BioID features in cells derived from the mycFKBP-BioID2 mouse. (B) Nucleoplasmic fusions on NL proteins often disrupt localisation, thus the FRB-Sun1 bait construct will be assessed in a range of cell types. (C) Work-flow for Sun1 interactome discovery. Myoblasts transduced with the FRB-Sun1 construct are cultured under the indicated conditions. After 24 hours of culture in biotin-containing media, cells are lysed and biotin-labelled proteins are isolated using immobilised streptavidin. Proteins are subsequently digested and peptides are identified by mass spectrometry. Following protein identification, candidates are scored and clustered for analysis of enriched groups. Finally, candidate interactors are validated using an independent method which is dealt with in Chapter 7.

A

B
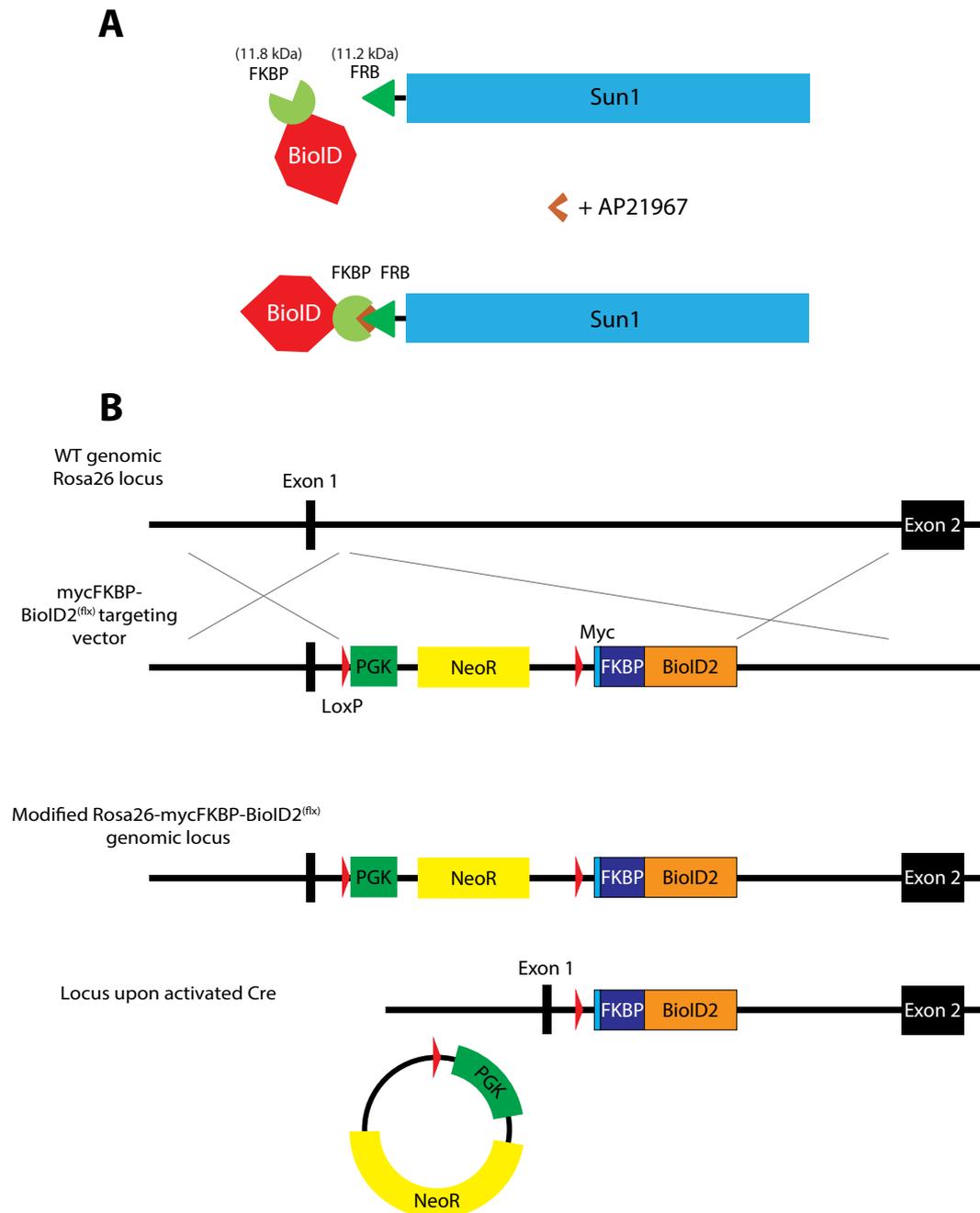
Figure 6-2 – The 2C-BioID principle and targeting strategy. The BioID biotin ligase enzyme and the protein of interest are fused to FKBP and FRB conditionally heterodimerising domains, respectively. The addition of rapalogue AP21967 induces dimerization, recruiting BioID to the protein of interest. (B) mycFKBP-BioID2(flx) construct was targeted to the Rosa26 locus, incorporating a floxed neomycin resistance cassette. Upon Cre expression, recombination of the LoxP sites allows transcription of FKBP-BioID2 transgene.

6.2. Results

6.2.1. *Myc-tagged FKBP-BioID2 can be successfully targeted to the Rosa26 locus, and heterozygous offspring are generated at the expected Mendelian ratio.*

Prior to this project, the mycFKBP-BioID2 construct was designed by Monash Genome Modification Platform and was integrated to Bruce4 embryonic stem cells by the Biological Resource Centre (BRC), A*STAR, Singapore by electroporation. After generating chimeric mice with a germline mycFKBP-BioID2 transgene, a mouse line was established by genotyping from tail biopsies. To determine the inheritance status of animal, primers were designed to specifically amplify either the wild-type, or the modified Rosa26 locus. A typical genotyping experiment involved amplifying the neomycin cassette and a region spanning the 5' junction between the endogenous and wild-type sequences to detect the knock-in, while a wildtype primer set amplified the WT allele to validate the reaction (Figure 6-3A). When
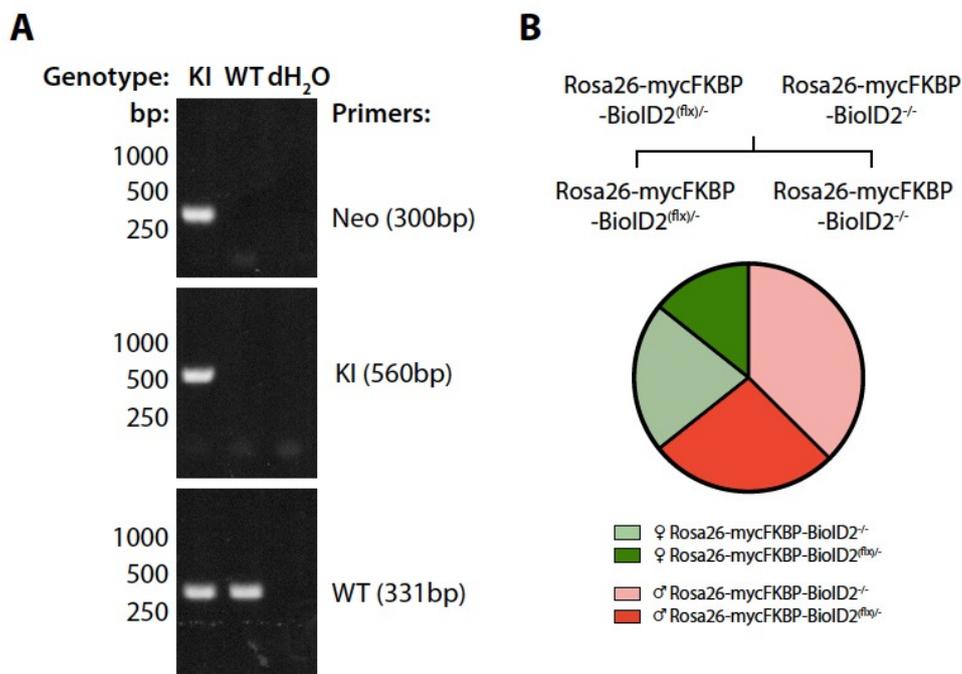
Figure 6-3 - C57BL/6 mice successfully inherit and propagate the Rosa26-mycFKBP-BioID2 allele. (A) PCR reactions used to determine the genotype of animals. (B) Schematic for genetic inheritance of FKBP-BioID2 and records of offspring genotype were recorded and tested for Mendelian inheritance (N=56; χ^2 test, p = 0.0925).

heterozygous Rosa26-mycFKBP-BioID2$^{(flx)/-}$ animals were crossed with WT mice, the offspring genotypes were detected at the expected Mendelian ratios (Figure 6-3B), indicating that a single modified Rosa26 allele does not affect viability.

6.2.2. *Myc-tagged FKBP-BioID2 expressed in a Cre-dependent manner from the Rosa26 locus in fibroblasts*

MycFKBP-BioID2$^{(flx)/-}$ mouse adult fibroblasts (MAFs) were isolated from a tail biopsy of a Rosa26-mycFKBP-BioID2$^{(flx)/-}$ mouse to assess Cre-dependent recombination of the locus *in vitro*. Doxycycline treatment of mycFKBP-BioID2$^{(flx)/-}$ MAFs stably transduced with a doxycycline inducible Cre-containing lentiviral vector was predicted to yield mycFKBP-BioID2$^{+/-}$ MAFs, expressing mycFKBP-BioID2. As expected, Cre-expression lead to efficient and complete recombination which was monitored by PCR (Figure 6-4, Table 6-1). It is notable that PCR products corresponding to both the cassette and recombined gene assemblies were detected in the non-dox-induced, Cre-integrated cells (Figure 6-4).
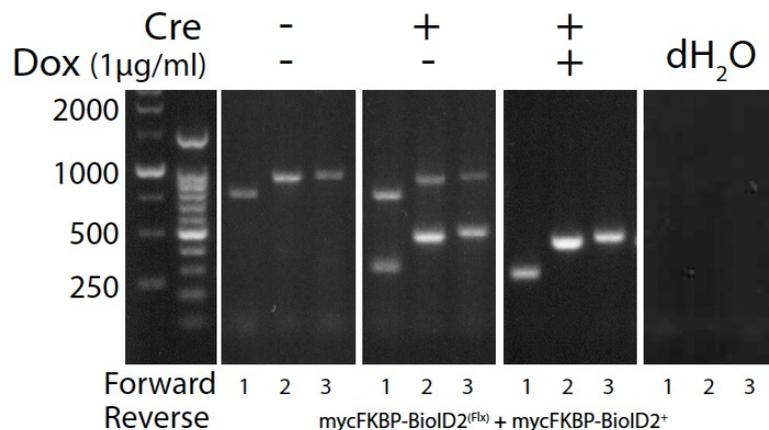
Figure 6-4 – mycFKBP-BioID2 is efficiently and universally activated mouse adult fibroblasts. MAFs were infected by doxycycline inducible cre-lentivirus. Genomic DNA was purified and amplified by PCR using a combination of three primers; one common forward primer, and two reverse primers specific for either the floxed, or recombined allele.

Forward Primer:	Reverse Primers:	Rosa26-mycFKBP-BioID2$^{(Flx)}$	Rosa26-mycFKBP-BioID2^{+}
		Expected PCR product (bp):	
1		749	289
2		904	444
3		930	470

Table 6-1 – The predicted size of PCR products using genotyping primers targeted to Rosa26-mycFKBP-BioID2$^{(Flx)}$ and Rosa26-mycFKBP-BioID2^{+} Rosa26 alleles.

Next, it was necessary to establish that the mycFKBP-BioID2 protein was expressed from the Rosa26 locus, and enzymatically functional. The predicted molecular weight of the mycFKBP-BioID2 protein is approximately 39 kDa, comprising the 26 kDa BioID2, 11.8 kDa FKBP tag and 1.2 kDa myc-tag. mycFKBP-BioID2 MAFs were analysed by western blotting, and in the presence of dox, antibodies directed against the myc-tag or the BioID2 enzyme revealed an immuno-reactive band at the expected molecular weight (Figure 6-5A). It was also important to verify localisation of the enzyme. mycFKBP-BioID2$^{+/-}$ MAFs were therefore analysed by immunofluorescence microscopy, using antibodies raised against the myc-epitope and BioID2. It was found that mycFKBP-BioID2 had a diffuse localisation throughout the cytoplasm and nucleoplasm (Figure 6-5B), an observation consistent with that described previously (Chojnowski *et al.*, 2018). It was also observed that each cell has a very similar expression level, most likely due to known and universal gene dosage in these targeted cells which is rarely observed using transfection or lentiviral systems. This feature is also likely to be of advantage since heterogeneity of interactome within a cell, or tissue population will not be diluted by differential mycFKBP-BioID2 expression levels.
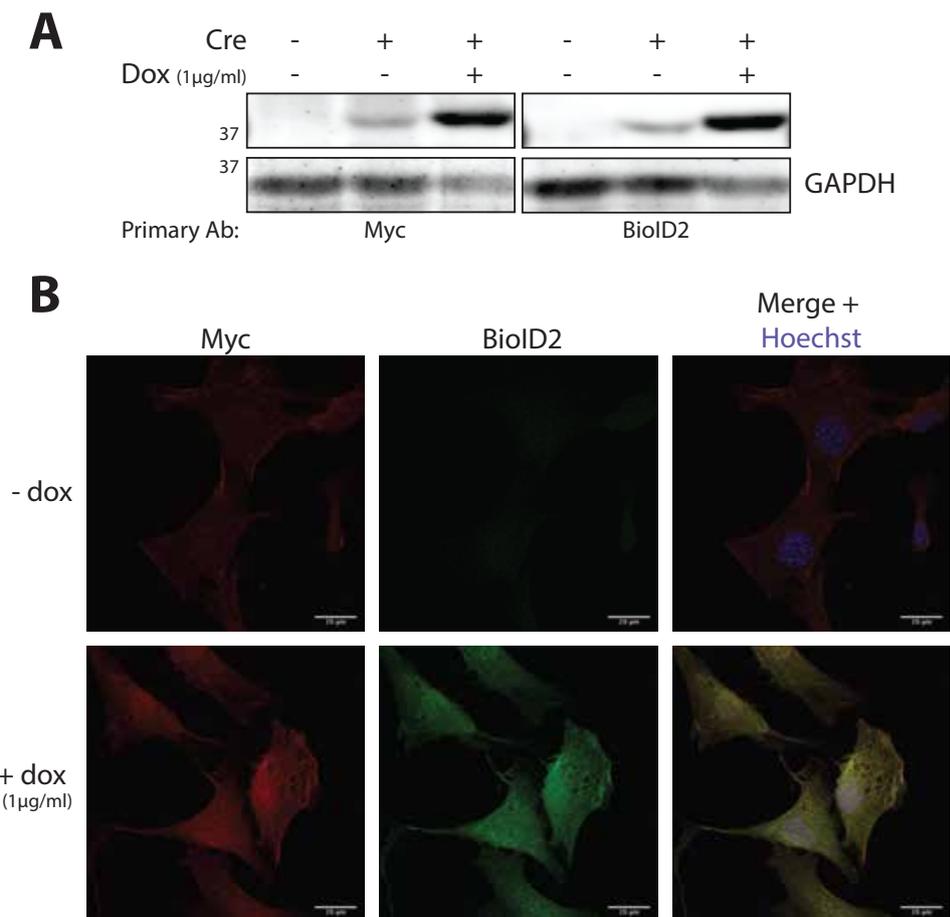
Figure 6-5 - mycFKBP-BioID2 enzyme is expressed upon doxycycline induction of Cre-recombination. FKBP-BioID2 expression was induced in MAFs by infection with Cre and doxycycline and whole cell lysates were analysed by western blotting (A) and immunofluorescence staining (B).

6.2.3. *MycFKBP-BioID2 expressed from the Rosa26 locus is functional in MAFs.*

An important parameter of the 2C-BioID system to determine was the ability of the enzyme to biotinylate proximal peptides. The undirected cytoplasmic localisation, combined with the constitutive enzymatic activity of mycFKBP-BioID2 should be sufficient to cause a global increase in peptide biotinylation, through its stochastic activity on neighbouring proteins. This was tested by culturing mycFKBP-BioID2$^{+/-}$ MAFs in media supplemented with 50 μM biotin for 16 hours. Biotinylated peptides are easily detected using streptavidin, and analysis by western blotting indicated a 6-7 fold increase in global biotinylation upon biotin supplementation (Figure 6-6). Biotinylation was also assessed *in situ* by immunofluorescence

microscopy. Streptavidin staining was consistent and co-localised with mycFKBP-BioID2 (Figure 6-7). Western blotting also revealed the presence of prominent streptavidin reactive bands at ~130 and ~75 kDa throughout all conditions which could be attributed to endogenously biotinylated carboxylases (Myers and Nickoloff, 1999; Tytgat *et al.*, 2015). Some of these are mitochondrial proteins and a characteristic staining pattern can be observed by microscopy in the absence of both biotin and doxycycline (Figure 6-7).
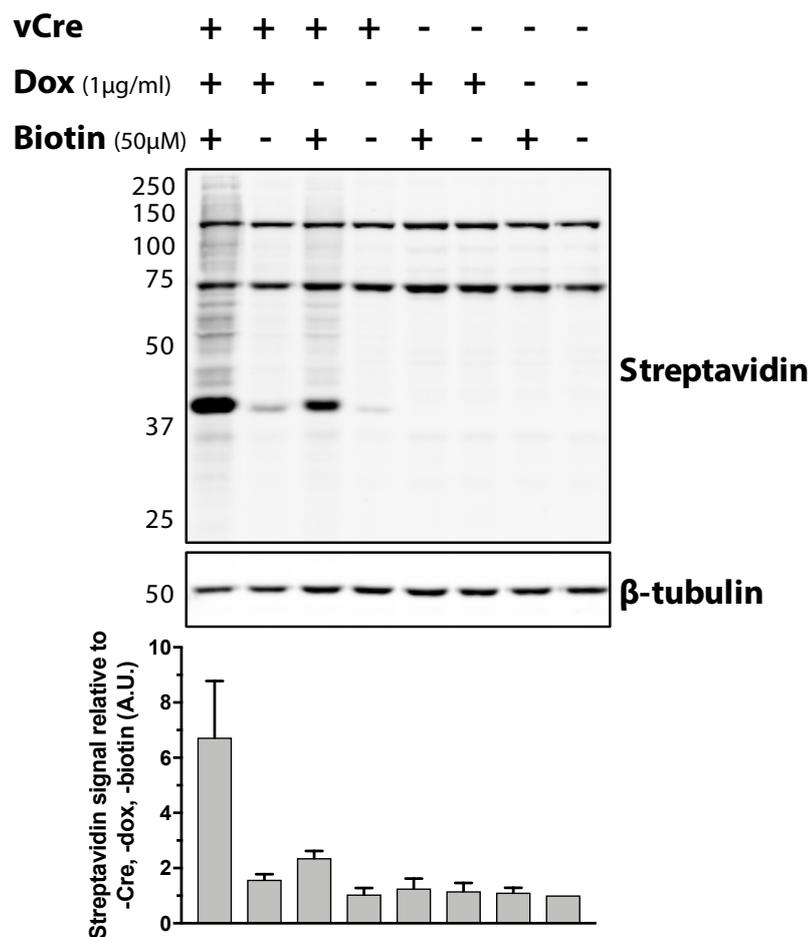
Figure 6-6 – The biotin ligase component, mycFKBP-BioID2 is active upon recombination induced by Cre virus. MAFs derived from Rosa26-mycFKBP-BioID2[+/-] mice were cultured in the presence or absence of 50μM biotin and analysed by western blotting. Biotinylated peptides were detected using streptavidin. (N=2).

Doxycyline (1μg/ml)	+	+	-	-
Biotin (50μM)	+	-	+	-
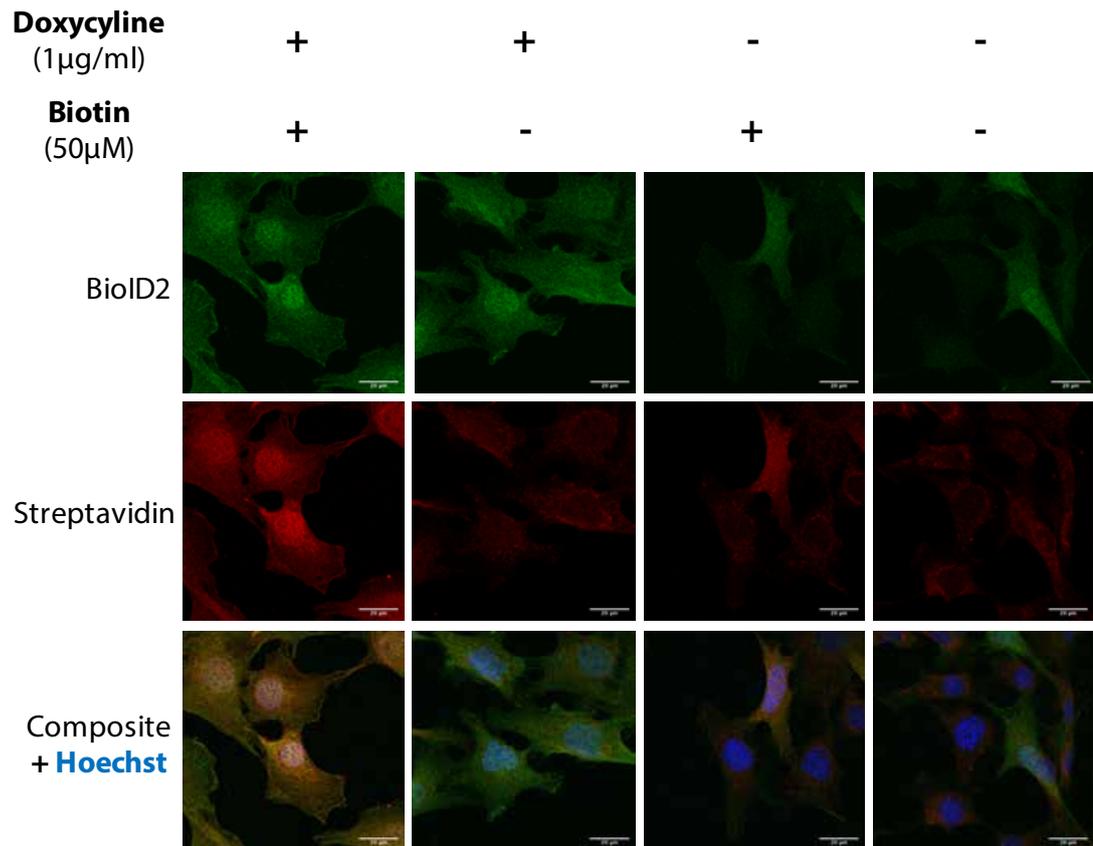
Figure 6-7 – Colocalisation of biotinylated peptides and mycFKBP-BioID2 enzyme as a result of stochastic activity. MAFs were infected with Cre-lentivirus, and expression induced by culturing with doxycycline. Biotinylated peptides were detected using Alexa-conjugated streptavidin.

An essential feature of the system to determine is the ability of mycFKBP-BioID2 to relocalise to an FRB-tagged PoI upon addition of the dimerising agent AP21597. For this, Rosa26-mycFKBP-BioID2$^{+/-}$ MAFs were transduced with doxycycline-inducible FRB-LAP2β, a fusion protein already established to correctly localise to the INM (Chojnowski *et al.*, 2018). In the presence of the dimerising compound, mycFKBP-BioID2 localises robustly to the FRB-PoI in a proportion of cells (arrowhead), however this recruitment is not detected ubiquitously (Figure 6-8). To test whether mycFKBP-BioID2 recruitment could be optimised, the dimeriser was titrated into the culture system to determine whether the original concentration was sufficient to visualise robust mycFKBP-BioID2 re-localisation to the NE. Across the experimental conditions, no difference could be observed in the re-localisation efficiency (Supplementary figure III). It seems unlikely that recruitment simply does

not occur given its routine use in other studies. It is possible, however, that the available V5FRB-LAP2β becomes saturated, and excess mycFKBP-BioID2 masks the signal from this.
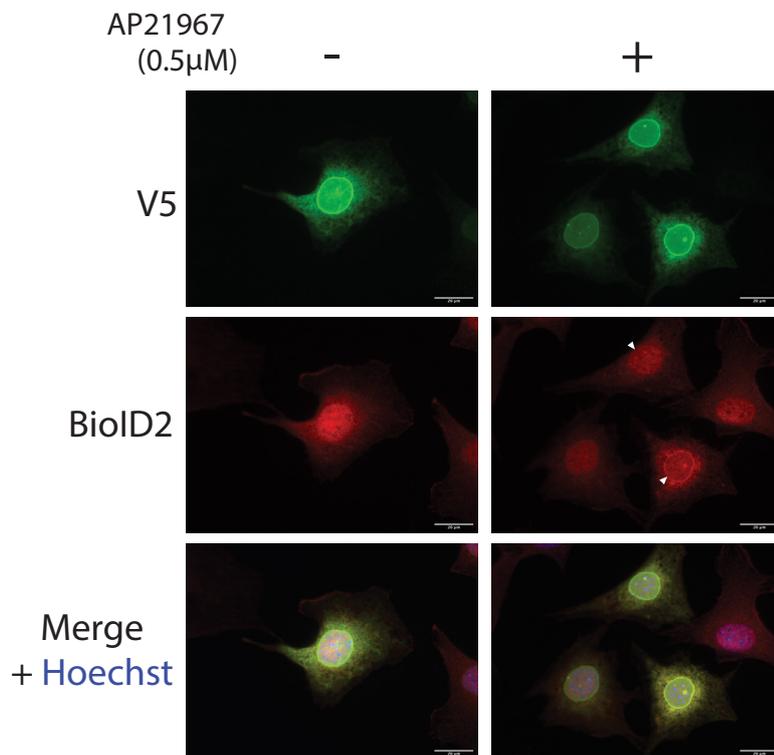
Figure 6-8 - Recruitment of mycFKBP-BioID2 to a protein of interest. MAFs were transduced with V5-LAP2β-FRB construct and recruitment was analysed by immunofluorescence microscopy upon the addition of dimerising compound AP21967. Localisation of the V5-LAP2β-FRB construct is detected using antibodies raised against V5 and shown in green. Immunofluorescence revealing BioID2 localisation is shown in red.
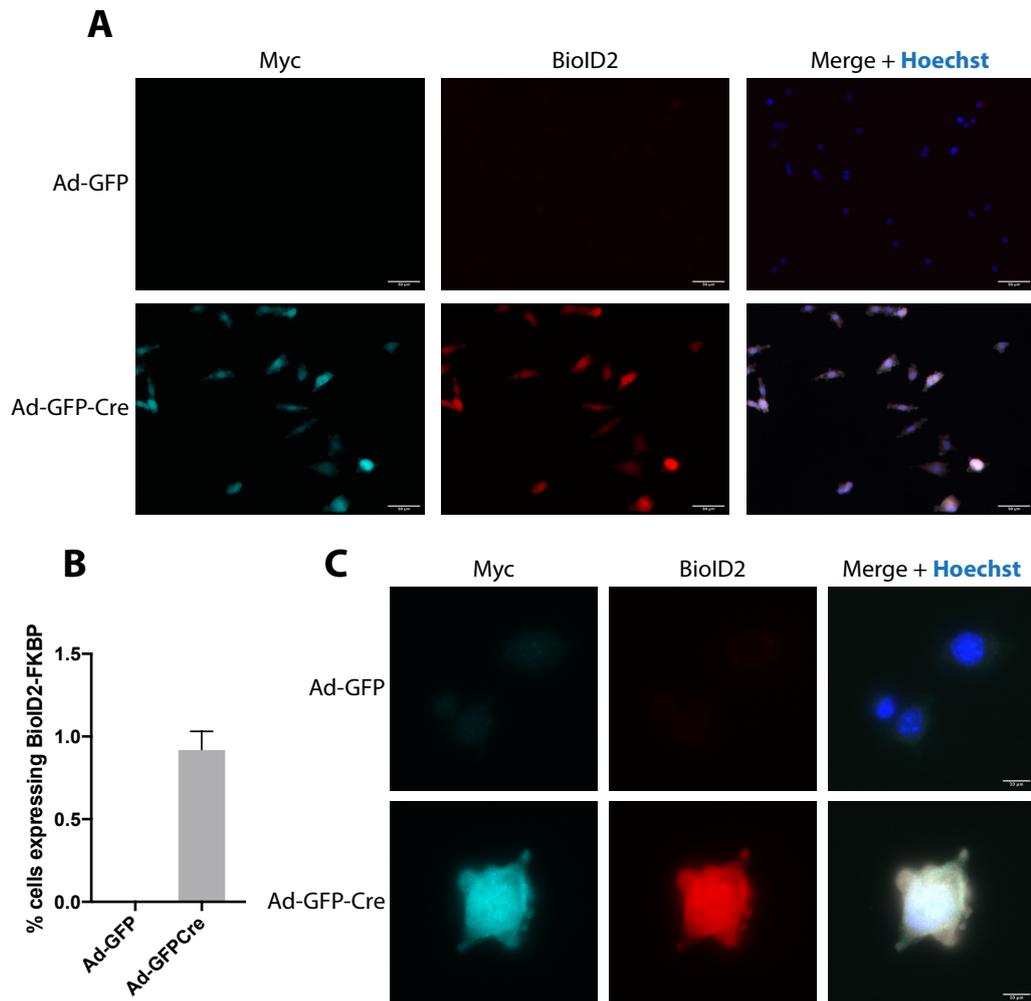
Figure 6-9 - mycFKBP-BioID is expressed in myoblasts upon GFP-Cre transduction. Myoblasts were derived from hindlimb muscle of 8 week old floxed 2C-BioID mice, transduced with adenovirus containing with GFP alone or GFP-Cre, sorted by FACS and stained with antibodies raised against the myc epitope or BioID2 enzyme. (A) 20x representative imaging of myoblasts selected by FACS for transduction with adenovirus containing either GFP or GFP-Cre, with percentage mycFKBP-BioID2 expressed cells quantified in (B), data gathered from 3 independent experiments. (C) 60x imaging shows that mycFKBP-BioID2 is localised diffusely throughout the cell.

6.2.4. *MycFKBP-BioID2 is also expressed and functional in myoblasts*

To further establish the Rosa26-mycFKBP-BioID2$^{+/-}$ KI mouse as a valid system to define the Sun1 interactome, it was important to verify the expression of the mycFKBP-BioID2 enzyme in other cell types. Given that muscle tissue is often affected by nuclear envelopathies, myoblasts and fibroblasts (MAFs) were derived from the hindlimb muscle of 8-week-old Rosa26-mycFKBP-BioID2$^{+/-}$ mice. Myoblasts were subsequently enriched by pre-plating. To induce mycFKBP-BioID2 in the experimental myoblasts, pure myoblasts were transduced using either GFP- or GFP-Cre-containing adenovirus, and expression was assessed. Adenoviruses have non-integrative genomes, thus reducing the risk of numerous genetic lesions caused by lentivirus transduction affecting cell behaviour. Multiplicity of infection (MOI) for myoblasts was determined empirically and MOI 100 was found to be the most efficient, balancing the ratio of transduced cells and viability (supplementary figure IV). Transduced, GFP-positive, cells were sorted by FACS and stained for myc and BioID2 for analysis by immunofluorescence microscopy. This revealed that infection using GFP-Cre yielded myoblasts expressing mycFKBP-BioID2 (Figure 6-9), distributed uniformly throughout the cell, comparable to the expression profile in MAFs, and as previously reported (Chojnowski *et al.*, 2018).

Despite the Rosa26 locus being known as a safe harbour for ubiquitous transgene expression (Friedrich and Soriano, 1991), it was unknown whether the expression of mycFKBP-BioID2 was persistent in different tissues and cell types. It is possible to differentiate myoblasts into myotubes *in vitro* by culturing in reduced serum medium. Although an *in vitro* phenomenon, myotubes possess many characteristics of myofibres, including the expression of genes associated with maturity, cell fusion together with nuclear positioning events, and spontaneous contractile events. In the absence of a global Cre-driver to assess tissue expression *in vivo*, myotubes were used as an *in vitro* tool to provide indication of the tolerance of another cell type to the expression of mycFKBP-BioID2 and expand its characterisation. Upon culturing under reduced serum, Rosa26-mycFKBP-BioID2$^{+/-}$ myoblasts fuse to form elongated multinucleated syncytia and expression of

230

mycFKBP-BioID2 is sustained (Figure 6-10). Together, these results indicate that myoblasts and myotubes can tolerate prolonged expression of mycFKBP-BioID2, and expand understanding of the characteristics of this transgene in cells of skeletal muscle lineage, valuable information for future *in vivo* studies.
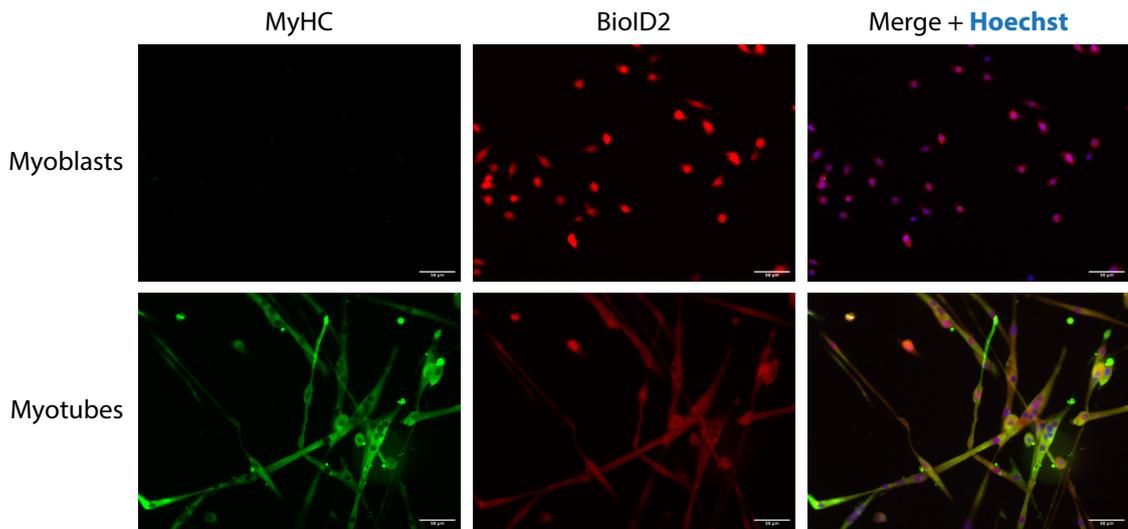
Figure 6-10 – Cre-recombined Rosa26-2C-BioID myoblasts retain competence for differentiation and show sustained expression or MycFKBP-BioID2. Cre-recombined myoblasts from Rosa26-mycFKBP-BioID2 mice were differentiated into myotubes and stained for Myosin heavy chain (MyHC) and the BioID2 enzyme.

6.2.5. *FRB-Sun1 is appropriately localised to the nuclear envelope in a variety of cell types.*

Despite the contribution of peptide fusions to our understanding of protein function, the modifications themselves risk altering the physiological function of the endogenous protein. Modified proteins of the INM have heightened predisposition to become non-functional, since their import to the nucleus is strictly governed by the NPC. The current notion maintains that a smaller modification should, in many cases, improve physiological function of a PoI. 2C-BioID was used to define the nucleoplasmic Sun1 interactome since the 11.2 kDa FRB modification is smaller than a direct fusion of the BioID2 enzyme. Upon expression in HeLa cells, V5FRB-Sun1 was found to localise appropriately to the NE (Figure 6-11A). Similarly, MAFs transduced with a V5FRB-Sun1 under doxycycline control also displayed robust localisation of the construct to the nuclear periphery (Figure 6-11B) upon culture under increasing doses of doxycycline (Figure 6-11C).

6.2.6. *Defining the myoblast Sun1 interactome*

With the aim to identify candidate nucleoplasmic interactors of Sun1 in myoblasts, cells transduced with V5FRB-Sun1 were first grown to control for its expression, appropriate localisation and biotinylation capabilities of the mycFKBP-BioID enzyme. V5FRB-Sun1 was detected using antibodies against V5 and Sun1 and expression was confirmed by analysis using western blotting and immunofluorescence microscopy. Immunofluorescence experiments indicated the peripheral nuclear localisation of V5FRB-Sun1, in line with that observed in HeLa cells and MAFs (Figure 6-12A), and analysis by SDS-PAGE revealed V5 and corresponding Sun1 immunoreactive bands at the expected molecular weight (Figure 6-12B). Probing the western blot using streptavidin showed that global biotinylation was increased when myoblasts were cultured in additional biotin (Figure 6-12B). Taken together, myoblasts from the BioID mouse transduced with V5FRB-Sun1 appeared to be a valid model in which to uncover nucleoplasmic interactors of Sun1.
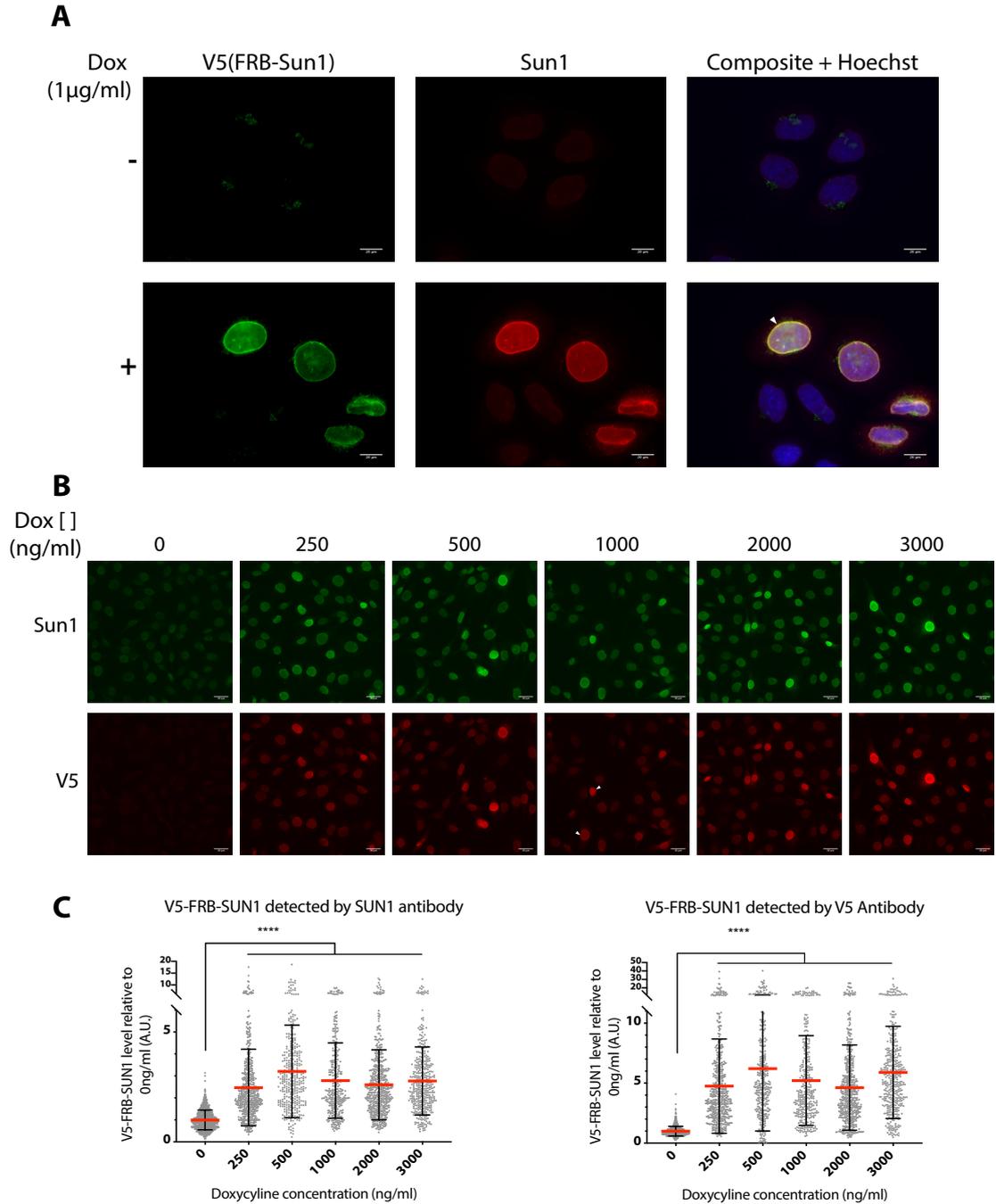
Figure 6-11 - V5FRB-SUN1 localises to the nuclear envelope. (A) HeLa cells stably transduced with TRIPZ-V5FRB-Sun1 stained for V5 and Sun1 in the presence and absence of 1µg/ml doxycycline. (B) Rosa26-mycFKBP-BioID2[+/-] stably transduced with TRIPZ-V5FRB-Sun1 were subject to various doxycycline concentrations. (C) Mean intensity quantification of (B) within nuclear mask given by Hoechst stain. N = >350 nuclei per condition. Kruskal-Wallis with Dunn's multiple comparisons test; **** = $p < 0.0001$.
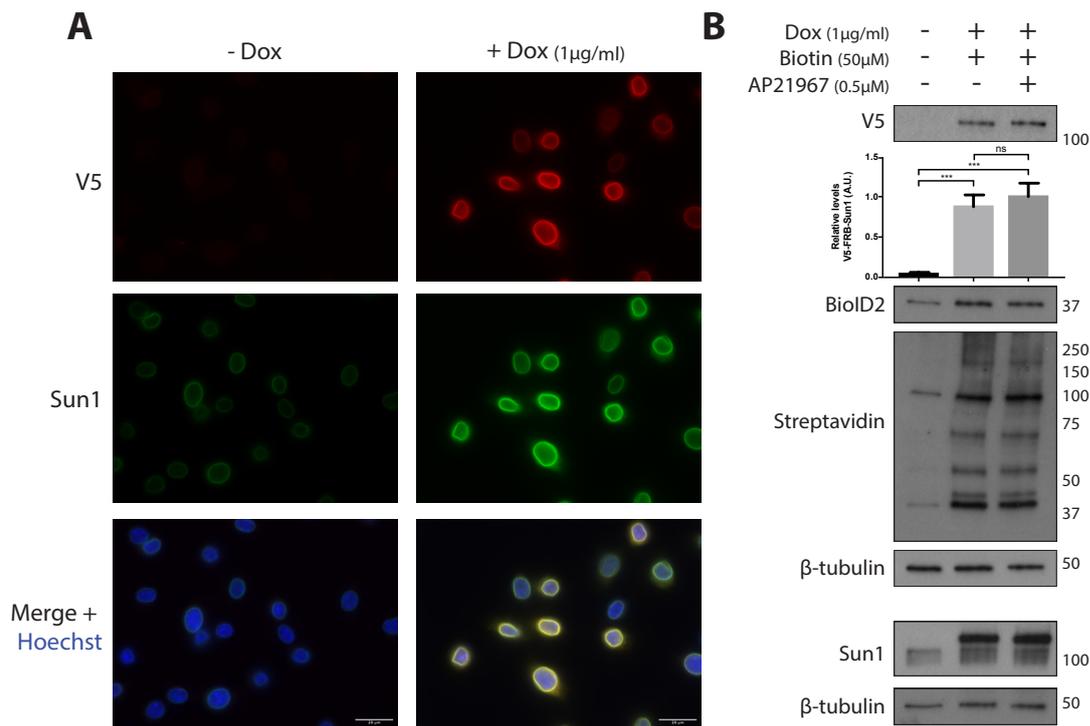
Figure 6-12 - V5FRB-Sun1 is expressed and localises properly to the nuclear envelope in mouse myoblasts. Samples of Rosa26-FKBP-BioID2 myoblasts were stably transduced with V5-FRB-Sun1, cultured in the presence or absence of doxycycline and stained using V5 and Sun1 antibodies for immunofluorescence microscopy (A). (B) Samples for proteomics analysis were or analysed by western blotting to control for bait expression, detected by V5 and Sun1 probing, and biotinylation, as shown by streptaviding blot. N=3; Two-tailed unpaired T-test, *** = p < 0.001.

6.2.7. *BioID results Sun1 interactors and selection of follow-up candidates*

Biotinylated peptides were purified from whole cell lysates made from the FRB-Sun1 expressing Rosa26-FKBP-BioID2 myoblasts using immobilised streptavidin and analysed by LC-MS. Proteins were then identified and quantified from MS samples using MaxQuant software. To highlight proteins of interest, the list of identified proteins was ranked using the BioID score, from highest potential interest (100), to lowest (0). The BioID score is a non-dimensional index built upon multiple parameters including protein abundance, peptide coverage, and queries against the CRAPome database (Chojnowski *et al.*, 2018). Compared with conventional BioID fused bait proteins, the 2C-BioID approach benefits from the fact that the addition / omission of the dimeriser provides a more accurate internal control, accounting for

234

background biotinylation, thus reducing false-positives (Chojnowski *et al.*, 2018). BioID were here calculated by taking into account the following control conditions: 1) with dox and biotin, but without dimeriser (AP21967), and 2) without dox, biotin or dimeriser (AP21967). The majority of the top 200 scorers for each analysis were shared (141 candidates), and these lists were then combined and categorised against the cellular components gene ontology (GO) database at a global and representative pathway level to identify enriched clusters of proteins (see appendix F, tables F.I and F.II). Global analysis revealed the enrichment of numerous GO terms associated with 'nuclear' (Figure 6-13A). Over half of the submitted genes are associated with the term 'nucleus' which is consistent with the localisation of the FRB-Sun1 bait construct. Given the often uninformative nature of peak hierarchy GO terms, further insight into the enriched groups of detected proteins was pursued. The analysis was therefore modified to return more specific cellular component GO terms. These analyses again indicated an enrichment in terms connected with current knowledge of Sun1, including 'nuclear envelope'. Interestingly, histone modifications were identified as enriched pathways which may indicate that components involved in genome regulation and chromatin remodelling are also detected. Thus, these observations are also in line with functions recently ascribed to LINC complexes (Le *et al.*, 2016; Tajik *et al.*, 2016) (Figure 6-13B).

To select candidates for further investigation, networks were assembled using both the global and more representative GO term hierarchy levels to assess identified groups of enriched proteins together. Candidates were mapped to their calculated BioID scores to give a visual interpretation of their relative significance. Encouragingly, the highest scoring candidates appeared on network maps constructed from both control datasets (Figure 6-14A-D). One puzzling observation is the identification of genes more commonly associated with the myofibril and sarcomere. In mature skeletal muscle fibres, Tnni1 is the slow twitch inhibitory subunit of the regulatory troponin complex (Cummins and Perry, 1978). Despite its canonical localisation in the cytosol as part of the actomysin assembly, confidence in
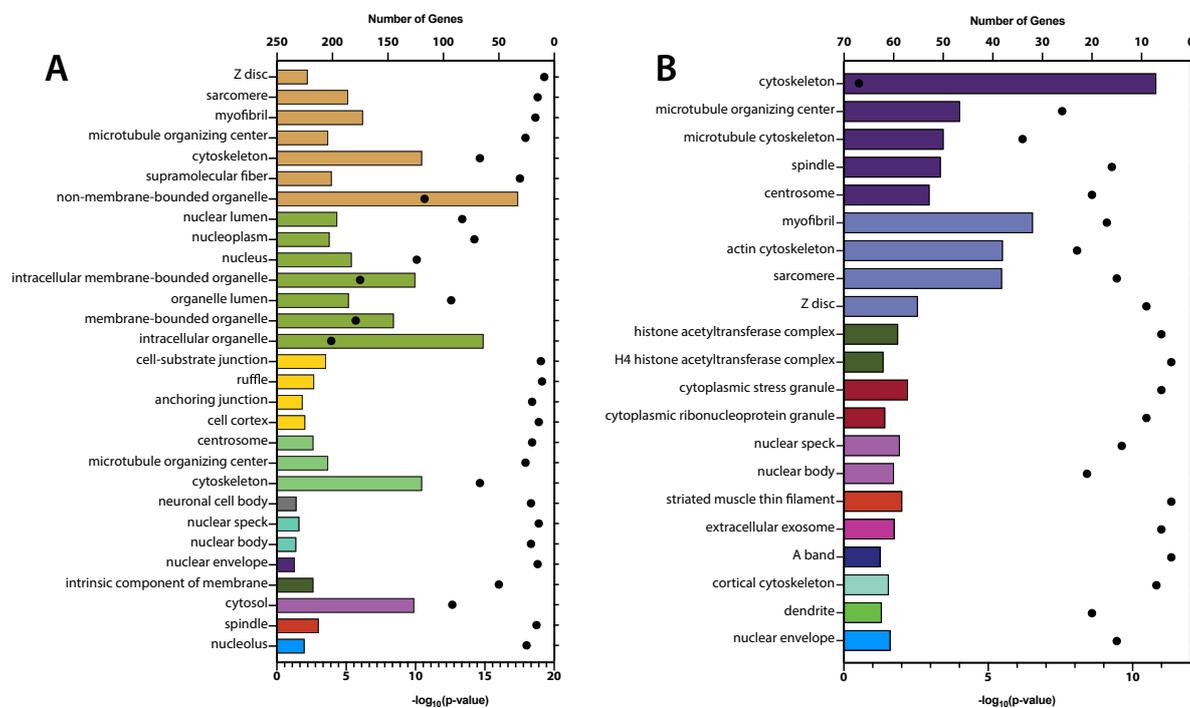
Figure 6-13 - Cellular compartment gene ontology (GO) term analysis of the combined top 200 Sun1 candidate interactors ranked according to BioID scoring based on background biotinylation in the absence of dimeriser, or in combination with control lacking of dox, biotin and dimeriser. (A) GO terms were restricted to the upper levels of the hierarchy, (B) GO terms were restricted to the mid-levels of the hierarchy to reflect more specific pathways and the number of genes associated with each term plotted (upper *x*-axis) in each analysis. In each graph, bars correspond to p-values, and points correspond to gene number. Analysis was performed using ClueGO, a plugin for Cytoscape software, and p-values were calculated for each term, with Bonferroni step-down correction.

its interaction with the nucleoplasmic terminus of Sun1 is ranked highly by BioID scoring (Figure 6-14, Table 6-2), providing rationale for further investigation. Since there is currently no consensus for optimal controls to use as background for BioID analyses, the candidates chosen for further investigation included those that ranked highly in both BioID scoring approaches, together with some of those that scored highly in one, where there was reasoning to build hypotheses from established biological mechanisms (Table 6-2).
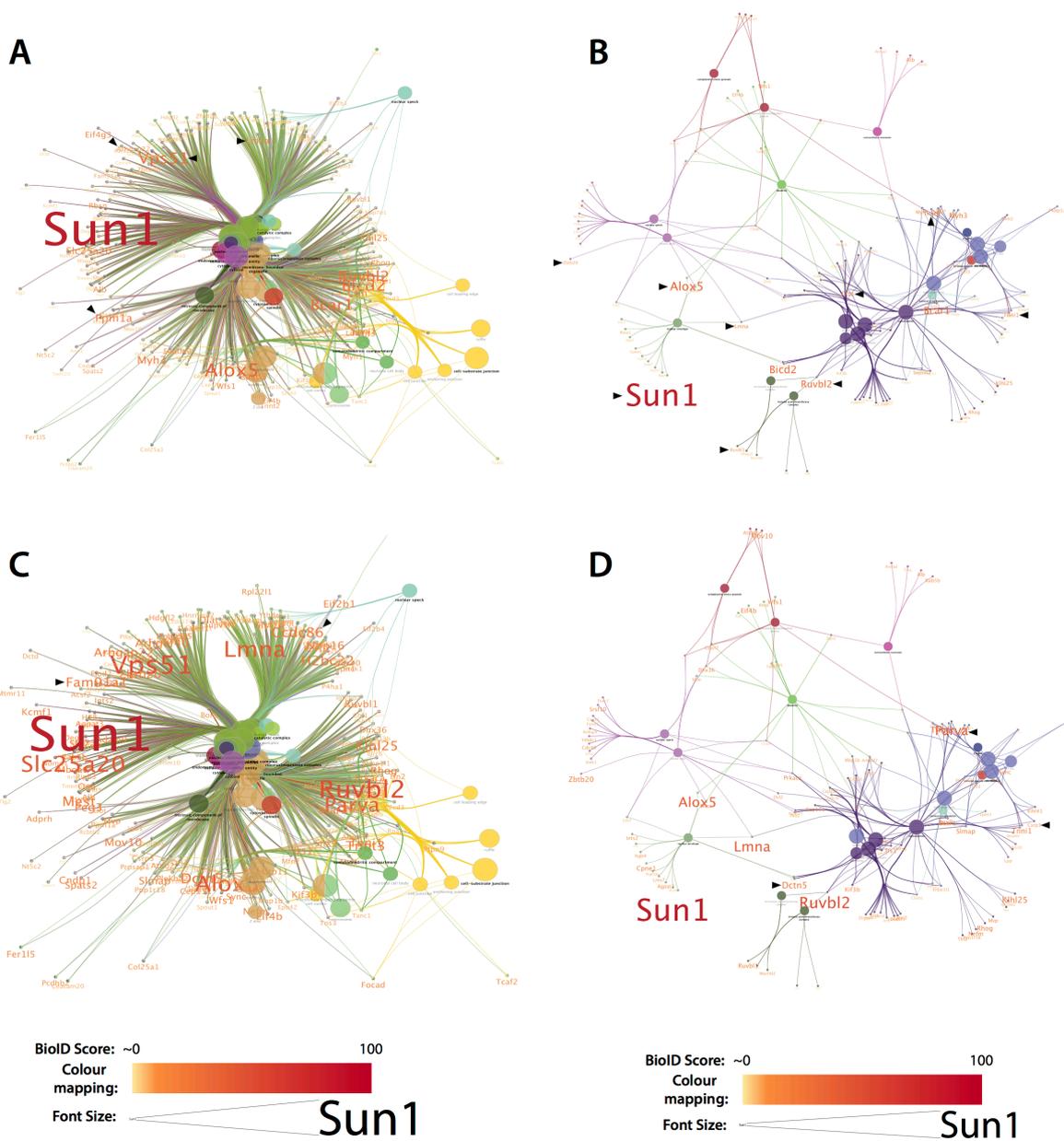
Figure 6-14 - Network maps assembled using the identified enriched GO terms and associated genes. Global (A) and specific (B) terms assembled using top 200 candidates ranked by BioID scores normalised to combined control conditons (-AP21967; and -Dox, -Biotin, -AP21967). Global (C) and specific (D) terms assembled using top 200 candidates ranked by BioID scores normalised to -AP21967 control. Networks were constructed using ClueGO, a plugin for Cytoscape software. In each network, gene symbol sizes and colours are mapped to their BioID score, as indicated. Arrows denote those genes subsequently selected for further analysis.

Gene	-Dox, -Biotin, -AP21967 and -AP21967		-AP21967	
	BioID Score	Rank	BioID Score	Rank
Sun1	100	1	100	1
Alox5	34.9511716	2	40.57759079	6
Vps51	29.74432972	4	49.47563875	2
Ruvbl2	26.48550826	6	47.75283878	3
Ppm1a	17.18972085	9	8.630845899	95
Cttn	13.98166773	11	5.418549273	199
Tnnt3	12.82129682	14	21.42576584	15
Tnni1	12.64314441	15	18.19687849	19
Lmna	11.73877239	17	41.68915717	5
Ruvbl1	8.495427662	28	14.37001639	35
Parva	6.779878256	39	37.97833333	7
Dctn5	4.612688849	77	24.09756427	9
Ccdc86	2.24449039	194	23.3335344	11
Hist1h2bp	0.311414743	498	22.18614439	12
Fam91a1	6.302523517	44	21.87402726	13
Arhgap22	8.200464459	30	18.83936366	17
Zbtb20	7.700207772	32	18.12638624	20

Table 6-2 - Sun1 interacting candidates selected for further characterisation showing their BioID scores and ranks calculated using background indicated.

6.3. Discussion

Studies presented in this chapter aimed to validate a mouse model genetically modified to ubiquitously express the 2C-BioID component mycFKBP-BioID2 *in vitro*, and generate a protein interactome for Sun1. Results indicate that the mycFKBP-BioID2 transgene is active and functional when knocked into the Rosa26 locus of the mouse. PCR analysis confirms that the induction of Cre expression results in the recombination of the LoxP sites, and this manipulation leads to the expression of the

mycFKBP-BioID2 protein as expected. Functional analyses concluded that the endogenous mycFKBP-BioID2 fusion protein behaved as predicted. Stochastic background activity of the biotin ligase was revealed by widespread streptavidin staining *in situ*, and a global ~7-fold increase in biotinylation could be observed in streptavidin probed western blotting of whole MAF cell lysates. Experiments verifying the recruitment of mycFKBP-BioID2 to an FRB-tagged protein of interest were equivocal. LC-MS analysis of streptavidin affinity purified biotinylated peptides in the experimental condition, however, robustly identified Sun1 and nuclear associated proteins.

6.3.1. *Cre induction*

One observation made during the verification experiments was the persistent, low-level expression of the mycFKBP-BioID2 enzyme in the cells transduced with Cre, but not dox induced. This trait could be detected at the genetic and protein levels, by PCR or western blotting and immunofluorescence microscopy, respectively. Tet-responsive systems have been used extensively over the past 25 years. They can be used to stringently regulate gene expression, though their basal leakiness has been reported (Shaikh and Nicholson, 2006; Pham *et al.*, 2008; Costello *et al.*, 2019). Doxycycline concentration can modulate expression levels of genes in Tet-ON systems in a dose dependent manner. However, in this experimental situation, while the basal leakiness may not result in detectable levels of Cre protein, its action on the single floxed genomic target at the Rosa26 locus is binary, regardless of the amount of Cre-expression. The increase in global biotinylation in cells transduced with lentiviral-Cre, supplemented with biotin, without dox induction appears to be rather small, indicating that the basal expression of Cre occurs only in a subset of cells.

6.3.2. *Detection of mycFKBP-BioID recruitment to bait protein*

Experiments aimed at demonstrating association of the mycFKBP-BioID2 enzyme to an FRB-LAP2β PoI were inconclusive as universal relocalisation could not be observed. To address this, the dimeriser (AP21967) was titrated into the culture, however, no differences were observed in the rate of re-localisation (Supplementary figure III). It is noted that chemically-induced dimerisation using the FKBP-FRB-

rapamycin system has been widely used in previous studies (Putyrski and Schultz, 2012), and the interaction is high affinity (Banaszynski, Liu and Wandless, 2005). It therefore seems unlikely that recruitment does not occur. Indeed, a previous study using a 2C-BioID system *in vitro* observed recruitment to an FRB-LAP2β (Chojnowski *et al.*, 2018). However, it seems that here, too, recruitment is not complete, as staining remains in the cytoplasm of the cell. It is therefore plausible that there is an excess of the mycFKBP-BioID2 in comparison to FRB-PoI. Thus, the signal from surplus mycFKBP-BioID2 would conceal the proportion localised at the NE upon the addition of AP21967. The previous study uses human fibroblasts and a lentiviral approach to constitutively express mycFKBP-BioID2. The random integration of lentiviral vectors and exogenous promoter elements introduce expression variables which render the system incomparable to the presented Rosa26-knock-in approach (Chojnowski *et al.*, 2018). Ultimately, appropriate relocalisation of mycFKBP-BioID2 to a nuclear PoI should be detected through mass spectrometric analyses. If relocalisation occurs efficiently, one would expect an enrichment of other proteins associated with the nucleus, as can be observed.

6.3.3. *On translation to in vivo studies*

Current work focussed on functional verification *in vitro*. How this scales to a whole organism is currently under investigation. To date, the Stewart laboratory has generated a globally Cre-recombined, homozygous mouse, which has reduced body weight (Hendrikje Werner, personal communication). As mentioned by Roux *et al.*, (2012), the addition of biotin to endogenous proteins by stochastic enzyme activity alters charges on endogenous proteins, thus potentially disrupting their physiological function. In the literature, the only reported mammalian *in vivo* BioID assay involved the BioID2 enzyme which was knocked into the Junctophilin 2 gene, a component of the cardiac dyad. Although this study used a protein fusion to generate novel interactors, it did find that the integration of the BioID2 gene was not detrimental to viability (Feng *et al.*, 2020). Limited studies in invertebrates also creates difficulty in gleaning insights into the tolerance of a whole organism to excess biotinylation. BioID-lineage ligases, including TurboID and miniTurbo perform optimally at 37°C (Branon *et al.*, 2018), meaning APEX tends to be the preferred

enzyme in flies and worms (Chen *et al.*, 2015; Mannix *et al.*, 2019). These studies also reported no such detrimental effects, although it is not clear that aberrant physiology resulting from APEX-fusions was studied in any depth. For future experimentation, it should be acknowledged that non-discrete expression and localisation of mycFKBP-BioID2 may have some detrimental effects in a mammalian system. Indeed, some tissues may have greater sensitivity to biotinylation than others. This is particularly important to recognise and rigorously control for when dealing with complex multi-tissue disease models.

Another consideration when moving towards an *in vivo* BioID system is how the protein of interest in expressed and tagged. To use Sun1 as an example, it is now established that Sun1 becomes extensively spliced in various tissues; indeed, maturation of muscle sees the full length isoform downregulated and multiple shorter versions become expressed (Göb *et al.*, 2011; Nishioka *et al.*, 2016; Loo *et al.*, 2019). It is therefore likely that the Sun1 interactome is developmentally dynamic. To investigate a comparative protein interactome for the full length Sun1 isoform in a model of mature muscle, myoblasts transduced with V5FRB-Sun1 were differentiated into myotubes. However, the addition of doxycycline to these cells did not induce expression, unlike the myoblast lines (Supplementary figure V). The CMV promoter has been previously reported to be suboptimal in cells of muscle lineage, so a constitutive EF1a promoter was also tested and found to robustly express V5-FRB-Sun1 in myoblasts, but these cells failed to differentiate into myotubes efficiently, and expression of the transgene was low and heterogeneous within myotubes. A PGK or muscle hybrid (MH) promoter may be a good replacement, since the latter has been specifically designed for *in vivo* delivery of gene therapies to mature muscle, and was developed in C2C12 myotubes (Jackson *et al.*, 2013; Piekarowicz *et al.*, 2019). Ultimately, however, it is likely preferable to directly modify the gene of interest to directly fuse the FRB moiety. This approach is more likely to ensure expression under endogenous control, and capture a more representative and physiological interactome, since in the case of Sun1, the FRB modification would be maintained on all isoforms.

242

6.3.4. *BioID identified Sun1 interactors*

Since its description in 2012, BioID has become a widespread starting point to discover novel PPIs. However there does not appear to be consensus over the type of control, or indeed the number of controls required for reliable identification of interactors, minimising false positives, while maximising detection of low abundance interactors without further experimentation for validation. For these reasons, proteins selected for further analysis were chosen using BioID scoring calculated using two different negative control conditions. While both of these analyses revealed proteins determined through GOterm analysis to be in the nuclear region, these were typically not specific to the NE, the position of Sun1. Some solace can be sought in the robust identification of the Sun1 bait, which homotrimerises, as well as the identification of Lamin A, a known Sun1 interactor. Other NE components known to interact with the nucleoplasmic domain of Sun1 via *in vitro* assays such as the B-type lamins (Nishioka *et al.*, 2016), emerin and short Nesprin-2 isoforms (Haque *et al.*, 2010; Meinke *et al.*, 2014) are notably absent from the interactome generated. Of course, some interactions, such as that between Sun1 and the B-type lamins remain controversial, as *in vitro* approaches specifically exclude this interaction (Crisp *et al.*, 2006; Haque *et al.*, 2006). The Sun1 interactome generated during the current study does not address whether some of the Sun1-NL interactions are cell type specific. The heterogeneity of Sun1-NL interactions could be answered either by analysing the Sun1 interactome in numerous cell types, or by an *in vivo* approach.

Overall, it is difficult to accurately and absolutely interpret raw BioID data which, in the Sun1 interactome experiments presented here, is borne out by the disparity between proteins identified between the two methods of analysis. For this reason, potential interacting candidates gathered from BioID experiments typically depend on at least one other independent experimental method to validate the predicted *de novo* interaction prior to further investigation. Therefore, experiments aimed at validating candidate Sun1 interactors are the subject of the next results chapter.

6.3.5. *Chapter summary*

In summary, the results presented in this chapter showed that the floxed mycFKBP-BioID2 transgene was efficiently activated and functional upon Cre expression in fibroblasts and myoblasts, and its expression persisted through myoblast differentiation *in vitro* into myotubes. Moreover, an FRB-tagged Sun1 construct appears to localise as expected at the NE in a variety of cell types, including primary myoblasts. Finally, using the validated FRB-Sun1 bait protein, the Sun1 protein interactome was defined using the 2C-BioID technique in primary myoblasts derived from the Rosa26-FKBP-BioID2 mouse, revealing nuclear proteins and known interactors of Sun1.

Chapter 7: Validation of Sun1 candidate interactors

7.1. Introduction

BioID has proved to be a valuable system to begin examining protein interactors. However, it is paramount to use an independent method to verify novel interaction to rapidly exclude aberrantly identified proteins. As a method based on proximity biotinylation, proteins discovered using BioID might not be relevant interactors; they are simply localised within the same protein neighbourhood. Moreover, due to the constitutive nature of the BioID enzyme, unrelated proteins within the translation or trafficking machinery may also be identified. While 2C-BioID and BioID scoring aims to minimise these problems, it still merely serves as a tool to highlight the most likely interactors, and scoring alone is not able to unequivocally determine binding partners since the technique is fundamentally proximity-based, mass spectrometry datasets are often incomplete, and the CRAPome is formed from varied experimental approaches and cell types. Investigations presented in this chapter aimed to verify genuine interactors of Sun1 from those proteins identified through the BioID assay.

7.1.1. *Methods to verify protein-protein interactions*

An extensive toolbox of biochemical and non-biochemical methods now exists to verify protein-protein interactions, and different methods were evaluated based on specific experimental requirements to assess multiple interacting proteins of Sun1. *In vitro* biochemical approaches which were considered include assaying direct interactions of recombinant proteins, or testing protein associations originating in cell lysates by probing an immunoprecipitated sample of Sun1 for putative interactors which are co-purified. Working with membrane bound or NL proteins, however, is complex due to insolubility. This can be overcome by using non-membrane bound protein fragments *in vitro*, though it is acknowledged that these

may not faithfully capture full biological context of Sun1, including concerns over protein localisation and whether truncated proteins are accurately folded. As an alternative strategy to evade issues of protein folding and solubility, there are now a number of techniques which can be used to test protein interactions *in vivo*, within living cells. *In vivo* interaction assays benefit from the use of fluorescent proteins to provide an easily assessable visual readout signifying the detection of PPI. Förster resonance energy transfer (FRET) describes a technique utilising complementary fluorescent proteins with overlapping emission and excitation spectra (De Los Santos *et al.*, 2015). When fused to proximal proteins, excitation of shorter wavelength protein results in emission at the level of excitation for the longer wavelength protein, thus resulting in emission spectra corresponding to the proximal protein. While this technique can be applied with ease to membrane proteins, applicability to NE proteins is complicated by the large modifications required and consequent difficulties in the NE targeting of these proteins. Moreover, FRET efficiency is dependent on a high degree of protein overexpression, which can lead to molecular crowding and false positives when examining proteins in small organelles (Snapp and Hegde, 2006). To screen numerous interactors of the NE-retained Sun1, it is therefore desirable to employ techniques which are not dependent on very high protein expression, which might be difficult to control for multiple candidates, and importantly, those strategies utilise small modifications.

Combining these requirements with an *in vivo* approach are protein complementation assays (PCAs), which are PPI detection systems whereby a functional protein is divided into two (or more) non-functional fragments. One such PCA is bimolecular fluorescence complementation (BiFC). BiFC divides a fluorescent protein into two non-fluorescent fragments, which display minimal self-association when unfused to proteins of interest (Hu, Chinenov and Kerppola, 2002; Ohashi *et al.*, 2012). Given the total modification to endogenous proteins is smaller than that used for FRET for instance, BiFC is well suited to verifying PPIs of the NE where protein targeting is an issue. Upon fusion to interacting proteins, fluorescent protein fragments are guided into close proximity where they become reconstituted, providing a fluorescent readout (Figure 7-1). Studies have also found BiFC to be more sensitive than other strategies to lower expression levels and lower affinity

interactions (Magliery *et al.*, 2005; Piehler, 2005; Xing *et al.*, 2016), a useful property for evaluating previously unknown interactions. Once associated, BiFC fragments remain stable which is suitable for the detection of transient interactors, though this feature does increase the likelihood for aberrantly detecting false positive interactors.

Nevertheless, BiFC was chosen as a strategy to validate new Sun1 interactors for its advantages over conventional biochemical assays which would require labour intensive experimentation to optimise solubilisation conditions for Sun1, as a membrane anchored protein which interacts with the NL. Moreover, the fluorescent readout in BiFC is easily assessed using standard equipment such as flow cytometers while fluorescence microscopy can gather some positional information for the interaction.

7.1.2. *Chapter aims and hypothesises*

The overall aim presented in this chapter was to verify BioID-identified interactors of Sun1 using BiFC. To do this, candidate interactors were characterised by expression with amino- and carboxy-terminal epitope tags, and these were correlated to BiFC-fusions to ensure reproducibility of protein localisation. Fluorescence microscopy was used to verify BiFC reconstitution, while flow cytometry (BiFC-FC) was used to quantify candidates with the strongest signals to identify follow-up candidates. Validated candidates would then be further evaluated to attempt to determine a functional relevance to their interaction with Sun1.

7.2. Results

7.2.1. *FLAG-tagged Sun1 candidate interactors do not localise exclusively to the nuclear periphery*.

Analysis of proteins identified as proximal to Sun1 through the 2C-BioID assay in primary myoblasts revealed that the function of many are not well established. Therefore, candidate genes were amplified by PCR from Rosa26-mycFKBP$^{+/-}$ cDNA, FLAG-epitope tagged separately at both their carboxy- and amino-termini to identify any artefacts arising as a result of the modification and cloned into an expression vector. Expression in HeLa cells revealed that many of the proteins were consistently localised regardless of which terminus was FLAG-tagged (Supplementary Figure VI, Table 7-1), providing confidence in their determined localisation, and that the modification does not affect this. However, the localisation of Sun1 candidates as determined experimentally often contradicted that recorded in Uniprot or Human Protein Atlas databases (Table 7-1). It was thought that some proteins might possess cell-specific functions and localisation, so FLAG-tagged constructs were also expressed in C2C12 cells to test this idea. Candidate protein localisation in C2C12 were overall comparable to that previously observed in HeLa cells (Supplementary Figure VII, Table 7-1), thus providing confidence in the experimentally determined localisations.

Gene	Uniprot localisation	Human Protein Atlas localisation	Construct	HeLa localisation	Consistency Nt/Ct	Consistency Data-base	C2C12 consistency with HeLa
Sun1	Nucleus	Nuclear membrane	Ct-HA	Nuclear membrane			
			HA-Nt				
Alox5	Nucleus	Nucleoplasm	Ct-FLAG	Cytoplasmic, nucleoplasmic			
			FLAG-Nt				
Vps51	Golgi, endosomes	Golgi, endosomes, nucleoli	Ct-FLAG	Cytoplasmic vesicles			Also nucleoplasmic
			FLAG-Nt				
Ruvbl2	Nucleus	Nucleoplasm, cytosol	Ct-FLAG	Cytoplasmic, nucleoplasmic			
			FLAG-Nt				
Ppm1a	Nucleus and cytosol	Cytosol	Ct-FLAG	Cytoplasmic, nucleoplasmic			
			FLAG-Nt				
Cttn	PM and cytoskeleton	PM, vesicles, Golgi, cytosol	Ct-FLAG	Plasma membrane, cytoplasmic			
			FLAG-Nt				
Tnnt3	Cytosol	No data	Ct-FLAG	Nucleoplasmic			Cytoplasmic
			FLAG-Nt	Nucleoplasmic, nucleolar			
Tnni1	Cytosol	Nucleoplasm, nuclear bodies	Ct-FLAG	Nucleolar, nucleoplasmic, cytoplasmic			No expression
			FLAG-Nt				
Lmna	Nucleus	Nucleus	Ct-FLAG				
			FLAG-Nt				
Ruvbl1	Nucleus	Nucleoplasm, cytosol	Ct-FLAG	Cytoplasmic, some nucleoplasmic			
			FLAG-Nt				
Parva	PM, cytoskeleton	FAs, cytosol, actin filaments	Ct-FLAG	Nucleoplasmic, cytoplasmic, PM			
			FLAG-Nt				
Dctn5	Nucleus, cytoskeleton	Nucleoplasm, nuc. membrane	Ct-FLAG	Cytoplasmic, nucleoplasmic, vesicular			
			FLAG-Nt				No expression
Ccdc86	nucleus	Nucleoplasm, nucleolar	Ct-FLAG	Nucleoplasmic; peripheral			
			FLAG-Nt	Nucleoplasmic, nucleolar			
Hist1h2bp	Nucleus	No data	Ct-FLAG	Nucleoplasmic; enriched at nuc. periphery			
			FLAG-Nt				
Fam91a1	Golgi	Nucleoplasm, MTs	Ct-FLAG	Cytoplasmic, nuclear periphery			
			FLAG-Nt				
Arhgap22	Nucleus	Nucleoplasm, cytosol	Ct-FLAG	Cytoplasmic, PM			
			FLAG-Nt				
Zbtb20	Nucleus	Nucleoplasm, nuclear bodies	Ct-FLAG	No expression			No expression
			FLAG-Nt	Nucleo- and cytoplasmic			No expression
Csrp3	Nucleus, cytoskeleton	No data	Ct-FLAG	PM, cytoplasmic, nucleoplasmic			Apparently not nucleoplasmic
			FLAG-Nt				

Table 7-1 - Localisation of Sun candidate interactors tagged at the N- and C-terminus with a FLAG epitope tag in HeLa and C2C12. Coloured boxes indicate the whether the determined localisation is consistent with the oppositely tagged construct, the localisations recorded in on Uniprot or Human Protein Atlas, and whether those determined in C2C12 myoblasts are consistent with HeLa cells. Green = consistent, Yellow = some discrepancies, Orange = inconsistent.

7.2.2. _Establishing biomolecular fluorescence complementation (BiFC) for validation of Sun1 interactors_

A striking observation from experiments to study the subcellular localisation of Sun1 candidate interactors was the absence of proteins with well-defined nuclear membrane localisation. Rather, many of the identified proteins localised to the nucleoplasm, or both nucleoplasm and cytoplasm. To independently verify Sun1 interactors identified in BioID, BiFC was used due to its ease of throughput for a number of candidate interactors, simple fluorescent readout, and ability to detect low affinity interactions between proteins in physiological conditions. The system used in this study is based on the yellow fluorescent protein-derivative, Venus, which is asymmetrically divided at amino acid 210 into two fragments which do not fluoresce without their counterpart (Ohashi _et al._, 2012). Briefly, the protein of interest and potential interactors are fused to complementary Venus fragments which do not fluoresce without their counterpart (Figure 7-1A). Upon close proximity, directed by the interactions of the fused protein of interest, the Venus fluorescent protein becomes reconstituted, regaining its fluorescent properties (Figure 7-1B), thus signalling pairs of proteins of further biological interest.
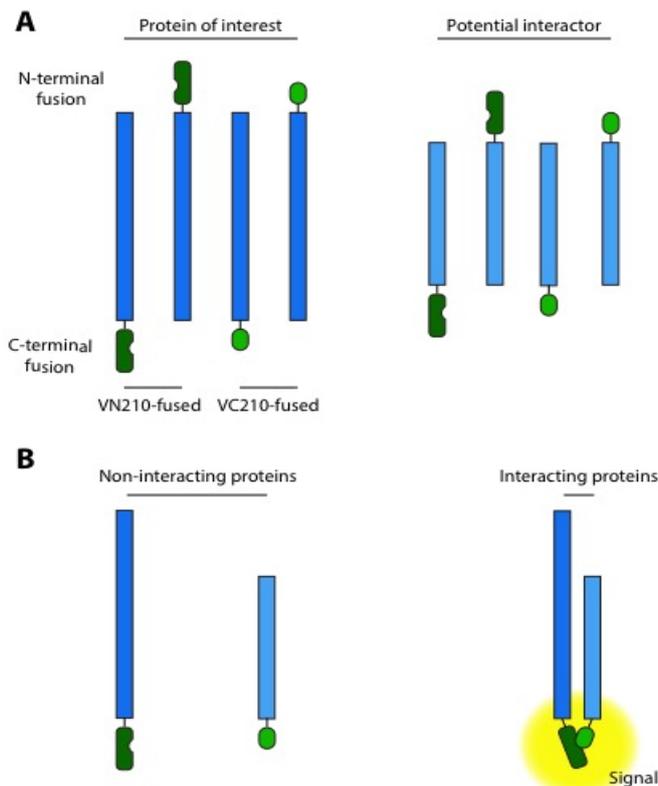
Figure 7-1 – Bimolecular fluorescence complementation. (A) Indicating a typical BiFC experimental set-up. VN210 and VC210 mVenus fragments are fused to both the N- and C-termini of query proteins. (B) Protein of interest and potential interactors fused to complementary fragments are co-transfected. Interacting proteins cause mVenus reconstitution and yellow fluorescence.

To establish a BiFC screen for Sun1, it was first sought to verify that Sun1 BiFC fusion proteins were expressed and localised to the nuclear periphery. Expression of Sun1 fused to either the VN210 or smaller VC210 BiFC fragments in HeLa cells, staining with integral HA- or myc-epitope tags, respectively, followed by analysis by fluorescence microscopy revealed that carboxyl-terminus tagged Sun1 fusions localised appropriately to the NE, as indicated by co-localisation with lamin B1 (Figure 7-2B, arrows). However, nucleoplasmic amino-terminus fusions did not localise well to the nuclear periphery, and were instead observed largely within an ER-like structure (Figure 7-2B, arrowheads). Topologically, a nucleoplasmic fusion was necessary to verify interactions within the nuclear interior. A common strategy to improve targeting to the INM involves reducing the size of the modification, however this is not possible for BiFC tags. Another approach is to modify linking peptides between the protein of interest and tag. Those originally encoded within the VC210-myc-Sun1 construct were optimised for Cold Fusion cloning efficiency rather than fusion protein function and amino acids encoded by this sequence have many bulky side chains, possibly reducing its flexibility. Therefore, the linker between the VC210-myc-tag and Sun1 was replaced by a more flexible sequence consisting of either one, two or three Gly_4Ser repeats (G_4S) (Figure 7-3A). To test whether G_4S linking sequences improve expression and localisation of VC210-tagged Sun1, VC210-myc-Sun1 and VC210-myc-x(G_4S)-Sun1 (where x is the number of G_4S repeats) constructs were transfected into HeLa cells. Immunofluorescence microscopy showed that flexible G_4S linkers greatly increased the tolerance of expression compared with VC210-myc-Sun1 (Figure 7-3). In addition, replacing the linking sequence with 1-3 G_4S repeats also improved the localisation of VC210-myc-tagged Sun1, which can be observed at the NE, aligned with a co-stain for lamin B1 (Figure 7-3B, arrowheads). Cells in which the VC210-myc-x(G_4S)-Sun1 is overexpressed display NE localisation, but also considerable cytoplasmic staining (Figure 7-3B, arrows). This is likely 'spill-over', which has been previously reported by others (Crisp *et al.*, 2006), and is thought to occur as binding sites at the NL become saturated with LINC complexes (Powell and Burke, 1990). Nevertheless, adding G_4S appears to significantly improve the expression and localisation of the VC210-myc-x(G_4S)-Sun1, demonstrating its feasibility in a BiFC assay.

To validate that VC210-myc-*x*(G$_4$S)-Sun1 were functional in BiFC assays, constructs were co-transfected into HEK293T cells alongside VN210-tagged lamin A, a known Sun1 interactor. Fluorescence microscopy showed that Sun1 constructs with the G$_4$S linking sequence gave rise to nuclear rim-like fluorescence comparable to a positive control where cells were transfected with lamin A constructs fused separately to the complementary VN210 and VC210 Venus fragments (Figure 7-4A). Additionally, co-expression of VC210-myc-*x*(G$_4$S)-Sun1 with VN210 tagged lamin A revealed a significantly BiFC signal compared to that obtained using the VC210-Sun1 construct using the original linker paired with VN210-Lamin A constructs (Figure 7-4A,B). These results further show that the G$_4$S linking sequence render the VC210-myc-*x*(G$_4$S)-Sun1 constructs useful for further validation of candidate interactors identified in from the Sun1 BioID experiment.
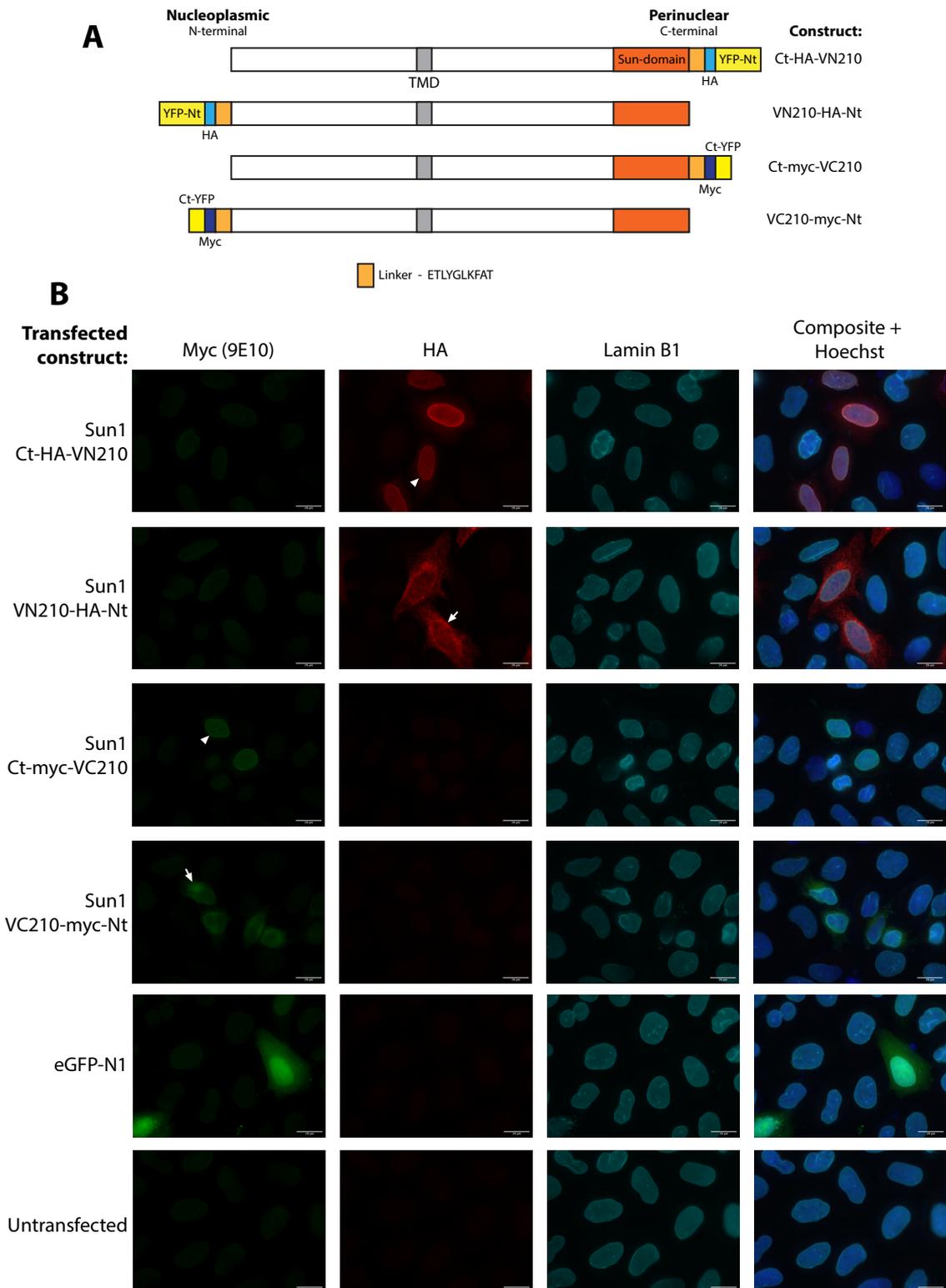
Figure 7-2 - N-terminal, but not C-terminal, BiFC tags fused to Sun1 prevent nuclear envelope localisation. (A) Schematic diagrams of Sun1 BiFC constructs. (B) Separate transfection of BiFC-tagged Sun1 constructs in HeLa cells and staining with antibodies raised against respective Myc and HA epitope tags reveals subcellular localisation. Arrowhead indicating representative nuclei where the transfected Sun1 construct is appropriately confined to nuclear periphery. Arrow denotes nuclei where Sun1-fusion constructs are not localised to the NE.

A

VC210-myc-**1**(G$_4$S)-Nt

G$_4$S

TMD

Sun-domain

Ct-YFP

Myc

VC210-myc-**2**(G$_4$S)-Nt

VC210-myc-**3**(G$_4$S)-Nt

Flexible linker - GGGGS

B

| | Myc (9B11) | Lamin B1 | Composite +Hoechst |

Sun1-myc-VC210

VC210-myc-Sun1

VC210-myc-1(G$_4$S)-Sun1

VC210-myc-2(G$_4$S)-Sun1

VC210-myc-3(G$_4$S)-Sun1
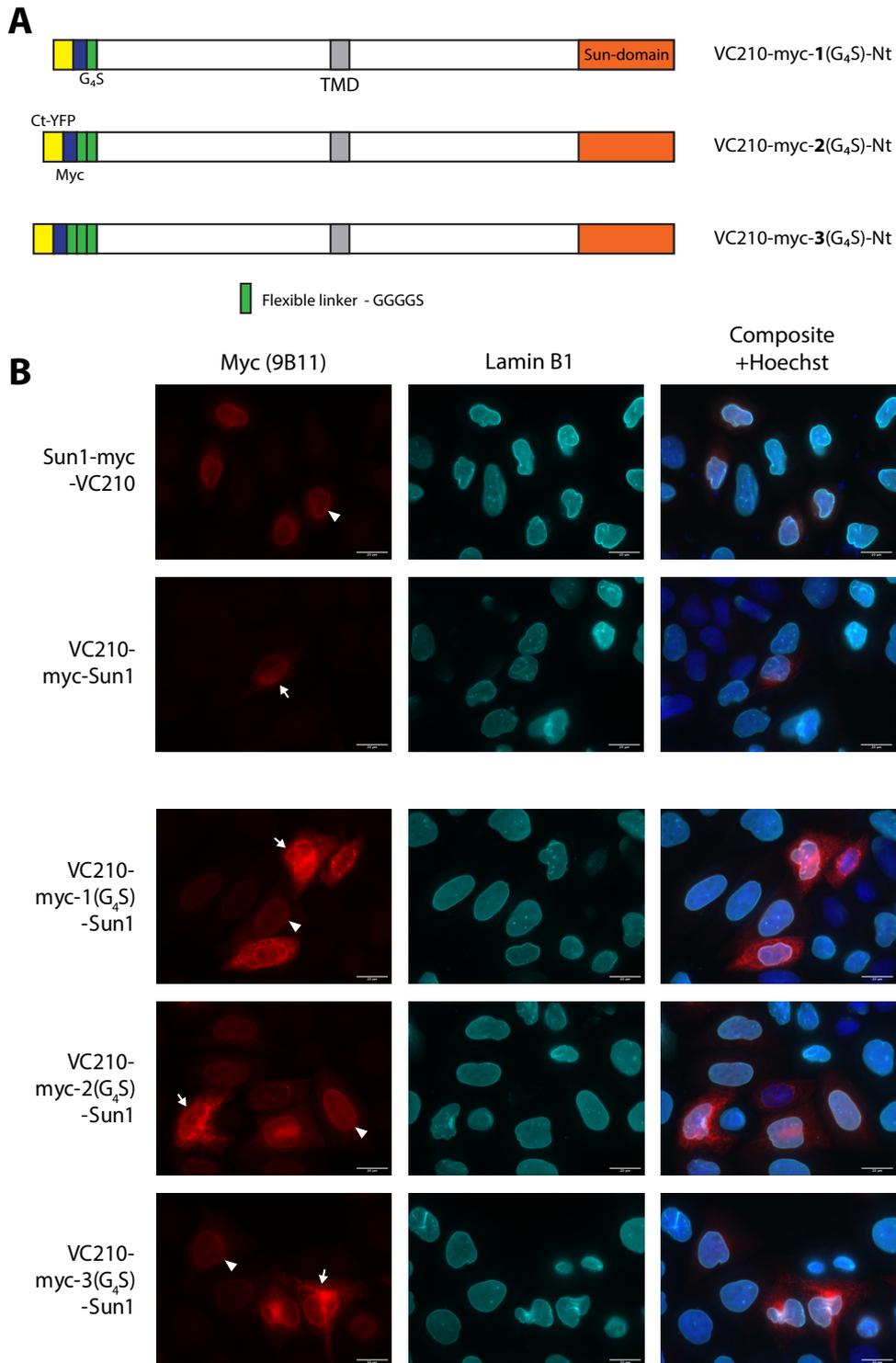
Figure 7-3 – Replacement of the cloning-optimised linker sequence between Sun1 and N-terminal VC210 Venus fragment with glycine-serine (G$_4$S) linkers improves expression and localisation in HeLa cells. (A) Schematic indicating constructs generated. (B) VC210-myc-x(G$_4$S)-Sun1 constructs were transfected into HeLa cells and stained with anti-myc antibody to reveal localisation. Arrowheads show construct localised at nuclear envelope; arrows show over expressed constructs.

254
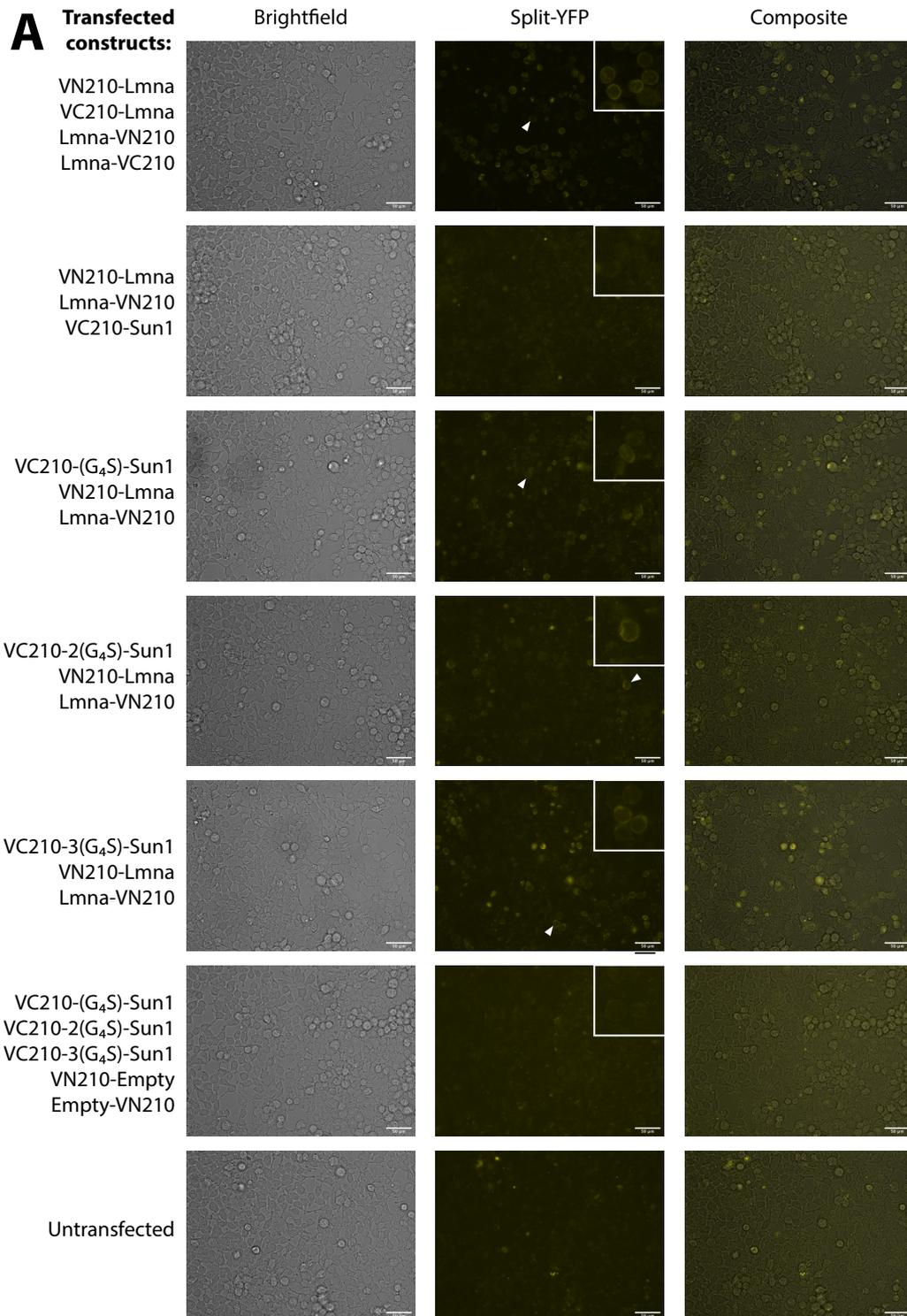
Figure 7-4 - The Sun1-Lamin A interaction can be detected using VC210-x(G$_4$S)-Sun1 in BiFC assays. HEK293T cells were co-transfected with VC210-x(G$_4$S)-Sun1 and Lamin A tagged with the complementary VN210 tag at its N- and C-terminal. The mVenus fluorescent protein becomes reconstituted and gives nuclear rim-like fluorescence (arrowheads), which is not evident when using the original VC210-myc-Sun1 construct (A). Quantification shown in (B) overleaf. Mean fluorescence intensity for at least 49 nuclei was measured in each condition expressed relative to background intensity. D'Agostino & Pearson testing revealed datasets were not normally distributed, so groups were tested for statistical difference using Kruskal-Wallis test with Dunn's multiple comparisons. **** = p<0.0001; ns = not significant.

B

Y-axis: YFP Intensity Fold Change (A.U.)

Transfected construct:						
VN210-Lmna	+	+	+	+	+	
VC210-Lmna	+					
Lmna-VN210	+	+	+	+	+	
Lmna-VC210	+					
VC210-Sun1		+				
VC210-(G$_4$S)-Sun1			+		+	
VC210-2(G$_4$S)-Sun1				+	+	
VC210-3(G$_4$S)-Sun1					+	+
VN210-Empty					+	
Empty-VN210					+	
Untransfected						+
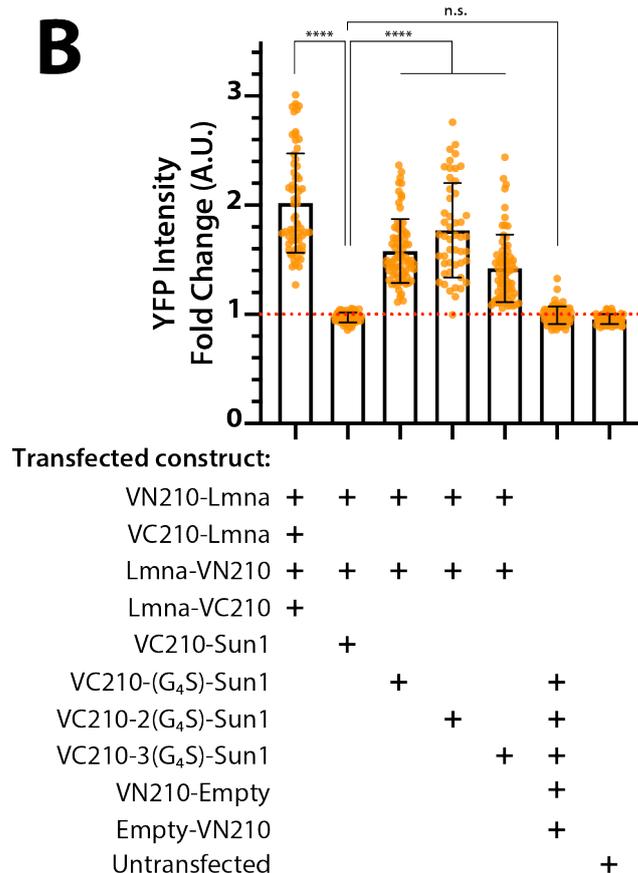
7.2.3. *BiFC using Sun1 and BioID candidate interactors confirms proximity to numerous queried proteins.*

Having developed a functional Sun1 BiFC system, candidate interactors of Sun1 gleaned from 2C-BioID were assessed for their interaction with Sun1 using this system. Initially, candidate interactors tagged separately at both their amino- and carboxy-terminus with the VN210 BiFC fragment were expressed in HeLa cells to verify that the modified versions of the proteins were localised comparably to FLAG-tagged versions. Most candidate Sun1 interacting proteins were generally consistently localised when tagged at either their N- or C-terminus with VN210 (Supplementary Figure VIII). Moreover, candidate proteins were overall consistent in their localisation compared to FLAG-tagged constructs (Table 7-2). One exception is Csrp3, which localised to filamentous structures, or diffusely throughout the nucleoplasm and cytoplasm when tagged at the C- and N-terminus, respectively (Supplementary Figure VIII, Table 7-2). To perform the BiFC assay, HEK293T cells were co-transfected with myc-VC210-3(G$_4$S)-Sun1 and both tagged versions of the

candidate interactors to take into account discrepancies in determined localisations between N- and C-terminal tagged constructs. A BiFC condition between lamin A tagged separately with both of the complementary BiFC fragments was included as a positive control. Upon analysis by fluorescence microscopy, oligomerisation between lamin A filaments (Figure 7-5S), and binding of Sun1 to lamin A (Figure 7-5Q) could be confirmed. In addition, the interactions of Sun1 with Csrp3, Cttn, Parva, Ppm1a, Ruvbl1 and Ruvbl2 were visualised (Figure 7-5D, E, I, J, K, L). To quantitate which candidate Sun1 interactors provide the strongest BiFC signal, narrowing which of the proteins to further investigate, BiFC was combined with flow cytometry (BiFC-FC). BiFC experiments performed using transient transfection from multiple vectors require a high sample size due to the lower number of cells containing each component thus taking advantage of the capacity of flow cytometry to rapidly analyse a high number of particles.

Gene	Construct	HeLa localisation	Consistency		
			Nt / Ct	HeLa FLAG	Database
Sun1	Ct-VC210	Nuclear membrane	(Replaced G$_4$S linker)		
	VC210-Nt	Endoplasmic reticulum			
Alox5	Ct-VN210	Cytoplasmic, little nucleoplasmic			
	VN210-Nt				
Vps51	Ct-VN210	Cytoplasmic, vesicular			
	VN210-Nt				
Ruvbl2	Ct-VN210	Cytoplasmic, nucleoplasmic, nuclear periphery			
	VN210-Nt				
Ppm1a	Ct-VN210	Cytoplasmic, nucleoplasmic			
	VN210-Nt				
Cttn	Ct-VN210	PM, filopodial			
	VN210-Nt				
Tnnt3	Ct-VN210	Nucleoplasmic, nuclear speckles			
	VN210-Nt	Nucleoplasmic, nucleolar			
Tnni1	Ct-VN210	No expression			
	VN210-Nt	Cytoplasmic, nucleoplasmic			
Lmna	Ct-VN210				
	VN210-Nt				
Ruvbl1	Ct-VN210	Cytoplasmic			
	VN210-Nt				
Parva	Ct-VN210	Nucleoplasmic, cytoplasmic			
	VN210-Nt				
Dctn5	Ct-VN210	Cytoplasmic, nucleoplasmic, diffuse			
	VN210-Nt				
Ccdc86	Ct-VN210	Nucleoplasmic			
	VN210-Nt	Nucleoplasmic, nucleolar			
Hist1h2bp	Ct-VN210	Cytoplasmic, nucleoplasmic			
	VN210-Nt				
Fam91a1	Ct-VN210	Cytoplasmic, nuclear periphery			
	VN210-Nt				
Arhgap22	Ct-VN210	Cytoplasmic			
	VN210-Nt	Cytoplasmic, punctate			
Zbtb20	Ct-VN210	No expression			
	VN210-Nt	Nucleoplasmic, cytoplasmic; punctate			
Csrp3	Ct-VN210	Filamentous			
	VN210-Nt	Nucleoplasmic, cytoplasmic			

Table 7-2 - Localisation of Sun1 candidate interactors in HeLa cells tagged at the N- and C-terminus with VN210 BiFC fragment. Coloured boxes indicate consistency of subcellular localisation with the indicated criteria; between the oppositely tagged construct of the same gene, to the subcellular localisation determined when the construct is FLAG-tagged, and to localisations reported in Uniprot and Human Protein Atlas databases (Table 1). Green = consistent, Yellow = some discrepancies, Orange = inconsistent.

258
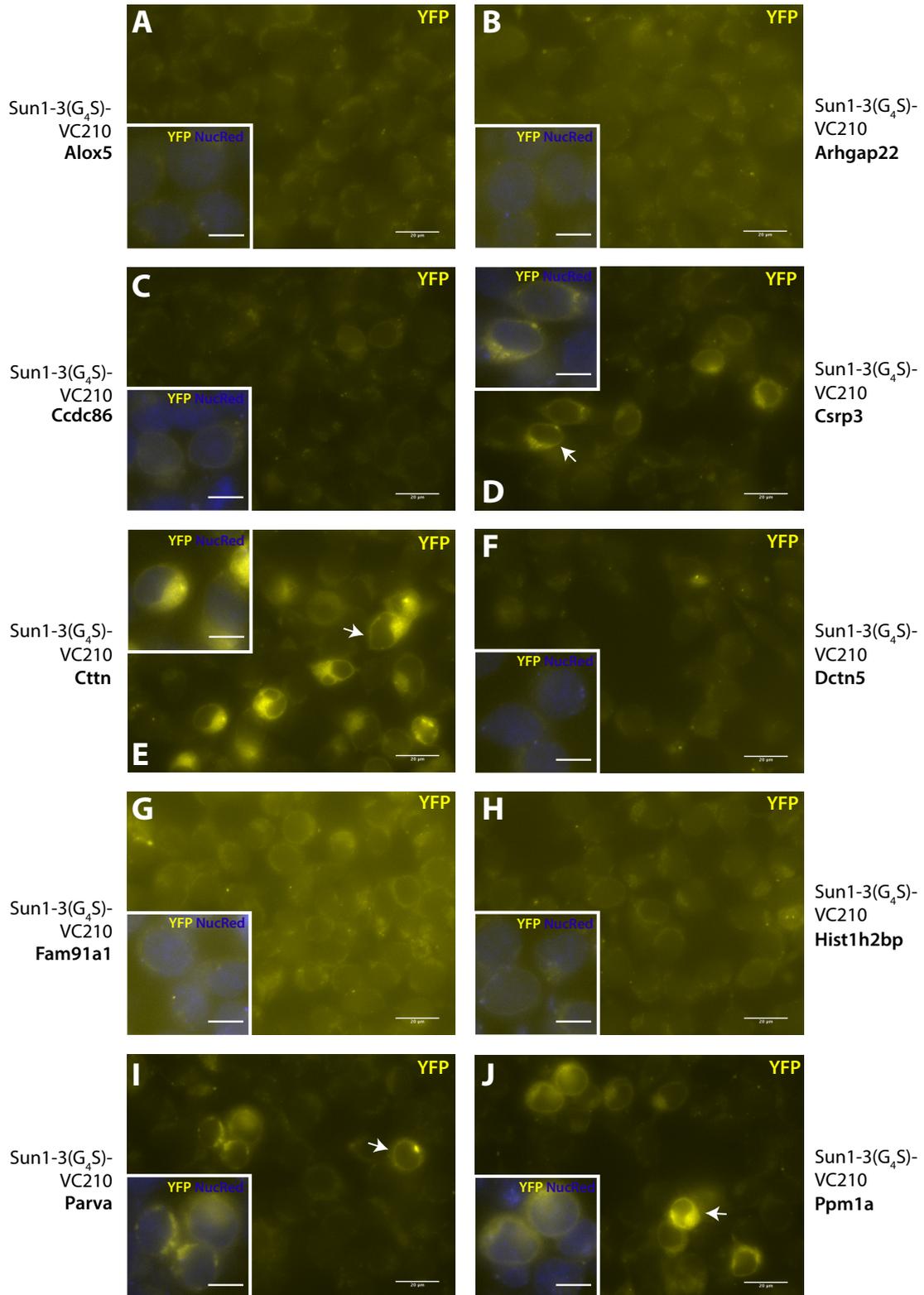
Figure 7-5 – Analysis of Sun1 BiFC assay by fluorescence microscopy. HEK 293T cells were co-transfected with myc-VC210-3(G$_4$S)-Sun1 and N- and C-terminally tagged candidate interactors (A-Q), VN210 only (R), or constructs encoding VN210 and (Figure 7-5 cont.) VC210 untagged (T), or tagged to lamin A (S). Cells were incubated for 48 hours before imaging by fluorescence microscopy. Positive signals are indicated with arrow, and inset scale bar = 10µm.

Sun1-3(G₄S)-VC210 **Ruvbl1**

Sun1-3(G₄S)-VC210 **Ruvbl2**

Sun1-3(G₄S)-VC210 **Tnni1**

Sun1-3(G₄S)-VC210 **Tnnt3**

Sun1-3(G₄S)-VC210 **Zbtb20**

Sun1-3(G₄S)-VC210 **Vps51**

Sun1-3(G₄S)-VC210 **Lmna**

Sun1-3(G₄S)-VC210 **Empty**

Lmna/Lmna

Empty/Empty
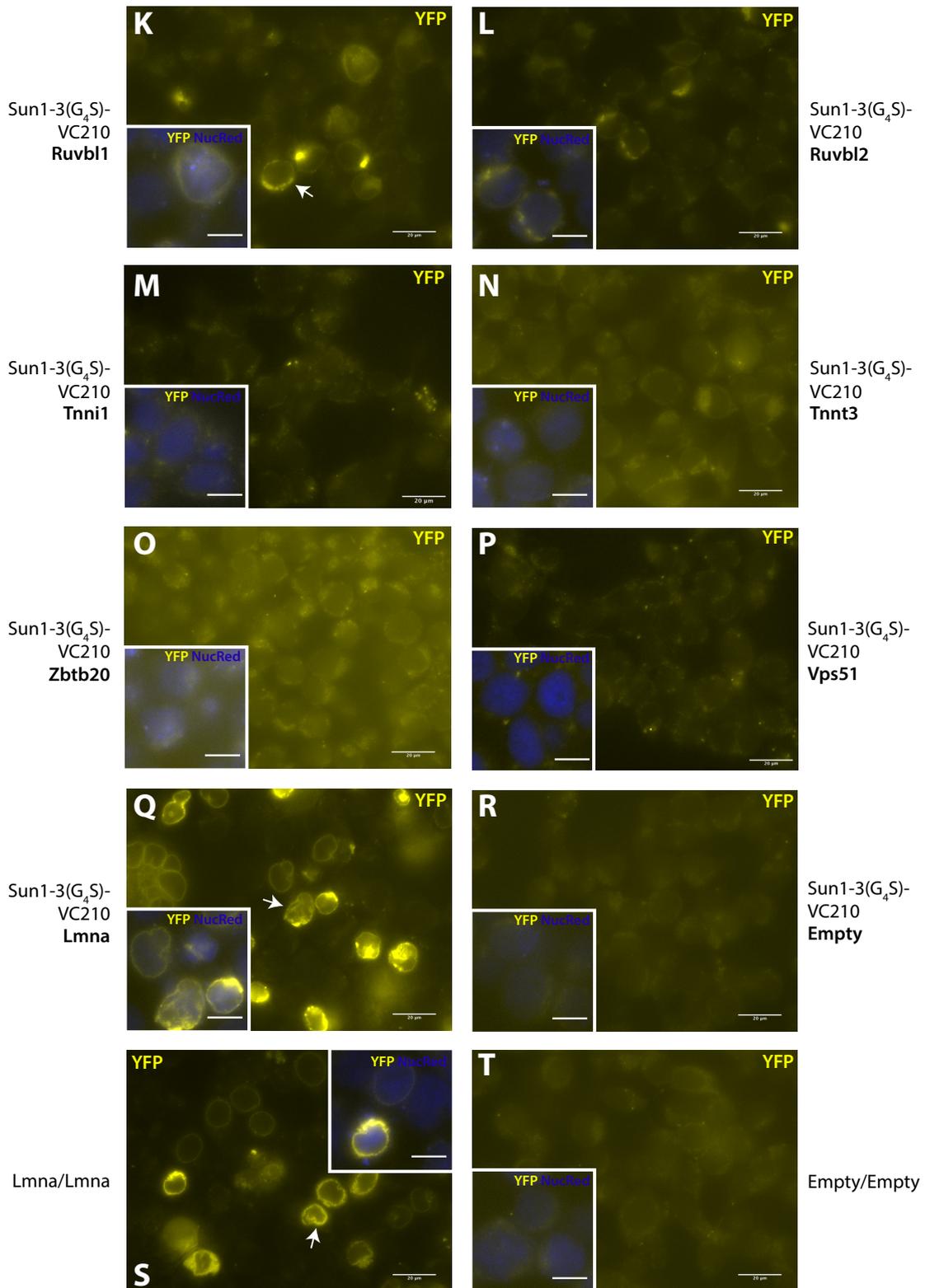
Similar to the previous BiFC experiment analysed by fluorescence microscopy, HEK293T cells were co-transfected with myc-VC210-3(G_4S)-Sun1 and both N- and C-terminal tagged versions of candidate interactors and flow cytometry was performed 48 hours post transfection. Consistent with imaging results, BiFC-FC revealed strong BiFC signal intensity upon myc-VC210-3(G_4S)-Sun1 co-expression with VN210-tagged versions of Lmna, Cttn and Ppm1a (Figure 7-6). Calculation of effect size revealed that weaker positive BiFC signals were present in conditions with Csrp3, Parva, Ruvbl1 and Ruvbl2 when compared to the negative control; myc-VC210-3(G_4S)-Sun1 co-expressed with unfused VN210 (Figure 7-6). Alox5 and Tnni1 were identified as negative interactors by microscopy, and were included to demonstrate that alternative fusion proteins do not generate a BiFC signal (Figure 7-6), despite robust expression (Supplementary Figure IX). In summary, BiFC-FC combined with BioID evidence indicate Csrp3, Cttn, Parva, Ppm1a, Ruvbl1 and Ruvbl2 are novel proteins of the Sun1 interactome in myoblasts. The strongest BiFC signals were induced upon co-expression of Cttn and Ppm1a with Sun1, indicating that further characterisation of these proteins was of interest.
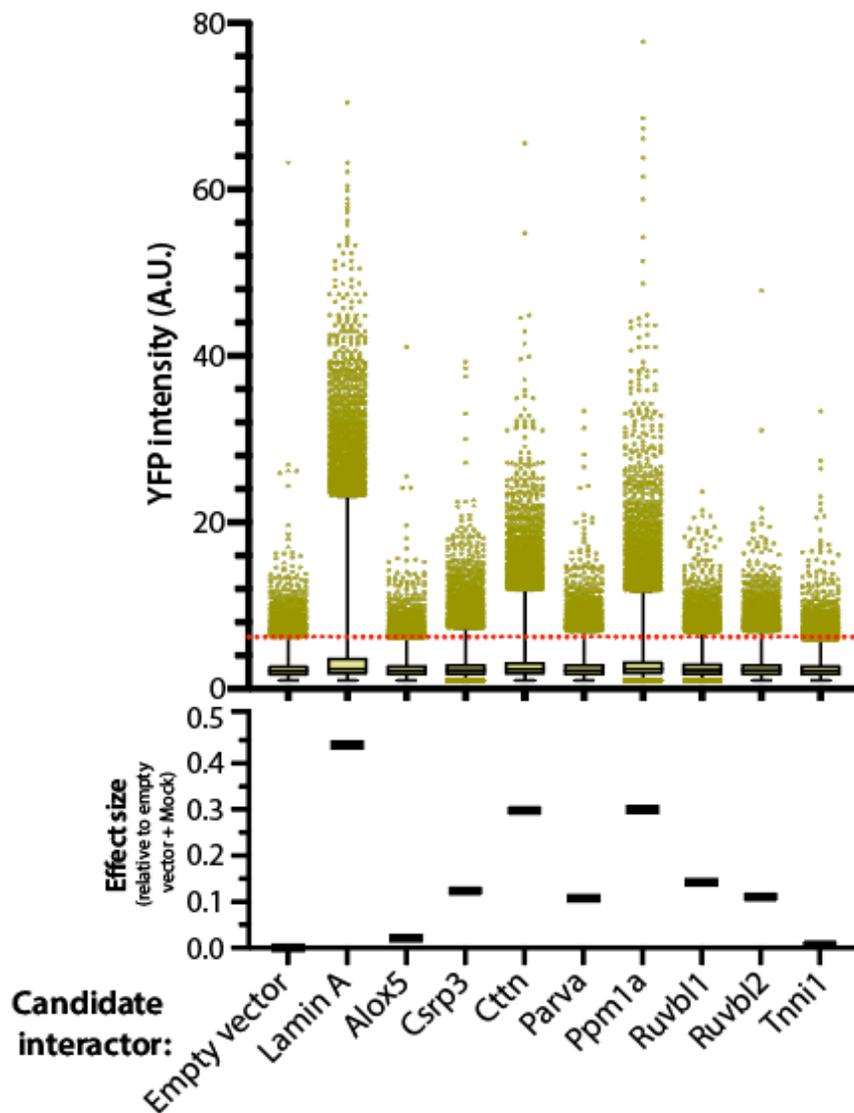
Figure 7-6 – BiFC coupled with flow cytometry. HEK293T cells co-transfected with myc-VC210-3(G_4S)-Sun1 and VN210-tagged candidate interactors were analysed by flow cytometry after 48 hours. Three independent experiments comprising a total of at least 118,149 cells for each condition were analysed, and the BiFC signal from the reconstituted YFP fluorescent protein is shown in the upper graph. Broken red line signifies the 99th percentile of the empty vector control. Lower panel shows a graphical representation of the effect size calculated in relation to the empty vector control to give a visual illustration of differences between conditions.

7.2.4. _FLAG-tagged candidate interactors are not recruited to overexpressed Sun1-HA in HeLa cells_

BioID combined with BiFC data refined the list of Sun1 interactors (Table 6-2). Though not universally reported in the literature (Kodama and Hu, 2012), BiFC is perhaps better described as another proximity based assay (Lönn and Landegren, 2017). Indeed, flexibility provided by linking peptides between the protein of interest and BiFC fragment are designed precisely to allow movement and optimal engagement of complementary BiFC fragments regardless of the orientation of interaction between proteins of interest. One implication in considering BiFC as a proximity-based interaction technique is that the assayed protein interaction may not occur directly and other, unidentified factors retain the proteins within proximity permitting fluorescent protein reconstitution. It was previously identified that many candidate interactors of Sun1 did not localise exclusively to the NE (Supplementary Figure VI, VII, Table 7-1). To address proximity-based caveats of BiFC, it was hypothesised that the overexpression of Sun1 would result in the accumulation of genuine interactors at the nuclear periphery, providing further support that BioID-identified proteins were direct interactors of Sun1. To test whether Ppm1a and Cttn, candidates which exhibited the strongest BiFC signal, could be detected at the NE, FLAG-tagged constructs were co-transfected in HeLa cells in the presence or absence of Sun1 tagged with the HA-epitope tag on the luminal C-terminus. Upon fixing and staining cells with antibodies raised against the HA- and FLAG-epitope tags after either 24 or 48 hours, there was no detectable accumulation of FLAG-tagged Cttn or Ppm1a to the nuclear peripheral position of Sun1 (Figure 7-7A,B). Notably, there was also no relocalisation of any other Sun1 candidate interactors which provided a positive BiFC signal (Supplementary Figure X, XI).
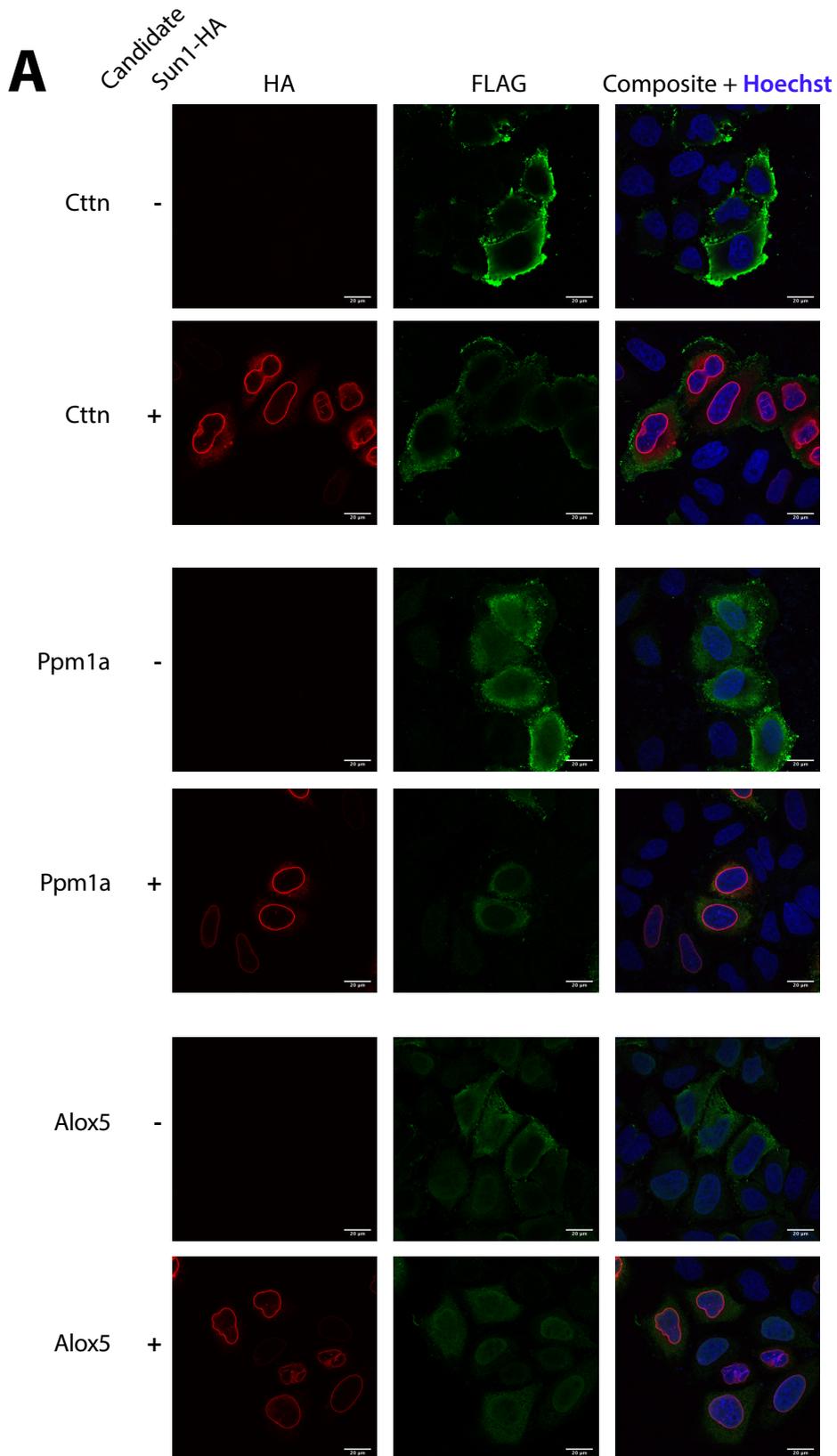
Figure 7-7 – Overexpression of Sun1-HA in HeLa cells does not recruit Sun1 candidate interactors Cttn or Ppm1a. HeLa cells were transfected with the indicated constructs and (Figure 7-7 cont.) fixed either 24 (A) or 48 (B) hours post-transfection. After fixing, cells were immunostained using antibodies against HA and FLAG tags, and lamin B1 (B) to delineate the nuclear envelope.

B

Candidate	Sun1-HA	HA	FLAG	Composite + **Hoechst**, **Lamin B1**
Cttn	−			
Cttn	+			
Ppm1a	−			
Ppm1a	+			
Alox5	−			
Alox5	+			

265
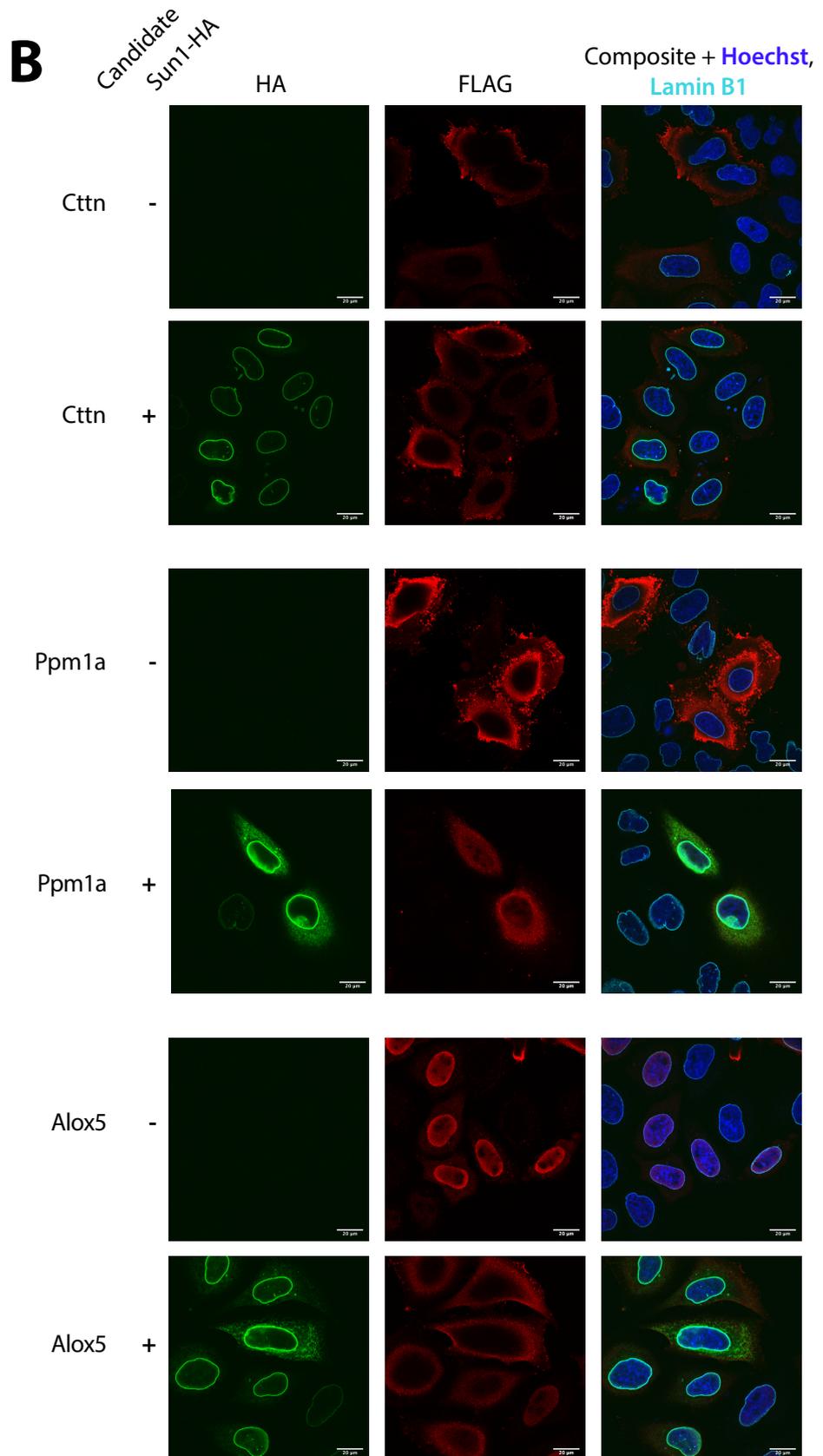
Experiments designed to assay protein relocalisation are dependent on the lack of interference from endogenous processes. As an example, it is plausible that overexpression of a downstream component within a signalling pathway would not lead to relocalisation of a candidate interactor since it may be modulated upstream, perhaps within a different cellular compartment. The notch signalling pathway neatly illustrates this idea. The notch receptor localises to the plasma membrane, where engagement by a notch ligand exposes a metalloprotease cleavage domain in the intracellular juxtamembrane region. Cleavage by ADAM metalloproteases liberates the notch intracellular domain (ICD) which subsequently translocates to the nucleus and interact with other proteins to direct diverse transcriptional outputs (Brou *et al.*, 2000; Mumm *et al.*, 2000; Gordon *et al.*, 2015; Siebel and Lendahl, 2017). Without inducing notch cleavage, relocalisation of the notch ICD to compartments or structures other than the plasma membrane would not be observed. Alternatively, specification of different cell types results from the alternative regulation of proteins and their interactions, which may be mediated directly, or indirectly. HeLa cells were initially chosen due to their high transfection efficiency, and ability to tolerate high exogenous protein expression. However, validation of Sun1 PPIs discovered in myoblasts by assaying recruitment to the nuclear periphery in HeLa cells would depend upon protein interactions, or signalling pathways, being consistently activated in both cell types. Since it is challenging to control all possible signalling inputs for poorly defined potential protein interactors, it was instead sought to explore whether myoblasts would result in co-localisation of Sun1 and candidate interactors at the nuclear periphery. To accomplish this, Sun1-HA and FLAG-tagged candidate interactors were expressed in C2C12 myoblasts which were subsequently stained using antibodies against the HA and FLAG tags. As with HeLa cells, immunofluorescence analysis of cells co-expressing Sun1-HA and FLAG-tagged versions of Ppm1a or Cttn did not result in co-localisation at the nuclear periphery, as delineated by lamin B1 co-staining (Figure 7-8). Moreover, C2C12 cells co-expressing Sun1 with other positive interactors of Sun1 gleaned from BiFC assays did not lead to any observable enrichment at the NE (Supplementary Figure XII). Co-expression of Sun1 interactors in C2C12 myoblasts only addresses the argument that protein interactions in question are cell-type dependent, and not that Sun1

266

interactors are sequestered from the NE by unidentified means. Even though results from C2C12 co-localisation were also negative, this experimental approach does not control for the possibility that other factors are responsible for permitting the accumulation of candidate interactor to the NE. This is not required for detection of PPIs by BiFC.
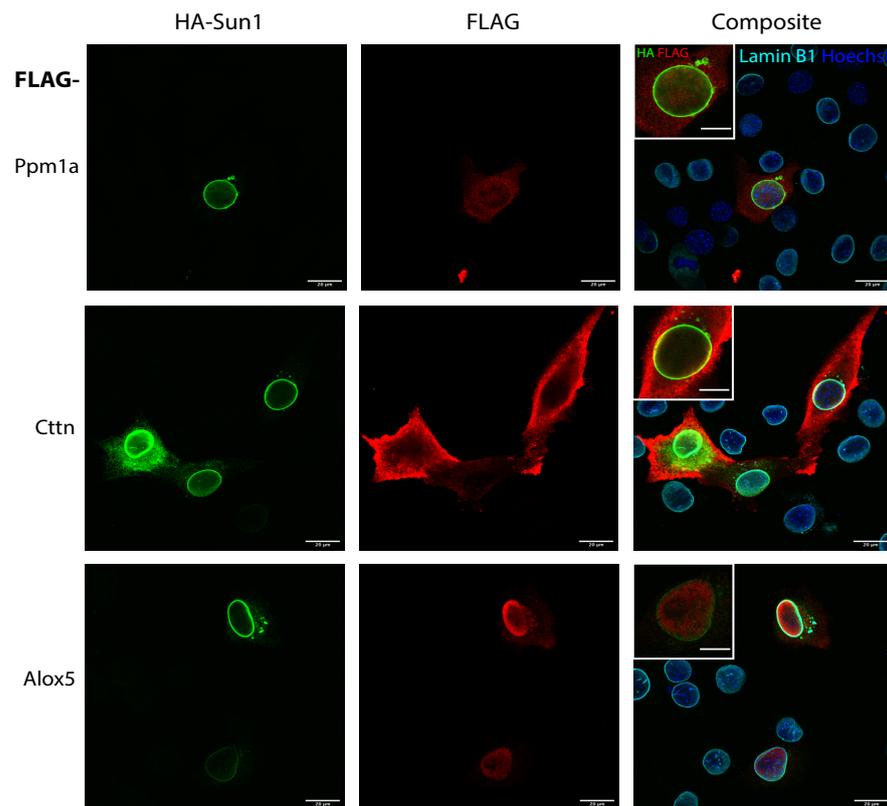
Figure 7-8 –Sun1-HA and FLAG-tagged Ppm1a or Cttn do not accumulate and co-localise at the nuclear periphery when co-expressed in C2C12 myoblasts. C2C12 myoblasts were co-transfected with constructs encoding Sun1-HA and FLAG-tagged versions of the indicated candidate interactors of Sun1. Cells were stained after 24 hours using antibodies against the HA and FLAG tags to show localisation of Sun1 and candidate interactor, respectively. Lamin B1 delineates the nuclear envelope. Samples were imaged by confocal microscopy to optically dissect high-contrast sections through the nucleus. Scale bars = 20μm; 10μm (inset panels).

7.2.5. *Ppm1a does not co-immunoprecipitate with the nucleoplasmic domain of Sun1.*

Biochemical assays to verify interactions between proteins within the NE have thus far been avoided due to the complexities in balancing protein extraction of membrane bound Sun1 with preservation of endogenous interactions. Nevertheless, *in vitro* strategies do exist which address these caveats, including *in vitro* translation (IVT) followed by co-immunoprecipitation. To circumvent issues with solubility of Sun1, only the nucleoplasmic region of Sun1, comprising amino acids 1-355 fused to HA (Sun1(1-355)-HA), was used to probe an interaction with myc-tagged Ppm1a, while lamin A was used as a positive control. As previously observed, and in concordance with BioID and BiFC data gathered in this study, myc-lamin A was co-immunoprecipitated with Sun1(1-355)-HA, when compared with the HA-antibody only control condition. However, myc-Ppm1a was not detected in the immunoprecipitated fraction (Figure 7-9). If the putative interaction between Sun1 and Ppm1a is weak or transient in nature, it is possible that the binding or washing conditions were too stringent to preserve this interaction. Alternatively, it is possible that other proteins of the NE are required to stabilise the Sun1-Ppm1a interaction which are not present in this IVT-co-immunoprecipitation assay. Time constraints meant that these possibilities could not be explored further.

7.2.6. _Preliminary findings indicate that Sun1 functions to at least partially regulate TGFβ signalling._

A literature search for Ppm1a reveals that it is a protein phosphatase involved in the modulation of TGFβ signalling in the nuclear compartment. The TGFβ pathway is central to numerous cellular functions including apoptosis, proliferation, and is involved in development of fibrosis of skeletal muscle (MacDonald and Cohn, 2012; Kharraz _et al._, 2014; Bernasconi _et al._, 2018). TGFβ itself is an extracellular secreted ligand that engages a heterodimeric TGF receptor which is situated on the plasma membrane (Schmierer and Hill, 2007). Integral tyrosine kinase activity activates secondary signalling molecules including the receptor-regulated Smads 2 and 3 transcription factors by phosphorylation. Subsequently, activated Smad2/3 complexes with Smad4 and translocates to the nucleus where it can direct
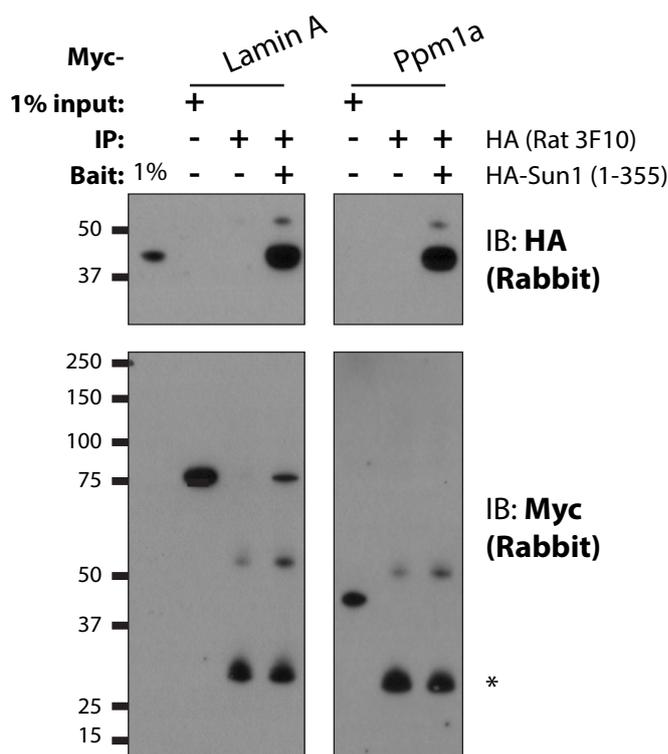
Figure 7-9 – In vitro translation co-immunoprecipitation assay shows that nucleoplasmic Sun1 binds lamin A, but not Ppm1a. IVT-produced Sun1-(1-355)-HA and Myc-tagged lamin A and Ppm1a were incubated together and immunoprecipitated using an antibody against the HA-epitope and protein G dynabeads. 30% of the total precipitate was then analysed by SDS-PAGE and western blotting. Western blots were probed with antibodies raised against the myc and HA epitopes to probe the immunoprecipitated fraction. Asterisk denotes a non-specific band thought to originate from protein G dynabeads.

transcriptional activity (Schmierer and Hill, 2007). The phosphatase activity of Ppm1a in involved in attenuating Smad2/3 mediated transcription by dephosphorylation, so permitting their exit from the nucleus. It was recently found that the INM protein Man1 could act as a scaffold to sequester Smad2/3 in proximity to Ppm1a for dephosphorylation, resulting in TGFβ activated transcription ceasing (Bourgeois *et al.*, 2013). The genetic ablation of Man1 in mice results in phenotypes associated with the over activation of TGFβ signalling, resulting in excess ECM protein expression, defective vasculogenesis with embryonic lethality (Cohen, Kosti and Stewart, 2007). Recent reports now implicate the LINC complex in the regulation TGFβ signalling. The genetic deletion of Sun2 may elevate Smad2 levels in the nucleus, although the downstream consequences, nor the effect of Sun1 was evaluated (Stewart, Rodriguez and King, 2019). The identification of Ppm1a as a novel interactor of Sun1 through the current study raises the possibility that Sun1- and/or Sun2-containing LINC complexes may function, in addition to Man1 in regulating TGFβ signalling. To investigate whether Sun1 has a function in TGFβ signalling, the nuclear translocation of Smad2 was tested in primary cultured mouse adult fibroblasts (MAFs) derived from WT and Sun1 KO mice. Measurement of the Smad2 intensity ratio between the nuclear and cytoplasmic compartments after TGFβ stimulation revealed that Sun1 KO MAFs have a lower basal nuclear Smad2 localisation than WT MAFs (Figure 7-10A,B). However, the extent of Smad2 nuclear translocation upon treatment with TGFβ was comparable in both WT and Sun1 KO MAFs. (Figure 7-10A,C).

Since the nuclear translocation of Smad2 appears unaffected by the ablation of Sun1, it was next sought to determine whether there were any differences in TGFβ-mediated transcriptional output. The expression of α-smooth muscle actin (α-SMA) is induced by the stimulation of cells with TGFβ (Desmouliere *et al.*, 1993), thus increasing the contractility of myofibroblasts (Hinz *et al.*, 2012) and contributing to fibrosis which is prevalent in many muscle wasting diseases, including laminopathies (Bernasconi *et al.*, 2018). To establish whether fibroblasts lacking Sun1 exhibit any differences in their expression of α-SMA, WT and Sun1 KO fibroblasts were cultured in the presence or absence of TGFβ for 2 days before lysis and analysis by SDS-PAGE.

Subsequent western blotting revealed that while there was no apparent induction of α-SMA in response to TGFβ, basal α-SMA levels were ~2 fold higher in cells lacking Sun1 (Figure 7-11A, B). Interestingly, re-introduction of V5FRB-Sun1 into Sun1 KO fibroblasts and stimulation with TGFβ appears to reduce the expressed levels of α-SMA, though further experiments are required to establish whether this is a statistically significant effect (Figure 7-11C, D) and is directly attributable to Sun1 or by Sun1 affecting Man1 activity or levels.
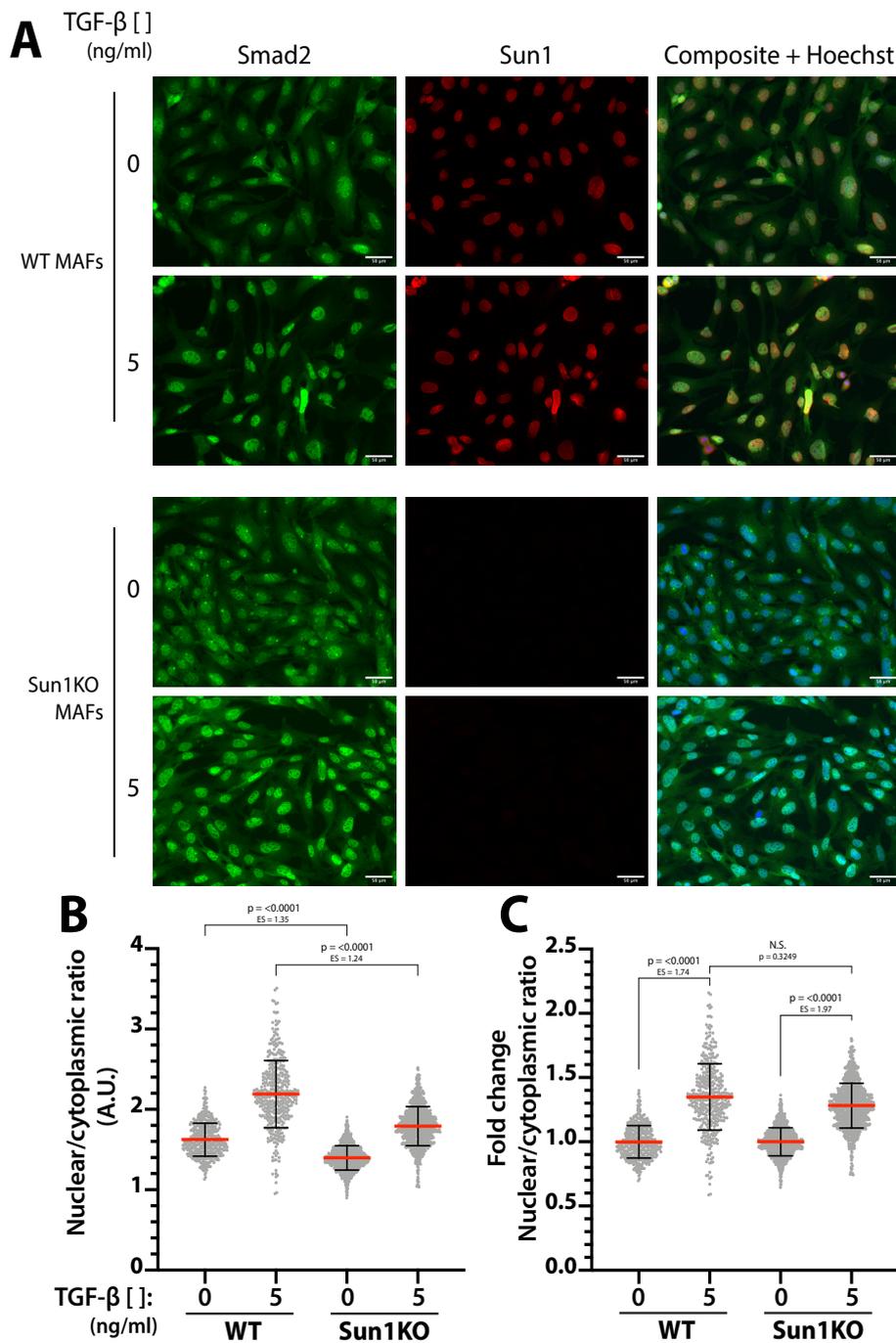
Figure 7-10 – The lack of Sun1 in MAFs does not impair Smad2 translocation to the nucleus upon TGFβ treatment. WT and Sun1KO MAFs were cultured in the presence or absence of 5ng/ml TGFβ for 24 hours prior to staining with antibodies raised against Smad2. Cells were then imaged using fluorescence microscopy (A). Images were analysed by a CellProfiler pipeline designed to calculate nucleus-cytoplasmic (N/C) ratio of Smad2. for each cell. (B) Raw N/C ratios, (C) N/C ratios normalised to 0ng/ml TGFβ. Images representative of results from one biological replicate with at least 414 cells analysed in each condition. (Figure 7-10 cont.) D'Agostino and Pearson's test for normality followed by Kruskal-Wallis test with multiple comparisons. Effect size calculation indicates scale of differences.
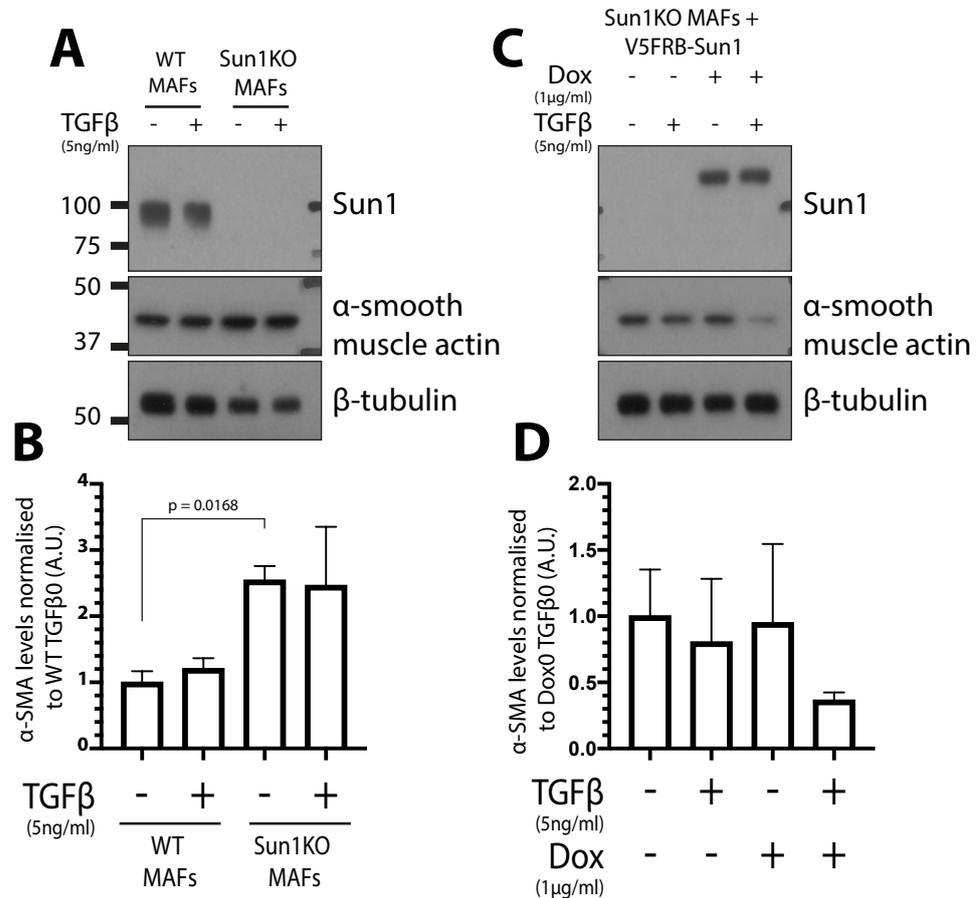
Figure 7-11 –Sun1KO fibroblasts appear to have increased levels of α-SMA. Fibroblasts derived from WT and Sun1KO mice were incubated with 5ng/ml TGFβ 48 hours prior to SDS-PAGE and western blotting. (A) Western blots were probed with antibodies raised against α-SMA and the levels normalised to the β-tubulin loading control. (B) Fibroblasts lacking Sun1 were transduced with an inducible construct encoding V5FRB-Sun1. Cells were either cultured in the presence or absence of 1μg/ml doxycycline for 3 days prior to TGFβ for a further 2 days. Cells were then lysed and analysed by SDS-PAGE and western blotting. Western blots were probed for α-SMA, which was normalised to a β-tubulin loading control. Blots shown are representative of three biological replicates, graphs indicate mean α-SMA levels with SD. Statistical significance was determined using a one-way ordinary ANOVA with Tukey's multiple comparisons.

7.3. Discussion

BioID is used for proximity labelling to identify protein interactions, although proteins identified in BioID experiments are not necessarily direct interactors. The biotinylation radius for BioID is thought to be around 10-20 nm (Kim *et al.*, 2014), meaning that neighbouring, but non-interacting proteins may also be detected. In this chapter, experiments were performed with the explicit aim to demonstrate the existence of Sun1 PPIs independently using BiFC, potentially differentiating direct and non-direct interactors gleaned from BioID.

7.3.1. *Bimolecular fluorescence complementation as a means to validate PPIs*

Three independent assays were used throughout experiments in this chapter to establish whether proteins identified through a Sun1 BioID experiment were genuine, novel interactors; BiFC, overexpression-recruitment assays and IVT co-immunoprecipitations. Notably, only BiFC provided results which were consistent with BioID data. Assaying recruitment to the NE as a result of Sun1-HA overexpression and IVT co-immunoprecipitation failed to detect interactions between Sun1 and candidate interactors. It is therefore worth considering the merits and limitations of each technique employed to evaluate the reliability of results.

BiFC was established in 2002, and has been extensively used to evaluate PPIs, particularly *in planta* (Kerppola, 2008). In light of results here, there still appears to be unaddressed questions with regards to the rate of identification of false positive interactors, as well as the fidelity of the system to confidently discern direct from indirect interactions.

As mentioned, it is likely that BiFC can detect proximal as well as direct interactions, and once in close proximity, BiFC fragments become irreversibly associated (Kerppola, 2008; Ohashi *et al.*, 2012). While this feature is useful for the detection of weak or transient interactions, it may also reveal non-specific and irrelevant interactions, particularly if the proteins of interest are highly expressed membrane proteins concentrated to a small area or organelle (Xing *et al.*, 2016). Stringent controls are therefore recommended to exclude the possibility of false

positives. Appropriate controls have been suggested to consist of a subtle mutation of one of the interacting partners in order to ablate the interaction (Hu, Chinenov and Kerppola, 2002; Kerppola, 2008; Kodama and Hu, 2012). Manipulation of interactors in this way, however, is not suitable when testing a suite of potential interactors which are not all well-characterised, as in this study. Significant prior knowledge of protein structure and interacting surfaces would be required to make such rational alterations. Alternatively, BiFC specificity can be tested using a competitive binding assay, whereby a non-BiFC-tagged version of one of the putative interactors is provided in addition to BiFC-fused forms with the hypothesis that BiFC efficiency is decreased due to its presence (Kodama and Hu, 2012; Fujii and Kodama, 2015; Fujii, Yoshimura and Kodama, 2018). Indeed, a competition assay was performed during this study to more firmly address that the BiFC signals between Sun1 and Ppm1a or Cttn were interaction-specific. Supplementing BiFC reactions with non-BiFC-tagged Sun1-HA decreased BiFC efficiency between Sun1-Lamin A, though this effect was not observed for the interactions of Sun1 with Cttn or Ppm1a (Supplementary Figure XIII). While this negative result might indicate the possibility that Cttn and Ppm1a are false positive interactors of Sun1, recent reports highlight that BiFC competition assays require temporal consideration (Xing *et al.*, 2016). Over time, the competitor protein, which binds to the BiFC-fusion protein based on endogenous kinetics, becomes permanently excluded from binding due to the stable reconstitution of the BiFC fragments (Xing *et al.*, 2016). The time taken to reach an out-competed state may also differ depending upon specific pairs of interacting proteins for reasons which are likely entwined with biology specific to the proteins of interest (Fujii, Yoshimura and Kodama, 2018). The BiFC competition assay performed in this study measured BiFC fluorescence at 48 hours, which may still be sufficient to visualise differences in stable protein interactions such as those between Sun1 and the NL. Other interactions may have a higher turnover rate, thus reaching an out-competed state over a shorter period. Therefore, to exclude the possibility that the competition effects are missed, it would be preferential to analyse BiFC efficiency over time.

Another consideration for BiFC assays is the confidence in differentiating direct from indirect protein interactions. It is thought that the orientation of the BiFC

fragments fused to interacting proteins is important for reconstitution of the fluorescent protein (Horstman *et al.*, 2014). Indeed, a flexible glycine-serine linking peptide repeated three times between Sun1 and the VC210 fragment was chosen specifically to accommodate a range positions, decreasing the potential for false negatives. On the other hand, BiFC fragments fused via long, flexible linkers may increase the likelihood of BiFC reconstitution in the absence of interaction between the assayed proteins. Although assayed for FRET not BiFC, one study exemplified the importance of rational linker peptide design. An artificial protein comprising the complementary cyan and yellow fluorescent proteins used in FRET linked by glycine-serine peptides of varying lengths and composition ratios found that FRET efficiency was increased for glycine-rich linkers around 20 amino acids in length, while linkers incorporating more serine produced less efficient FRET (Van Rosmalen, Krom and Merkx, 2017). How this translates to the study of physiological protein interactions, however, is unclear. A definitive study balancing the requirement for flexibility, to reliably capture genuine PPIs, and length, to minimise the detection of vicinal proteins is lacking. Such research would be valuable for application to BioID datasets where interacting and neighbourhood proteins are difficult to distinguish.

Using the current experimental system, direct comparison of different BiFC conditions is complicated by the unknown and variable expression level of fusion proteins. Although expression of proteins was confirmed by immunofluorescence staining (Supplementary Figure IX), expression level of Sun1 relative to its potential interacting partners was not controlled, with the possible implication that some interactors are not strongly expressed and therefore display reduced BiFC signal. To circumvent this, ratiometric expression could be achieved by expression of the interacting partners from the same construct (Grefen and Blatt, 2012; Xing *et al.*, 2016), though gene sizes may hinder this approach.

BiFC enabled the refining of BioID identified Sun1 interactors, though with caveats of BiFC in mind, a complementary assay was preferable to reinforce conclusions. However, neither results from immunofluorescence analysis of tagged constructs nor co-immunoprecipitation of *in vitro* translated proteins was consistent with BiFC. It is notable that these techniques do not exclude the possibility that other proteins are involved in mediating the interactions. Proteins that can be induced to

accumulate at the NE in response to the overexpression of a NE-localised binding partner, such as Sun1, must be independent of other regulation precluding a binding and relocalisation event. Such proteins, would only be predicted to co-immunoprecipitate *in vitro* providing no other proteins are involved in the stabilisation of that complex. Still, unfaithful recreation of physiological binding conditions may destabilise an interaction, thus leading to lack of detection.

Finally, the existence of tags, rather than endogenous proteins may interfere with PPIs. When used for relocalisation or co-immunoprecipitation assays, Sun1 was tagged exclusively at the C-terminus, to not disrupt any potential binding assayed at the nucleoplasmic domain. However, candidate interacting proteins were typically characterised using tags at both N- and C-termini to control for the possibility of tag interference. Candidate interactors for co-immunoprecipitation assays were tagged at the amino-terminus only. From the results, it is not possible to exclude the possibility that the epitope modification affects interaction with Sun1. Therefore, in addition to testing Sun1 co-immunoprecipitations using different binding conditions, future research would benefit from addressing whether epitope tags do indeed affect interaction detection.

Despite inconsistencies in the results of interaction assays, and the drawbacks of BiFC in unequivocal identification of genuine interactors of Sun1, it could be reasonably argued that within the context of refining BioID hits, BiFC and BiFC-FC are satisfactory techniques where the aim was to increase confidence in identifying potential interactors. Indeed, the robust fluorescence between Sun1 and Ppm1a prompted further hypotheses well supported by previous literature.

7.3.2. *Implications for the interaction of Sun1 with other BioID-identified and BiFC-validated neighbours*

7.3.2.1. *Sun1 and Cttn*

Cortactin is recruited to regions of actin polymerisation at the plasma membrane, where it directs cell migration and other cell behaviours related to actin dynamics (Lua and Low, 2005). Interestingly, it appears that cortactin translocates between the nucleus and cytoplasm. Acetylated cortactin localises to the nucleus, precluding its binding to Keap1, which, in turn, retains cortactin at the cell cortex to promote cell migration by actin rearrangement (Ito *et al.*, 2015). Quite what function cortactin has within the nucleus, complexed with components of the nuclear periphery such as Sun1, is not known. It may be that nuclear localisation simply represents a means to sequester cortactin away from filamentous actin, inhibiting its function in cell migration. Alternatively, it may have myriad, yet unknown, nuclear functions in transcription or genome organisation.

7.3.2.2. *Sun1 and Ruvbl1/Ruvbl2*

Ruvbl1 and Ruvbl2 are members of the ATPase associated with diverse cellular activities+ (AAA+) protein family. Ruvbl1/2 complexes as a dodecamer containing equal amounts of each component (Puri *et al.*, 2007), and X-ray crystallography revealed that the complex is formed from two stacked rings of heterohexamers (Gorynia *et al.*, 2011). As AAA+ family proteins, Ruvbl1 and Ruvbl2 have been identified in numerous processes, including within large nuclear protein complexes involved in transcription, DNA damage and chromatin remodelling as part of the SWR1 complex in yeast and mammalian cells (Shen *et al.*, 2000; Jónsson *et al.*, 2001; Jin *et al.*, 2005). Within the INO80 family of chromatin remodelling complexes, the SWR1 complex promotes the exchange of histone H2A.Z for H2A within nucleosomes, resulting in a change of the activation state of chromatin (Mizuguchi *et al.*, 2004), or the rate of diffusion of the DNA region within the nucleus (Neumann *et al.*, 2012). Interestingly, these functions are consistent with some of those previously assigned to Sun1 including in DDR and genome organisation (Le *et al.*, 2016). The role for Sun1 in DNA damage repair is unclear since it has been implicated in both HDR and NHEJ recombination events (Lei *et al.*, 2012; Lawrence *et al.*, 2016;

278

Marnef *et al.*, 2019), depending upon the DNA damage model used. One consistent feature of the DDR is the relocalisation of DSBs to the nuclear periphery, either to LINC complexes or NPCs, when undergoing NHEJ (Oza and Peterson, 2010; Lottersberger *et al.*, 2015). Studies in yeast have found that the Sun-domain homologue, Mps3, and the Ruvbl1/2-containing SWR1 complex are both required for double stranded break relocalisation in a cell-cycle dependent manner (Horigome *et al.*, 2014). However, it is not yet known whether a comparable system exists in mammalian cells, although the identification of Ruvbl1 and 2 in the vicinity of Sun1 using BioID indicates that this is indeed a possible avenue for further research.

7.3.2.3. *Other Sun1 interactors: Csrp3 and Parva*

Two other Sun1 interactors validated using BiFC were Csrp3 and Parva. When Csrp3 is localised to the nucleus, it acts as a pro-myogenic transcriptional co-factor; driving expression of genes including MyoD1 (Arber, Halder and Caroni, 1994; Kong *et al.*, 1997). As such, Csrp3 has been previously implicated in hypertrophic cardiomyopathy in humans (Geier *et al.*, 2003, 2008; Bos *et al.*, 2006; Salazar-Mendiguchía *et al.*, 2020), and its ablation in mice leads to hypertrophic dilated cardiomyopathy (Sun *et al.*, 2020). While DCM is a common phenotype of the laminopathy diseases and mouse models, concomitant hypertrophy is notably absent (Cupesi *et al.*, 2010). Given the observation that Sun1 deletion in the presence of lamin A mutations subdues related lamin phenotypes, the significance of an interaction between Sun1 and Csrp3 is not clear without further study. It might first be useful to establish whether there are any changes to levels, or subcellular localisation of Csrp3 in cells harbouring lamin A or Sun1 mutations. In the eventuality that these are altered, experimenting with double Lamin A, Sun1 mutant cells might reveal whether Csrp3 has a function in the aetiology of laminopathies.

The *Parva* gene encodes the α-parvin protein. It was first identified as a component of focal adhesions (FA) and binds to the LD-domain of FA protein paxillin. Furthermore, α-parvin contains CH-domains which confer actin binding (Nikolopoulos and Turner, 2000). Similar to many other proteins identified proximal to Sun1, α-parvin is not conventionally associated with the nucleus, and its functions

in this compartment are vague. Despite this, α-parvin does contain a potential NLS at its N-terminus (Olski, Noegel and Korenbaum, 2001), though it has not been isolated and tested for functionality. That said, α-parvin also localises strongly to the nucleus in keratinocytes (Olski, Noegel and Korenbaum, 2001), which is consistent with expression studies carried out in HeLa cells in this study (Supplemenary Figure VII, VIII). Observations of nuclear-localised α-parvin might indicate that the NLS is functional, while differential nuclear accumulation in fibroblast, epidermal cells, keratinocytes, HeLa cells, and C2C12 myoblasts signifies cell-type specific localisations and roles (Olski, Noegel and Korenbaum, 2001).

7.3.3. *Potential roles for Sun1 in the regulation of TGFβ signalling*

TGFβ signalling is a master regulator of fibrosis in many tissues, including in diseased striated muscles (Chen *et al.*, 2005). Fibrosis of skeletal muscle tissues is problematic as it displaces contractile myofibres, increases tissue stiffness thus preventing regeneration (MacDonald and Cohn, 2012). Fibrosis is driven by the activation of genes involved in ECM deposition and organisation, such as collagen I (Smith *et al.*, 2011; Meyer and Lieber, 2012; Gillies *et al.*, 2017), and genes which inhibit ECM breakdown, such as tissue inhibitor of matrix metalloproteases (TIMPs) (Kim, Sheppard and Chapman, 2018). Through its ability to bind and sequester Smad2/3 and Ppm1a, Man1 negatively regulates TGFβ signalling (Bourgeois *et al.*, 2013). Indeed, genetic ablation of Man1 causes embryonic lethality in mice due increased deposition of ECM due to activated Smad2/3, leading to failed vascularisation of the yolk sac (Ishimura *et al.*, 2006; Cohen, Kosti and Stewart, 2007). There are some recent reports that Sun2-containing LINC complexes may contribute to TGFβ regulation, given its absence results in upregulated levels of Man1, together with increased nuclear localisation of Smad2. However, despite cardiac hypertrophy driven through the AKT/MAPK pathway when Sun2 is ablated, fibrosis, and genes associated with fibrosis are unchanged (Stewart, Rodriguez and King, 2019). In contrast, preliminary results from this study indicate that nuclear translocation of a TGFβ regulated transcription factor, Smad2, appear unchanged in fibroblasts, even though basal levels of α-SMA appear increased compared with WT. Interestingly, and in opposition to Sun2, whose ablation has no impact on fibrosis,

increased α-SMA levels would implicate Sun1 within the fibrotic pathway. α-SMA contributes to increased contractility of fibroblasts (Hinz *et al.*, 2001), and increased α-SMA levels have been found in cardiac muscle of *Lmna* deficient mice (Nikolova-Krstevski *et al.*, 2011). These observations are summarised in the diagram shown in Figure 7-12. Previous results revealed that Sun1 ablation in *Lmna* knockout mice improves pathology of the cardiac tissue (Chen *et al.*, 2012). How observations that individual ablation of Lamin A or Sun1 increase, yet double knockout reduces markers of fibrosis, is not clear. However, it appears that Sun1 knockout alone does not lead to increased fibrosis of muscular tissues (Lei *et al.*, 2009; Chen *et al.*, 2012), perhaps indicating that turnover of α-SMA or other driver of fibrosis is affected by the absence of Lamin A. Although Smad2 and Smad3 are commonly assessed together when probing the TGFβ pathway, α-SMA is in fact a specific transcriptional target of Smad3 (Hu, Wu and Phan, 2003). Increased α-SMA levels with no change in Smad2 translocation might therefore indicate Smad3 translocation is disrupted in Sun1 deficient cells. These results were obtained using primary, but SV40 immortalised fibroblasts, and so these observations should first be reproduced using earlier passage, primary cells. Providing that elevated α-SMA is consistently observed, understanding of the mechanism behind this would benefit from experiments focussed on Man1 levels on a Sun1 KO background, as well as establishing whether Sun1 and Sun2 interact with Man1, perhaps to influence its function in Smad regulation.
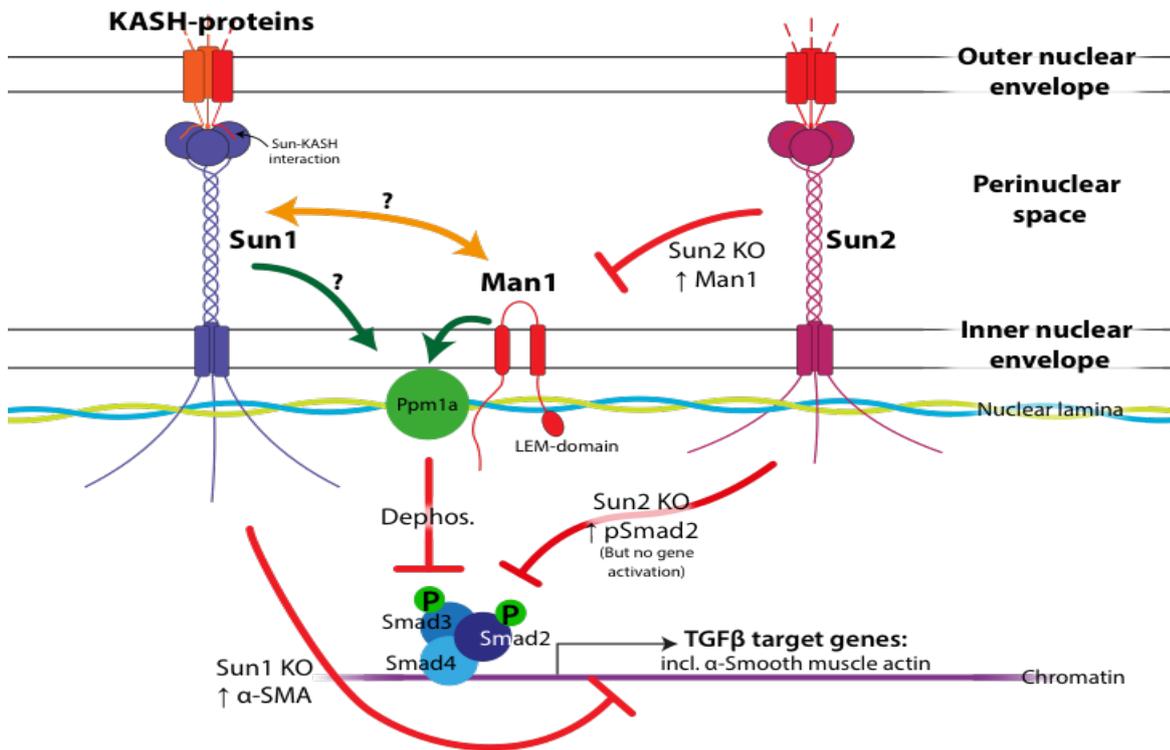
Figure 7-12 – Working model for the potential influence of LINC complexes on Smad2/3 signalling and target gene expression. Man1 is known to sequester Smads2/3 to the NE where their action as transcription factors is inhibited by dephosphorylation by Ppm1a. Sun2 KO displays an inconsistent phenotype whereby Man1 levels are increased, pSmad2 accumulate in the nucleus, yet target genes are not activated, indicating other components are required for Smad2-dependent gene-activation. Conversely, ablation of Sun1 appears to increase basal levels of SMA, indicating that it normally exists to suppress Smad-dependent gene activity. However, the direct relationship between Sun1 and Ppm1a or Man1 is not clear.

282

7.3.4. *Summary and key findings*

Here, a BiFC strategy was devised for the further characterisation of BioID-identified Sun1 interactors. To do this, the nucleoplasmic domain of Sun1 was linked to the short VC210 mVenus fragment via a flexible glycine- and serine-rich peptide sequence. Along with lamin A, a known Sun1 interactor, the BiFC approach revealed Csrp3, Cttn, Parva, Ppm1a, Ruvbl1 and Ruvbl2 to be at least proximal to Sun1 and quantitative analysis using BiFC-FC showed that Cttn and Ppm1a provided the strongest BiFC signal. However, studies aimed at confirming direct interactions including recruitment and IVT co-immunoprecipitation were inconclusive. Nevertheless, there is precedent for Sun1 function in TGFβ regulation via its association with Ppm1a. Although the nuclear localisation of Smad2 does not appear to be affected in Sun1 deficient MAFs, α-SMA levels are significantly increased. These observations warrant further investigation given the contribution of fibrosis to the aetiology of laminopathy-caused muscle wasting conditions.

Chapter 8: Discussion

8.1. Summary

8.1.1. *Summary: Dystroglycan and the nucleus*

Overall, evidence gathered and presented in this thesis suggests that the function of dystroglycan within the nucleus of human myoblasts remains equivocal. An initial aim for research in this project was to confirm a putative PPI between Sun1 and β-dystroglycan previously identified in a mass spectrometry screen. Biochemical and cell biological strategies used to address this, however, were not able to detect an interaction. However, it is not possible to fully exclude the existence of an interaction since it may be that levels of interacting proteins present were below the detection sensitivity of the techniques employed. In the case of biochemical assays, binding conditions may have been too stringent to preserve the interactions assayed. Therefore, research scope was expanded to examine the effects of the loss of dystroglycan on nuclei. Human myoblasts devoid of dystroglycan were generated using CRISPR/Cas9, though the analysis of nuclei in these cells also showed negligible effects on the nuclear morphology; a feature previously reported to be disrupted. Some evidence was found however, to support the function of dystroglycan in the tethering of the centrosome to the nucleus, though the findings observed in this study requires verification using further cell lines.

In many reports of nuclear β-dystroglycan, the contribution of its post-translational processing seems under-appreciated. Published results indicate myriad modifications of dystroglycan including phosphorylation, ubiquitination, glycosylation and proteolytic cleavage (Sotgia *et al.*, 2001; Mitchell *et al.*, 2013; Endo, 2015; Leocadio, Mitchell and Winder, 2016; Lipscomb *et al.*, 2016). Evidence suggests that proteolytic fragments of β-dystroglycan can be trafficked to the nucleus in a number of cell types (Mathew *et al.*, 2013; Azuara-Medina *et al.*, 2019). To better understand how dystroglycan is regulated and positioned at the subcellular level, epitope tags were inserted throughout the coding sequence of dystroglycan to permit detection of fragmented β-dystroglycan *in situ*. This approach showed that

β-dystroglycan fragments can be separately trafficked throughout the cell. Again, in conflict with previous observations, though in concordance with results throughout this thesis, β-dystroglycan fragments were not observed within the nucleus. This is interesting given the extensive reports of nuclear β-dystroglycan in multiple cell types. Therefore, observations made during this project raise the possibility that β-dystroglycan is differentially regulated between cell types; an area which, to date, has not been directly addressed. Indeed, the glycosylation signature of α-dystroglycan varies between cell types (Ervasti, Burwell and Geissler, 1997), revealing the possibility that aspects of β-dystroglycan regulation are similarly controlled according to the cell type. While not addressed during the current study, the involvement of other post-translational modifications in conjunction with fragmentation, may function to further diversify roles for β-dystroglycan.

8.1.2. *Summary: the myoblast Sun1 interactome*

Sun1 is a component of the LINC complex which is responsible for mechanosensitisation of the nucleus to extra-nuclear and extracellular stimuli. Moreover, Sun1 mutations have been shown to be both causative and modifiers of other laminopathy conditions. Counterintuitively, the deficiency of Sun1 appears to suppress the disease phenotypes of mice harbouring laminopathy-causing mutations of lamin A. Thus, a second aim of research presented within this thesis was to uncover nucleoplasmic interactors of Sun1 in muscle to shed light on Sun1 function in health and disease. To do this, an *in vivo* model for proximity biotinylation of nuclear proteins was validated *ex vivo* using fibroblasts and myoblasts from a transgenic animal modified for 2C-BioID analyses. Mass spectrometry revealed a number of known and potential interactors, while GO term analysis revealed that many of the unknown genes were associated with the nuclear compartment, consistent with the localisation of Sun1. Proximity biotinylation is a powerful tool for unbiased discovery of neighbouring proteins, however, robust independent validation of interactors is crucial for further research. Consequently, BiFC was established for use with nucleoplasmic Sun1 interactors, which confirmed the presence of an interaction between Sun1 and Lamin A. BiFC confirmed that 6 other proteins identified through 2C-BioID were at least vicinal to Sun1. Among those was

Ppm1a, the identification of which implicates Sun1 in the regulation of TGFβ signalling. What the function of a Sun1-Ppm1a-TGFβ axis is in the context of muscle remains to be elucidated, however the prevalence of TGFβ malfunction in many muscle wasting conditions makes this an interesting area for further investigation.

8.2. Nuclear functions for β-dystroglycan

8.2.1. *The putative interaction between Sun1 and β-dystroglycan in myoblasts*

Results presented in chapter 3 show that an interaction between Sun1 and β-dystroglycan could not be confirmed *in vitro* in a human myoblast cell line using multiple strategies. There are a number of possible biochemical reasons behind this (discussed in 3.3.1), however the generation of the Sun1 interactome in an analogous system, primary murine myoblasts, provided a unique opportunity to further probe the putative Sun1-β-dystroglycan interaction. Proximity-dependent biotinylation is considered to have greater sensitivity than conventional protein complex purification for the detection of protein interactions within membranes (Roux, Kim and Burke, 2013). In addition, BioID assays isolate the detection of PPIs from the caveats involved with dependence upon the detection of antibody epitopes within crowded microenvironments. Nonetheless, consistent with results in chapter 3, proximity-dependent biotinylation by 2C-BioID using a Sun1 bait did not label β-dystroglycan.

8.2.2. *Is β-dystroglycan in the nucleus of myoblasts?*

The absence of a detectable interaction between Sun1 and β-dystroglycan does not of course preclude the concept that β-dystroglycan resides within nuclear structures. Despite extensive papers reporting the position of dystroglycan in the nuclei of multiple cell types, data presented in this thesis, together with previous observations appears to challenge this notion. Using strategies including confocal immunofluorescence microscopy and cell fractionation, in conjunction with antibody detection of the endogenous and tagged versions of the protein, β-dystroglycan could not be unequivocally localised to nuclei or any nuclear structure

in human myoblasts or HeLa cells. Interestingly, and consistent with observations made during the current study, previous proteomic studies defining proteins of the NE also did not detect dystroglycan in a variety of tissue and cell types, including lymphocytes, liver and muscle tissue (Schirmer *et al.*, 2003; Wilkie *et al.*, 2011; Korfali *et al.*, 2012). Other evidence purportedly displaying β-dystroglycan within the nuclei of C2C12 cells cannot be relied upon since some crucial experiments seem to have been performed using irregular techniques which appear not to be well controlled (Martínez-Vieyra *et al.*, 2013). In the absence of other robust evidence, cytoplasmic observation of β-dystroglycan, together with its absence in proteomic studies perhaps signal towards the conclusion that dystroglycan is not present in the nuclei of muscle cells, though other factors could contribute to its nuclear localisation in other cell or tissue types (Mathew *et al.*, 2013; Leocadio, Mitchell and Winder, 2016). To confirm the idea that β-dystroglycan segregates to the nucleus only in certain cell types, it would be useful to express a construct in multiple cell types where β-dystroglycan is thought to differently localise to the nucleus. The inclusion of a phosphomimetic mutation at Y890 would also be useful, as it appears to be one determinant for nuclear localisation (Gracida-Jiménez *et al.*, 2017).

The idea that dystroglycan is not present in nuclei or nuclear structures of myoblasts, however, does not explain the increase in nuclear-centrosome distance observed between myoblasts with a wild-type or dystroglycan-deficient genotype. Though this result needs to be verified, given the modest increase in distance observed in the control myoblast clone (see section 4.2.6.2). Indeed, nuclear-centrosome aberrations have been previously observed in cells deficient in dystroglycan (Martínez-Vieyra *et al.*, 2013; Vélez-Aguilera *et al.*, 2018), thought to be mediated through concomitant disruption of emerin (Salpingidou *et al.*, 2007). In addition to centrosome tethering, more contemporary work also reveals that emerin is a key determiner of nuclear polarity (Nastały *et al.*, 2020). Experiments performed during the current study found that overall emerin levels were consistent between dystroglycan-deficient and wild-type myoblasts and did not further probe for any differences in sub-nuclear emerin enrichment in polarised cells. Nevertheless, an increased nuclear-centrosome distance in dystroglycan-deficient cells, combined

with previous reports indicating β-dystroglycan-emerin binding (Gómez-Monsivais *et al.*, 2020), and the contribution of emerin relocalisation for cell polarity potentially implicate β-dystroglycan in the subtleties of NE protein polarisation. If β-dystroglycan does not accumulate within the nuclear region of human myoblasts, however, then it is not clear how β-dystroglycan co-operates with emerin in order to control nuclear-centrosome distance. There is evidence that a population of emerin on the ONM can control nuclear-centrosome distance (Salpingidou *et al.*, 2007). However, to establish whether β-dystroglycan collaborates with emerin in the context of the ONM would require further, highly resolved microscopy studies considering β-dystroglycan, emerin and centrosomal markers together.

8.3. Regulation of the subcellular distribution of β-dystroglycan

Through the current study, an epitope-tagging strategy revealed that, upon proteolytic cleavage, β-dystroglycan fragments become separately organised within the cytosol. This observation supports the initial hypothesis that fragmented β-dystroglycan becomes spatially organised. However, the latter hypothesis, which postulated that distinct pools of protein have separate functions, was not tested in any depth. Fates for intracellular dystroglycan range from degradation to nuclear import (Lara-Chacón *et al.*, 2010; Lipscomb *et al.*, 2016; Cho *et al.*, 2018), though it is not possible to exclude further signalling functions as it transits through the cell, as the promiscuous carboxy-terminus is exposed to the cytoplasm. Possible fates for dystroglycan are outlined in Figure 8-1. The separate organisation of dystroglycan fragments and the existence of several known post-translational modifications together indicates that β-dystroglycan populations have the potential for further specification to undertake alternative functions. This idea has precedent given that dystroglycan exists in a number of other complexes in addition to the DGC (Johnson *et al.*, 2013), though its identity in this study appears to be limited to the 43 kDa species. Moreover, multiple phosphorylation and ubiquitination sites have been identified on dystroglycan. Phosphorylation at Y890 appears not to be an exclusive signal for degradation given the observation of pY892 β-dystroglycan in nuclei (Gracida-Jiménez *et al.*, 2017). Hence, there must be alternative mechanisms to

prevent all pY892 β-dystroglycan being targeted for degradation. Indeed, whether each fragment of β-dystroglycan is equally susceptible to these post-translational modifications is unclear, as is the possibility that β-dystroglycan can be labelled by combinations of modifications.

The endocytic trafficking pathway for β-dystroglycan has not been fully resolved (Figure 8-1), and the idea that β-dystroglycan, when subjected to abundant modifications, has diverging functions may have precedent since there are inconsistent observations in attempts to dissect its subcellular trafficking. β-dystroglycan has been reported to both co-localise and not co-localise with EEA1-positive early endosomes. Staining of total β-dystroglycan shows some co-staining with EEA1 (Gracida-Jiménez *et al.*, 2017), whereas specific staining for non-Y892 phosphorylated β-dystroglycan, or β-dystroglycan phosphorylated on Y892 displays clear separation between signals, as does an alkaline phosphatase fused β-dystroglycan construct (Sotgia *et al.*, 2003; Miller *et al.*, 2012). The disparities in these observations potentially indicate that other factors, or modifications, are involved in β-dystroglycan subcellular sorting.

β-dystroglycan has been shown to be ubiquitinated and subsequently degraded (Gazzerro *et al.*, 2010; Lipscomb *et al.*, 2011, 2016; Cho *et al.*, 2018). A recent report identified WWP1, a NEDD4 family ubiquitin ligase as responsible for β-dystroglycan degradation, which in turn appears to be dependent upon the presence of the PPxY motif at its extreme C-terminus (Cho *et al.*, 2018). This observation is consistent with previous reports linking β-dystroglycan degradation via the ubiquitin-proteasome system to its phosphorylation at Y892, (Lipscomb *et al.*, 2016). However, little is known about whether other ubiquitination modifications could specify further functions for dystroglycan other than degradation. Again, the handling of fragmented forms of dystroglycan by the ubiquitin-proteasome system is not known.

To date, there are no reports of conditions directly linked to aberrant fragmentation of dystroglycan, except for cancer (Losasso *et al.*, 2000; Singh *et al.*, 2004). However, primary dystroglycanopathies are thought to be exceptionally rare as a result of the essential role of dystroglycan during early embryonic development

(Williamson *et al.*, 1997; Brancaccio, 2019). Therefore, mutations which affect normal proteolytic regulation may also be embryonic lethal and not prevail in human populations. Genetic experiments specifically mutating sites of proteolysis would further refine our understanding of the consequences of non-functional dystroglycan.
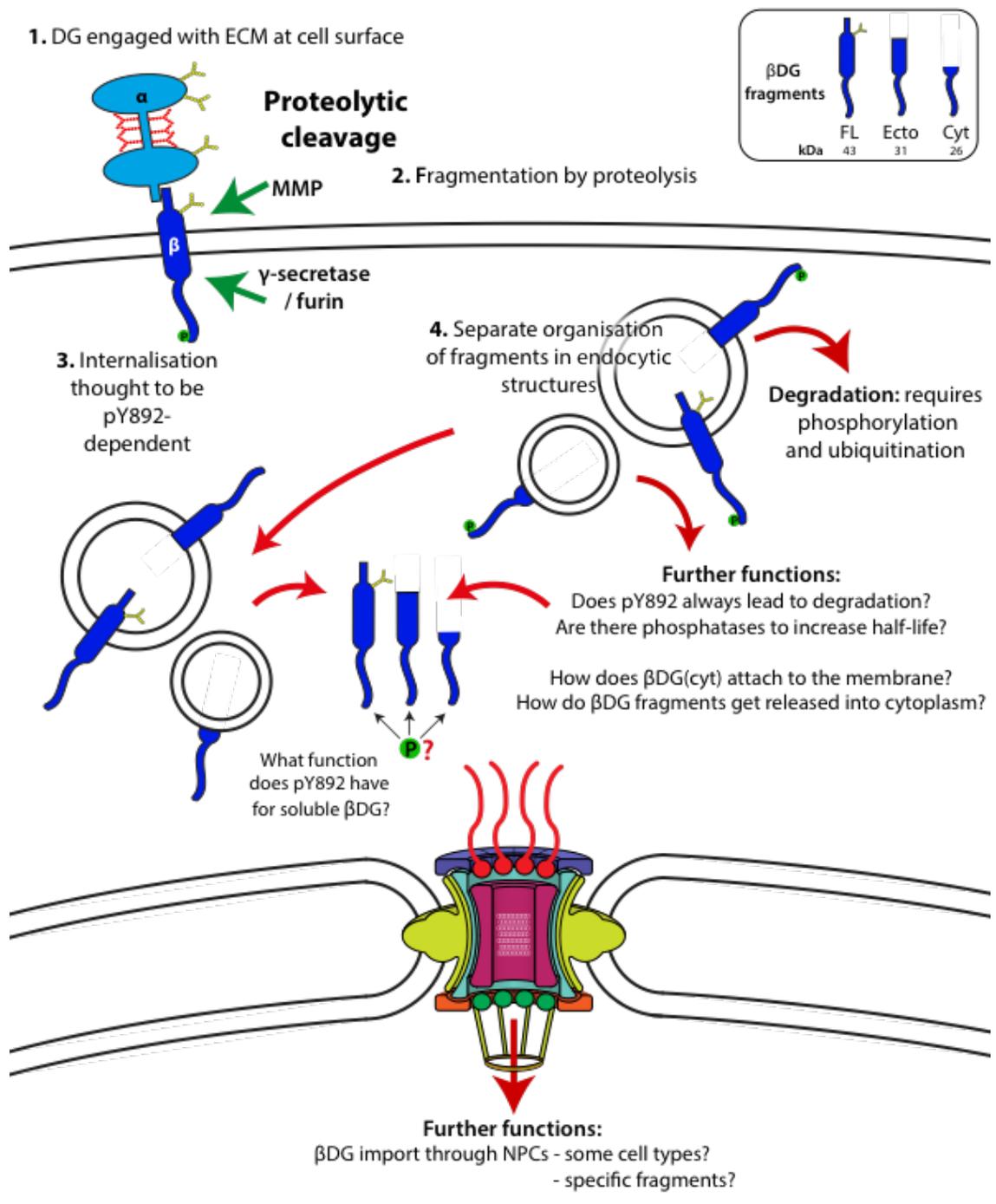
Figure 8-1 – Working model and outstanding questions for membrane trafficking of dystroglycan. β-dystroglycan (β-DG) can be cleaved by multiple enzymes resulting in the existence of multiple fragments. Previous work suggests that internalisation of β-dystroglycan (β-DG) is dependent on phosphorylation at Y892 (pY892). Data in the current study indicates that at least the cytoplasmic fragment can be separately organised from full length within the cytosol. One fate for β-DG is ubiquitination and degradation, which also appears Y892 dependent, though it is not known if this event is exclusive.

8.4. Dystroglycan as a moonlighting protein?

The observation of dystroglycan within several subcellular compartments, which may be dependent upon a number of modifications leads one to consider whether dystroglycan could be considered a moonlighting protein. A moonlighting protein is sometimes referred to as gene-sharing due to the multifunctional nature of these proteins challenging the original paradigm that one function is performed by one protein arising from a single gene (Jeffery, 1999). Indeed, like β-dystroglycan, the functions of moonlighting proteins are not dispersed into different protein modules and are commonly regulated using post-translational modifications, expression in different cell types or subcellular localisations (Jeffery, 2003; Copley, 2012). A high degree of evolutionary conservation is also characteristic of moonlighting proteins, and this is also featured in β-dystroglycan (Figure 5-2) (Huberts and van der Klei, 2010). However, an appreciation that different dystroglycan species possibly function in alternative pathways is likely to aid advances in unveiling its biology by refining coarse experiments to analyse specific pools of dystroglycan. Thus, the development of new tools to probe dystroglycan is essential. A possible starting point may be to employ a comparative proximity-dependent biotinylation technique where the promiscuous biotin ligase enzyme is recruited, or included (depending whether a 2C-BioID system is optimal), within different positions through the β-subunit of dystroglycan, perhaps combined with phosphomimetic mutants. Differences between datasets would establish specific interactors for each β-dystroglycan species, clarifying known functions while uncovering further details for subcellular trafficking routes and novel interactors for variably modified β-dystroglycan. Stable expression of dystroglycan is required for BioID studies (Roux, Kim and Burke, 2013). However, maintenance of dystroglycan transduction originating from numerous constructs in multiple cell types was not uniformly achieved throughout this study. Previous reports indicate that cells tightly regulate the amount of truncated dystroglycan within nuclei of some cell types (Vélez-Aguilera *et al.*, 2018), raising the possibility that exogenous dystroglycan may become silenced over time, thus requiring adjustments to the system for stable, low-level expression.

Of course, the rather disparate functions currently thought to exist for dystroglycan may not be significant of moonlighting functions, but rather gaps in knowledge of how dystroglycan undergoes sequential signalling, or functions within pathways which accounts for observations made throughout the cell to date.

8.5. Portability of 2C-BioID2 into other systems

Several features of the 2C-BioID system employed in this study render it applicable to the study of other proteins *in vivo*. The benefits of Cre-dependence and FRB-FKBP chemical dimerisation have been previously discussed (see section 6.1.3), however the use of BioID is also well suited to mammalian *in vivo* studies over other proximity dependent biotinylation approaches. One such technique is the engineered ascorbate peroxidase (APEX) (Hung *et al.*, 2014). APEX has been extensively used in similar manner to BioID, though its kinetics are faster (Hung *et al.*, 2014) making it more useful than BioID for the study of dynamic processes. However, its dependence on toxic compounds including hydrogen peroxide and biotin-phenol precludes its administration to animals. On the other hand, BioID can utilise dietary biotin which is easily controlled (Hendrikje Werner, personal communication) while the requirement for only one compound is likely to improve reproducibility between samples. Dynamic processes are likely to be challenging to examine using BioID *in vivo* given the constitutive activity of all biotin ligase enzymes known to date. However, time resolved analyses might be useful to probe protein interactome during disease progression, and so the combination of BioID or derivatives such as TurboID, which combines the convenience of BioID with increased biotinylation kinetics (Branon *et al.*, 2018), with inducible gene regulation may be of use. However, time resolution using either enzyme is likely to be coarse due to the induction time required for genetic recombination. Differences in dynamic molecular events *in vivo* would require the development of a biotin ligase enzyme which can be rapidly induced, targeted or activated.

8.6. The Sun1 interactome in muscle

The Sun1 interactome in myoblasts was initially characterised using 2C-BioID, though the top scoring candidate interactors were not typically NE proteins (see section 6.2.7). This would indicate that although the lamin A network appears to be a common anchor for many NE proteins (Haque *et al.*, 2010; Wilson and Foisner, 2010), proteins of the NE are distinctly localised. Like the lamin A interactome derived using BioID, the Sun1 interactome uncovered during this study does not detect an enrichment of histones, which have abundant lysine residues which can be readily biotinylated (Roux *et al.*, 2012), further reinforcing the validity of identified proteins. Instead, the enrichment of histone modifying and chromatin remodelling complexes resonates with functions ascribed to Sun1 and the LINC complex in genome regulation (Guilluy *et al.*, 2014; Le *et al.*, 2016; Tajik *et al.*, 2016). Emerin, lamin A and the LINC complex appear to act in synergy to elicit changes in stiffness to the NE (Guilluy *et al.*, 2014). While Sun1 BioID detects lamin A, and the interaction can be shown independently, the absence of emerin within the Sun1 interactome is perplexing given that *in vitro* immunoprecipitation studies and PLA between Sun1 and emerin indicate that such an interaction exists (Haque *et al.*, 2010) (Figure 3-13). Interestingly, B-type lamins were also not observed as a top scoring interactor of Sun1, despite its observation using PLA as a positive control (Figure 3-8), however this interaction remains contentious in the literature (Crisp *et al.*, 2006; Haque *et al.*, 2006; Nishioka *et al.*, 2016). One explanation behind these inconsistencies could be that interaction networks amongst NE proteins are differentially regulated between cell types. Indeed, NE proteins between tissues are highly variable (Kavanagh *et al.*, 2007; Korfali *et al.*, 2012). Sun-domain proteins appear indispensable for mechanosensing through emerin (Guilluy *et al.*, 2014) The indirect association of key players like Sun1 and emerin, as is observed here, may indicate that a direct interaction is dispensable, and mechanosensing can occur through pathways that are linked to the NL which ultimately functions as a nexus to integrate signals.

One aspect of interactome studies which is sometimes neglected is that of the contribution of splice variants. Unfortunately, due to difficulties in verifying the FRB-

knock in to the Sun1 genomic locus and time constraints, the Sun1 interactome generated during this study also does not address this. The impact on the Sun1 interactome through using a single construct within myoblasts is thought to be minimal since the predominant splice isoform of Sun1 is the full length version (Loo *et al.*, 2019). How the Sun1 interactome changes upon differentiation from muscle progenitors into syncytial myofibres would be of particular interest given that other components of the LINC complex play a fundamental role in the relocation of the MTOC during muscle development (Gimpel *et al.*, 2017), and since Sun1 which harbours EDMD mutations appears to affect the ability of C2C12 myoblasts to efficiently differentiate in culture (Meinke *et al.*, 2014). However, using a full length Sun1 isoform for comparative studies across differentiation stages between myoblasts and *in vitro* myotubes, would be physiologically inappropriate. Similar to normal myofibres, myotubes downregulate the full length and upregulate a variety of splice isoforms (Nishioka *et al.*, 2016; Loo *et al.*, 2019), and identification of interactors using only one Sun1 splice variant increases the likelihood that false interactors are discovered, while decreasing those which are real. Attempts were made to conserve the Sun1 splicing architecture by inserting a FRB-tag to Sun1 using CRISPR, though despite its detection in the genome by PCR, expression of the V5-tagged version could not be detected. Moreover, the decline in myogenic potential of myoblast lines of both human and murine origin was continually encountered during clonal expansion possibly because cells with the ability to differentiate are diluted (Beauchamp *et al.*, 1999). A very large number of myoblast lines would require screening for their ability to differentiate, making addressing the Sun1 interactome in myotubes technically challenging using a CRISPR system in myoblasts.

8.6.1. *The Sun1 interactome in the context of the nuclear laminopathies*

One of the most promising Sun1 interacting candidate was Ppm1a, a protein phosphatase responsible for dephosphorylating Smads 2 and 3, thus inactivating TGFβ signalling. As mentioned (see section 7.3.3), Smads the principal effectors of the TGFβ signalling pathway. Fibrosis is a common feature of muscular dystrophies resulting from laminopathies (Chatzifrangkeskou *et al.*, 2016; Bernasconi *et al.*,

2018), as fibroblasts differentiate into myofibroblasts, replace muscle tissue and decreasing muscle function (Hinz *et al.*, 2001). However, within myoblasts, TGFβ inhibits transcription factors crucial for myogenesis (Liu, Black and Derynck, 2001). Whether deregulated TGFβ signalling in muscle progenitors is partially responsible for muscle wasting observed in laminopathies is not known. However, some lamin A mutations markedly decrease the differentiation efficiency of myoblasts (Favreau *et al.*, 2004; Markiewicz, Ledran and Hutchison, 2005; Frock *et al.*, 2006; Bertrand *et al.*, 2020). Whether lamin A mutations in muscle precursor cells elicit these effects through dysfunction of TGFβ handling at the NE is not clear, however there is precedent for the requirement of lamin A in the regulation of TGFβ signalling. For differentiation, myoblasts must exit the cell cycle; a process which is dependent on the activation of pRB by hypophosphorylation (Pekovic *et al.*, 2007). Cells lacking lamin A show accelerated proliferation since pRB is normally dephosphorylated by the lamin A-bound phosphatase, PP2a (Van Berlo *et al.*, 2005). Lamin A KO myoblasts also display an increase in Smads 2 and 3 levels, which contribute to cell death during differentiation into myotubes (Cohen *et al.*, 2013). These observations, combined with data generated during the current study, would further implicate Sun1 and the LINC complex further into the gene expression hypothesis for the aetiology of the laminopathy diseases.

To date, the NE regulation of TGFβ signalling is limited to Smads 2 and 3. However, BMP engagement by TGFβ receptors causes the activation and nuclear localisation of Smads 1, 5 and 8, which are antagonised by Smad6 and have alternative gene targets (Guo and Wang, 2009). Embryonic myogenesis is dependent upon BMP signalling to ensure proper proliferation of muscle precursors (Ono *et al.*, 2011). Dysregulation of BMP signalling has also be implicated in a model of laminopathies (Janin *et al.*, 2018). In lamin A deficient mice, BMP4 is downregulated, while levels of the inhibitory Smad6 are increased. While the RNA interference of Smad6 is sufficient to rescue differentiation of myoblasts, it is not known whether the NL can normally function to do this.

The striking observation that the depletion of Sun1 reduces the disease phenotypes of several laminopathy disease models led to the hypothesis that

nucleoplasmic Sun1 interactors might change depending on the specific lamin mutation present. In this way, the broad range of known laminopathy diseases were postulated to be provoked by variably deleterious Sun1 protein interactomes. The research presented in this thesis did not have scope to fully address this hypothesis, though advances for the *in vivo* applications of 2C-BioID and resulting Sun1 interactome in myoblasts have made inroads to make further testing of the hypothesis more accessible. Moreover, BioID data reveals candidate Sun1 interactors which contribute to pathways already implicated within the laminopathies. Ultimately, *in vitro* studies must be transferred *in vivo* to further unravel how the Sun1 interactome changes between tissue-types to capture full physiological context; both with respect to the splicing of Sun1, and tissue microenvironments which cannot be faithfully replicated *in vitro*.

8.7. Concluding remarks

β-dystroglycan is an important component of the DGC, though distinct functions within other subcellular environments and the nucleus remain elusive. Epitope-tagging of β-dystroglycan provides a small glimpse into an assortment of β-dystroglycan façades created by fragmentation and post-translational modifications. An appreciation of the existence of versions of β-dystroglycan will surely accelerate our understanding of its diverse functions, and evidence indicates that subcellular populations can be dissected for further investigation.

Results from this study also add confidence in an *in vivo* 2C-BioID system. The observation of expression in numerous cell types including fibroblasts, myoblasts and myotubes provides reassurance that the system should be applicable to other tissues as well, which permits portability into further proteins. The generation of the Sun1 interactome further reveals the influence of Sun1 reaching beyond the NE, and potentially into the regulation of signalling pathways. Exactly how LINC complexes engage with signalling components or modulators remain to be elucidated, but these observations provide an interesting avenue for further research which may lead to therapeutic strategies for laminopathy diseases.

Supplementary information

Antibody controls:

Transfected construct(s):	V5	HA	Myc	Merge + **Hoechst**
αV5-β-DG				
α-HAβ-DG				
α-βmyc-DG				
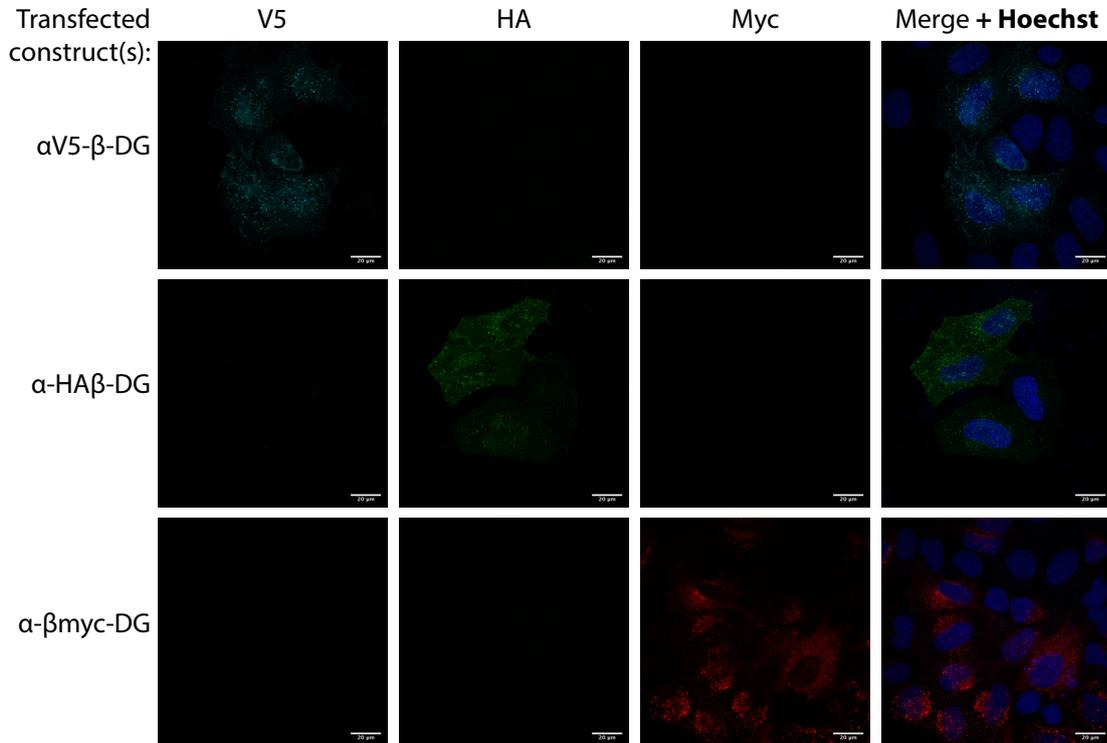
Supplemental Figure I – Relating to Figure 5-4. Antibody controls for staining of epitopes interspersed throughout the dystroglycan protein. HeLa cells were transfected with the αV5-HAβmyc-DG construct. Samples were stained using one of the primary antibodies against the epitope tags, which is indicated. Primary antibodies were detected using the same mixture of secondary antibodies used for co-staining to ensure the absence of cross reactivity. Hoechst was used to stain nuclei.

| Myc | HA | Lamin A | Composite + Hoechst |

Untransfected

| Myc | HA | Calnexin | Composite + Hoechst |
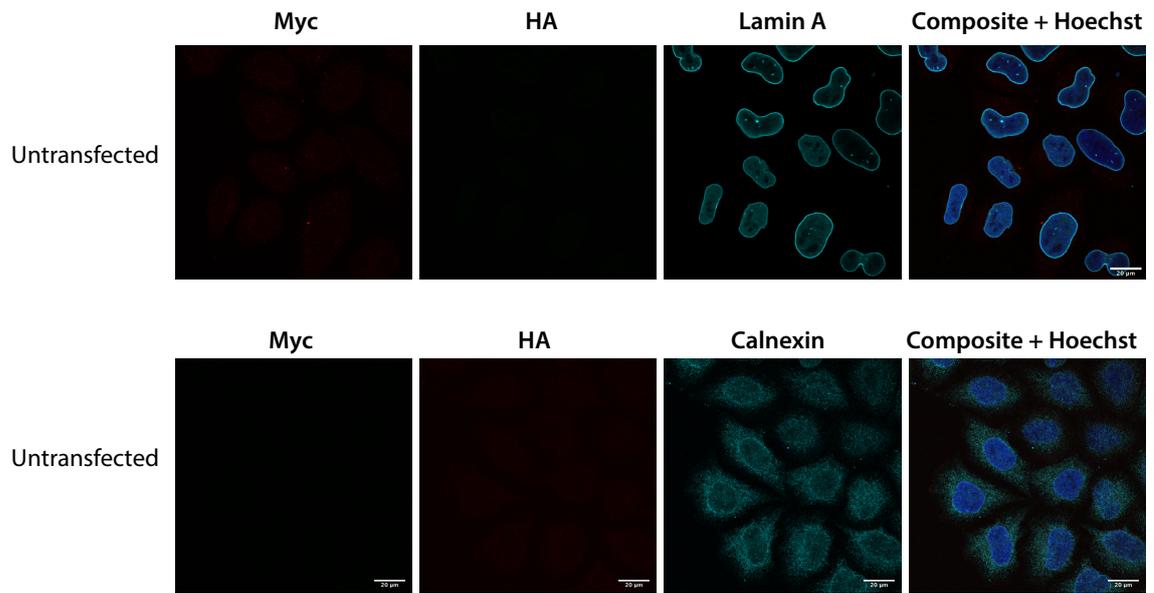
Untransfected

Supplemental Figure II – Relating to Figure 5-9. Secondary antibody controls for Myc and HA antibodies. Untransfected HeLa cells were stained with Myc and HA antibodies to ensure the absence of non-specific staining.
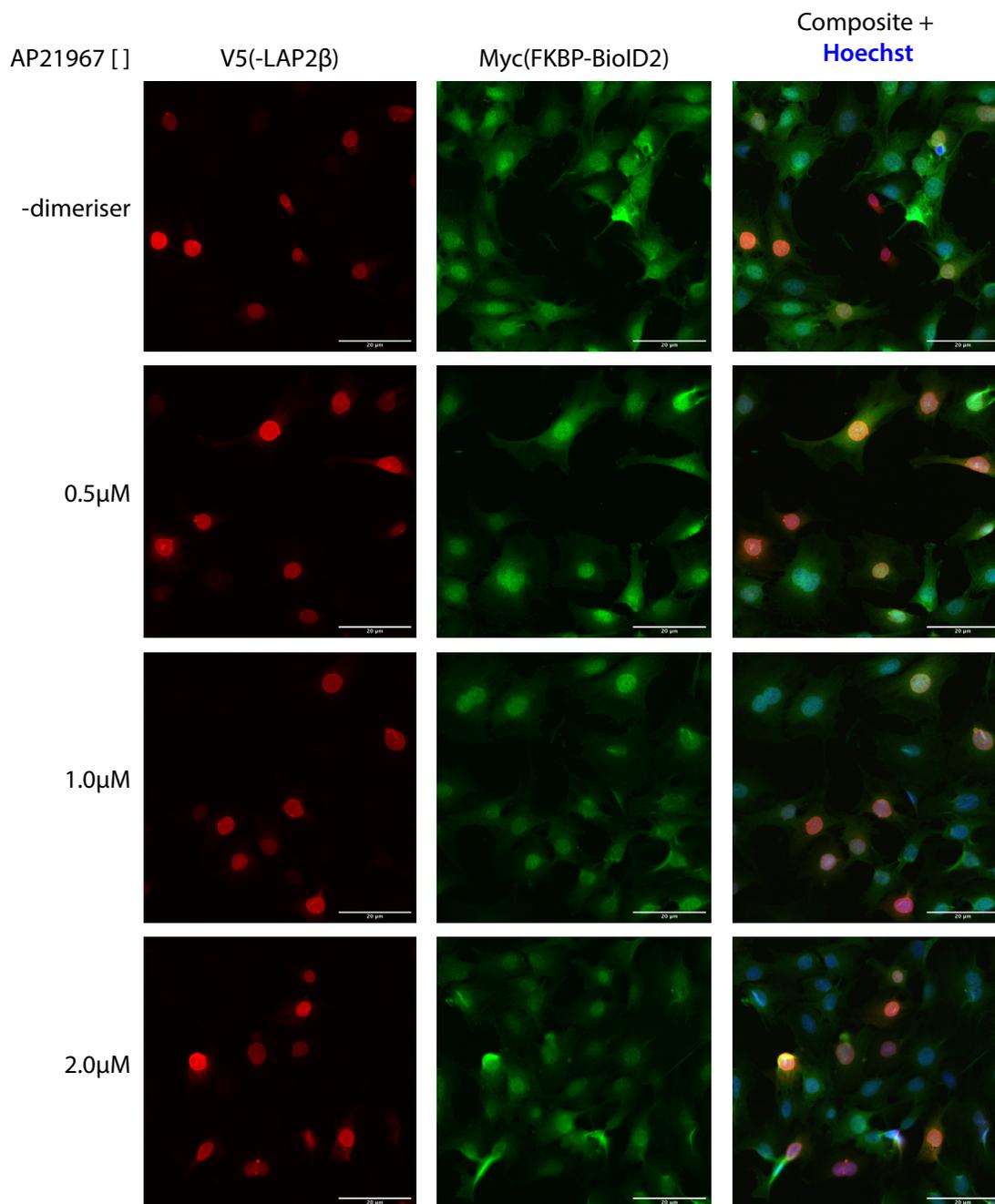
Supplemental Figure III – Optimising the relocalisation of mycFKBP-BioID2 to FRB-tagged protein of interest (FRB-Lap2β) in MAFs. Increasing concentrations of FKBP-FRB dimeriser AP21967 were titrated into cultures of Rosa26-mycFKBP-BioID2$^{+/-}$ MAFs and recruitment assessed by immunofluorescence microscopy.
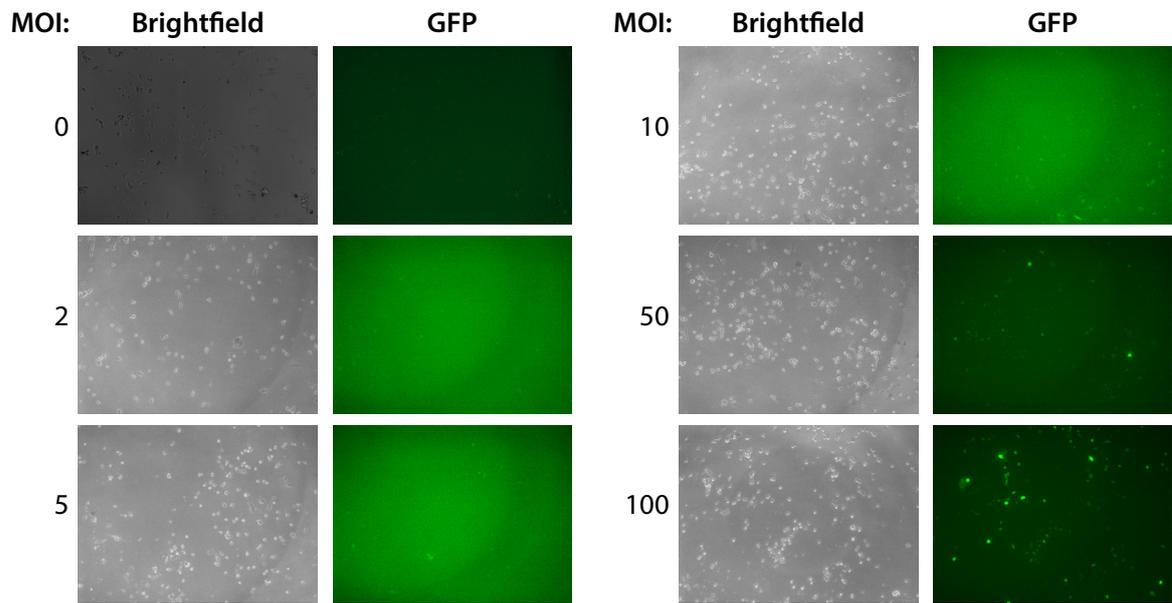
Supplemental Figure IV - Optimal multiplicity of infection (MOI) for adenovirus GFP in myoblasts is 100. Myoblasts were infected with the indicated MOI of GFP containing adenovirus and cultured for three days before analysis by fluorescence microscopy.
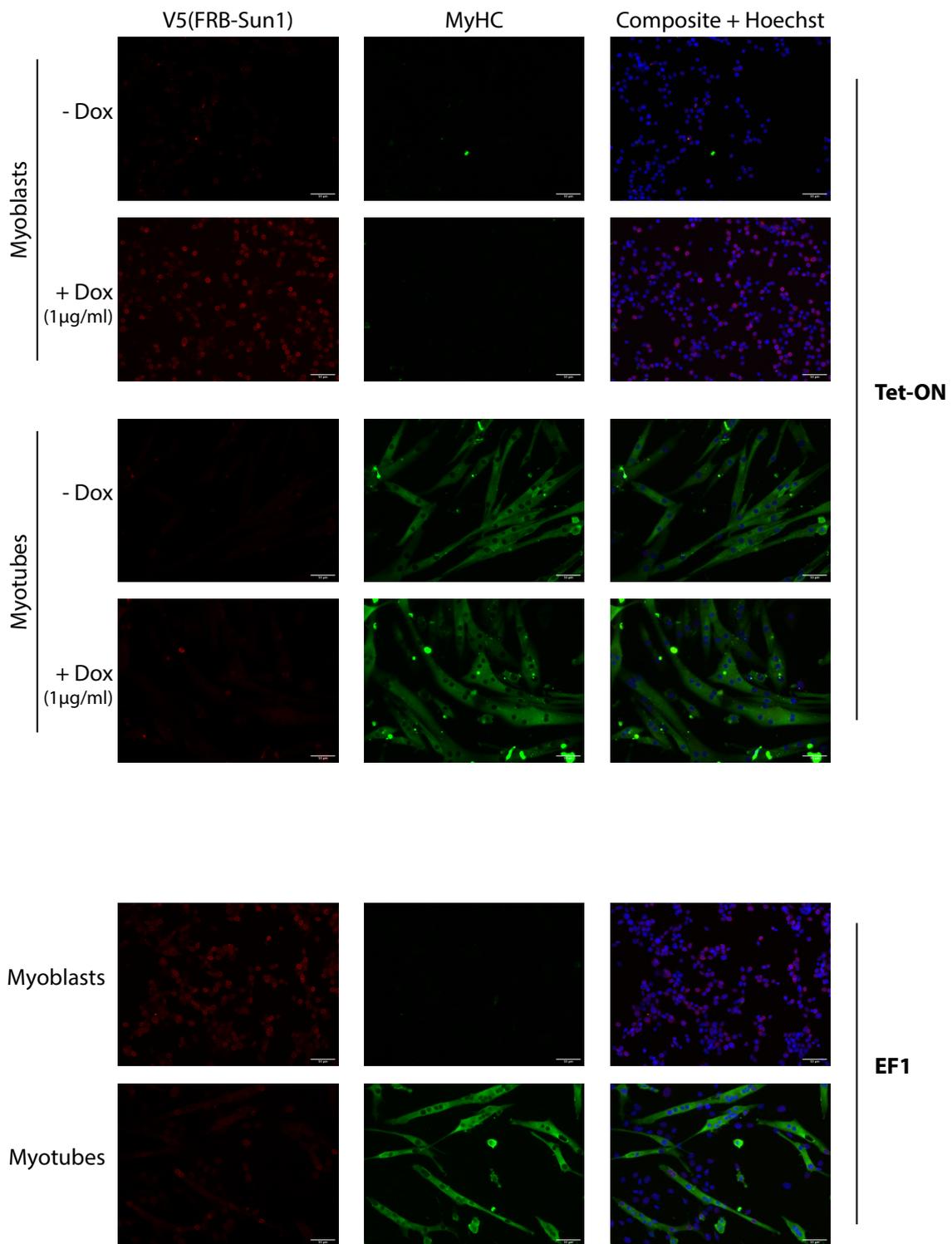
Supplemental Figure V - Sun1 expression is not sustained upon myotube differentiation. Myoblasts were transduced with an inducible V5FRB-Sun1 lentiviral construct with either a Tet-ON minimal CMV or EF1 constitutive promoter and selected for stable integration. Myoblasts were and cultured for 5 days or differentiated into myotubes over 5 days. V5FRB-Sun1 expression was induced by adding 1μg/ml doxycycline to the culturing medium for three days prior to fixation and immunofluorescence analysis. Myoblasts and myotubes were stained using antibodies against V5 to detect the FRB-Sun1 transgene and myosin heavy chain to indicate differentiated myotubes.

	HA	Lamin B1	Composite + Hoechst

Sun1-HA

HA-**Sun1**

	FLAG	Lamin B1	Composite + Hoechst

Alox5-FLAG

FLAG-**Alox5**

Arhgap22-FLAG

FLAG-**Arhgap22**

	FLAG	**Lamin B1**	**Composite + Hoechst**
Tnnt3-FLAG			
FLAG-**Tnnt3**			
Vps51-FLAG			
FLAG-**Vps51**			
Zbtb20-FLAG			
FLAG-**Zbtb20**			

	FLAG	**Lamin B1**	**Composite + Hoechst**
Ccdc86-FLAG			
FLAG-**Ccdc86**			
Dctn5-FLAG			
FLAG-**Dctn5**			
Parva-FLAG			
FLAG-**Parva**			

	FLAG	**Lamin B1**	**Composite + Hoechst**

Hist1h2bp-FLAG

FLAG-**Hist1h2bp**

Ppm1a-FLAG

FLAG-**Ppm1a**

Csrp3-FLAG

FLAG-**Csrp3**

	FLAG	**Lamin B1**	**Composite + Hoechst**
Ruvbl2-FLAG			
FLAG-**Ruvbl2**			
Tnni1-FLAG			
FLAG-**Tnni1**			
Fam91a1-FLAG			
FLAG-**Fam91a1**			

	FLAG	Lamin B1	Composite + Hoechst
Ruvbl1-FLAG			
FLAG-**Ruvbl1**			

488 exposure:	FLAG	HA	Lamin B1	Composite + Hoechst
30 ms				
60 ms				

Supplemental Figure VI – Expression of Sun1 candidate interacting proteins in HeLa cells. Protein coding sequences were amplified from cDNA originating from Rosa26-mycFKBP-BioID2+/- myoblasts, and FLAG-tagged at the amino and carboxy-terminus. HeLa cells were transfected with constructs and fixed and stained for the FLAG-tag and lamin B1 to delineate the nuclear lamina 24 hours later.

	HA	Lamin B1	Composite + Hoescht
HA-Sun1			
Sun1-HA			
	FLAG		
FLAG-Alox5			
Alox5-FLAG			
FLAG-Arhgap22			
Arhgap22-FLAG			

312

	FLAG	Lamin B1	Composite + Hoescht

FLAG-Tnnt3

Tnnt3-FLAG

FLAG-Vps51

Vps51-FLAG

Parva-FLAG

	FLAG	Lamin B1	Composite + Hoescht
FLAG-Ccdc86			
Ccdc86-FLAG			
FLAG-Csrp3			
Csrp3-FLAG			
FLAG-Dctn5			

	FLAG	Lamin B1	Composite + Hoescht
FLAG-Fam91a1			
Fam91a1-FLAG			
FLAG-Hist1h2bp			
Hist1h2bp-FLAG			
Ppm1a-FLAG			

	FLAG	Lamin B1	Composite + Hoescht
FLAG-Ruvbl1			
Ruvbl1-FLAG			
FLAG-Ruvbl2			
Ruvbl2-FLAG			
FLAG-Tnni1			

	FLAG	Lamin B1	Composite + Hoescht
FLAG-Zbtb20			

	FLAG	HA	Lamin B1	Composite + Hoechst
Untransfected				
No primary Ab				

Supplemental Figure VII – Expression of Sun1 candidate interacting proteins in C2C12 myoblasts. FLAG-tagged constructs were transfected into C2C12 cells and fixed 24 hours later. Cells were stained with antibodies against the FLAG-epitope tag and lamin B1 to delineate the nuclear envelope.

HA	Lamin B1	Merge + Hoechst

Alox5
Ct-VN210

Alox5
VN210-Nt

Arhgap22
Ct-VN210

Arhgap22
VN210-Nt

Ccdc86
Ct-VN210

Ccdc86
VN210-Nt

	HA	Lamin B1	Merge + Hoechst

Dctn5
Ct-VN210

Dctn5
VN210-Nt

Parva
Ct-VN210

Parva
VN210-Nt

Ruvbl2
Ct-VN210

Ruvbl2
VN210-Nt

Tnni1
VN210-Nt

	HA	**Lamin B1**	**Merge + Hoechst**

Tnnt3
Ct-VN210

Tnnt3
VN210-Nt

Zbtb20
VN210-Nt

Fam91a1
Ct-VN210

Fam91a1
VN210-Nt

Hist1h2bp
Ct-VN210

Hist1h2bp
VN210-Nt

	HA	**Lamin B1**	**Merge + Hoechst**
Ppm1a Ct-VN210			
Ppm1a VN210-Nt			
Cttn Ct-VN210			
Cttn VN210-Nt			
Csrp3 Ct-VN210			
Csrp3 VN210-Nt			

	HA	Lamin B1	Merge + Hoechst
Ruvbl1 Ct-VN210			
Ruvbl1 VN210-Nt			
Empty Vector Ct-VN210			
Empty Vector VN210-Nt			
Not transfected			

Supplemental Figure VIII – Localisation of Sun1 candidate interactors with BiFC tags is consistent with that determined with FLAG epitope tags. HeLa cells were transfected with constructs encoding Sun1 candidate interactors tagged separately at the N- and C-terminus with the VN210 BiFC fragment. Cells were fixed with 4% PFA 24 hours post-transfection with antibodies raised against the HA epitope tag for detection of the candidate interactor and lamin B1 to delineate the nuclear envelope.

| | HA | 3(G₄S)-Myc-Sun1 | Composite + Hoechst |

Lamin A, Parva, Alox5, Ruvbl2, Ruvbl1, Tnni1, Ppm1a, Csrp3, Cttn, Untransfected
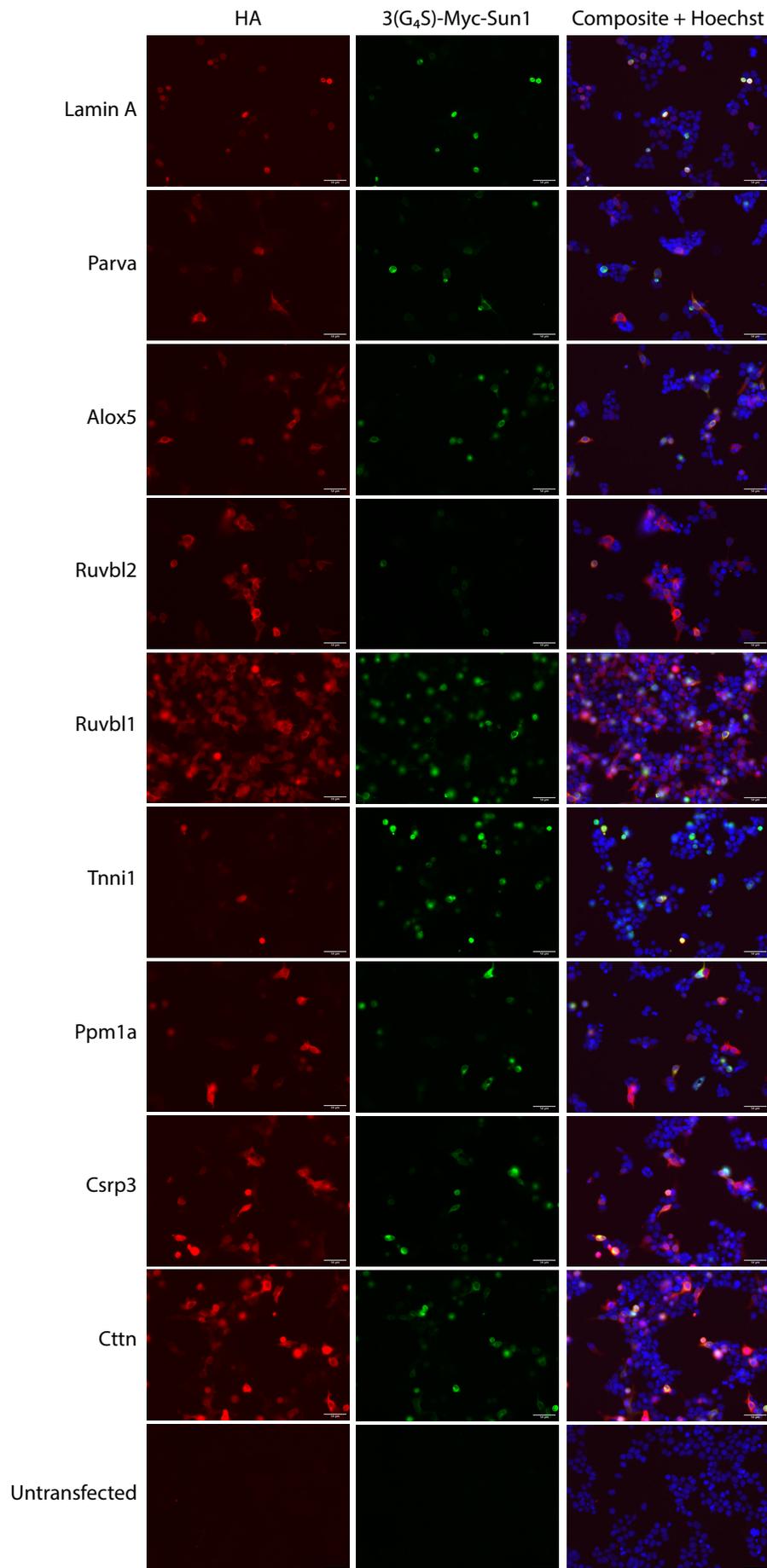
Supplemental Figure IX – Controlling for the expression of BiFC constructs in samples for BiFC-FC. Double immunofluorescence staining of HEK293T transfected with BiFC constructs. Probing with HA and Myc antibodies reveals the expression of BiFC constructs used in BiFC-FC experiments.

323

		HA	FLAG	Composite + **Hoechst**
Ccdc86	**-**			
Ccdc86	**+**			
Parva	**-**			
Parva	**+**			
Ruvbl2	**-**			
Ruvbl2	**+**			

Candidate interactor / Sun1-HA

324
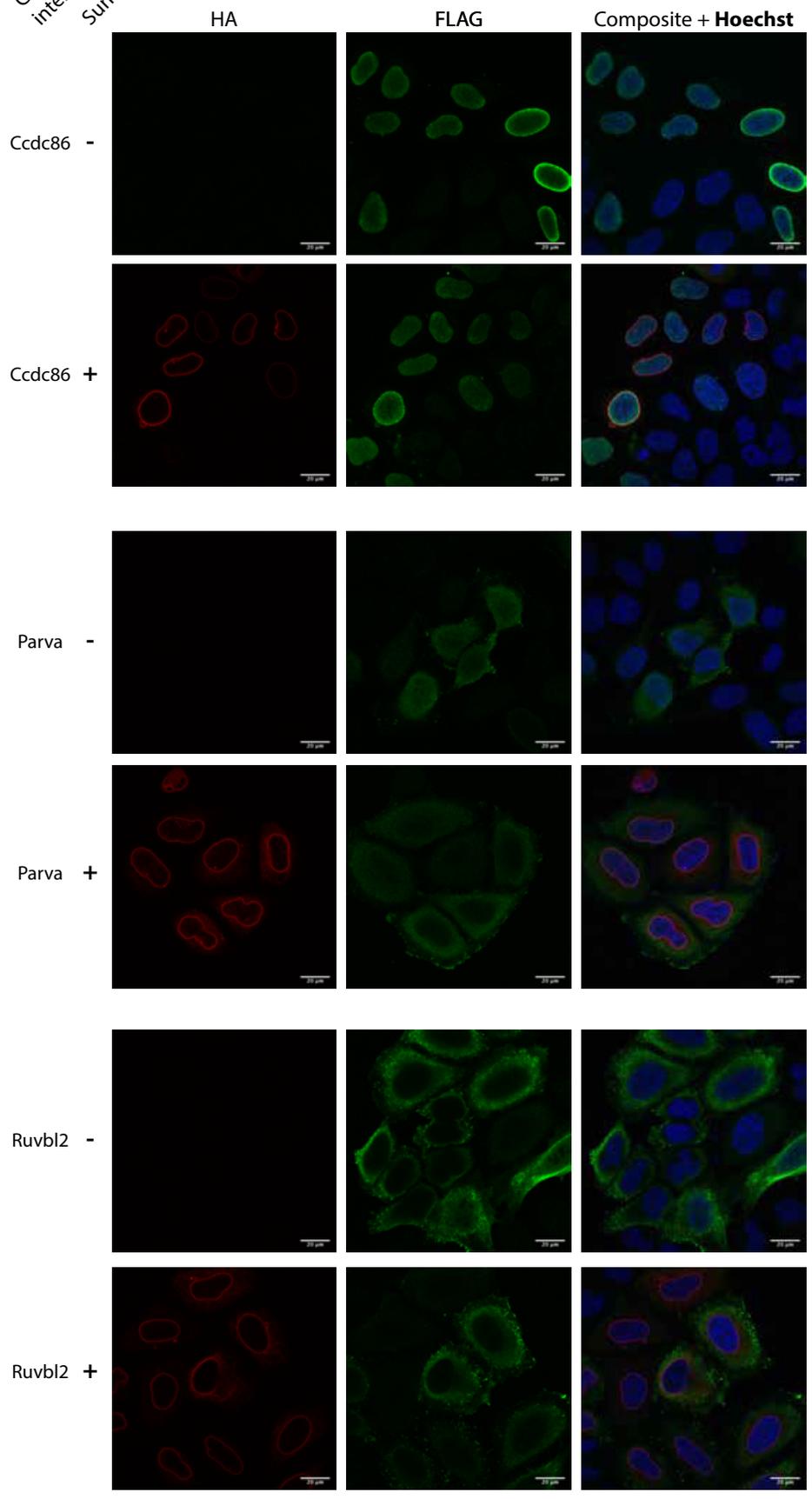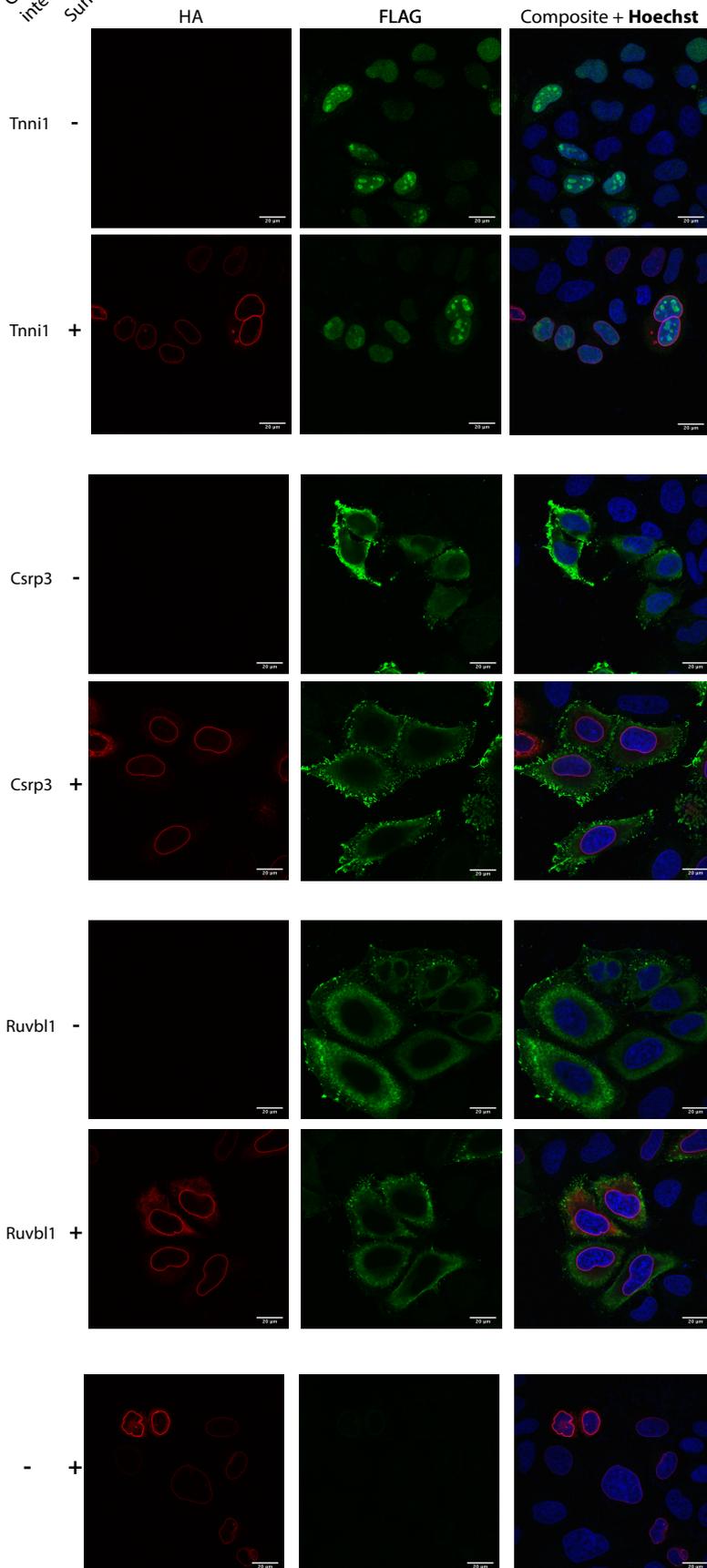
Supplemental Figure X – FLAG-tagged Sun1 candidate interactors are not recruited to a co-expression of Sun1-HA protein 24 hours post-transfection in HeLa cells. HeLa cells were transfected with Sun1-HA and the indicated potential interactor. After 24 hours, HeLa cells were fixed and stained with antibodies against HA and FLAG.

		FLAG	HA	Composite + Hoechst
Ccdc86	−			
Ccdc86	+			
Parva	−			
Parva	+			
Ruvbl2	−			
Ruvbl2	+			

Candidate interactor   Sun1-HA

326
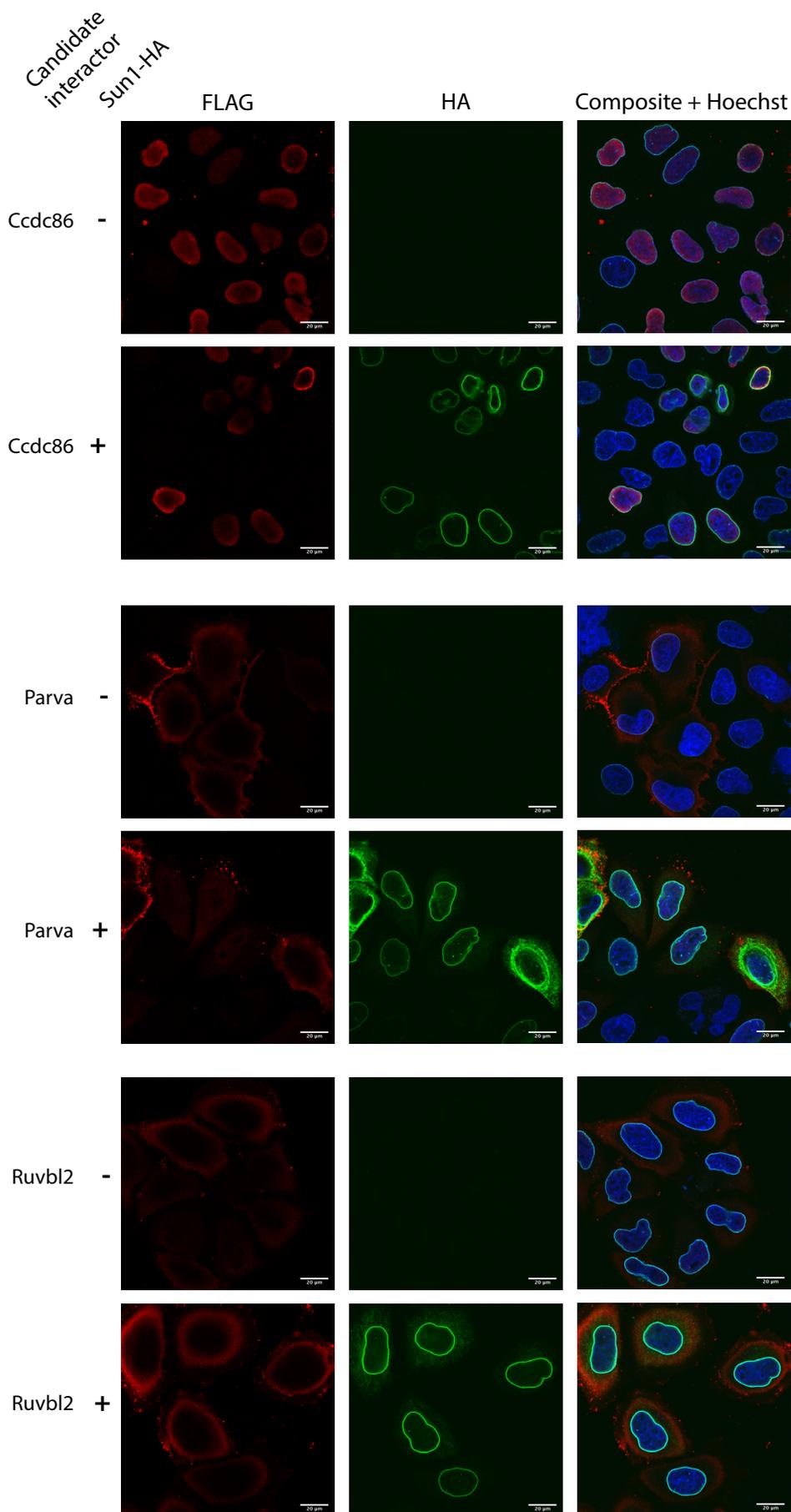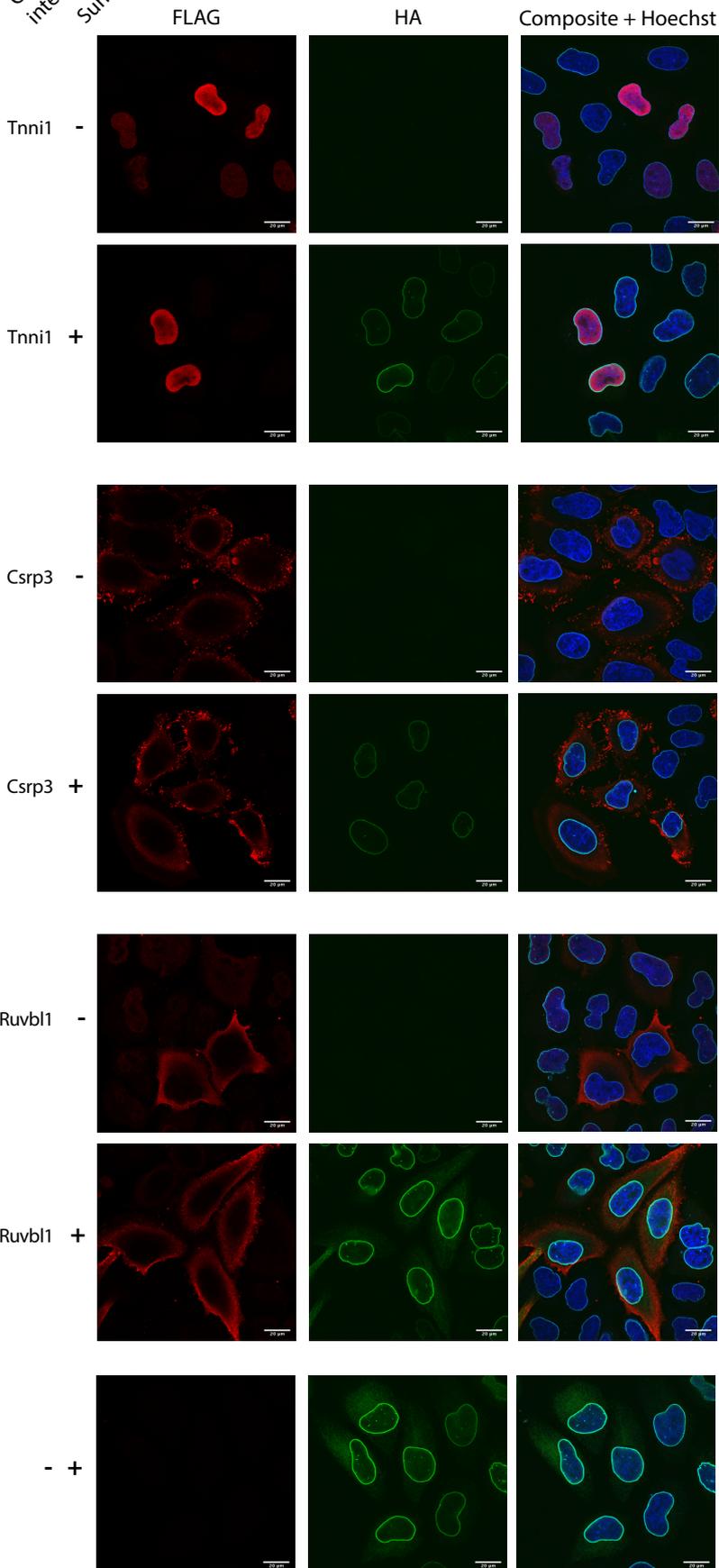
Supplemental Figure XI – FLAG-tagged Sun1 candidate interactors are not recruited to a co-expression of Sun1-HA protein 48 hours post-transfection in HeLa cells. HeLa cells were co-transfected with Sun1-HA and a potential Sun1 interactor, and fixed 48 hours post-transfection. Cells were then stained with antibodies raised against the HA and FLAG-epitope tags and lamin B1 to demarcate the nuclear envelope
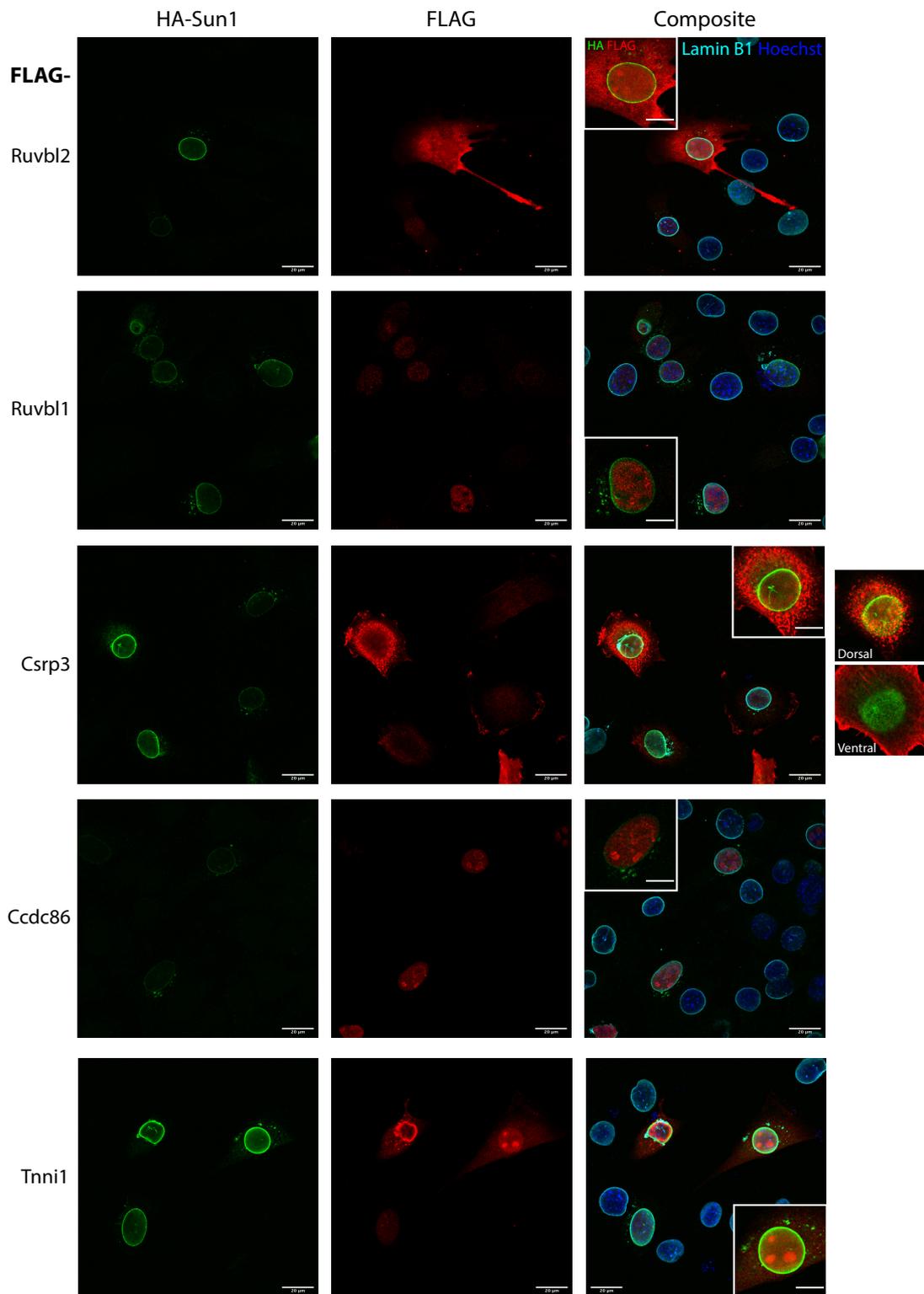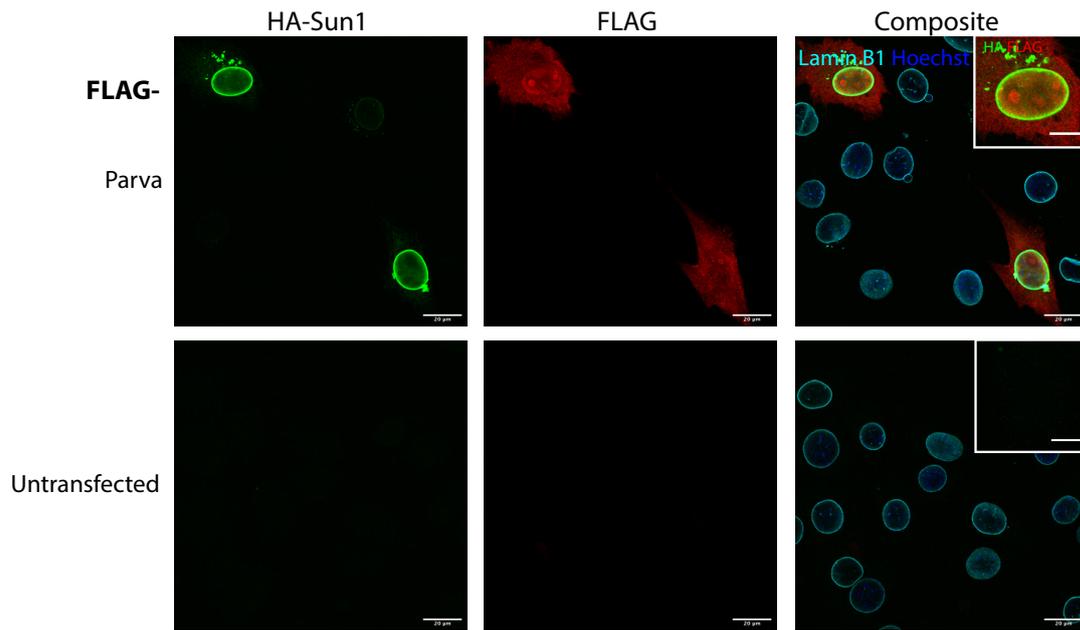
Supplemental Figure XII – Co-expressed of Sun1-HA and FLAG tagged candidate Sun1 interactors in C2C12 myoblasts do not co-localise at the nuclear periphery. C2C12 myoblasts were co-transfected with Sun1-HA and the indicated protein and fixed 24 hours later. Cells were stained with antibodies against the HA and FLAG-epitope tags to reveal the subcellular location of transfected constructs in relation to each other and the nuclear envelope, demarcated by lamin B1 staining (cyan).
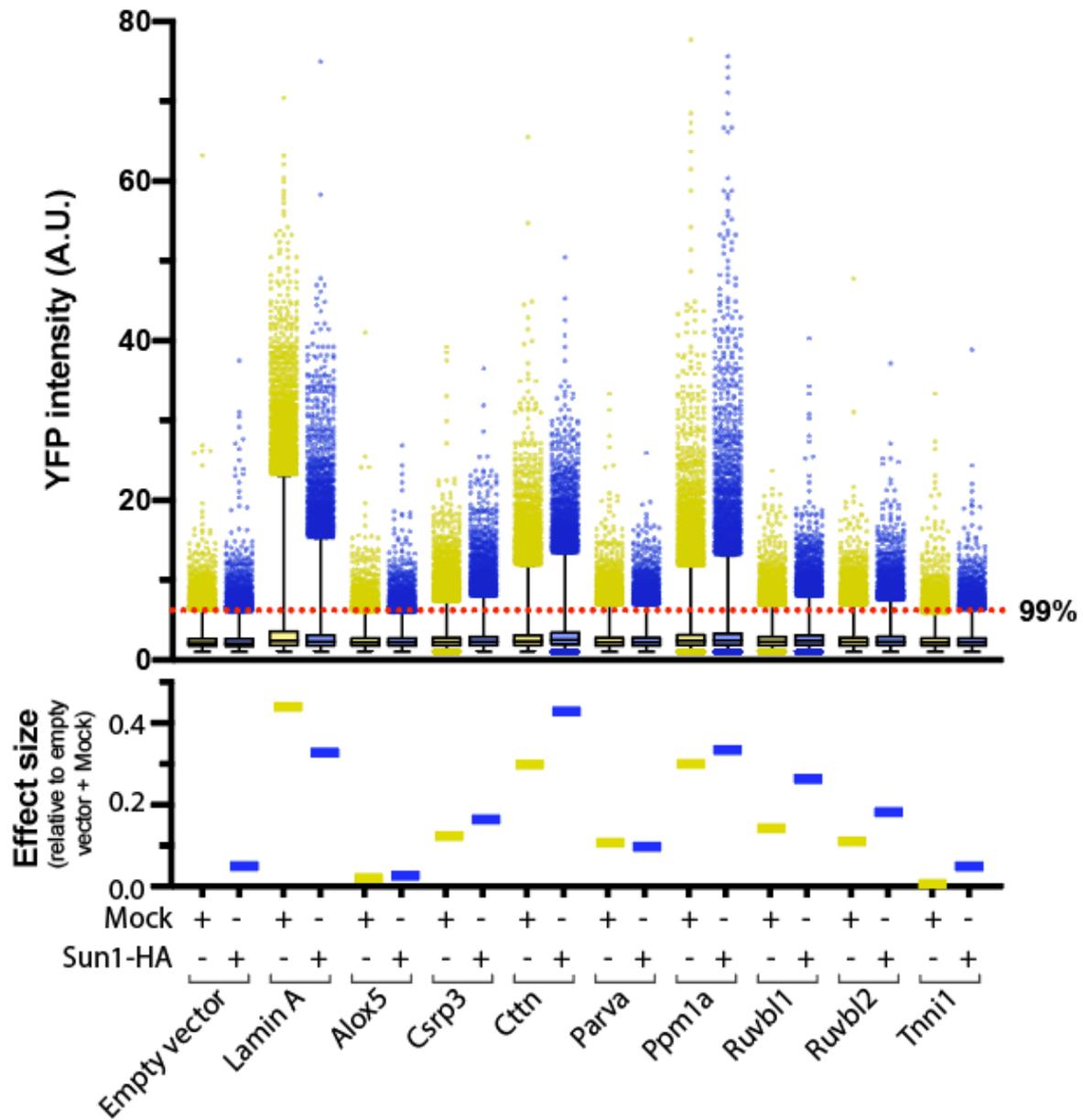
Supplemental Figure XIII – related to Figure 7-6. Sun1 BiFC competition assay. Data from figure 7-6 is shown here and represents BiFC intensity in HEK 293T cells which were co-transfected with Sun1-3(G_4S)-VC210 and the indicated VN210-tagged candidate interactor with an empty pcDNA3.1 plasmid (Mock). Datasets coloured blue replace the mock construct for pcDNA3.1-Sun1-HA in order to compete for binding of candidate interactor. Cells were grown for 48 hours before analysis by flow cytometry. Graph representative of three independent experiments comprising a total of at least 118,149 cells for each condition. BiFC intensity from reconstituted mVenus is shown in the upper graph. Broken red line signifies 99% percentile for intensities of samples in the empty vector control with mock plasmid. Lower graph shows graphical representation of effect size calculated in relation to empty vector control with mock to give visual illustration of differences between conditions.

Appendices

I.   Primary antibodies

Antibody name	Antigen	Species	Isotype	Clone	Dilution IF	Dilution WB	Dilution IP	Reference
Mouse Sun1	Sun1 Sun-domain	Mouse	IgG1	X12.11	20			In house
Sun1	Sun1 Sun-domain	Rabbit	Polyclonal		200		50	(Haque *et al.*, 2006)
Lamin B1		Mouse	IgG2b	C7	20	50		In house
Mandag2	βDG Ct	Mouse	IgG1	7D11	100	500	50	(Pereboev *et al.*, 2001)
1709	p892 βDG	Rabbit	Polyclonal		200	500	50	(Ilsley, Sudol and Winder, 2001)
LG5	βDG Ct	Rabbit	Polyclonal		200	500	50	D. Mornet Lab
JAF1	βDG Ct	Rabbit	Polyclonal		200	500	50	(FRANÇOIS Rivier *et al.*, 1999)
MF20	Myosin Heavy Chain	Mouse	IgG2b		20	200		(Bader, Masaki and Fischman, 1982)
F5D	Myogenin	Mouse	IgG1		20	200		(Wright, Binder and Funk, 1991)
V5	V5-tag	Mouse	IgG2a		1000	5000		Invitrogen R960-25
V5	V5-tag	Rabbit			200			Sigma V8237
HA	HA-tag	Rat	IgG1	3F10	500	1000	40	Sigma
HA	HA-tag	Rabbit	IgG	C29F4	800	800		CST
Myc	Myc-tag	Mouse	IgG2a	9B11	1000			CST, 2276
Myc	Myc-tag	Mouse	IgG1	9E10		1000		Millipore
BioID2	BioID2	Mouse	IgG1	SS 3A5-E2	10			In house
GAPDH	GAPDH	Rabbit	Polyclonal			2500		Abcam Ab9485
γ-tubulin	γ-tubulin	Mouse	IgG1	GTU-88				Sigma
α-tubulin	α-tubulin	Mouse	IgG1	DM1a		3000		Sigma
β-tubulin	β-tubulin	Mouse	IgG1	TUB2.1		1000		Sigma
Calnexin	Calnexin	Rabbit	Polyclonal		200	1000		Abcam Ab22595
Lamin A	Lamin A Ct	Rabbit	Polyclonal		200			Abcam Ab26300
Lamin A/C		Mouse	IgG2a	4C11	200			CST
Lamin B	Lamin B1 Ct	Goat	Polyclonal	M-20	50			SCBT (sc-6217)
FLAG	FLAG-tag	Mouse	IgG2b	M2	1000			Sigma
Emerin	Emerin aa100-200	Rabbit	Polyclonal			1000		Abcam Ab40688
Emerin		Rabbit			100			SCBT (H-12)
α-SMA	α-SMA Nt	Mouse	IgG2a	1A4		3000		Abcam Ab7817
Smad2	Smad2	Rabbit		EP567y	300			Abcam Ab33875

II. Secondary antibodies

Antibody	Conjugation	Use	Dilution IF	Dilution WB	Source
Anti-mouse	Alexa Fluor	IF	500		Invitrogen
Anti-mouse IgG1	Alexa Fluor	IF	500		Invitrogen
Anti-mouse IgG2a	Alexa Fluor	IF	500		Invitrogen
Anti-mouse IgG2b	Alexa Fluor	IF	500		Invitrogen
Anti-rabbit	Alexa Fluor	IF	500		Invitrogen
Anti-rat	Alexa Fluor	IF	500		Invitrogen
Anti-goat	Alexa Fluor	IF	500		Invitrogen
Streptavidin	Alexa Fluor 568	IF	500		Invitrogen
Anti-mouse HRP	HRP	WB		10,000	Dako
Anti-Rabbit HRP	HRP	WB		10,000	Dako
Anti-Rat HRP	HRP	WB		10,000	Dako
Anti-mouse	IRDye	WB – LiCor		10,000	LiCor Biosciences
Anti-rabbit	IRDye	WB – LiCor		10,000	LiCor Biosciences
Streptavidin	IRDye	WB – LiCor		10,000	LiCor Biosciences
Phalloidin	Alexa Fluor	IF	40		Invitrogen

Appendix B – Table of oligonucleotides

I. Genotyping primers

Code	Function	Sequence (5'-3')
M006A	Rosa26-mycFKBP–BioID2 (KI-LoxP)	AGTCGACTTCGAATAACTTCGTATAGCATAC
M006A2	Rosa26-mycFKBP–BioID2 (KI-SA)	TCTTTCTATCTGTAGGGCGCAGTAGTC
M006A3	Rosa26-mycFKBP–BioID2 (KI/WT)	TCTTCCCTCGTGATCTGCAACTC
M006B2	Rosa26-mycFKBP–BioID2 (KI-CreRec.)	ATCACCTCCTGCTTGCCTAGC
M006B3	Rosa26-mycFKBP–BioID2 (WT)	ATCTCAAGCAGGAGAGTATAAAACTCGG
M006B4	Rosa26-mycFKBP–BioID2 (WT)	TTGATAAGGCTGCAGAAGGAGCG
M006B5	Rosa26-mycFKBP–BioID2 (KI-CreRec.)	TAGAAGCTCCACATCGAAGACGAG
M006B6	Rosa26-mycFKBP–BioID2 (KI-Cassette)	TGTCTGTTGTGCCCAGTCATAGC
M013A3	Rosa26-mycFKBP–BioID2 (KI)	CAGAAAACTGGCCCTTGCCAT
M013B3	Rosa26-mycFKBP–BioID2 (KI)	CTGCGCCCTACAGATAGAAAGACT
M023A	FRB-Sun1 KI, upstream Exon 2 start	CTTCAAGAAATGGGAATTCACAGG
M023B	FRB-Sun1 KI, downstream Exon 2	ACAACCTGTCTGAAAACATAAGACTGC
349A	FRB forward	GAGATTTAATGGAGGCCCAAGA
349B	FRB reverse	CCGTTCCATCATAGCATGCA
Neo_For	Neomycin forward	
Neo_Rev	Neomycin reverse	
M066A	Amplify FRB-Sun1 knock-in for sequencing	GGAGTCACTGTTAAGTTTTTCTTAGAAAGTGC
M066B		GGTAAGAACAAAGATACCCACTTGTCCA
M067A		GAAATGCTGTTCCCAAGTATCACCT

II. Cloning primers

Name	Function	Sequence	Construct
M007A2	αDG fwd.	GGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGGTGGGTGACCCGGCACTGGTGTTG	αV5-HAβmyc-DG
M007B	βDG rev PacI NotI	TTGCGCAGCGGCCGCTGCGCAATTAATTAACCTAGGTTAAGGTGGGACA	
M007B2	αDG rev V5	CGTAGAATCGAGACCGAGGAGAGGGTTAGGGATAGGCTTACCAAACTTGGCCTTGAACCTTGCAG	
M008A	KpnI-HA-βDG	TGCGCAAGGTACCACCCTACCCATACGATGTTC	
M008A2	myc-βDG	GAACAAAAACTCATCTCAGAAGAGGATCTGATGGGAGAGTACACGCCCCTG	
M008B2	myc-βDG	CAGATCCTCTTCTGAGATGAGTTTTTGTTCGGTGTCCTGGTTCAGAGGAGTG	
M009A	βDG fwd.	TGCGCAAGGTACCACCCAGGAGAGTGC	Δα-HAβmyc-DG
M012A	βDG(cyto) fwd.	TGATCGAGGCGCGCCTTGCGCAATGCGCAAGAAGCGGAAGGGC	βmyc-DG
M046A	VC210 fwd.	TGCGCAACATATGCCAAGTACGCCCCCT	pGBS-VC210-x(G₄S)-Sun1
M046A2	Sun1 1(G₄S) for	GGTGGAGGCGGTTCAGCCACCATGGACTTTTCTCGGCT	
M046A3	Sun1 2(G₄S) for	GGTGGAGGCGGTTCAGGTGGAGGCGGTTCAGCCACCATGGACTTTTCTCGGCT	
M046A4	Sun1 3(G₄S) for	GGTGGAGGCGGTTCAGGTGGAGGCGGTTCAGGTGGAGGCGGTTCAGCCACCATGGACTTTTCTCGGCT	
M046B	Sun1 rev.	TTGCGCAAAGCTTCTCTAATAGCCACTCGAGGGA	
M047B2	Sun1 1(G₄S) rev	TGAACCGCCTCCACCCAGCGTAATCTGGAACATCGTATGGGT	
M047B3	Sun1 2(G₄S) rev	TGAACCGCCTCCACCTGAACCGCCTCCACCCAGCGTAATCTGGAACATCGTATGGGT	
M047B4	Sun1 3(G₄S) rev	TGAACCGCCTCCACCTGAACCGCCTCCACCTGAACCGCCTCCACCCAGCGTAATCTGGAACATCGTATGGGT	
M048A	AscI HA-Sun1 for	TGCGCAAGGCGCGCCGCCACCATGTACCCATACGATGTTCCAGATTACGCTGACTTTTCTCGGCTGCACACGT	pcDNA3.1-HA-Sun1

M048B	Sun1 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTGGATGGGCTCTC CGTGGA	
M048A2	AscI NdeI Sun1 for	TGCGCAAGGCGCGCCACATATGACCATGGACTTTTCTCG GCTGCA	pcDNA3.1-Sun1-HA
M048B2	Sun1-HA PacI rev	TTGCGCAGCGGCCGCCCTAGGTTAAGCGTAATCTGGAA CATCGTATGGGTACTGGATGGGCTCTCCGTGGA	
M049A	AscI FLAG-Alox5 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGA CGACAAGCCCTCCTACACGGTCACCGT	pcDNA3.1-FLAG-Alox5
M049B	Alox5 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTAGATGGCTACGCTGT TGGGAAT	
M049A2	AscI NdeI Alox5 for	TGCGCAAGGCGCGCCACATATGACCATGCCCTCCTACAC GGT	pcDNA3.1-Alox5-FLAG
M049B2	Alox5-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTC TTTGTAGTCGATGGCTACGCTGTTGGGAAT	
M050A	AscI FLAG-Arhgap22 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGA CGACAAGCTGCCGACGGCTTCCAGCA	pcDNA3.1-FLAG-Arhgap22
M050B	Arhgap22 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTCCGGGGCTCTGG ACCCCT	
M050A2	AscI NdeI Arhgap22 for	TGCGCAAGGCGCGCCACATATGACCATGCTGCCGACGG CTT	pcDNA3.1-Arhgap22 -FLAG
M050B2	Ccdc86 -FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTC TTTGTAGTCCTCCGGGGCTCTGGACCCCT	
M051A	AscI FLAG-Ccdc86 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGA CGACAAGGATACGCCGCTGAGGCGCA	pcDNA3.1-FLAG-Ccdc86
M051B	Ccdc86 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTAGACCTTGGCTACAG GCCTCT	
M051A2	AscI NdeI Ccdc86 for	TGCGCAAGGCGCGCCACATATGACCATGGATACGCCGC TGA	pcDNA3.1-Ccdc86-FLAG
M051B2	Ccdc86 -FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTC TTTGTAGTCGACCTTGGCTACAGGCCTCT	
M052A	AscI FLAG-Dctn5 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGA CGACAAGGAGTTGGGCGAACTGCTGTACA	pcDNA3.1-FLAG-Dctn5
M052B	Dctn5 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTAGACTTGTGTTAGGG GTAGAAACTTCTGGT	
M052A2	AscI NdeI Dctn5 for	TGCGCAAGGCGCGCCACATATGACCATGGAGTTGGGCG AACTGCT	pcDNA3.1-Dctn5-FLAG
M052B2	Parva-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTC TTTGTAGTCGACTTGTGTTAGGGGTAGAAACTTCTGGT	
M053A	AscI FLAG-Parva for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGA CGACAAGGCCACATCCCCACAGAAGTCG	pcDNA3.1-FLAG-Parva
M053B	Parva NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTCCACATTCCGGTA CTTGGTGA	
M053A2	AscI NdeI Parva for	TGCGCAAGGCGCGCCACATATGACCATGGCCACATCCC CACAGA	pcDNA3.1-Parva-FLAG
M053B2	Parva-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTC TTTGTAGTCCTCCACATTCCGGTACTTGGTGA	
M054A	AscI FLAG-Ruvbl2 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGA CGACAAGGCAACCGTGGCAGCCACCACCA	pcDNA3.1-FLAG-Ruvbl2
M054B	Ruvbl2 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTAGGAGGTGTCCATTG TTTCGCCTT	
M054A2	AscI NdeI Ruvbl2 for	TGCGCAAGGCGCGCCACATATGACCATGGCAACCGTGG CAGCCA	pcDNA3.1-Ruvbl2-FLAG
M054B2	Ruvbl2-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTC TTTGTAGTCGGAGGTGTCCATTGTTTCGCCTT	
M055A	AscI FLAG-Tnni1 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGA CGACAAGCCGGAAGTTGAGAGGAAATCCAAGAT	pcDNA3.1-FLAG-Tnni1
M055B	Tnni1 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTGGGAGGTCGGGG ACTTAGCA	
M055A2	AscI NdeI Tnni1 for	TGCGCAAGGCGCGCCACATATGACCATGCCGGAAGTTG AGAGGA	pcDNA3.1-Tnni1-FLAG
M055B2	Tnni1-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTC TTTGTAGTCCTGGGAGGTCGGGGACTTAGCA	

M056A	AscI FLAG-Tnnt3 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGACGACAAGTCTGACGAGGAAACTGAACAAGTTGA	pcDNA3.1-FLAG-Tnnt3
M056B	Tnnt3 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTCCAGCGCCCACCGACTT	
M056A2	AscI NdeI Tnnt3 for	TGCGCAAGGCGCGCCACATATGACCATGTCTGACGAGGAAACTGAA	pcDNA3.1-Tnnt3-FLAG
M056B2	Tnnt3-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTCTTTGTAGTCCTTCCAGCGCCCACCGACTT	
M057A	AscI FLAG-Zbtb20 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGACGACAAGCTAGAACGGAAGAAACCCAAGACAGCT	pcDNA3.1-FLAG-Zbtb20
M057B	Zbtb20 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTATCCGTCAGACACATGCATCCT	
M057A2	AscI NdeI Zbtb20 for	TGCGCAAGGCGCGCCACATATGACCATGCTAGAACGGAAGAAACCCA	pcDNA3.1-Zbtb20-FLAG
M057B2	Fam91a1-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTCTTTGTAGTCTCCGTCAGACACATGCATCCT	
M058A	AscI FLAG-Fam91a1 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGACGACAAGAACATCGACGTTGAGTTCCACAT	pcDNA3.1-FLAG-Fam91a1
M058B	Fam91a1 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACAAATGGAGACTAGCGATGAGCAGG	
M058A2	AscI NdeI Fam91a1 for	TGCGCAAGGCGCGCCACATATGACCATGAACATCGACGTTGAGTT	pcDNA3.1-Fam91a1-FLAG
M058B2	Fam91a1-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTCTTTGTAGTCCAAATGGAGACTAGCGATGAGCAGG	
M059A	AscI FLAG-Hist1h2bp for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGACGACAAGCCTGAGCCTGTTAAGTCCGTTCC	pcDNA3.1-FLAG-Hist1h2bp
M059B	Hist1h2bp NotI rev	TTGCGCAGCGGCCGCCCTAGGTTAAAAGGAAGGAAGATAATAGAATTTGTTCCACAGGA	
M059A2	AscI NdeI Hist1h2bp for	TGCGCAAGGCGCGCCACATATGACCATGCCTGAGCCTGTTAAGT	pcDNA3.1-Hist1h2bp -FLAG
M059B2	Hist1h2bp-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTCTTTGTAGTCAAAGGAAGGAAGATAATAGAATTTGTTCCACAGGA	
M060A	AscI FLAG-Ppm1a for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGACGACAAGGGAGCATTTTTAGACAAGCCAAAGATGGA	pcDNA3.1-FLAG-Ppm1a
M060B	Ppm1a NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACCACATATCATCGGTTGACGCAGA	
M060A2	AscI NdeI Ppm1a for	TGCGCAAGGCGCGCCACATATGACCATGGGAGCATTTTTAGACAAGCCA	pcDNA3.1-Ppm1a-FLAG
M060B2	Ppm1a-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTCTTTGTAGTCCCACATATCATCGGTTGACGCAGA	
M061A	AscI FLAG-Csrp3 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGACGACAAGCCAAACTGGGGTGGAGGTGCA	pcDNA3.1-FLAG-Csrp3
M061B	Csrp3 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTCCTTCTTTTCCACTTGCTGTGTAA	
M061A2	AscI NdeI Csrp3 for	TGCGCAAGGCGCGCCACATATGACCATGCCAAACTGGGGTGGA	pcDNA3.1-Csrp3-FLAG
M061B2	Csrp3-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTCTTTGTAGTCCTCCTTCTTTTCCACTTGCTGTGTAA	
M062A	AscI FLAG-Ruvbl1 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGACGACAAGAAGATTGAGGAGGTGAAGAGCACCA	pcDNA3.1-FLAG-Ruvbl1
M062B	Ruvbl1 NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTCATGTACTTGTCCTGCTGGTCA	
M062A2	AscI NdeI Ruvbl1 for	TGCGCAAGGCGCGCCACATATGACCATGAAGATTGAGGAGGTGAAGAGCA	pcDNA3.1-Ruvbl1-FLAG
M062B2	Ruvbl1-FLAG NotI rev	TTGCGCAGCGGCCGCCCTAGGTTACTTGTCGTCATCGTCTTTGTAGTCCTTCATGTACTTGTCCTGCTGGTCA	
M063A	AscI FLAG-Vps51 for	TGCGCAAGGCGCGCCACCATGGACTACAAAGACGATGACGACAAGGCGGCTGCGGCAGCTGTGG	pcDNA3.1-FLAG-Vps51
M063B	Vps51 NotI rev	TTGCGCATTAATTAACCTAGGTTAGCCGCGCTCGCAGATGACCT	

M063A2	AscI NdeI Vps51 for	TGCGCAAGGCGCGCCACATATGACCATGGCGGCTGCGGCAGCT	pcDNA3.1-Vps51-FLAG
M063B2	Vps51-FLAG NotI rev	TTGCGCATTAATTAACCTAGGTTACTTGTCGTCATCGTCTTTGTAGTCGCCGCGCTCGCAGATGACCT	
M024A	LNKa-Alox5	AAACGCTCTATGGTCTAAAGTTTGCCACCATGCCCTCCTACACGGTCACCGT	pGBS-VN210-Alox5 / pGBS-Alox5-VN210
M024B	Alox5-LNKb	AAACCCCGATTGAGATATAGTTTGATGGCTACGCTGTTGGGAATCCTGTCTGGTGACAGGT	
M025A2	LNKa-Arhgap22	AAACGCTCTATGGTCTAAAGTTTGCCACCATGCTGCCGACGGCTTCCAGCAAAAGAAGAACCTTCGCT	pGBS-VN210-Arhgap22/ pGBS-Arhgap22-VN210
M025B2	Arhgap22-LNKb	AAACCCCGATTGAGATATAGTTTCTCCGGGGCTCTGGACCCCTTGGTCCCTGTAGTCAA	
M026A2	LNKa-Ccdc86	AAACGCTCTATGGTCTAAAGTTTGCCACCATGGATACGCCGCTGAGGCGCAGCCGGCGGCTGGAA	pGBS-VN210-Ccdc86/ pGBS-Ccdc86-VN210
M026B2	Ccdc86-LNKb	AAACCCCGATTGAGATATAGTTTGACCTTGGCTACAGGCCTCTGGGGGTGGCTGTTTCTGGA	
M027A	LNKa-Dctn5	AAACGCTCTATGGTCTAAAGTTTGCCACCATGGAGTTGGGCGAACTGCTGTACAA	pGBS-VN210-Dctn5/ pGBS-Dctn5-VN210
M027B	Dctn5-LNKb	AAACCCCGATTGAGATATAGTTTGACTTGTGTTAGGGGTAGAAACTTCTGGT	
M029A	LNKa-Parva	AAACGCTCTATGGTCTAAAGTTTGCCACCATGGCCACATCCCCACAGAAGTCGCCCTT	pGBS-VN210-Parva / pGBS-Parva -VN210
M029B	Parva-LNKb	AAACCCCGATTGAGATATAGTTTCTCCACATTCCGGTACTTGGTGAAGAGGTT	
M030A	LNKa-Ruvbl2	AAACGCTCTATGGTCTAAAGTTTGCCACCATGGCAACCGTGGCAGCCACCACCAA	pGBS-VN210-Ruvbl2/ pGBS-Ruvbl2-VN210
M030B	Ruvbl2-LNKb	AAACCCCGATTGAGATATAGTTTGGAGGTGTCCATTGTTTCGCCTTTGAGCT	
M031A	LNKa-Tnni1	AAACGCTCTATGGTCTAAAGTTTGCCACCATGCCGGAAGTTGAGAGGAAATCCA	pGBS-VN210-Tnni1/ pGBS-Tnni1-VN210
M031B	Tnni1-LNKb	AAACCCCGATTGAGATATAGTTTCTGGGAGGTCGGGGACTTAGCAGCGT	
M032A	LNKa-Tnnt3	AAACGCTCTATGGTCTAAAGTTTGCCACCATGTCTGACGAGGAAACTGAACAAGT	pGBS-VN210-Tnnt3/ pGBS-Tnnt3-VN210
M032B	Tnnt3-LNKb	AAACCCCGATTGAGATATAGTTTCTTCCAGCGCCCACCGACTTTGCCCTT	
M033A	LNKa-Vps51	AAACGCTCTATGGTCTAAAGTTTGCCACCATGGCGGCTGCGGCAGCTGTGGGGCCT	pGBS-VN210-Vps51/ pGBS-Vps51-VN210
M033B	Vps51-LNKb	AAACCCCGATTGAGATATAGTTTGCCGCGCTCGCAGATGACCTCGACAA	
M034A2	LNKa-Zbtb20	AAACGCTCTATGGTCTAAAGTTTGCCACCATGCTAGAACGGAAGAAACCCAAGACAGCTGAAAACCA	pGBS-VN210-Zbtb20/ pGBS-Zbtb20-VN210
M034B2	Zbtb20-LNKb	AAACCCCGATTGAGATATAGTTTTCCGTCAGACACATGCATCCTCATGTGGTCGTTGAACTGCT	
M035A	LNKa-Sun1	AAACGCTCTATGGTCTAAAGTTTGCCACCATGGACTTTTCTCGGCTGCACACGTA	pGBS-VN210-Sun1/ pGBS-Sun1-VN210
M035B	Sun1-LNKb	AAACCCCGATTGAGATATAGTTTCTGGATGGGCTCTCCGTGGACT	
M037A	LNKa-Fam91a1	AAACGCTCTATGGTCTAAAGTTTGCCACCATGAACATCGACGTTGAGTTCCACAT	pGBS-VN210-Fam91a1/ pGBS- Fam91a1-VN210
M037B	Fam91a1-LNKb	AAACCCCGATTGAGATATAGTTTCAAATGGAGACTAGCGATGAGCAGGGA	
M038A	LNKa-Hist1h2bp	AAACGCTCTATGGTCTAAAGTTTGCCACCATGCCTGAGCCTGTTAAGTCCGTT	pGBS-VN210-Hist1h2bp / pGBS-Hist1h2bp -VN210
M038B	Hist1h2bp -LNKb	AAACCCCGATTGAGATATAGTTTAAAGGAAGGAAGATAATAGAATTTGTTCCA	
M039A	LNKa-Ppm1a	AAACGCTCTATGGTCTAAAGTTTGCCACCATGGGAGCATTTTTAGACAAGCCAAAGA	pGBS-VN210-Ppm1a / pGBS-Ppm1a -VN210
M039B	Ppm1a-LNKb	AAACCCCGATTGAGATATAGTTTCCACATATCATCGGTTGACGCAGAATCA	

M041A	LNKa-Cttn	AAACGCTCTATGGTCTAAAGTTTGCCACCATGTGGAAAG CCTCTGCAGGCCATGCT	pGBS-VN210-Cttn / pGBS-Cttn -VN210
M041B	Cttn-LNKb	AAACCCCGATTGAGATATAGTTTCTGCCGCAGCTCCACA TAGTTGGCT	
M042A	LNKa-Csrp3	AAACGCTCTATGGTCTAAAGTTTGCCACCATGCCAAACT GGGGTGGAGGTGCAA	pGBS-VN210-Csrp3/ pGBS-Csrp3-VN210
M042B	Csrp3-LNKb	AAACCCCGATTGAGATATAGTTTCTCCTTCTTTTCC ACTTGCTGTGTAA	
M043A	LNKa-Ruvbl1	AAACGCTCTATGGTCTAAAGTTTGCCACCATGAAG ATTGAGGAGGTGAAGAGCACCA	pGBS-VN210-Ruvbl1/ pGBS-Ruvbl1-VN210
M043B	Ruvbl1-LNKb	AAACCCCGATTGAGATATAGTTTCTTCATGTACTTG TCCTGCTGGTCA	
M068A	Sun1 (IVT)	GTGGTTTTCCTTTGAAAAACACGATGATAATATGG ACTTTTCTCGGCTGCACA	T7CFE-CHA-Sun1
M068B	Sun1 (IVT)	TTGCGCAAGATCTCAGGCGTAGTCGGGCACGTCG TAGGGGTACTCGAGTGCGGCCGCCTGGATGGGCT CTCCGTGG	
M068A2	Sun1-HA	CTCTGGAAGCTTCTTGAAGACAAACAA	
M068B2	Sun1-HA	TGTGCAGCCGAGAAAAGTCCATATTATCATCGTGT TTTTCAAAGGAAAACCAC	
M069A	Sun1-HA 1-355 (IVT)	AAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACA AC	T7CFE-CHA-Sun1 (1-355)
M071B	Sun1-HA 1-355 (IVT)	GTAGGGGTACTCGAGTGCGGCCGCTCTAGTCCTTC GCAGTGCTTG	

338

III. Sequencing primers

Name	Function	Sequence (5'-3')
M002A	DAG1 gene sequence	CTATGTGGAGCCTACTGCAGTTGCT
M002A2	DAG1 gene sequence	GAACCAGAACAGTGTGCCTGACATT
M002A3	DAG1 gene sequence	GCCGGTGGTGAATAACAGACTATTT
M002A4	DAG1 gene sequence	CCCAGAGCTCAAGAACCATATTGA
M002A5	DAG1 gene sequence	CGAGTATTTCATGCATGCCACAGA
M002B	DAG1 gene sequence	GGAATGACTGTGTGCAGGTAGACA
M002B2	DAG1 gene sequence	GATAGGCACCCCCTTCTTGATGAA
M013A2	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CACAAATGGCGTGTTTTGGTTGG
M013A3	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CAGAAAACTGGCCCTTGCCAT
M013A4	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CCAAAGTCGCTCTGAGTTGTTATCAGTA
M013A5	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CTACTCCTCCCCTAGTCAGGAAGTT
M013A6	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	GCTCCTGCCGAGAAAGTATCCAT
M013A7	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	GATTCCACCGCCGCCTTCTAT
M013A8	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CAGAGAGCCAAACTGACTATATCTCCA
M013A9	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	GATCGAGAGCAAGATGCTGTACCT
M013B	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	GAGCCAATCAGACGACGAGGC
M013B2	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	GAGGCAACTCAAGTCGGAAACGT
M013B3	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CTGCGCCCTACAGATAGAAAGACT
M013B4	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CCATTTGTCACGTCCTGCACGA
M013B5	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	GTAGCCGGATCAAGCGTATGCA
M013B6	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	GAACAAACGACCCAACACCGT
M013B7	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CTGTCCACCTCCTTCAGCCA
M013B8	Rosa26-mycFKBP-BioID2$^{(flx)}$ seq	CGAGGATCCGAGCTCGAATTCGA
M015A	FRB-Sun1 targeting vector seq.	TAGGACAGCAGAGCTACACAGAG
M015A2	FRB-Sun1 targeting vector seq.	TCTCTTCTGCAGCATGGTCTGAG
M015A3	FRB-Sun1 targeting vector seq.	TGTCTCCTTCCGTGTTTCAGTTAGC
M015A4	FRB-Sun1 targeting vector seq.	ATTGTCTGTTGTGCCCAGTC
M015A5	FRB-Sun1 targeting vector seq.	TAGCTTTAATAGTTTGAGCATCCACCAGAG
M015A6	FRB-Sun1 targeting vector seq.	TCTCTGAGCCTGTCTGGGAAC
M015A7	FRB-Sun1 targeting vector seq.	AGATAGGAGGAGAGGGGTGTGTATG
M015B	FRB-Sun1 targeting vector seq.	TCCACACACATGAGCAGCAG
349A	FRB forward	GAGATTTAATGGAGGCCCAAGA
349B	FRB reverse	CCGTTCCATCATAGCATGCA
143A	Em7 promoter sequence	GGCCCTCCTATAGTGAGTCGTATTATA
129B	Ubiquitin C promoter reverse seq	CGCTCTCTGGAAAGAAAACCA
70A	CMV promoter forward sequence	CATCAATGGGCGTGGATAGC
70B	CMV promoter reverse sequence	GCTTATATAGACCTCCACCGTACA
185A	IRES NB forward sequence	TGCAAAGGCGGCACAAC
185B2	IRES NB reverse sequence	CCCCTAGGAATGCTCGTCAA
406A	BiFC Linker A sequence (forward)	AAACGCTCTATGGTCTAAAGT
406B	BiFC Linker B Sequence (reverse)	AAACCCCGATTGAGATATAGT

IV. Guide RNA sequences for CRISPR/Cas9

Targeted gene	Exon	Name	Sequence (5'-3')
DAG1	1	DAG1_gRNA-001	AGTGCTCTCAGACCTCCACG
DAG1	1	DAG1_gRNA-002	TACTTAGCAAGACTATCGAC
Sun1	2	Sun1_gRNA-001	CTGGCTACACTTACGCACTC
Sun1	2	Sun1_gRNA-002	CCTGTTAACCCTCTAACACC
Sun1	2	Sun1_gRNA-003	CTCCGGCACACACTGGGGTG
Sun1	2	Sun1_gRNA-004	TTCTGCAGCATGGTCTGAGA
Sun1	2	Sun1_gRNA-005	CTTCTGTCTCTTCTGCAGCA

I. CRISPR plasmids

Construct Name	Source
pX458	Addgene (Ran, Hsu, Wright, *et al.*, 2013)
pX458_DAG1-gRNA_001	This study
pX458_DAG1-gRNA_002	This study
pX458_Sun1-gRNA_001	This study
pX458_Sun1-gRNA_002	This study
pX458_Sun1-gRNA_003	This study
pX458_Sun1-gRNA_004	This study
pX458_Sun1-gRNA_005	This study

II. Mammalian expression constructs

Construct Name	Source
pcDNA3.1-αV5-β-DG	This study
pcDNA3.1-α-HAβ-DG	This study
pcDNA3.1-α- βmyc-DG	This study
pcDNA3.1-αV5-HAβmyc-DG	This study
pcDNA3.1-α-HAβmyc-DG	This study
TRIBLA-αV5-HAβmyc-DG	This study
TRIBLA-α-HAβmyc-DG	This study
TRIBLA-$\Delta\alpha$-HAβmyc-DG	This study
TRIBLA-SP-$\Delta\alpha$-HAβmyc-DG	This study
TRIBLA-βmyc-DG(cyto)	This study
pCIneo-myc-mSun1 (FL)	(Haque *et al.*, 2006)
pCIneo-myc-mSun1 (1-355)	(Haque *et al.*, 2006)
pCMV-myc-mSUN1 (1-432)	(Haque *et al.*, 2010)
pCMV-myc-mSUN1 (355-913)	(Haque *et al.*, 2010)
TRIPZ-V5FRB-Sun1	Alexandre Chojnowski
TRIPZ-vCre	Alexandre Chojnowski
TRIBLA-vCre	This study
TRIPZ-V5FRB-LAP2β	(Chojnowski *et al.*, 2018)

III. Epitope tagged Sun1 candidate interactors

Construct Name	Source
pcDNA3.1-HA-Sun1	This study
pcDNA3.1-Sun1-HA	This study
pcDNA3.1 - FLAG-Alox5	This study
pcDNA3.1 - Alox5-FLAG	This study
pcDNA3.1 - FLAG-Arhgap22	This study
pcDNA3.1 - Arhgap22-FLAG	This study
pcDNA3.1 - FLAG-Ccdc86	This study
pcDNA3.1 - Ccdc86-FLAG	This study
pcDNA3.1 - FLAG-Csrp3	This study
pcDNA3.1 - Csrp3-FLAG	This study
pcDNA3.1 - FLAG-Dctn5	This study
pcDNA3.1 - Dctn5-FLAG	This study
pcDNA3.1 - FLAG-Fam91a1	This study
pcDNA3.1 - Fam91a1-FLAG	This study
pcDNA3.1 - FLAG-Hist1h2bp	This study
pcDNA3.1 - Hist1h2bp-FLAG	This study
pcDNA3.1 - FLAG-Ppm1a	This study
pcDNA3.1 - Ppm1a-FLAG	This study
pcDNA3.1 - FLAG-Ruvbl1	This study
pcDNA3.1 - Ruvbl1-FLAG	This study
pcDNA3.1 - FLAG-Ruvbl2	This study
pcDNA3.1 - Ruvbl2-FLAG	This study
pcDNA3.1 - FLAG-Tnni1	This study
pcDNA3.1 - Tnni1-FLAG	This study
pcDNA3.1 - FLAG-Tnnt3	This study
pcDNA3.1 - Tnnt3-FLAG	This study
pcDNA3.1 - FLAG-Vps51	This study
pcDNA3.1 - Vps51-FLAG	This study
pcDNA3.1 - FLAG-Parva	This study
pcDNA3.1 - Parva-FLAG	This study
pcDNA3.1 - FLAG-Zbtb20	This study
pcDNA3.1 - Zbtb20-FLAG	This study
pcDNA3.1 - FLAG-Cttn	This study
pcDNA3.1 - Cttn-FLAG	This study

IV. Constructs for bimolecular fluorescence complementation

Construct Name	Source
Sun1 - pGBS - Ct-VC210	This study
Sun1 - pGBS - VC210-Nt	This study
Sun1 1GS_Linker VC210-Nt	This study
Sun1 2GS_Linker VC210-Nt	This study
Sun1 3GS_Linker VC210-Nt	This study
Alox5 - pGBS - Ct-VN210	This study
Alox5 - pGBS - VN210-Nt	This study
Arhgap22 - pGBS - Ct-VN210	This study
Arhgap22 - pGBS - VN210-Nt	This study
Ccdc86 - pGBS - Ct-VN210	This study
Ccdc86 - pGBS - VN210-Nt	This study
Dctn5 - pGBS - Ct-VN210	This study
Dctn5 - pGBS - VN210-Nt	This study
Parva - pGBS - Ct-VN210	This study
Parva - pGBS - VN210-Nt	This study
Ruvbl2 - pGBS - Ct-VN210	This study
Ruvbl2 - pGBS - VN210-Nt	This study
Tnni1 - pGBS - Ct-VN210	This study
Tnni1 - pGBS - VN210-Nt	This study
Tnnt3 - pGBS - Ct-VN210	This study
Tnnt3 - pGBS - VN210-Nt	This study
Zbtb20 - pGBS - Ct-VN210	This study
Zbtb20 - pGBS - VN210-Nt	This study
Fam91a1 - pGBS - Ct-VN210	This study
Fam91a1 - pGBS - VN210-Nt	This study
Hist1h2bp - pGBS - Ct-VN210	This study
Hist1h2bp - pGBS - VN210-Nt	This study
Ppm1a - pGBS - Ct-VN210	This study
Ppm1a - pGBS - VN210-Nt	This study
Cttn - pGBS - Ct-VN210	This study
Cttn - pGBS - VN210-Nt	This study
Csrp3 - pGBS - Ct-VN210	This study
Csrp3 - pGBS - VN210-Nt	This study
Ruvbl1 - pGBS - Ct-VN210	This study
Ruvbl1 - pGBS - VN210-Nt	This study
Vps51 - pGBS - Ct-VN210	This study
Vps51 - pGBS - VN210-Nt	This study

V. IVT plasmids

Construct Name	Source
T7CFE-CHA-Sun1 (1-355)	This study
T7CFE-NHA-Ppm1a	This study
T7CFE-NHA-Lamin A	Hendrikje Werner, CS Lab

VI. Targeting plasmids

Construct Name	Source
Rosa26-SA-PGKneo-FKBP-BioID2-PA #10/2	Monash Genome Modification Platform
pDTA-FRB-SUN1-LneoL #6/2	Monash Genome Modification Platform

VII.   Recombinant expression plasmids

Construct Name	Source
pGEX-4T-mSUN1 (1-208)	(Haque, 2011)
pGEX-4T-mSUN2 (1-224)	(Haque, 2011)
pGEX-4T-mSUN2 (224-729)	(Haque, 2011)
pGEX-2TK-bDGcyt	Winder
pSJW1-bDGcyt	Winder

Appendix D – Recipes for solutions

I. Common buffers

<u>PBS</u>

1x Dulbecco's phosphate buffered saline without calcium, magnesium, phenol red (v/v) (Hyclone)

(137mM NaCl

2.7mM KCl

10mM Na_2HPO_4

1.8mM KH_2PO_4

pH 7.4)

<u>Tris-Acetate-EDTA (TAE)</u>

50x TAE buffer diluted to 1x (v/v) (1st Base)

(40mM Tris-Acetate

1mM EDTA)

<u>5x DNA gel loading buffer</u>

39% glycerol (w/v)

0.5% SDS (w/v)

10mM EDTA

0.25% bromophenol blue (w/v)

<u>HotSHOT alkaline lysis reagent</u>

25mM NaOH

0.2mM EDTA

pH 12

<u>HotSHOT neutralisation buffer</u>

40mM Tris-HCl

pH 5

II. Protein methods, SDS-PAGE gels and western blotting

<u>25x Protease inhibitor cocktail</u>

    1 protease inhibitor cocktail tablet (EDTA-free) (Roche) in 2ml dH$_2$O

<u>Lysis (RIPA) buffer</u>

    50mM Tris-HCl pH 7.5

    150mM NaCl

    1mM EGTA

    1mM EDTA

    1% Triton X-100 (v/v)

    0.5% Sodium deoxycholate (w/v)

    0.1% Sodium dodecylsulphate (w/v)

    1mM Sodium azide

    1x Protease Inhibitor (Roche), added fresh before use

<u>Hypotonic lysis buffer</u>

    5mM PIPES pH 8.0

    85mM KCl

    1% NP-40, added fresh before use (v/v)

    1x Protease Inhibitor (Roche), added fresh before use

<u>Resolving gel</u>

Gel percentage	7.5%	10.0%	12.5%	15%
dH$_2$O (ml)	3.55	2.75	1.90	1.10
Tris pH 8.8 (1M) (ml)	3.75	3.75	3.75	3.75
30% Acryl-bisacrylamide (ml)	2.50	3.35	4.15	5.00
10% SDS (w/v) (µl)	100	100	100	100
10% APS (w/v) (µl)	100	100	100	100
TEMED (µl)	5	5	5	5

5% Stacking gel

    3.48ml dH$_2$O

    0.63ml Tris pH 6.8 (1M stock)

    0.8ml 30% Acryl-bisacrylamide (Bio-Rad cat# 1610156)

    50μl 10% SDS (w/v)

    50μl 10% APS (w/v)

    15μl TEMED

Tris-Glycine Running buffer

    25mM Tris pH 8.3

    192mM Glycine

    0.1% SDS (w/v)

Coomassie stain solution

    0.125% Coomassie Blue (w/v)

    50% methanol (v/v)

    10% glacial acetic acid (v/v)

De-stain solution

    50% methanol (v/v)

    10% glacial acetic acid (v/v)

Transfer buffer

    25mM Tris pH 8.3

    192mM Glycine

    20% Methanol (v/v)

Ponceau stain

    (Sigma-Aldrich Cat # P7170-1L)

    0.1% (w/v) in 5% acetic acid

Blocking buffers

    5% skimmed dried milk powder in TBST or;

    Odyssey PBS blocking buffer (LI-COR Biosciences Cat#927-40000)

Tris buffered saline with Tween-20 (TBST)

    20mM Tris pH 7.4

    150mM NaCl

    0.1% Tween-20 (v/v)

2x immunoprecipitation buffer (IP) buffer

    20mM Tris pH 7.4

    200mM NaCl

    2mM Dtt

    0.2% Triton X-100 (v/v)

    1x Protease inhibitor cocktail (added fresh before use)

    For binding buffer, add 1% BSA (w/v)

III.   Immunofluorescence microscopy

4% Formaldehyde

    4% Formaldehyde (16% w/v Thermo Scientific Cat# 28908)

    1x PBS

Blocking buffer

    5% FBS (v/v)

    3% BSA (w/v)

    1x PBS

Mounting medium

    72% Glycerol (v/v) (80% stock v/v)

    1% DABCO (v/v) (10% DABCO in 10x PBS w/v stock)

IV.   Bacterial culture

<u>Transformation and storage buffer</u>

70mM $CaCl_2$

15% glycerol (v/v)

<u>2xYT medium</u>

1% Yeast extract (w/v)

1.6% Tryptone (w/v)

5g/L NaCl

(For plates, 16g/L agar)

<u>LB (Miller) medium</u>

1% Yeast extract (w/v)

1% Tryptone (w/v)

5g/L NaCl

(For plates, 15g/L agar)

V.   Solutions for cell culture

V.a. Media formulations

<u>ZLZ-M</u>

DMEM with high glucose (GE Healthcare Life Sciences Hyclone Cat# SH30022.01)

10% Heat-inactivated fetal bovine serum (v/v) (FBS –Life Technologies Cat# 10437028)

<u>Human myoblast growth medium</u>

Promocell C-23060

<u>Human myoblast differentiation medium</u>

Promocell C23061

<u>Ham's 20% primary myoblast growth medium</u>

Ham's F10 medium (Pan-Biotech Cat# P04-12500)

20% Heat-inactivated fetal bovine serum (v/v)

Ham's 2% primary myoblast differentiation medium

Ham's F10 medium (Pan-Biotech Cat# P04-12500)

2% horse serum (v/v)

FACS buffer

10% FBS (v/v)

0.1% Sodium azide (w/v)

In 1xPBS

V.b. Cell isolation solutions

V.*b.i, Fibroblast isolation*

Anti-microbial solution

1x Anti-bacterial/Anti-mycotic (v/v) (100x; Thermo Fisher Scientific Cat# 15240062)

Collagenase solution

1% collagenase II (w/v) (Gibco Cat# 17101-015) in 1x HBSS and 25mM HEPES

Dispase solution

1% Dispase II (w/v) (Roche Cat# 925825) in 1x HBSS and 25mM HEPES

Dissociation solution

40% collagenase solution (v/v)

40% dispase solution (v/v)

20% DMEM (v/v)

V.b.ii. Myoblast isolation

Hank's Balanced Salt Solution (HBSS)

$CaCl_2$ $MgCl_2$ and $MgSO_4$ free. (Invitrogen Cat# 14170-112)

HEPES-buffered HBSS

1x HBSS

50mM HEPES (1M; Gibco Cat#15620-080)

Enzyme solution

1.2 U/ml dispase

0.5% collegase II (w/v)

2.5mM CaCl$_2$

<u>0.1% gelatin solution</u>

0.1% (v/v) 2% gelatin solution (Sigma) in H$_2$O

VI. Solutions for BioID

<u>Cell culture media containing 50μM biotin</u>

0.025% 200mM biotin in DMSO (v/v)

Cell culture medium

<u>Lysis buffer</u>

50mM Tris

500mM NaCl

0.4% SDS (w/v)

Before use, add:

1x protease inhibitor (v/v) (Roche EDTA-free)

1% Triton X-100 (v/v)

1mM DTT

Buffer A

0.5% Sodium Deoxycholate (w/v)

150mM NaCl

1% Nonidet P40 (Sigma) (v/v)

0.1% SDS (w/v)

50mM Tris pH 7.4

Before use, add:

1x protease inhibitor (v/v) (Roche EDTA-free)

Buffer B

20mM Tris pH 7.4

Before use, add:

1x protease inhibitor (v/v) (Roche EDTA-free)

Buffer D

80% acetonitrile (v/v)

0.5% acetic acid (v/v)

Buffer E

0.5% acetic acid (v/v)

Triethylammonium bicarbonate (TEAB)

100mM in water, pH 8.5

Trifluoro acetic acid (TFA) (Sigma)

10% TFA in water (v/v)

Buffer ED1

50mM triethylammonium bicarbonate (TEAB)

50% trifluoroethanol (v/v) (Merck)

LysC preparation (WAKO Cat#129-02541)

10AU vial resuspended in 4.4ml H_2O. Stored at -20°C

Trypsin preparation (Promega, Cat#V5111)

20μg sequencing grade trypsin was resuspended as per manufacturers

instruction in 20μl 50mM acetic acid.

352

I.   Bacterial selection

Antibiotic	Working concentration (µg/ml)	Notes
Ampicillin	100	
Kanamycin	30	
Blasticidin	50	Used with Ampicillin for cloning and propagation of TRIBLA backbone plasmids

II.   Eukaryotic selection

Cell type	Puromycin (µg/ml)	Blasticidin (µg/ml)
Rosa26-FKBP-BioID2 MAFs	1	2.5
Rosa26-FKBP-BioID2 Myoblasts	0.5	
HeLa	1	7.5
HEK 293T	0.5	7.5
C2C12	5	2.5
KM155 Human myoblasts	Resistant	7.5

I.   Top 200 candidate Sun1 interacting proteins ranked using BioID score calculated with all negative controls

See Appendix F Sheet Ap.F.I All NEG.pdf

The top 200 BioID score ranked Sun1 interacting candidates ranked according to their BioID score calculated by combining control conditions where the dimeriser is expressed, but lacking the dimeriser, combined with that lacking dimeriser, doxycycline (and therefore FRB-Sun1 bait expression), and supplemented biotin. AvLFQi – averaged label free quantification, AvSC – averaged spectral count.

II.   Top 200 candidate Sun1 interacting proteins ranked using BioID score calculated with control lacking dimeriser only

See Appendix F Sheet Ap.F.II NoDZ.pdf

The top 200 BioID score ranked Sun1 interacting candidates ranked according to their BioID score calculated against the internal control condition where the dimeriser absent. AvLFQi – averaged label free quantification, AvSC – averaged spectral count.

References

Aaronson, R. P. and Blobel, G. (1975) 'Isolation of nuclear pore complexes in association with a lamina', *Proceedings of the National Academy of Sciences of the United States of America*, 72(3), pp. 1007–1011. doi: 10.1073/pnas.72.3.1007.

Adamo, C. M. *et al.* (2010) 'Sildenafil reverses cardiac dysfunction in the mdx mouse model of Duchenne muscular dystrophy', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1013077107.

Aebi, U. *et al.* (1986) 'The nuclear lamina is a meshwork of intermediate-type filaments', *Nature*, 323(6088), pp. 560–564. doi: 10.1038/323560a0.

Agircan, F. G., Schiebel, E. and Mardin, B. R. (2014) 'Separate to operate: Control of centrosome positioning and separation', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1650). doi: 10.1098/rstb.2013.0461.

Agrawal, S. *et al.* (2006) 'Dystroglycan is selectively cleaved at the parenchymal basement membrane at sites of leukocyte extravasation in experimental autoimmune encephalomyelitis', *Journal of Experimental Medicine*, 203(4), pp. 1007–1016. doi: 10.1084/jem.20051342.

Ahn, J. *et al.* (2019) 'Structural basis for lamin assembly at the molecular level', *Nature Communications*, 10(1), pp. 1–12. doi: 10.1038/s41467-019-11684-x.

Akhavan, A. *et al.* (2008) 'SEA domain proteolysis determines the functional composition of dystroglycan', *FASEB J*, 22, pp. 612–621. doi: 10.1096/fj.07-8354com.

Alam, M. S. (2018) 'Proximity Ligation Assay (PLA)', *Current Protocols in Immunology*, 123(1), pp. 1–8. doi: 10.1002/cpim.58.

Alberts, B. *et al.* (2008) *Molecular Biology of the Cell*. Fifth Edit. New York: Garland Science.

Anderson, E. D. *et al.* (2002) 'The ordered and compartment-specific autoproteolytic removal of the furin intramolecular chaperone is required for enzyme activation', *Journal of Biological Chemistry*, 277(15), pp. 12879–12890. doi: 10.1074/jbc.M108740200.

Antoku, S. *et al.* (2019) 'ERK1/2 Phosphorylation of FHOD Connects Signaling and Nuclear Positioning Alternations in Cardiac Laminopathy', *Developmental Cell*, 51(5), pp. 602-616.e12. doi: 10.1016/j.devcel.2019.10.023.

Araishi, K. *et al.* (1999) 'Loss of the sarcoglycan complex and sarcospan leads to muscular dystrophy in β-sarcoglycan-deficient mice', *Human Molecular Genetics*, 8(9), pp. 1589–1598. doi: 10.1093/hmg/8.9.1589.

Arber, S., Halder, G. and Caroni, P. (1994) 'Muscle LIM protein, a novel essential regulator of myogenesis, promotes myogenic differentiation', *Cell*, 79(2), pp. 221–231. doi: 10.1016/0092-8674(94)90192-9.

Asai, A. *et al.* (2007) 'Primary role of functional ischemia, quantitative evidence for the two-hit mechanism, and phosphodiesterase-5 inhibitor therapy in mouse

muscular dystrophy', *PLoS ONE*, 2(8). doi: 10.1371/journal.pone.0000806.

Attali, R. *et al.* (2009) 'Mutation of SYNE-1, encoding an essential component of the nuclear lamina, is responsible for autosomal recessive arthrogryposis', *Human Molecular Genetics*, 18(18), pp. 3462–3469. doi: 10.1093/hmg/ddp290.

Azuara-Medina, P. M. *et al.* (2019) 'The intracellular domain of β-dystroglycan mediates the nucleolar stress response by suppressing UBF transcriptional activity', *Cell Death and Disease*, 10(3). doi: 10.1038/s41419-019-1454-z.

Bader, D., Masaki, T. and Fischman, D. A. (1982) 'Immunochemical analysis of myosin heavy chain during avian myogenesis in vivo and in vitro', *Journal of Cell Biology*, 95(3), pp. 763–770. doi: 10.1083/jcb.95.3.763.

Banaszynski, L. A., Liu, C. W. and Wandless, T. J. (2005) 'Characterization of the FKBP-rapamycin-FRB ternary complex', *Journal of the American Chemical Society*, 127(13), pp. 4715–4721. doi: 10.1021/ja043277y.

Barascu, A. *et al.* (2012) 'Oxidative stress induces an ATM-independent senescence pathway through p38 MAPK-mediated lamin B1 accumulation', *EMBO Journal*, 31(5), pp. 1080–1094. doi: 10.1038/emboj.2011.492.

Barateau, A. *et al.* (2017) 'A novel lamin A mutant responsible for congenital muscular dystrophy causes distinct abnormalities of the cell nucleus', *PLoS ONE*, 12(1), pp. 1–18. doi: 10.1371/journal.pone.0169189.

Barrangou, R. *et al.* (2007) 'CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes', *Science*, 315(March), pp. 1709–1712. doi: 10.1126/science.1138140.

Batchelor, C. L. *et al.* (2007) 'Recruitment of Dbl by ezrin and dystroglycan drives membrane proximal Cdc42 activation and filopodia formation', *Cell Cycle*, 6(3), pp. 353–363. doi: 10.4161/cc.6.3.3819.

Baumann, M. *et al.* (2017) 'Homozygous SYNE1 mutation causes congenital onset of muscular weakness with distal arthrogryposis: A genotype-phenotype correlation', *European Journal of Human Genetics*, 25(2), pp. 262–266. doi: 10.1038/ejhg.2016.144.

Beauchamp, J. R. *et al.* (1999) 'Dynamics of myoblast transplantation reveal a discrete minority of precursors with stem cell-like properties as the myogenic source', *Journal of Cell Biology*, 144(6), pp. 1113–1121. doi: 10.1083/jcb.144.6.1113.

Behrens, T. W. *et al.* (1994) 'Jaw1, a lymphoid-restricted membrane protein localized to the endoplasmic reticulum', *Journal of Immunology*, 153(2), pp. 682–690.

Beltrán, D. *et al.* (2019) 'Exogenous expression of the glycosyltransferase LARGE1 restores α-dystroglycan matriglycan and laminin binding in rhabdomyosarcoma', *Skeletal Muscle*, 9(1), pp. 1–10. doi: 10.1186/s13395-019-0195-0.

Ben-Harush, K. *et al.* (2009) 'The Supramolecular Organization of the C. elegans Nuclear Lamin Filament', *Journal of Molecular Biology*, 386(5), pp. 1392–1402. doi: 10.1016/j.jmb.2008.12.024.

Van Berlo, J. H. *et al.* (2005) 'A-type lamins are essential for TGF-β1 induced PP2A to dephosphorylate transcription factors', *Human Molecular Genetics*, 14(19), pp. 2839–2849. doi: 10.1093/hmg/ddi316.

Bernasconi, P. *et al.* (2018) 'Elevated TGFb2 serum levels in Emery-Dreifuss Muscular Dystrophy: Implications for myocyte and tenocyte differentiation and fibrogenic processes', *Nucleus*, 9(1), pp. 337–349. doi: 10.1080/19491034.2018.1467722.

Bertini, E. *et al.* (2011) 'Congenital Muscular Dystrophies: A Brief Review', *Seminars in Pediatric Neurology*, 18(4), pp. 277–288. doi: 10.1016/j.spen.2011.10.010.

Bertrand, A. T. *et al.* (2020) 'Lamin A/C Assembly Defects in LMNA-Congenital Muscular Dystrophy Is Responsible for the Increased Severity of the Disease Compared with Emery–Dreifuss Muscular Dystrophy', *Cells*, 9(4), p. 844. doi: 10.3390/cells9040844.

Bhat, H. F., Adams, M. E. and Khanday, F. A. (2013) 'Syntrophin proteins as Santa Claus: Role(s) in cell signal transduction', *Cellular and Molecular Life Sciences*, 70(14), pp. 2533–2554. doi: 10.1007/s00018-012-1233-9.

Bindea, G. *et al.* (2009) 'ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks', *Bioinformatics*, 25(8), pp. 1091–1093. doi: 10.1093/bioinformatics/btp101.

Bione, S. *et al.* (1994) 'Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy.', *Nature genetics*, 8(4), pp. 323–7. doi: 10.1038/ng1294-323.

Bitoun, M. *et al.* (2005) 'Mutations in dynamin 2 cause dominant centronuclear myopathy', *Nature Genetics*, 37(11), pp. 1207–1209. doi: 10.1038/ng1657.

Blake, D. J. *et al.* (1995) 'Coiled-coil regions in the carboxy-terminal domains of dystrophin and related proteins: potentials for protein-protein interactions', *Trends in Biochemical Sciences*, pp. 3–5. doi: 10.1016/S0968-0004(00)88986-0.

Blake, D. J. *et al.* (2002) 'Function and Genetics of Dystrophin and Dystrophin-Related Proteins in Muscle', *Physiological Reviews*, 82(2), pp. 291–329. doi: 10.1152/physrev.00028.2001.

Bolotin, A. *et al.* (2005) 'Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin', *Microbiology*, 151(8), pp. 2551–2561. doi: 10.1099/mic.0.28048-0.

Bonne, G. *et al.* (1999) 'Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy', *Nature Genetics*, 21(march), pp. 285–288. doi: 10.1038/6799.

Bonne, G. and Levy, N. (2003) 'LMNA mutations in atypical Werner's syndrome', *The Lancet*, 362, pp. 1585–1586.

Bonneman, C. G. *et al.* (1996) 'β-sarcoglycan (A3b) mutations cause autosomal recessive muscular dystrophy with loss of the sarcoglycan complex', *Nature Genetics*, pp. 266–273.

Bos, J. M. *et al.* (2006) 'Genotype-phenotype relationships involving hypertrophic cardiomyopathy-associated mutations in titin, muscle LIM protein, and telethonin', *Molecular Genetics and Metabolism*, 88(1), pp. 78–85. doi: 10.1016/j.ymgme.2005.10.008.

Bourgeois, B. *et al.* (2013) 'Inhibition of TGF-β signaling at the nuclear envelope: Characterization of interactions between MAN1, Smad2 and Smad3, and PPM1A', *Science Signaling*, 6(280), pp. 1–11. doi: 10.1126/scisignal.2003411.

Bozzi, M. *et al.* (2009) 'Enzymatic processing of beta-dystroglycan recombinant ectodomain by MMP-9: Identification of the main cleavage site', *IUBMB Life*, 61(12), pp. 1143–1152. doi: 10.1002/iub.273.

Brancaccio, A. (2019) 'A molecular overview of the primary dystroglycanopathies', *Journal of Cellular and Molecular Medicine*, pp. 1–5. doi: 10.1111/jcmm.14218.

Branon, T. C. *et al.* (2018) 'Efficient proximity labeling in living cells and organisms with TurboID', *Nature biotechnology*, 36(9), pp. 880–887. doi: 10.1038/nbt.4201.

Briand, N. *et al.* (2018) 'The lipodystrophic hotspot lamin A p.R482W mutation deregulates the mesodermal inducer T/Brachyury and early vascular differentiation gene networks', *Human Molecular Genetics*, 27(8), pp. 1447–1459. doi: 10.1093/hmg/ddy055.

Bridger, J. M. and Kill, I. R. (2004) 'Aging of Hutchinson-Gilford progeria syndrome fibroblasts is characterised by hyperproliferation and increased apoptosis', *Experimental Gerontology*, 39(5), pp. 717–724. doi: 10.1016/j.exger.2004.02.002.

Broach, J. R., Guarascio, V. R. and Jayaram, M. (1982) 'Recombination within the yeast plasmid 2μ circle is site-specific', *Cell*, 29(1), pp. 227–234. doi: 10.1016/0092-8674(82)90107-6.

Brodsky, G. L. *et al.* (2001) 'Lamin A/C gene mutation associated with dilated cardiomyopathy with variable skeletal muscle involvement.', *Circulation*, 103(4), pp. 473–477. doi: 10.1161/01.CIR.103.4.e20.

Broers, J. L. *et al.* (1999) 'Dynamics of the nuclear lamina as monitored by GFP-tagged A-type lamins', *J Cell Sci*, 112 (Pt 2, pp. 3463–3475. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10504295%5Cnhttp://jcs.biologists.org/content/112/20/3463.full.pdf.

Brou, C. *et al.* (2000) 'A novel proteolytic cleavage involved in Notch signaling: The role of the disintegrin-metalloprotease TACE', *Molecular Cell*, 5(2), pp. 207–216. doi: 10.1016/S1097-2765(00)80417-7.

Bunnell, T. M. *et al.* (2008) 'Destabilization of the dystrophin-glycoprotein complex without functional deficits in α-dystrobrevin null muscle', *PLoS ONE*, 3(7), pp. 1–6. doi: 10.1371/journal.pone.0002604.

Burakov, A. V. and Nadezhdina, E. S. (2013) 'Association of nucleus and centrosome: Magnet or velcro?', *Cell Biology International*, 37(2), pp. 95–104. doi:

10.1002/cbin.10016.

Burke, B. and Stewart, C. L. (2013) 'The nuclear lamins: Flexibility in function', *Nature Reviews Molecular Cell Biology*, 14(1), pp. 13–24. doi: 10.1038/nrm3488.

Butler, M. H. *et al.* (1992) 'Association of the M(r) 58,000 postsynaptic protein of electric tissue with Torpedo dystrophin and the M(r) 87,000 postsynaptic protein', *Journal of Biological Chemistry*, 267(9), pp. 6213–6218.

Cabanillas, R. *et al.* (2011) 'Néstor-Guillermo progeria syndrome: A novel premature aging condition with early onset and chronic development caused by BANF1 mutations', *American Journal of Medical Genetics, Part A*, 155(11), pp. 2617–2625. doi: 10.1002/ajmg.a.34249.

Cadot, B., Gache, V. and Gomes, E. R. (2015) 'Moving and positioning the nucleus in skeletal muscle-one step at a time', *Nucleus*, 6(5), pp. 373–381. doi: 10.1080/19491034.2015.1090073.

Cain, N. E. *et al.* (2014) 'The SUN protein UNC-84 is required only in force-bearing cells to maintain nuclear envelope architecture', *Journal of Cell Biology*, 206(2), pp. 163–172. doi: 10.1083/jcb.201405081.

Calderwood, D. A. (2004) 'Integrin activation', *Journal of Cell Science*, 117(5), pp. 657–666. doi: 10.1242/jcs.01014.

Calvi, A. *et al.* (2015) 'SUN4 is essential for nuclear remodeling during mammalian spermiogenesis', *Developmental Biology*, 407(2), pp. 321–330. doi: 10.1016/j.ydbio.2015.09.010.

Capell, B. C. *et al.* (2005) 'Inhibiting farnesylation of progerin prevents the characteristic nuclear blebbing of Hutchinson-Gilford progeria syndrome', *Proceedings of the National Academy of Sciences of the United States of America*, 102(36), pp. 12879–12884. doi: 10.1073/pnas.0506001102.

Capell, B. C. and Collins, F. S. (2006) 'Human laminopathies: Nuclei gone genetically awry', *Nature Reviews Genetics*, 7(12), pp. 940–952. doi: 10.1038/nrg1906.

Carlson, C. G. (1998) 'The dystrophinopathies: An alternative to the structural hypothesis', *Neurobiology of Disease*, 5(1), pp. 3–15. doi: 10.1006/nbdi.1998.0188.

Carlson, C. G. and Roshek, D. M. (2001) 'Adult dystrophic (mdx) endplates exhibit reduced quantal size and enhanced quantal variation', *Pflugers Archiv European Journal of Physiology*, 442(3), pp. 369–375. doi: 10.1007/s004240100561.

Cartwright, S. and Karakesisoglou, I. (2014) 'Nesprins in health and disease', *Seminars in Cell and Developmental Biology*, 29, pp. 169–179. doi: 10.1016/j.semcdb.2013.12.010.

Cattin, M. E., Muchir, A. and Bonne, G. (2013) '"State-of-the-heart" of cardiac laminopathies', *Current Opinion in Cardiology*, 28(3), pp. 297–304. doi: 10.1097/HCO.0b013e32835f0c79.

Celetti, G. *et al.* (2020) 'The liquid state of FG-nucleoporins mimics permeability barrier properties of nuclear pore complexes', *The Journal of cell*

biology, 219(1). doi: 10.1083/jcb.201907157.

Cesarini, E. *et al.* (2015) 'Lamin A/C sustains PcG protein architecture, maintaining transcriptional repression at target genes', *Journal of Cell Biology*, 211(3), pp. 533–551. doi: 10.1083/jcb.201504035.

Chan, Y. M. *et al.* (1998) 'Molecular organization of sarcoglycan complex in mouse myotubes in culture', *Journal of Cell Biology*, 143(7), pp. 2033–2044. doi: 10.1083/jcb.143.7.2033.

Chang, N. C. *et al.* (2018) 'The Dystrophin Glycoprotein Complex Regulates the Epigenetic Activation of Muscle Stem Cell Commitment', *Cell Stem Cell*, 22(5), pp. 755-768.e6. doi: 10.1016/j.stem.2018.03.022.

Chang, N. C., Chevalier, F. P. and Rudnicki, M. A. (2016) 'Satellite Cells in Muscular Dystrophy - Lost in Polarity', *Trends in Molecular Medicine*, 22(6), pp. 479–496. doi: 10.1016/j.molmed.2016.04.002.

Chang, W. *et al.* (2013) 'Emerin organizes actin flow for nuclear movement and centrosome orientation in migrating fibroblasts', *Molecular Biology of the Cell*, 24(24), pp. 3869–3880. doi: 10.1091/mbc.E13-06-0307.

Chang, W. *et al.* (2015) 'Linker of nucleoskeleton and cytoskeleton (Linc) complex-mediated actin-dependent nuclear positioning orients centrosomes in migrating myoblasts', *Nucleus*, 6(1), pp. 77–88. doi: 10.1080/19491034.2015.1004947.

Chang, W. *et al.* (2019) 'Imbalanced nucleocytoskeletal connections create common polarity defects in progeria and physiological aging', *Proceedings of the National Academy of Sciences of the United States of America*, 116(9), pp. 3578–3583. doi: 10.1073/pnas.1809683116.

Chatzifrangkeskou, M. *et al.* (2016) 'ERK1/2 directly acts on CTGF/CCN2 expression to mediate myocardial fibrosis in cardiomyopathy caused by mutations in the lamin A/C gene', *Human Molecular Genetics*, 25(11), pp. 2220–2233. doi: 10.1093/hmg/ddw090.

Cheedipudi, S. M. *et al.* (2019) 'Genomic Reorganization of Lamin-Associated Domains in Cardiac Myocytes Is Associated With Differential Gene Expression and DNA Methylation in Human Dilated Cardiomyopathy', *Circulation research*, 124(8), pp. 1198–1213. doi: 10.1161/CIRCRESAHA.118.314177.

Chen, C. L. *et al.* (2015) 'Proteomic mapping in live Drosophila tissues using an engineered ascorbate peroxidase', *Proceedings of the National Academy of Sciences of the United States of America*, 112(39), pp. 12093–12098. doi: 10.1073/pnas.1515623112.

Chen, C. Y. *et al.* (2012) 'Accumulation of the inner nuclear envelope protein Sun1 is pathogenic in progeric and dystrophic laminopathies', *Cell*, 149(3), pp. 565–577. doi: 10.1016/j.cell.2012.01.059.

Chen, I. H. B. *et al.* (2006) 'Nuclear envelope transmembrane proteins (NETs) that are up-regulated during myogenesis', *BMC Cell Biology*, 7, pp. 1–16. doi: 10.1186/1471-2121-7-38.

Chen, L. *et al.* (2003) 'LMNA mutations in atypical Werner's syndrome', *Lancet*, 362(9382), pp. 440–445. doi: 10.1016/S0140-6736(03)14069-X.

Chen, Y.-J. *et al.* (2003) 'Direct interaction of beta-dystroglycan with F-actin', *The Biochemical journal*, 375(Pt 2), pp. 329–337. doi: 10.1042/BJ20030808.

Chen, Y. W. *et al.* (2005) 'Early onset of inflammation and later involvement of TGFβ in Duchenne muscular dystrophy', *Neurology*, 65(6), pp. 826–834. doi: 10.1212/01.wnl.0000173836.09176.c4.

Chen, Z. J. *et al.* (2014) 'Dysregulated interactions between lamin A and SUN1 induce abnormalities in the nuclear envelope and endoplasmic reticulum in progeric laminopathies', *Journal of Cell Science*, 127(8), pp. 1792–1804. doi: 10.1242/jcs.139683.

Chi, Y. H. *et al.* (2009) 'Requirement for Sun1 in the expression of meiotic reproductive genes and piRNA', *Development*, 136(6), pp. 965–973. doi: 10.1242/dev.029868.

Chien, C. T. *et al.* (1991) 'The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest', *Proceedings of the National Academy of Sciences of the United States of America*, 88(21), pp. 9578–9582. doi: 10.1073/pnas.88.21.9578.

Chikashige, Y. *et al.* (2006) 'Meiotic Proteins Bqt1 and Bqt2 Tether Telomeres to Form the Bouquet Arrangement of Chromosomes', *Cell*, 125(1), pp. 59–69. doi: 10.1016/j.cell.2006.01.048.

Cho, E. *et al.* (2018) 'β-dystroglycan is regulated by a balance between WWP1-mediated degradation and protection from WWP1 by dystrophin and utrophin', *BBA - Molecular Basis of Disease*, 1864(6), pp. 2199–2213. doi: 10.1016/j.bbadis.2018.04.001.

Chojnowski, A. *et al.* (2018) '2C-BioID: An advanced two component BioID system for precision mapping of protein interactomes', *iScience*, 0(0), pp. 40–52. doi: 10.1016/j.isci.2018.11.023.

Clarke, R. *et al.* (2018) 'Enhanced Bacterial Immunity and Mammalian Genome Editing via RNA-Polymerase-Mediated Dislodging of Cas9 from Double-Strand DNA Breaks', *Molecular Cell*, 71(1), pp. 42-55.e8. doi: 10.1016/j.molcel.2018.06.005.

Coffinier, C. *et al.* (2010) 'Abnormal development of the cerebral cortex and cerebellum in the setting of lamin B2 deficiency', *Proceedings of the National Academy of Sciences*, 107(11), pp. 5076–5081. doi: 10.1073/pnas.0908790107.

Coffinier, C. *et al.* (2011) 'Deficiencies in lamin B1 and lamin B2 cause neurodevelopmental defects and distinct nuclear shape abnormalities in neurons.', *Molecular biology of the cell*, 22(23), pp. 4683–93. doi: 10.1091/mbc.E11-06-0504.

Cohen, J. (1988) *Statistical Power Analysis for the Behavioural Science (2nd Edition)*, *Statistical Power Anaylsis for the Behavioral Sciences*.

Cohen, T. V. *et al.* (2013) 'Defective skeletal muscle growth in lamin A/C-deficient mice is rescued by loss of lap2α', *Human Molecular Genetics*, 22(14), pp. 2852–2869. doi: 10.1093/hmg/ddt135.

Cohen, T. V., Kosti, O. and Stewart, C. L. (2007) 'The nuclear envelope protein MAN1 regulates TGFβ signaling and vasculogenesis in the embryonic yolk sac', *Development*, 134(7), pp. 1385–1395. doi: 10.1242/dev.02816.

Cohn, R. D. *et al.* (2002) 'Disruption of Dag1 in differentiated skeletal muscle reveals a role for dystroglycan in muscle regeneration', *Cell*, 110(5), pp. 639–648. doi: 10.1016/S0092-8674(02)00907-8.

Cong, L. *et al.* (2013) 'Multiplex genome engineering using CRISPR/Cas systems', *Science*, 339(6121), pp. 819–823. doi: 10.1126/science.1231143.

Conti, E. *et al.* (1998) 'Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin α', *Cell*, 94(2), pp. 193–204. doi: 10.1016/S0092-8674(00)81419-1.

Conti, E. and Kuriyan, J. (2000) 'Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin α', *Structure*, 8(3), pp. 329–338. doi: 10.1016/S0969-2126(00)00107-6.

Conti, E., Müller, C. W. and Stewart, M. (2006) 'Karyopherin flexibility in nucleocytoplasmic transport', *Current Opinion in Structural Biology*, 16(2), pp. 237–244. doi: 10.1016/j.sbi.2006.03.010.

Cook, A. *et al.* (2007) 'Structural Biology of Nucleocytoplasmic Transport', *Annual Review of Biochemistry*, 76(1), pp. 647–671. doi: 10.1146/annurev.biochem.76.052705.161529.

Copley, S. D. (2012) 'Moonlighting is mainstream: Paradigm adjustment required', *BioEssays*, 34(7), pp. 578–588. doi: 10.1002/bies.201100191.

Corrigan, D. P. *et al.* (2005) 'Prelamin A endoproteolytic processing in vitro by recombinant Zmpste24', *Biochemical Journal*, 387(1), pp. 129–138. doi: 10.1042/BJ20041359.

Costello, A. *et al.* (2019) 'Leaky Expression of the TET-On System Hinders Control of Endogenous miRNA Abundance', *Biotechnology Journal*, 14(3), pp. 1–10. doi: 10.1002/biot.201800219.

Côté, P. D. *et al.* (1999) 'Chimaeric mice deficient in dystroglycans develop muscular dystrophy and have disrupted myoneural synapses', *Nature Genetics*, 23(3), pp. 338–342. doi: 10.1038/15519.

Côté, P. D., Moukhles, H. and Carbonetto, S. (2002) 'Dystroglycan is not required for localization of dystrophin, syntrophin, and neuronal nitric-oxide synthase at the sarcolemma but regulates integrin α7B expression and caveolin-3 distribution', *Journal of Biological Chemistry*, 277(7), pp. 4672–4679. doi: 10.1074/jbc.M106879200.

Crisp, M. *et al.* (2006) 'Coupling of the nucleus and cytoplasm: Role of the LINC complex', *Journal of Cell Biology*, 172(1), pp. 41–53. doi: 10.1083/jcb.200509124.

Crosbie, R. H. *et al.* (1997) 'Sarcospan, the 25-kDa transmembrane component of the dystrophin- glycoprotein complex', *Journal of Biological Chemistry*, 272(50), pp. 31221–31224. doi: 10.1074/jbc.272.50.31221.

Crosbie, R. H. *et al.* (1999) 'Membrane targeting and stabilization of sarcospan is mediated by the sarcoglycan subcomplex', *Journal of Cell Biology*, 145(1), pp. 153–165. doi: 10.1083/jcb.145.1.153.

Cullen, M. J. *et al.* (2001) 'Immunogold confirmation that utrophin is localized to the normal position of dystrophin in dystrophin-negative transgenic mouse muscle', *Histochemical Journal*, 33(9–10), pp. 579–583. doi: 10.1023/A:1014964127156.

Cummins, P. and Perry, S. V. (1978) 'Troponin I from human skeletal and cardiac muscles', *Biochemical Journal*, 171(1), pp. 251–259. doi: 10.1042/bj1710251.

Cupesi, M. *et al.* (2010) 'Attenuated hypertrophic response to pressure overload in a lamin A/C haploinsufficiency mouse', *Journal of Molecular and Cellular Cardiology*, 48(6), pp. 1290–1297. doi: 10.1016/j.yjmcc.2009.10.024.

Dae In, K. *et al.* (2016) 'An improved smaller biotin ligase for BioID proximity labeling', *Molecular biology of the cell*, 27, pp. 1188–1196. doi: 10.1091/mbc.E15-12-0844.

Dahl, K. N., Ribeiro, A. J. S. and Lammerding, J. (2008) 'Nuclear shape, mechanics, and mechanotransduction', *Circulation Research*, 102(11), pp. 1307–1318. doi: 10.1161/CIRCRESAHA.108.173989.

Dai, Y. *et al.* (2018) 'Whole exome sequencing identified a novel *DAG1* mutation in a patient with rare, mild and late age of onset muscular dystrophy-dystroglycanopathy', *Journal of Cellular and Molecular Medicine*, (July). doi: 10.1111/jcmm.13979.

Daigle, N. *et al.* (2001) 'Nuclear pore complexes form immobile networks and have a very low turnover in live mammalian cells', *Journal of Cell Biology*, 154(1), pp. 71–84. doi: 10.1083/jcb.200101089.

Deconinck, A. E. *et al.* (1997) 'Postsynaptic abnormalities at the neuromuscular junctions of utrophin- deficient mice', *Journal of Cell Biology*, 136(4), pp. 883–894. doi: 10.1083/jcb.136.4.883.

Denais, C. M. *et al.* (2016) 'Nuclear envelope rupture and repair during cancer cell migration', *Science*, 352(6283), pp. 353–358. doi: 10.1126/science.aad7297.

Desmouliere, A. *et al.* (1993) 'Transforming growth factor-β1 induces α-smooth muscle actin expression in granulation tissue myofibroblasts and in quiescent and growing cultured fibroblasts', *Journal of Cell Biology*, 122(1), pp. 103–111. doi: 10.1083/jcb.122.1.103.

Dhe-Paganon, S. *et al.* (2002) 'Structure of the globular tail of nuclear lamin', *Journal of Biological Chemistry*, 277(20), pp. 17381–17384. doi: 10.1074/jbc.C200038200.

Dinçer, P. *et al.* (2003) 'A novel form of recessive limb girdle muscular dystrophy with mental retardation and abnormal expression of α-dystroglycan', *Neuromuscular Disorders*, 13(10), pp. 771–778. doi: 10.1016/S0960-8966(03)00161-5.

Ding, X. *et al.* (2007) 'SUN1 Is Required for Telomere Attachment to Nuclear

Envelope and Gametogenesis in Mice', *Developmental Cell*, 12(6), pp. 863–872. doi: 10.1016/j.devcel.2007.03.018.

Dittmer, T. and Misteli, T. (2011) 'The lamin protein family', *Genome Biology*, 12(5), p. 222. doi: 10.1186/gb-2011-12-5-222.

Dong, M. *et al.* (2015) 'DAG1 mutations associated with asymptomatic hyperCKemia and hypoglycosylation of α-dystroglycan', *Neurology*, 84(3), pp. 273–279. doi: 10.1212/WNL.0000000000001162.

Dreesen, O. *et al.* (2013) 'Lamin B1 fluctuations have differential effects on cellular proliferation and senescence', *Journal of Cell Biology*, 200(5), pp. 605–617. doi: 10.1083/jcb.201206121.

Driscoll, M. K. *et al.* (2012) 'Automated image analysis of nuclear shape: What can we learn from a prematurely aged cell?', *Aging*, 4(2), pp. 119–132. doi: 10.18632/aging.100434.

Duclos, F. *et al.* (1998) 'Progressive muscular dystrophy in α-sarcoglycan-deficient mice', *Journal of Cell Biology*, 142(6), pp. 1461–1471. doi: 10.1083/jcb.142.6.1461.

Dumont, N. A. *et al.* (2015) 'Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division', *Nature Medicine*, 21(12), pp. 1455–1463. doi: 10.1038/nm.3990.

Earle, A. J. *et al.* (2020) 'Mutant lamins cause nuclear envelope rupture and DNA damage in skeletal muscle cells', *Nature Materials*, 19(4), pp. 464–473. doi: 10.1038/s41563-019-0563-5.

Ellenberg, J. *et al.* (1997) 'Nuclear membrane dynamics and reassembly in living cells: Targeting of an inner nuclear membrane protein in interphase and mitosis', *Journal of Cell Biology*, 138(6), pp. 1193–1206. doi: 10.1083/jcb.138.6.1193.

Endo, T. (2007) 'Dystroglycan glycosylation and its role in α-dystroglycanopathies', *Acta Myologica*, 26(3), pp. 165–170.

Endo, T. (2015) 'Glycobiology of α-dystroglycan and muscular dystrophy', *Journal of Biochemistry*, 157(1), pp. 1–12. doi: 10.1093/jb/mvu066.

Eriksson, M. *et al.* (2003) 'Recurrent de novo point mutations in lamin A cause Hutchinson–Gilford progeria syndrome', *Nature*, 423, pp. 293–298. doi: 10.1177/002199839502901205.

Erpapazoglou, Z., Walker, O. and Haguenauer-Tsapis, R. (2014) 'Versatile Roles of K63-Linked Ubiquitin Chains in Trafficking', *Cells*, 3(4), pp. 1027–1088. doi: 10.3390/cells3041027.

Ervasti, J. M. *et al.* (1990) 'Deficiency of a glycoprotein component of the dystrophin complex in dystrophic muscle', *Nature*. doi: 10.1038/345315a0.

Ervasti, J. M., Burwell, A. L. and Geissler, A. L. (1997) 'Tissue-specific Heterogeneity in alpha-Dystroglycan Sialoglycosylation', 272(35), pp. 22315–22321.

Ervasti, J. M. and Campbell, K. P. (1991) 'Membrane organization of the dystrophin-glycoprotein complex', *Cell*, 66(6), pp. 1121–1131. doi: 10.1016/0092-

8674(91)90035-W.

Ervasti, J. M. and Campbell, K. P. (1993) 'A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin', *Journal of Cell Biology*, 122(4), pp. 809–823. doi: 10.1083/jcb.122.4.809.

Esapa, C. T. *et al.* (2003) 'The effects of post-translational processing on dystroglycan synthesis and trafficking', *FEBS Letters*, 555(2), pp. 209–216. doi: 10.1016/S0014-5793(03)01230-4.

Esvelt, K. M. *et al.* (2013) 'RNA-Guided Human Genome Engineering via Cas9', *Science*, 339(823), pp. 823–826. Available at: http://www.sciencemag.org.cyber.usask.ca/content/339/6121/823.full.pdf%5Cnfile:///Articles/2013/Mali/Science 2013 Mali.pdf.

Evan, G. I. *et al.* (1985) 'Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product.', *Molecular and Cellular Biology*, 5(12), pp. 3610–3616. doi: 10.1128/mcb.5.12.3610.

Fatkin, D. *et al.* (1999) 'Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease', *New England Journal of Medicine*, pp. 1715–1724. doi: 10.1056/NEJM199912023412302.

Favreau, C. *et al.* (2004) 'Expression of a Mutant Lamin A That Causes Emery-Dreifuss Muscular Dystrophy Inhibits In Vitro Differentiation of C2C12 Myoblasts', *Molecular and Cellular Biology*, 24(4), pp. 1481–1492. doi: 10.1128/mcb.24.4.1481-1492.2004.

Fawcett, D. W. (1966) 'On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates', *American Journal of Anatomy*, 119(1), pp. 129–145. doi: 10.1002/aja.1001190108.

Feng, W. *et al.* (2020) 'Identifying the Cardiac Dyad Proteome In Vivo by a BioID2 Knock-In Strategy', *Circulation*, 141, pp. 940–942. doi: 10.1161/CIRCULATIONAHA.119.043434.

Fields, S. (2005) 'High-throughput two-hybrid analysis: The promise and the peril', *FEBS Journal*, 272(21), pp. 5391–5399. doi: 10.1111/j.1742-4658.2005.04973.x.

Fisher, D. Z., Chaudhary, N. and Blobel, G. (1986) 'cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins', *Proceedings of the National Academy of Sciences of the United States of America*, 83(17), pp. 6450–6454. doi: 10.1073/pnas.83.17.6450.

Foeger, N. *et al.* (2006) 'Solubility properties and specific assembly pathways of the B-type lamin from Caenorhabditis elegans', *Journal of Structural Biology*, 155(2), pp. 340–350. doi: 10.1016/j.jsb.2006.03.026.

Foisner, R. and Gerace, L. (1993) 'Integral membrane proteins of the nuclear envelope interact with lamins and chromosomes, and binding is modulated by mitotic phosphorylation', *Cell*, 73(7), pp. 1267–1279. doi: 10.1016/0092-8674(93)90355-T.

Freund, A. *et al.* (2012) 'Lamin B1 loss is a senescence-associated biomarker',

Molecular Biology of the Cell, 23(11), pp. 2066–2075. doi: 10.1091/mbc.E11-10-0884.

Friedrich, G. and Soriano, P. (1991) 'Promoter traps in embryonic stem cells: A genetic screen to identify and mutate developmental genes in mice', *Genes and Development*, 5(9), pp. 1513–1523. doi: 10.1101/gad.5.9.1513.

Frock, R. L. *et al.* (2006) 'Lamin A/C and emerin are critical for skeletal muscle satellite cell differentiation', *Genes and Development*, 20(4), pp. 486–500. doi: 10.1101/gad.1364906.

Frohnert, C., Schweizer, S. and Hoyer-Fender, S. (2011) 'SPAG4L/SPAG4L-2 are testis-specific SUN domain proteins restricted to the apical nuclear envelope of round spermatids facing the acrosome', *Molecular Human Reproduction*, 17(4), pp. 207–218. doi: 10.1093/molehr/gaq099.

Fuentes-Mera, L. *et al.* (2006) 'Characterization of a novel Dp71 dystrophin-associated protein complex (DAPC) present in the nucleus of HeLa cells: Members of the nuclear DAPC associate with the nuclear matrix', *Experimental Cell Research*, 312(16), pp. 3023–3035. doi: 10.1016/j.yexcr.2006.06.002.

Fujii, Y. and Kodama, Y. (2015) 'In planta comparative analysis of improved green fluorescent proteins with reference to fluorescence intensity and bimolecular fluorescence complementation ability', *Plant Biotechnology*, 32(1), pp. 81–87. doi: 10.5511/plantbiotechnology.15.0120a.

Fujii, Y., Yoshimura, A. and Kodama, Y. (2018) 'A novel orange-colored bimolecular fluorescence complementation (BiFC) assay using monomeric Kusabira-Orange protein', *BioTechniques*, 64(4), pp. 153–161. doi: 10.2144/BTN-2017-0121.

Furukawa, K. and Hotta, Y. (1993) 'cDNA cloning of a germ cell specific lamin B3 from mouse spermatocytes and analysis of its function by ectopic expression in somatic cells.', *The EMBO Journal*, 12(1), pp. 97–106. doi: 10.1002/j.1460-2075.1993.tb05635.x.

Furukawa, K., Inagaki, H. and Hotta, Y. (1994) 'Identification and cloning of an mRNA coding for a germ cell-specific A-type lamin in mice', *Experimental Cell Research*, pp. 426–430. doi: 10.1006/excr.1994.1164.

Gant, T. M., Harris, C. a and Wilson, K. L. (1999) 'Roles of LAP2 Proteins in Nuclear Assembly and DNA Replication: Truncated LAP2 Proteins Alter Lamina Assembly, Envelope Formation, Nuclear Size, and DNA Replication Efficiency in Xenopus laevis Extracts', *Journal of Cell Biology*, 144(6), pp. 1083–1096.

Gao, Q. *et al.* (2020) 'The testis-specific LINC component SUN3 is essential for sperm head shaping during mouse spermiogenesis', *Journal of Biological Chemistry*, 295(19), pp. 6289–6298. doi: 10.1074/jbc.RA119.012375.

Gao, Q. Q. and McNally, E. M. (2015) 'The dystrophin complex: Structure, function, and implications for therapy', *Comprehensive Physiology*, 5(3), pp. 1223–1239. doi: 10.1002/cphy.c140048.

Gasiunas, G. *et al.* (2012) 'Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria', *Proceedings of the National*

Academy of Sciences of the United States of America, 109(39), pp. 2579–2586. doi: 10.1073/pnas.1208507109.

Gazzerro, E. *et al.* (2010) 'Therapeutic potential of proteasome inhibition in Duchenne and Becker muscular dystrophies', *American Journal of Pathology*, 176(4), pp. 1863–1877. doi: 10.2353/ajpath.2010.090468.

Geier, C. *et al.* (2003) 'Mutations in the human muscle LIM protein gene in families with hypertrophic cardiomyopathy', *Circulation*, 107(10), pp. 1390–1395. doi: 10.1161/01.CIR.0000056522.82563.5F.

Geier, C. *et al.* (2008) 'Beyond the sarcomere: CSRP3 mutations cause hypertrophic cardiomyopathy', *Human Molecular Genetics*, 17(18), pp. 2753–2765. doi: 10.1093/hmg/ddn160.

Geis, T. *et al.* (2013) 'Homozygous dystroglycan mutation associated with a novel muscle-eye-brain disease-like phenotype with multicystic leucodystrophy', *Neurogenetics*, 14(3–4), pp. 205–213. doi: 10.1007/s10048-013-0374-9.

Gerace, L., Blum, A. and Blobel, G. (1978) 'Immunocytochemical localization of the major polypeptides of the nuclear pore complex-lamina fraction: Interphase and mitotic distribution', *Journal of Cell Biology*, 79(2), pp. 546–566. doi: 10.1083/jcb.79.2.546.

Gibson, T. J., Seiler, M. and Veitia, R. A. (2013) 'The transience of transient overexpression', *Nature Methods*, 10(8), pp. 715–721. doi: 10.1038/nmeth.2534.

Gillies, A. R. *et al.* (2017) 'High resolution three-dimensional reconstruction of fibrotic skeletal muscle extracellular matrix', *Journal of Physiology*, 595(4), pp. 1159–1171. doi: 10.1113/JP273376.

Gimpel, P. *et al.* (2017) 'Nesprin-1α-Dependent Microtubule Nucleation from the Nuclear Envelope via Akap450 Is Necessary for Nuclear Positioning in Muscle Cells', *Current Biology*, 27(19), pp. 2999-3009.e9. doi: 10.1016/j.cub.2017.08.031.

Gingras, A. C., Abe, K. T. and Raught, B. (2019) 'Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles', *Current Opinion in Chemical Biology*, 48, pp. 44–54. doi: 10.1016/j.cbpa.2018.10.017.

Gingras, J. *et al.* (2016) 'α-Dystrobrevin-1 recruits Grb2 and α-catulin to organize neurotransmitter receptors at the neuromuscular junction', *Journal of Cell Science*, 129(5), pp. 898–911. doi: 10.1242/jcs.181180.

Ginsberg, M. H., Partridge, A. and Shattil, S. J. (2005) 'Integrin regulation', *Current Opinion in Cell Biology*, 17(5 SPEC. ISS.), pp. 509–516. doi: 10.1016/j.ceb.2005.08.010.

Göb, E. *et al.* (2010) 'Mammalian sperm head formation involves different polarization of two novel LINC complexes', *PLoS ONE*, 5(8). doi: 10.1371/journal.pone.0012072.

Göb, E. *et al.* (2011) 'Expression of individual mammalian Sun1 isoforms depends on the cell type', *Communicative & Integrative Biology*, 4(4), pp. 440–442. doi: 10.4161/cib.15369.

Goldman, A. E. *et al.* (1986) 'Keratin-like proteins that coisolate with intermediate filaments of BHK-21 cells are nuclear lamins', *Proceedings of the National Academy of Sciences of the United States of America*, 83(11), pp. 3839–3843. doi: 10.1073/pnas.83.11.3839.

Goldman, R. D. *et al.* (2004) 'Accumulation of mutant lamin A progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome', *Proceedings of the National Academy of Sciences of the United States of America*, 101(24), pp. 8963–8968. doi: 10.1073/pnas.0402943101.

Gomes, E. R. and Cadot, B. (2017) 'Molecular motors and nuclear movements in muscle', *Communicative & Integrative Biology*, 10(3), p. e1319537. doi: 10.1080/19420889.2017.1319537.

Gomes, E. R., Jani, S. and Gundersen, G. G. (2005) 'Nuclear movement regulated by Cdc42, MRCK, myosin, and actin flow establishes MTOC polarization in migrating cells', *Cell*, 121(3), pp. 451–463. doi: 10.1016/j.cell.2005.02.022.

Gómez-Monsivais, W. L. *et al.* (2020) 'The Molecular Basis and Biologic Significance of the β-Dystroglycan-Emerin Interaction', *International Journal of Molecular Sciences*, 21(5944), pp. 1–20.

Gonçalves, J. C. *et al.* (2020) 'Nesprin-2 Recruitment of BicD2 to the Nuclear Envelope Controls Dynein/Kinesin-Mediated Neuronal Migration In Vivo', *Current Biology*, pp. 1–14. doi: 10.1016/j.cub.2020.05.091.

González-Ramírez, R. *et al.* (2008) 'Nuclear and nuclear envelope localization of dystrophin Dp71 and dystrophin-associated proteins (DAPs) in the C2C12 muscle cells: DAPs nuclear localization is modulated during myogenesis', *Journal of Cellular Biochemistry*, 105(3), pp. 735–745. doi: 10.1002/jcb.21870.

González, E. *et al.* (2000) 'Alternative splicing regulates the nuclear or cytoplasmic localization of dystrophin Dp71', *FEBS Letters*, 482(3), pp. 209–214. doi: 10.1016/S0014-5793(00)02044-5.

Gonzàlez, J. M. *et al.* (2008) 'Fast regulation of AP-1 activity through interaction of lamin A/C, ERK1/2, and c-Fos at the nuclear envelope', *Journal of Cell Biology*, 183(4), pp. 653–666. doi: 10.1083/jcb.200805049.

Gordon, W. R. *et al.* (2015) 'Mechanical Allostery: Evidence for a Force Requirement in the Proteolytic Activation of Notch', *Developmental Cell*, 33(6), pp. 729–736. doi: 10.1016/j.devcel.2015.05.004.

Görlich, D. *et al.* (1996) 'A 41 amino acid motif in importin-alpha confers binding to importin-beta and hence transit into the nucleus.', *The EMBO Journal*, 15(8), pp. 1810–1817. doi: 10.1002/j.1460-2075.1996.tb00530.x.

Gorynia, S. *et al.* (2011) 'Structural and functional insights into a dodecameric molecular machine - The RuvBL1/RuvBL2 complex', *Journal of Structural Biology*, 176(3), pp. 279–291. doi: 10.1016/j.jsb.2011.09.001.

Gracida-Jiménez, V. *et al.* (2017) 'Retrograde trafficking of β-dystroglycan from the plasma membrane to the nucleus', *Scientific Reports*, 7(1), pp. 1–14. doi: 10.1038/s41598-017-09972-x.

Grady, R. M. *et al.* (1999) 'Role for α-dystrobrevin in the pathogenesis of dystrophin-dependent muscular dystrophies', *Nature Cell Biology*, 1(4), pp. 215–220. doi: 10.1038/12034.

Grady, R. M. *et al.* (2000) 'Maturation and maintenance of the neuromuscular synapse: Genetic evidence for roles of the dystrophin-glycoprotein complex', *Neuron*, 25(2), pp. 279–293. doi: 10.1016/S0896-6273(00)80894-6.

Graumann, K., Runions, J. and Evans, D. E. (2010) 'Characterization of SUN-domain proteins at the higher plant nuclear envelope', *Plant Journal*, 61(1), pp. 134–144. doi: 10.1111/j.1365-313X.2009.04038.x.

Green, N. *et al.* (1982) 'Immunogenic structure of the influenza virus hemagglutinin', *Cell*, 28(3), pp. 477–487. doi: 10.1016/0092-8674(82)90202-1.

Grefen, C. and Blatt, M. R. (2012) 'A 2in1 cloning system enables ratiometric bimolecular fluorescence complementation (rBiFC)', *BioTechniques*, 53(5), pp. 311–314. doi: 10.2144/000113941.

Gruenbaum, Y. *et al.* (2005) 'The nuclear lamina comes of age', *Nature Reviews Molecular Cell Biology*, 6(1), pp. 21–31. doi: 10.1038/nrm1550.

Guelen, L. *et al.* (2008) 'Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions', *Nature*, 453(7197), pp. 948–951. doi: 10.1038/nature06947.

Guilluy, C. *et al.* (2014) 'Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus.', *Nature Cell Biology*, 16(4), pp. 376–81. doi: 10.1038/ncb2927.

Guo, X. and Wang, X. F. (2009) 'Signaling cross-talk between TGF-β/BMP and other pathways', *Cell Research*, 19(1), pp. 71–88. doi: 10.1038/cr.2008.302.

Guschin, D. Y. *et al.* (2010) *A rapid and general assay for monitoring endogenous gene modification*. Edited by J. M. Walker. New York, Dordrecht, Heidelberg, London: Springer Protocols.

Hagan, I. and Yanagida, M. (1995) 'The product of the spindle formation gene sad1+ associates with the fission yeast spindle pole body and is essential for viability', *Journal of Cell Biology*, 129(4), pp. 1033–1047. doi: 10.1083/jcb.129.4.1033.

Hanke, T., Szawlowski, P. and Randall, R. E. (1992) 'Construction of solid matrix-antibody-antigen complexes containing simian immunodeficiency virus p27 using tag-specific monoclonal antibody and tag-linked antigen', *Journal of General Virology*, 73(3), pp. 653–660. doi: 10.1099/0022-1317-73-3-653.

Haque, F. *et al.* (2006) 'SUN1 interacts with nuclear lamin A and cytoplasmic nesprins to provide a physical connection between the nuclear lamina and the cytoskeleton.', *Molecular and cellular biology*, 26(10), pp. 3738–51. doi: 10.1128/MCB.26.10.3738-3751.2006.

Haque, F. *et al.* (2010) 'Mammalian SUN protein interaction networks at the inner nuclear membrane and their role in laminopathy disease processes', *Journal of Biological Chemistry*, 285(5), pp. 3487–3498. doi: 10.1074/jbc.M109.071910.

Haque, F. (2011) 'Characterization of the topology, targeting and binding properties of SUN proteins and their involvement in laminopathies', (September). Available at: http://hdl.handle.net/2381/8984.

Hara, Y. *et al.* (2011) 'A dystroglycan mutation associated with limb-girdle muscular dystrophy.', *N Engl J Med*, 364(10), pp. 939–946. doi: 10.1056/NEJMoa1006939.

Hasan, S. *et al.* (2006) 'Nuclear envelope localization of human UNC84A does not require nuclear lamins', *FEBS Letters*, 580(5), pp. 1263–1268. doi: 10.1016/j.febslet.2006.01.039.

Hatten, M. E. (1999) 'Central nervous system neuronal migration', *Annual Review of Neuroscience*, 22, pp. 511–39.

Heitlinger, E. *et al.* (1991) 'Expression of Chicken Lamin B2 in Escherichia coli: Characterization of its Structure, Assembly, and Molecular Interactions E.', *Cell*, 113(3), pp. 485–495.

Heitlinger, E. *et al.* (1992) 'The role of the head and tail domain in lamin structure and assembly: Analysis of bacterially expressed chicken Lamin A and truncated B2 lamins', *Journal of Structural Biology*, 108(1), pp. 74–91. doi: 10.1016/1047-8477(92)90009-Y.

Hennen, J. *et al.* (2018) 'Fluorescence Fluctuation Spectroscopy Reveals Differential SUN Protein Oligomerization In Living Cells', *Molecular Biology of the Cell*, 29, p. mbc.E17-04-0233. doi: 10.1091/mbc.E17-04-0233.

Hicke, L. and Dunn, R. (2003) 'Regulation of Membrane Protein Transport by Ubiquitin and Ubiquitin-Binding Proteins', *Annual Review of Cell and Developmental Biology*, 19, pp. 141–172. doi: 10.1146/annurev.cellbio.19.110701.154617.

Higginson, J. R., Thompson, O. and Winder, S. J. (2008) 'Targeting of dystroglycan to the cleavage furrow and midbody in cytokinesis', *International Journal of Biochemistry and Cell Biology*, 40(5), pp. 892–900. doi: 10.1016/j.biocel.2007.10.019.

Hinz, B. *et al.* (2001) 'Alpha-smooth muscle actin expression upregulates fibroblast contractile activity', *Molecular Biology of the Cell*, 12(9), pp. 2730–2741. doi: 10.1091/mbc.12.9.2730.

Hinz, B. *et al.* (2012) 'Recent developments in myofibroblast biology: Paradigms for connective tissue remodeling', *American Journal of Pathology*, 180(4), pp. 1340–1355. doi: 10.1016/j.ajpath.2012.02.004.

Ho, C. Y. *et al.* (2013) 'Lamin A/C and emerin regulate MKL1-SRF activity by modulating actin dynamics', *Nature*, 497(7450), pp. 507–513. doi: 10.1038/nature12105.

Ho, C. Y. and Lammerding, J. (2012) 'Lamins at a glance', *Journal of Cell Science*, 125(9), pp. 2087–2093. doi: 10.1242/jcs.087288.

Hodzic, D. M. *et al.* (2004) 'Sun2 is a novel mammalian inner nuclear membrane protein', *Journal of Biological Chemistry*, 279(24), pp. 25805–25812. doi: 10.1074/jbc.M313157200.

Hoelz, A. and Blobel, G. (2004) 'Popping out of the nucleus', *Nature*, 432, pp. 815–816. doi: 10.4103/2347-9906.184181.

Hoelz, A., Debler, E. W. and Blobel, G. (2011) 'The Structure of the Nuclear Pore Complex', *Annual Review of Biochemistry*, 80(1), pp. 613–643. doi: 10.1146/annurev-biochem-060109-151030.

Hoffman, E. P. *et al.* (1988) 'Characterization of Dystrophin in Muscle-Biopsy Specimens from Patients with Duchenne's or Becker's Muscular Dystrophy', *New England Journal of Medicine*. doi: 10.1056/NEJM198805263182104.

Hoffman, E. P., Brown, R. H. and Kunkel, L. M. (1987) 'Dystrophin : The Protein Product of the Duchenne Muscular Dystrophy Locus', *Cell*, 51, pp. 919–928.

Höger, T. H. *et al.* (1990) 'Characterization of a second highly conserved B-type lamin present in cells previously thought to contain only a single B-type lamin', *Chromosoma*, 99(6), pp. 379–390. doi: 10.1007/BF01726689.

Holt, I. *et al.* (2003) 'Effect of pathogenic mis-sense mutations in lamin A on its interaction with emerin in vivo', *Journal of Cell Science*, 116(14), pp. 3027–3035. doi: 10.1242/jcs.00599.

Holt, I. *et al.* (2019) 'Nesprin-1-alpha2 associates with kinesin at myotube outer nuclear membranes, but is restricted to neuromuscular junction nuclei in adult muscle', *Scientific Reports*, 9(1), pp. 1–12. doi: 10.1038/s41598-019-50728-6.

Holt, K. H. *et al.* (2000) 'Biosynthesis of dystroglycan: Processing of a precursor propeptide', *FEBS Letters*, 468(1), pp. 79–83. doi: 10.1016/S0014-5793(00)01195-9.

Holt, K. H. and Campbell, K. P. (1998) 'Assembly of the Sarcoglycan Complex', *Journal of Biological Chemistry*, 273(52), pp. 34667–34670. doi: 10.1074/jbc.273.52.34667.

Horigome, C. *et al.* (2014) 'SWR1 and INO80 chromatin remodelers contribute to DNA double-strand break perinuclear anchorage site choice', *Molecular Cell*, 55(4), pp. 626–639. doi: 10.1016/j.molcel.2014.06.027.

Horn, H. F., Kim, D. I., *et al.* (2013) 'A mammalian KASH domain protein coupling meiotic chromosomes to the cytoskeleton', *Journal of Cell Biology*, 202(7), pp. 1023–1039. doi: 10.1083/jcb.201304004.

Horn, H. F., Brownstein, Z., *et al.* (2013) 'The LINC complex is essential for hearing', *Journal of Clinical Investigation*, 123(2), pp. 740–750. doi: 10.1172/JCI66911.

Horstman, A. *et al.* (2014) 'A Cautionary note on the use of split-YFP/BiFC in plant protein-protein interaction studies', *International Journal of Molecular Sciences*, 15(6), pp. 9628–9643. doi: 10.3390/ijms15069628.

Hosokawa, H. *et al.* (2002) 'Vascular endothelial cells that express dystroglycan are involved in angiogenesis', *Journal of Cell Science*, 115(7), pp. 1487–1496.

Hu, B., Wu, Z. and Phan, S. H. (2003) 'Smad3 mediates transforming growth factor-β-induced α-smooth muscle actin expression', *American Journal of Respiratory Cell and Molecular Biology*, 29(3 I), pp. 397–404. doi:

10.1165/rcmb.2003-0063OC.

Hu, C. D., Chinenov, Y. and Kerppola, T. K. (2002) 'Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation', *Molecular Cell*, 9(4), pp. 789–798. doi: 10.1016/S1097-2765(02)00496-3.

Huang, X. *et al.* (2000) 'Structure of a WW domain containing fragment of dystrophin in complex with β-dystroglycan', *Nature Structural Biology*, 7(8), pp. 634–638. doi: 10.1038/77923.

Huard, J. *et al.* (1991) 'Dystrophin expression in myotubes formed by the fusion of normal and dystrophic myoblasts', *Muscle & Nerve*, 14(2), pp. 178–182. doi: 10.1002/mus.880140213.

Huberts, D. H. E. W. and van der Klei, I. J. (2010) 'Moonlighting proteins: An intriguing mode of multitasking', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1803(4), pp. 520–525. doi: 10.1016/j.bbamcr.2010.01.022.

Humphrey, E. L. *et al.* (2015) 'A new monoclonal antibody DAG-6F4 against human alpha-dystroglycan reveals reduced core protein in some, but not all, dystroglycanopathy patients', *Neuromuscular Disorders*, 25(1), pp. 32–42. doi: 10.1016/j.nmd.2014.09.005.

Hung, V. *et al.* (2014) 'Resource Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging', *Molecular Cell*, 55(2), pp. 332–341. doi: 10.1016/j.molcel.2014.06.003.

Ibraghimov-beskrovnaya, O. *et al.* (1993) 'Human dystroglycan: Skeletal muscle cDNA, genomic structure, origin of tissue specific isoforms and chromosomal localization', *Human Molecular Genetics*, 2(10), pp. 1651–1657. doi: 10.1093/hmg/2.10.1651.

Ibraghimov-Beskrovnaya, O. *et al.* (1992) 'Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix', *Nature*, 355(6362), pp. 696–702. doi: 10.1038/355696a0.

Ilsley, J. L., Sudol, M. and Winder, S. J. (2001) 'The interaction of dystrophin with beta-dystroglycan is regulated by tyrosine phosphorylation', *Cellular signalling*, 13, pp. 625–632.

Inamori, K. I. *et al.* (2012) 'Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE', *Science*, 335(6064), pp. 93–96. doi: 10.1126/science.1214115.

Ishikawa-Sakurai, M. *et al.* (2004) 'ZZ domain is essentially required for the physiological binding of dystrophin and utrophin to beta-dystroglycan', *Human Molecular Genetics*, 13(7), pp. 693–702. doi: 10.1093/hmg/ddh087.

Ishimura, A. *et al.* (2006) 'Man1, an inner nuclear membrane protein, regulates vascular remodeling by modulating transforming growth factor β signaling', *Development*, 133(19), pp. 3919–3928. doi: 10.1242/dev.02538.

Ito, A. *et al.* (2015) 'The subcellular localization and activity of cortactin is regulated by acetylation and interaction with Keap1', *Science Signaling*, 8(404), pp.

1–12. doi: 10.1126/scisignal.aad0667.

Ivorra, C. *et al.* (2006) 'A mechanism of AP-1 suppression through interaction of c-Fos with lamin A/C', *Genes and Development*, 20(6), p. 747. doi: 10.1101/gad.349506.5.

Jackson, M. F. *et al.* (2013) 'Genetic manipulation of myoblasts and a novel primary myosatellite cell culture system: Comparing and optimizing approaches', *FEBS Journal*, 280(3), pp. 827–839. doi: 10.1111/febs.12072.

Jackson, S. N. J. *et al.* (1997) 'Dunnigan-Kobberling syndrome: An autosomal dominant form of partial lipodystrophy', *QJM - Monthly Journal of the Association of Physicians*, 90(1), pp. 27–36. doi: 10.1093/qjmed/90.1.27.

Jahed, Z. *et al.* (2018) 'Molecular insights into the mechanisms of SUN1 oligomerization in the nuclear envelope', *Biophysical Journal*, 114(5), pp. 1190–1203. doi: 10.1016/j.bpj.2018.01.015.

James, M. *et al.* (2000) 'Adhesion-dependent tyrosine phosphorylation of β-dystroglycan regulates its interaction with utrophin', *Journal of Cell Science*, 113(10), pp. 1717–1726.

Janin, A. *et al.* (2017) 'Nuclear envelopathies: A complex LINC between nuclear envelope and pathology', *Orphanet Journal of Rare Diseases*, 12(1), pp. 1–16. doi: 10.1186/s13023-017-0698-x.

Janin, A. *et al.* (2018) 'SMAD6 overexpression leads to accelerated myogenic differentiation of LMNA mutated cells', *Scientific Reports*, 8(1), pp. 1–15. doi: 10.1038/s41598-018-23918-x.

Janin, A. and Gache, V. (2018) 'Nesprins and lamins in health and diseases of cardiac and skeletal muscles', *Frontiers in Physiology*, 9(SEP), pp. 1–12. doi: 10.3389/fphys.2018.01277.

Jeffery, C. J. (1999) 'Moonlighting proteins', *Trends in Biochemical Sciences*, 24(1), pp. 8–11.

Jeffery, C. J. (2003) 'Multifunctional proteins: Examples of gene sharing', *Annals of Medicine*, 35(1), pp. 28–35. doi: 10.1080/07853890310004101.

Jimenez-Gutierrez, G. E. *et al.* (2020) 'Loss of dystroglycan drives cellular senescence via defective mitosis-mediated genomic instability', *International Journal of Molecular Sciences*, 21(14), pp. 1–18. doi: 10.3390/ijms21144961.

Jimenez-Mallebrera, C. *et al.* (2009) 'A comparative study of α-dystroglycan glycosylation in dystroglycanopathies suggests that the hypoglycosylation of α-dystroglycan does not consistently correlate with clinical severity', *Brain Pathology*, 19(4), pp. 596–611. doi: 10.1111/j.1750-3639.2008.00198.x.

Jin, J. *et al.* (2005) 'A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex', *Journal of Biological Chemistry*, 280(50), pp. 41207–41212. doi: 10.1074/jbc.M509128200.

Jinek, M. *et al.* (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science*, 337(6096), pp. 816–821. doi:

10.1126/science.1225829.

Johnson, E. K. *et al.* (2013) 'Identification of New Dystroglycan Complexes in Skeletal Muscle', *PLoS ONE*, 8(8). doi: 10.1371/journal.pone.0073224.

Jónsson, Z. O. *et al.* (2001) 'Rvb1p and Rvb2p are Essential Components of a Chromatin Remodeling Complex that Regulates Transcription of over 5% of Yeast Genes', *Journal of Biological Chemistry*, 276(19), pp. 16279–16288. doi: 10.1074/jbc.M011523200.

Jung, D. *et al.* (1995) 'Identification and characterization of the dystrophin anchoring site on ??-dystroglycan', *Journal of Biological Chemistry*, 270(45), pp. 27305–27310. doi: 10.1074/jbc.270.45.27305.

Jung, D. *et al.* (1996) 'Characterization of δ-sarcoglycan, a novel component of the oligomeric sarcoglycan complex involved in limb-girdle muscular dystrophy', *Journal of Biological Chemistry*, 271(50), pp. 32321–32329. doi: 10.1074/jbc.271.50.32321.

Kalderon, D., Roberts, B. L., *et al.* (1984) 'A short amino acid sequence able to specify nuclear location', *Cell*, 39(3 PART 2), pp. 499–509. doi: 10.1016/0092-8674(84)90457-4.

Kalderon, D., Richardson, W. D., *et al.* (1984) 'Sequence requirements for nuclear location of simian virus 40 large-T antigen', *Nature*, 311(5981), pp. 33–38. doi: 10.1038/311033a0.

Kanagawa, M. and Toda, T. (2006) 'The genetic and molecular basis of muscular dystrophy: Roles of cell-matrix linkage in the pathogenesis', *Journal of Human Genetics*, 51(11), pp. 915–926. doi: 10.1007/s10038-006-0056-7.

Kanagawa, M. and Toda, T. (2017) 'Muscular Dystrophy with Ribitol-Phosphate Deficiency: A Novel Post-Translational Mechanism in Dystroglycanopathy', *Journal of Neuromuscular Diseases*, 4(4), pp. 259–267. doi: 10.3233/JND-170255.

Karabinos, A. *et al.* (2003) 'The single nuclear lamin of Caenorhabditis elegans forms in vitro stable intermediate filaments and paracrystals with a reduced axial periodicity', *Journal of Molecular Biology*, 325(2), pp. 241–247. doi: 10.1016/S0022-2836(02)01240-8.

Kavanagh, D. M. *et al.* (2007) 'Organelle proteome variation among different cell types: Lessons from nuclear membrane proteins', *Subcellular Biochemistry*, 43, pp. 51–76. doi: 10.1007/978-1-4020-5943-8_5.

Keays, D. A. *et al.* (2007) 'Mutations in α-Tubulin Cause Abnormal Neuronal Migration in Mice and Lissencephaly in Humans', *Cell*, 128(1), pp. 45–57. doi: 10.1016/j.cell.2006.12.017.

Kerppola, T. K. (2008) 'Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells', *Annual Review of Biophysics*, 37, pp. 465–487. doi: 10.1146/annurev.biophys.37.032807.125842.

Ketema, M. *et al.* (2013) 'Nesprin-3 connects plectin and vimentin to the nuclear envelope of Sertoli cells but is not required for Sertoli cell function in spermatogenesis', *Molecular Biology of the Cell*, 24(15), pp. 2454–2466. doi:

10.1091/mbc.E13-02-0100.

Khairallah, M. *et al.* (2008) 'Sildenafil and cardiomyocyte-specific cGMP signaling prevent cardiomyopathic changes associated with dystrophin deficiency', *Proceedings of the National Academy of Sciences of the United States of America*, 105(19), pp. 7028–7033. doi: 10.1073/pnas.0710595105.

Kharraz, Y. *et al.* (2014) 'Understanding the process of fibrosis in duchenne muscular dystrophy', *BioMed Research International*, 2014. doi: 10.1155/2014/965631.

Khurana, T. S. *et al.* (1991) 'Immunolocalization and developmental expression of dystrophin related protein in skeletal muscle', *Neuromuscular Disorders*, 1(3), pp. 185–194. doi: 10.1016/0960-8966(91)90023-L.

Kidd, S., Lieber, T. and Young, M. W. (1998) 'Ligand-induced cleavage and regulation of nuclear entry of Notch in Drosophila melanogaster embryos', *Genes and Development*, 12(23), pp. 3728–3740. doi: 10.1101/gad.12.23.3728.

Kim, D. I. *et al.* (2014) 'Probing nuclear pore complex architecture with proximity-dependent biotinylation', *Proceedings of the National Academy of Sciences*, 111(24), pp. E2453–E2461. doi: 10.1073/pnas.1406459111.

Kim, H. and Kim, J. S. (2014) 'A guide to genome engineering with programmable nucleases', *Nature Reviews Genetics*, 15(5), pp. 321–334. doi: 10.1038/nrg3686.

Kim, K. K., Sheppard, D. and Chapman, H. A. (2018) 'TGF-β1 signaling and tissue fibrosis', *Cold Spring Harbor Perspectives in Biology*, 10(4). doi: 10.1101/cshperspect.a022293.

Kim, S. H. *et al.* (2004) 'Evidence that assembly of an active γ-secretase complex occurs in the early compartments of the secretory pathway', *Journal of Biological Chemistry*, 279(47), pp. 48615–48619. doi: 10.1074/jbc.C400396200.

Kim, Y. H., Han, M. E. and Oh, S. O. (2017) 'The molecular mechanism for nuclear transport and its application', *Anatomy and Cell Biology*, 50(2), pp. 77–85. doi: 10.5115/acb.2017.50.2.77.

King, M. C., Lusk, C. P. and Blobel, G. (2006) 'Karyopherin-mediated import of integral inner nuclear membrane proteins', *Nature*, 442(7106), pp. 1003–1007. doi: 10.1038/nature05075.

Kisseberth, W. C. *et al.* (1999) 'Ubiquitous expression of marker transgenes in mice and rats', *Developmental Biology*, 214(1), pp. 128–138. doi: 10.1006/dbio.1999.9417.

Kitten, G. T. and Nigg, E. A. (1991) 'The CaaX motif is required for isoprenylation, carboxyl methylation, and nuclear membrane association of lamin B2', *Journal of Cell Biology*, 113(1), pp. 13–23. doi: 10.1083/jcb.113.1.13.

Kodama, Y. and Hu, C. D. (2012) 'Bimolecular fluorescence complementation (BiFC): A 5-year update and future perspectives', *BioTechniques*, 53(5), pp. 285–298. doi: 10.2144/000113943.

Koenig, M. *et al.* (1987) 'Complete cloning of the duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals', *Cell*, 50(3), pp. 509–517. doi: 10.1016/0092-8674(87)90504-6.

Kohwi, M. *et al.* (2013) 'Developmentally regulated subnuclear genome reorganization restricts neural progenitor competence in Drosophila', *Cell*, 152(1–2), pp. 97–108. doi: 10.1016/j.cell.2012.11.049.

Koncicka, M. *et al.* (2020) 'Expression of lamin C2 in mammalian oocytes', *PLoS ONE*, 15(4), pp. 9–11. doi: 10.1371/journal.pone.0229781.

Kong, Y. *et al.* (1997) 'Muscle LIM protein promotes myogenesis by enhancing the activity of MyoD.', *Molecular and Cellular Biology*, 17(8), pp. 4750–4760. doi: 10.1128/mcb.17.8.4750.

König, E. *et al.* (2017) 'Exploring digenic inheritance in arrhythmogenic cardiomyopathy', *BMC Medical Genetics*, 18(1), pp. 1–12. doi: 10.1186/s12881-017-0503-7.

Korfali, N. *et al.* (2010) 'The leukocyte nuclear envelope proteome varies with cell activation and contains novel transmembrane proteins that affect genome architecture', *Molecular and Cellular Proteomics*, 9(12), pp. 2571–2585. doi: 10.1074/mcp.M110.002915.

Korfali, N. *et al.* (2012) 'The nuclear envelope proteome differs notably between tissues', *Nucleus (United States)*, 3(6). doi: 10.4161/nucl.22257.

Kottlors, M. and Kirschner, J. (2010) 'Elevated satellite cell number in Duchenne muscular dystrophy', *Cell and Tissue Research*, 340(3), pp. 541–548. doi: 10.1007/s00441-010-0976-6.

Kozono, T. *et al.* (2018) 'Jaw1/LRMP has a role in maintaining nuclear shape via interaction with SUN proteins', *Journal of Biochemistry*, 164(4), pp. 303–311. doi: 10.1093/jb/mvy053.

Kracklauer, M. P. *et al.* (2007) 'Drosophila klaroid encodes a SUN domain protein required for klarsicht localization to the nuclear envelope and nuclear migration in the eye', *Fly*, 1(2), pp. 75–85. doi: 10.4161/fly.4254.

Krohne, G. *et al.* (1987) 'Nuclear lamin LI of Xenopus laevis: cDNA cloning, amino acid sequence and binding specificity of a member of the lamin B subfamily.', *The EMBO Journal*, 6(12), pp. 3801–3808. doi: 10.1002/j.1460-2075.1987.tb02716.x.

Krohne, G., Dabauvalle, M. C. and Franke, W. W. (1981) 'Cell type-specific differences in protein composition of nuclear pore complex-lamina structures in oocytes and erythrocytes of Xenopus laevis', *Journal of Molecular Biology*, 151(1), pp. 121–141. doi: 10.1016/0022-2836(81)90224-2.

Kucherenko, M. M. *et al.* (2008) 'Genetic modifier screens reveal new components that interact with the Drosophila Dystroglycan-Dystrophin complex', *PLoS ONE*, 3(6). doi: 10.1371/journal.pone.0002418.

Kutscheidt, S. *et al.* (2014) 'FHOD1 interaction with nesprin-2G mediates TAN line formation and nuclear movement', *Nature Cell Biology*, 16(7), pp. 708–715. doi:

378

10.1038/ncb2981.

Kwon, M. *et al.* (2008) 'Mechanisms to suppress multipolar divisions in cancer cells with extra centrosomes', *Genes and Development*, 22(16), pp. 2189–2203. doi: 10.1101/gad.1700908.

Lammerding, J. *et al.* (2004) 'Lamin A / C deficiency causes defective nuclear mechanics and mechanotransduction', *The Journal of Clinical Investigation*, 113(3), pp. 370–378. doi: 10.1172/JCI200419670.Introduction.

Lammerding, J. *et al.* (2006) 'Lamins A and C but not lamin B1 regulate nuclear mechanics', *Journal of Biological Chemistry*, 281(35), pp. 25768–25780. doi: 10.1074/jbc.M513511200.

Lammich, S. *et al.* (2002) 'Presenilin-dependent intramembrane proteolysis of CD44 leads to the liberation of its intracellular domain and the secretion of an Aβ-like peptide', *Journal of Biological Chemistry*, 277(47), pp. 44754–44759. doi: 10.1074/jbc.M206872200.

Lara-Chacón, B. *et al.* (2010) 'Characterization of an importin in α/β-recognized nuclear localization signal in β-dystroglycan', *Journal of Cellular Biochemistry*, 110(3), pp. 706–717. doi: 10.1002/jcb.22581.

de las Heras, J. I. *et al.* (2017) 'Tissue-specific NETs alter genome organization and regulation even in a heterologous system', *Nucleus*, 8(1), pp. 81–97. doi: 10.1080/19491034.2016.1261230.

Lau, K. S. *et al.* (1998) 'Skeletal muscle contractions stimulate cGMP formation and attenuate vascular smooth muscle myosin phosphorylation via nitric oxide', *FEBS Letters*, 431(1), pp. 71–74. doi: 10.1016/S0014-5793(98)00728-5.

Lawrence, K. S. *et al.* (2016) 'LINC complexes promote homologous recombination in part through inhibition of nonhomologous end joining', *Journal of Cell Biology*, 215(6), pp. 801–821. doi: 10.1083/jcb.201604112.

Le, H. Q. *et al.* (2016) 'Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment', *Nature Cell Biology*, 18(8), pp. 864–875. doi: 10.1038/ncb3387.

Lebakken, C. S. *et al.* (2000) 'Sarcospan-Deficient Mice Maintain Normal Muscle Function', *Molecular and Cellular Biology*, 20(5), pp. 1669–1677. doi: 10.1128/mcb.20.5.1669-1677.2000.

Lecker, S. H., Goldberg, A. L. and Mitch, W. E. (2006) 'Protein degradation by the ubiquitin-proteasome pathway in normal and disease states', *Journal of the American Society of Nephrology*, 17(7), pp. 1807–1819. doi: 10.1681/ASN.2006010083.

Lee, K. K. *et al.* (2002) 'Lamin-dependent localization of UNC-84, a protein required for nuclear migration in Caenorhabditis elegans', *Molecular Biology of the Cell*, 13(3), pp. 892–901. doi: 10.1091/mbc.01-06-0294.

Legate, K. R., Wickström, S. A. and Fässler, R. (2009) 'Genetic and cell biological analysis of integrin outside-in signaling', *Genes and Development*, 23(4), pp. 397–418. doi: 10.1101/gad.1758709.

Lei, K. *et al.* (2009) 'SUN1 and SUN2 play critical but partially redundant roles in anchoring nuclei in skeletal muscle cells in mice.', *Proceedings of the National Academy of Sciences of the United States of America*, 106(25), pp. 10207–10212. doi: 10.1073/pnas.0812037106.

Lei, K. *et al.* (2012) 'Inner nuclear envelope proteins SUN1 and SUN2 play a prominent role in the DNA damage response', *Current Biology*, 22(17), pp. 1609–1615. doi: 10.1016/j.cub.2012.06.043.

Leibovitz, Z. *et al.* (2018) 'Walker-Warburg syndrome and tectocerebellar dysraphia: A novel association caused by a homozygous DAG1 mutation', *European Journal of Paediatric Neurology*, 22(3), pp. 525–531. doi: 10.1016/j.ejpn.2017.12.012.

Leocadio-Victoria, D. (2015) *Post-translational modification and nuclear targeting of beta-dystroglycan.* The University of Sheffield.

Leocadio, D., Mitchell, A. and Winder, S. (2016) 'γ-Secretase Dependent Nuclear Targeting of Dystroglycan', *J Cell Biochem*, 117(9), pp. 2149–2157.

Leung, D. G. *et al.* (2014) 'Sildenafil does not improve cardiomyopathy in Duchenne/Becker muscular dystrophy', *Annals of Neurology*, 76(4), pp. 541–549. doi: 10.1002/ana.24214.

Levy, D. L. and Heald, R. (2010) 'Nuclear Size Is Regulated by Importin α and Ntf2 in Xenopus', *Cell*, 143(2), pp. 288–298. doi: 10.1016/j.cell.2010.09.012.

Li, P. *et al.* (2014) 'Contribution of SUN1 Mutations to the Pathomechanism in Muscular Dystrophies', *Human Mutation*, 35(4), pp. 452–461. doi: 10.1002/humu.22504.

Libotte, T. *et al.* (2005) 'Lamin A/C–dependent Localization of Nesprin-2, a Giant Scaffolder at the Nuclear Envelope', *Mol Biol Cell*, 16(December), pp. 3411–3424. doi: 10.1091/mbc.E04.

Lim, L. E. *et al.* (1995) 'β–sarcoglycan: Characterization and role in limb–girdle muscular dystrophy linked to 4q12', *Nature Genetics*, 11(3), pp. 257–265. doi: 10.1038/ng1195-257.

Lin, D. H. *et al.* (2016) 'Architecture of the symmetric core of the nuclear pore', *Science*, 352(6283). doi: 10.1126/science.aaf1015.

Lindeman, R. E. and Pelegri, F. (2012) 'Localized products of futile cycle/lrmp promote centrosome-nucleus attachment in the zebrafish zygote', *Current Biology*, 22(10), pp. 843–851. doi: 10.1016/j.cub.2012.03.058.

Link, J. *et al.* (2013) 'The Meiotic Nuclear Lamina Regulates Chromosome Dynamics and Promotes Efficient Homologous Recombination in the Mouse', *PLoS Genetics*, 9(1). doi: 10.1371/journal.pgen.1003261.

Link, J. *et al.* (2014) 'Analysis of Meiosis in SUN1 Deficient Mice Reveals a Distinct Role of SUN2 in Mammalian Meiotic LINC Complex Formation and Function', *PLoS Genetics*, 10(2), pp. 14–21. doi: 10.1371/journal.pgen.1004099.

Lipscomb, L. *et al.* (2011) 'The proteasomal inhibitor MG132 prevents muscular

dystrophy in zebrafish', *PLoS Currents*. doi: 10.1371/currents.RRN1286.

Lipscomb, L. *et al.* (2016) 'Dasatinib as a treatment for Duchenne muscular dystrophy', *Human Molecular Genetics*, 25(2), pp. 266–274. doi: 10.1093/hmg/ddv469.

Liu, D., Black, B. L. and Derynck, R. (2001) 'TGF-β inhibits muscle differentiation through functional repression of myogenic transcription factors by Smad3', *Genes and Development*, 15(22), pp. 2950–2966. doi: 10.1101/gad.925901.

Liu, Q. *et al.* (2007) 'Functional association of Sun1 with nuclear pore complexes', *Journal of Cell Biology*, 178(5), pp. 785–798. doi: 10.1083/jcb.200704108.

Lloyd, D. J., Trembath, R. C. and Shackleton, S. (2002) 'A novel interaction between lamin A and SREBP1: Implications for partial lipodystrophy and other laminopathies', *Human Molecular Genetics*, 11(7), pp. 769–777. doi: 10.1093/hmg/11.7.769.

Lönn, P. and Landegren, U. (2017) 'Close Encounters – Probing Proximal Proteins in Live or Fixed Cells', *Trends in Biochemical Sciences*, 42(7), pp. 504–515. doi: 10.1016/j.tibs.2017.05.003.

Loo, T. H. *et al.* (2019) 'The mammalian LINC complex component SUN1 regulates muscle regeneration by modulating drosha activity', *eLife*, 8, pp. 1–25. doi: 10.7554/eLife.49485.

De Los Santos, C. *et al.* (2015) 'FRAP, FLIM, and FRET: Detection and analysis of cellular dynamics on a molecular scale using fluorescence microscopy', *Molecular Reproduction and Development*, pp. 587–604. doi: 10.1002/mrd.22501.

Losasso, C. *et al.* (2000) 'Anomalous dystroglycan in carcinoma cell lines', *FEBS Letters*, 484, pp. 194–198. doi: 10.1016/S0014-5793(00)02157-8.

Lössl, P., Waterbeemd, M. and Heck, A. J. (2016) 'The diverse and expanding role of mass spectrometry in structural and molecular biology', *The EMBO Journal*, 35(24), pp. 2634–2657. doi: 10.15252/embj.201694818.

Lott, K. and Cingolani, G. (2011) 'The importin β binding domain as a master regulator of nucleocytoplasmic transport', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1813(9), pp. 1578–1592. doi: 10.1016/j.bbamcr.2010.10.012.

Lottersberger, F. *et al.* (2015) '53BP1 and the LINC Complex Promote Microtubule-Dependent DSB Mobility and DNA Repair', *Cell*, 163(4), pp. 880–893. doi: 10.1016/j.cell.2015.09.057.

Lu, W. *et al.* (2008) 'Sun1 forms immobile macromolecular assemblies at the nuclear envelope', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1783(12), pp. 2415–2426. doi: 10.1016/j.bbamcr.2008.09.001.

Lu, W. *et al.* (2012) 'Nesprin interchain associations control nuclear size', *Cellular and Molecular Life Sciences*, 69(20), pp. 3493–3509. doi: 10.1007/s00018-012-1034-1.

Lua, B. L. and Low, B. C. (2005) 'Cortactin phosphorylation as a switch for actin

cytoskeletal network and cell dynamics control', *FEBS Letters*, 579(3), pp. 577–585. doi: 10.1016/j.febslet.2004.12.055.

Lusk, C. P., Blobel, G. and King, M. C. (2007) 'Highway to the inner nuclear membrane: rules for the road', *Nature Reviews Molecular Cell Biology*, 8, pp. 414–420.

Luxton, G. W. G. *et al.* (2010) 'Linear Arrays of Nuclear Envelope Proteins Harness Retrograde Actin Flow for Nuclear Movement', *Science*, 956(August), pp. 956–960. doi: 10.1126/science.1189072.

MacDonald, E. M. and Cohn, R. D. (2012) 'TGFβ signaling: Its role in fibrosis formation and myopathies', *Current Opinion in Rheumatology*, 24(6), pp. 628–634. doi: 10.1097/BOR.0b013e328358df34.

Machiels, B. M. *et al.* (1996) 'An alternative splicing product of the lamin A/C gene lacks exon 10', *Journal of Biological Chemistry*, 271(16), pp. 9249–9253. doi: 10.1074/jbc.271.16.9249.

Magliery, T. J. *et al.* (2005) 'Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: Scope and mechanism', *Journal of the American Chemical Society*, 127(1), pp. 146–157. doi: 10.1021/ja046699g.

Mahamid, J. *et al.* (2016) 'Visualizing the molecular sociology at the HeLa cell nuclear periphery', *Science*, 351(6276), pp. 969–972. doi: 10.1126/science.aad8857.

Mallampalli, M. P. *et al.* (2005) 'Inhibiting farnesylation reverses the nuclear morphology defect in a HeLa cell model for Hutchinson-Gilford progeria syndrome', *Proceedings of the National Academy of Sciences of the United States of America*, 102(40), pp. 14416–14421. doi: 10.1073/pnas.0503712102.

Malone, C. J. *et al.* (1999) 'UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during C. elegans development', *Development*, 126(14), pp. 3171–3181.

Mamchaoui, K. *et al.* (2011) 'Immortalized pathological human myoblasts: Towards a universal tool for the study of neuromuscular disorders', *Skeletal Muscle*, 1(1). doi: 10.1186/2044-5040-1-34.

Manning, J. and O'Malley, D. (2015) 'What has the mdx mouse model of duchenne muscular dystrophy contributed to our understanding of this disease?', *Journal of Muscle Research and Cell Motility*, 36(2), pp. 155–167. doi: 10.1007/s10974-015-9406-4.

Mannix, K. M. *et al.* (2019) 'Proximity labeling reveals novel interactomes in live Drosophila tissue', *Development (Cambridge, England)*, 146(14). doi: 10.1242/dev.176644.

Manya, H. *et al.* (2004) 'Demonstration of mammalian protein O-mannosyltransferase activity: Coexpression of POMT1 and POMT2 required for enzymatic activity', *Proceedings of the National Academy of Sciences of the United States of America*, 101(2), pp. 500–505. doi: 10.1073/pnas.0307228101.

Manzur, A. Y. *et al.* (2008) 'Glucocorticoid corticosteroids for Duchenne muscular dystrophy', *Cochrane Database of Systematic Reviews*, (1). doi:

10.1002/14651858.CD003725.pub3.

Markiewicz, E., Ledran, M. and Hutchison, C. J. (2005) 'Remodelling of the nuclear lamina and nucleoskeleton is required for skeletal muscle differentiation in vitro', *Journal of Cell Science*, 118(2), pp. 409–420. doi: 10.1242/jcs.01630.

Marnef, A. *et al.* (2019) 'A cohesin/HUSH- And LINC-dependent pathway controls ribosomal DNA double-strand break repair', *Genes and Development*, 33(17–18), pp. 1175–1190. doi: 10.1101/gad.324012.119.

Marshall, J. L., Chou, E., *et al.* (2012) 'Dystrophin and utrophin expression require sarcospan: Loss of α7 integrin exacerbates a newly discovered muscle phenotype in sarcospan-null mice', *Human Molecular Genetics*, 21(20), pp. 4378–4393. doi: 10.1093/hmg/dds271.

Marshall, J. L., Holmberg, J., *et al.* (2012) 'Sarcospan-dependent Akt activation is required for utrophin expression and muscle regeneration', *Journal of Cell Biology*, 197(7), pp. 1009–1027. doi: 10.1083/jcb.201110032.

Marshall, J. L. and Crosbie-Watson, R. H. (2013) 'Sarcospan: A small protein with large potential for Duchenne muscular dystrophy', *Skeletal Muscle*, 3(1). doi: 10.1186/2044-5040-3-1.

Martin, D. K. and Whitelaw, E. (1996) 'The vagaries of variegating transgenes', *BioEssays*, 18(11), pp. 919–923.

Martin, P. T. (2005) 'The dystroglycanopathies: The new disorders of o-linked glycosylation', *Seminars in Pediatric Neurology*, 12(3), pp. 152–158. doi: 10.1016/j.spen.2005.10.003.

Martínez-Vieyra, I. A. *et al.* (2013) 'A role for β-dystroglycan in the organization and structure of the nucleus in myoblasts', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1833(3), pp. 698–711. doi: 10.1016/j.bbamcr.2012.11.019.

Mathew, G. *et al.* (2013) 'Nuclear targeting of dystroglycan promotes the expression of androgen regulated transcription factors in prostate cancer.', *Scientific reports*, 3, p. 2792. doi: 10.1038/srep02792.

Matsumoto, A. *et al.* (2016) 'Loss of the integral nuclear envelope protein SUN1 induces alteration of nucleoli', *Nucleus*, 7(1), pp. 68–83. doi: 10.1080/19491034.2016.1149664.

Matunis, M. J., Wu, J. and Blobel, G. (1998) 'SUMO-1 modification and its role in targeting the Ran GTPase-activating protein, RanGAP1, to the nuclear pore complex', *Journal of Cell Biology*, 140(3), pp. 499–509. doi: 10.1083/jcb.140.3.499.

Mauro, A. (1961) 'Satellite cell of skeletal muscle fibers', *The Journal of Biophysical and Biochemical Cytology*, 9, pp. 493–495. doi: 10.1083/jcb.9.2.493.

McCord, R. P. *et al.* (2013) 'Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome', *Genome Research*, 23(2), pp. 260–269. doi: 10.1101/gr.138032.112.

McGee, M. D. *et al.* (2006) 'UNC-83 Is a KASH Protein Required for Nuclear Migration and Is Recruited to the Outer Nuclear Membrane by a Physical Interaction

with the SUN Protein UNC-84', *Molecular Biology of the Cell*, 17, pp. 1790–1801.

Mckeon, F. D., Kirschner, M. W. and Caput, D. (1986) 'Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins', *Nature*, 319(6053), pp. 463–468. doi: 10.1038/319463a0.

Meinema, A. C. *et al.* (2011) 'Long unfolded linkers facilitate membrane protein import through the nuclear pore complex', *Science*, 333(6038), pp. 90–93. doi: 10.1126/science.1205741.

Meinke, P. *et al.* (2014) 'Muscular Dystrophy-Associated SUN1 and SUN2 Variants Disrupt Nuclear-Cytoskeletal Connections and Myonuclear Organization', *PLoS Genetics*, 10(9). doi: 10.1371/journal.pgen.1004605.

Mellacheruvu, D. *et al.* (2013) 'The CRAPome: A contaminant repository for affinity purification-mass spectrometry data', *Nature Methods*, 10(8), pp. 730–736. doi: 10.1038/nmeth.2557.

Mendell, J. R. and Lloyd-Puryear, M. (2013) 'Report Of MDA Muscle Disease Symposium On Newborn Screening For Duchenne Muscular Dystrophy', *Muscle and Nerve*, 48(1), pp. 21–26. doi: 10.1002/mus.23810.

Mercuri, E., Bönnemann, C. G. and Muntoni, F. (2019) 'Muscular dystrophies', *The Lancet*, 394(10213), pp. 2025–2038. doi: 10.1016/S0140-6736(19)32910-1.

Metzger, T. *et al.* (2012) 'MAP and kinesin-dependent nuclear positioning is required for skeletal muscle function', *Nature*, 484(7392), pp. 120–124. doi: 10.1038/nature10914.

Meyer, G. A. and Lieber, R. L. (2012) 'Skeletal muscle fibrosis develops in response to desmin deletion', *American Journal of Physiology - Cell Physiology*, 302(11), pp. 1609–1620. doi: 10.1152/ajpcell.00441.2011.

Michaluk, P. *et al.* (2007) 'β-Dystroglycan as a target for MMP-9, in response to enhanced neuronal activity', *Journal of Biological Chemistry*, 282(22), pp. 16036–16041. doi: 10.1074/jbc.M700641200.

Miller, G. *et al.* (2012) 'Preventing phosphorylation of dystroglycan ameliorates the dystrophic phenotype in mdx mouse', *Human Molecular Genetics*, 21(20), pp. 4508–4520. doi: 10.1093/hmg/dds293.

Minetti, C. *et al.* (1992) 'Dystrophin at the plasma membrane of human muscle fibers shows a costameric localization', *Neuromuscular Disorders*, 2(2), pp. 99–109. doi: 10.1016/0960-8966(92)90041-4.

Minn, I. *et al.* (2009) 'SUN-1 and ZYG-12, Mediators of Centrosome-Nucleus Attachment, Are a Functional SUN/KASH Pair in Caenorhabditis elegans', *Molecular biology of the cell*, 20, pp. 4586–4595. doi: 10.1091/mbc.E08.

Miranda, A. F. *et al.* (1988) 'Immunocytochemical study of dystrophin in muscle cultures from patient with Duchenne muscular dystrophy and unaffected control patients', *American Journal of Pathology*, 132(3), pp. 410–416.

Mitchell, A. *et al.* (2013) 'Dystroglycan function is a novel determinant of tumor growth and behavior in prostate cancer', *Prostate*, 73(4), pp. 398–408. doi:

10.1002/pros.22581.

Mizuguchi, G. *et al.* (2004) 'ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex', *Science*, 303(5656), pp. 343–348. doi: 10.1126/science.1090701.

Moir, R. D. *et al.* (2000) 'Nuclear lamins A and B1: Different pathways of assembly during nuclear envelope formation in living cells', *Journal of Cell Biology*, 151(6), pp. 1155–1168. doi: 10.1083/jcb.151.6.1155.

Montoliu, L., Chávez, S. and Vidal, M. (2000) 'Variegation associated with lacZ in transgenic animals: A warning note', *Transgenic Research*, 9(3), pp. 237–239. doi: 10.1023/A:1008995730285.

Moore, C. J. and Winder, S. J. (2010) 'Dystroglycan versatility in cell adhesion: A tale of multiple motifs', *Cell Communication and Signaling*, 8, pp. 1–12. doi: 10.1186/1478-811X-8-3.

Moore, S. a *et al.* (2002) 'Deletion of brain dystroglycan recapitulates aspects of congenital muscular dystrophy.', *Nature*, 418(6896), pp. 422–425. doi: 10.1038/nature00838.

Morgan, J. E. *et al.* (1992) 'Formation of skeletal muscle in vivo from the mouse C2 cell line', *Journal of Cell Science*, 102(4), pp. 779–787.

Morgan, J. T. *et al.* (2011) 'Nesprin-3 regulates endothelial cell morphology, perinuclear cytoskeletal architecture, and flow-induced polarization', *Molecular Biology of the Cell*, 22(22), pp. 4324–4334. doi: 10.1091/mbc.E11-04-0287.

Morimoto, A. *et al.* (2012) 'A conserved KASH domain protein associates with telomeres, SUN1, and dynactin during mammalian meiosis', *Journal of Cell Biology*, 198(2), pp. 165–172. doi: 10.1083/jcb.201204085.

Moroianu, J. and Blobel, G. (1995) 'Protein export from the nucleus requires the GTPase Ran and GTP hydrolysis', *Proceedings of the National Academy of Sciences of the United States of America*, 92(10), pp. 4318–4322. doi: 10.1073/pnas.92.10.4318.

Mosley-Bishop, K. L. *et al.* (1999) 'Molecular analysis of the klarsicht gene and its role in nuclear migration within differentiating cells of the Drosophila eye', *Current Biology*, 9(21), pp. 1211–1220. doi: 10.1016/S0960-9822(99)80501-6.

Moujaber, O. *et al.* (2017) 'Data on the association of the nuclear envelope protein Sun1 with nucleoli', *Data in Brief*, 13, pp. 115–123. doi: 10.1016/j.dib.2017.05.028.

Mounkes, L. C. *et al.* (2003) 'A progeroid syndrome in mice is caused by defects in A-type lamins', *Nature*, 423(6937), pp. 298–301. doi: 10.1038/nature01631.

Muchir, A. *et al.* (2000) 'Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B)', *Human Molecular Genetics*, 9(9), pp. 1453–1459. doi: 10.1093/hmg/9.9.1453.

Muchir, A. *et al.* (2003) 'Nuclear envelope alterations in fibroblasts from

LGMD1B patients carrying nonsense Y259X heterozygous or homozygous mutation in lamin A/C gene', *Experimental Cell Research*, 291(2), pp. 352–362. doi: 10.1016/j.yexcr.2003.07.002.

Muchir, A. *et al.* (2004) 'Nuclear envelope alterations in fibroblasts from patients with muscular dystrophy, cardiomyopathy, and partial lipodystrophy carrying lamin A/C gene mutations', *Muscle and Nerve*, 30(4), pp. 444–450. doi: 10.1002/mus.20122.

Muchir, A. *et al.* (2007) 'Activation of MAPK in hearts of EMD null mice: Similarities between mouse models of X-linked and autosomal dominant Emery - Dreifuss muscular dystrophy', *Human Molecular Genetics*, 16(15), pp. 1884–1895. doi: 10.1093/hmg/ddm137.

Muchir, A. *et al.* (2009) 'Inhibition of extracellular signal-regulated kinase signaling to prevent cardiomyopathy caused by mutation in the gene encoding A-type lamins', *Human Molecular Genetics*, 18(2), pp. 241–247. doi: 10.1093/hmg/ddn343.

Mudumbi, K. C. *et al.* (2020) 'Nucleoplasmic signals promote directed transmembrane protein import simultaneously via multiple channels of nuclear pores', *Nature Communications*, 11(1), pp. 1–14. doi: 10.1038/s41467-020-16033-x.

Mumm, J. S. *et al.* (2000) 'A ligand-induced extracellular cleavage regulates γ-secretase-like proteolytic activation of Notch1', *Molecular Cell*, 5(2), pp. 197–206. doi: 10.1016/S1097-2765(00)80416-5.

Myers, T. A. and Nickoloff, J. A. (1999) 'Modification of Enzyme- Conjugated Streptavidin- Biotin Western Blot Tech- nique to Avoid Detection of Endogenous Biotin- Containing Proteins BioTechniques', *BioTechniques*, 26, pp. 854–858.

Myshrall, T. D. *et al.* (2012) 'Dystroglycan on Radial Glia Endfeet Is Required For Pial Basement Membrane Integrity and Columnar Organization of the Developing Cerebral Cortex', *Journal Neuropathol Exp Neurol*, 71(12), pp. 1047–1063. doi: 10.1038/nature11130.Reduced.

Nagel, A., Lehmann-Horn, F. and Engel, A. G. (1990) 'Neuromuscular transmission in the mdx mouse', *Muscle & Nerve*, 13(8), pp. 742–749. doi: 10.1002/mus.880130813.

Nakamori, M. and Takahashi, M. P. (2011) 'The role of alpha-dystrobrevin in striated muscle', *International Journal of Molecular Sciences*, 12(3), pp. 1660–1671. doi: 10.3390/ijms12031660.

Nastały, P. *et al.* (2020) 'Role of the nuclear membrane protein Emerin in front-rear polarity of the nucleus', *Nature Communications*, 11(1), pp. 1–12. doi: 10.1038/s41467-020-15910-9.

Neumann, F. R. *et al.* (2012) 'Targeted INO80 enhances subnuclear chromatin movement and ectopic homologous recombination', *Genes and Development*, 26(4), pp. 369–383. doi: 10.1101/gad.176156.111.

Newey, S. E. *et al.* (2001) 'Syncoilin, a Novel Member of the Intermediate Filament Superfamily That Interacts with α-Dystrobrevin in Skeletal Muscle', *Journal*

of Biological Chemistry, 276(9), pp. 6645–6655. doi: 10.1074/jbc.M008305200.

Newport, J. W., Wilson, K. L. and Dunphy, W. G. (1990) 'A lamin-independent pathway for nuclear envelope assembly', *Journal of Cell Biology*, 111(6 I), pp. 2247–2259. doi: 10.1083/jcb.111.6.2247.

Nigro, V. *et al.* (1996) 'Autosomal recessive limb-girdle muscular dystrophy, LGMD2F, is caused by a mutation in the δ-sarcoglycan gene', *Nature Genetics*, 14(2), pp. 195–198. doi: 10.1038/ng1096-195.

Nikolopoulos, S. N. and Turner, C. E. (2000) 'Actopaxin, a new focal adhesion protein that binds paxillin LD motifs and actin and regulates cell adhesion', *Journal of Cell Biology*, 151(7), pp. 1435–1447. doi: 10.1083/jcb.151.7.1435.

Nikolova-Krstevski, V. *et al.* (2011) 'Nesprin-1 and actin contribute to nuclear and cytoskeletal defects in lamin A/C-deficient cardiomyopathy', *Journal of Molecular and Cellular Cardiology*, 50(3), pp. 479–486. doi: 10.1016/j.yjmcc.2010.12.001.

Nishioka, Y. *et al.* (2016) 'SUN1 splice variants, SUN1_888, SUN1_785, and predominant SUN1_916, variably function in directional cell migration', *Nucleus*, 7(6), pp. 572–584. doi: 10.1080/19491034.2016.1260802.

Noguchi, S. *et al.* (1995) 'Mutations in the dystrophin-associated protein γ-sarcoglycan in chromosome 13 muscular dystrophy', *Science*, 270(5237), pp. 819–822. doi: 10.1126/science.270.5237.819.

Novelli, G. *et al.* (2002) 'Mandibuloacral dysplasia is caused by a mutation in LMNA-encoding lamin A/C', *American Journal of Human Genetics*, 71(2), pp. 426–431. doi: 10.1086/341908.

O'Gorman, S., Fox, D. T. and Wahl, G. M. (1991) 'Recombinase-mediated gene activation and site-specific integration in mammalian cells', *Science*, 251(4999), pp. 1351–1355. doi: 10.1126/science.1900642.

Ohashi, K. *et al.* (2012) 'Visualization of cofilin-actin and Ras-Raf interactions by bimolecular fluorescence complementation assays using a new pair of split Venus fragments', *BioTechniques*, 52(1), pp. 45–50. doi: 10.2144/000113777.

Ohba, T. *et al.* (2004) 'Energy- and temperature-dependent transport of integral proteins to the inner nuclear membrane via the nuclear pore', *Journal of Cell Biology*, 167(6), pp. 1051–1062. doi: 10.1083/jcb.200409149.

Ohlendieck, K. *et al.* (1991) 'Dystrophin-related protein is localized to neuromuscular junctions of adult skeletal muscle', *Neuron*, 7(3), pp. 499–508. doi: 10.1016/0896-6273(91)90301-F.

Ohtsubo, M., Okazaki, H. and Nishimoto, T. (1989) 'The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA', *Journal of Cell Biology*, 109(4 I), pp. 1389–1397. doi: 10.1083/jcb.109.4.1389.

Okamoto, I. *et al.* (2001) 'Proteolytic release of CD44 intracellular domain and its role in the CD44 signaling pathway', *Journal of Cell Biology*, 155(5), pp. 755–762. doi: 10.1083/jcb.200108159.

Olski, T. M., Noegel, A. A. and Korenbaum, E. (2001) 'Parvin, a 42 kDa focal adhesion protein, related to the α-actinin superfamily', *Journal of Cell Science*, 114(3), pp. 525–538.

Ono, Y. *et al.* (2011) 'BMP signalling permits population expansion by preventing premature myogenic differentiation in muscle satellite cells', *Cell Death and Differentiation*, 18(2), pp. 222–234. doi: 10.1038/cdd.2010.95.

Oppizzi, M. L. *et al.* (2008) 'Nuclear translocation of β-dystroglycan reveals a distinctive trafficking pattern of autoproteolyzed mucins', *Traffic*, 9(12), pp. 2063–2072. doi: 10.1111/j.1600-0854.2008.00822.x.

Osmanagic-Myers, S., Dechat, T. and Foisner, R. (2015) 'Lamins at the crossroads of mechanosignaling', *Genes and Development*, pp. 225–237. doi: 10.1101/gad.255968.114.

Östlund, C. *et al.* (2009) 'Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton (LINC) complex proteins', *Journal of Cell Science*, 122(22), pp. 4099–4108. doi: 10.1242/jcs.057075.

Oza, P. and Peterson, C. L. (2010) 'Opening the DNA repair toolbox: Localization of DNA double strand breaks to the nuclear periphery', *Cell Cycle*, 9(1), pp. 43–49. doi: 10.4161/cc.9.1.10317.

Padmakumar, V. C. *et al.* (2004) 'Enaptin, a giant actin-binding protein, is an element of the nuclear membrane and the actin cytoskeleton', *Experimental Cell Research*, 295(2), pp. 330–339. doi: 10.1016/j.yexcr.2004.01.014.

Partridge, T. A. (2013) 'The mdx mouse model as a surrogate for Duchenne muscular dystrophy', *FEBS Journal*, 280(17), pp. 4177–4186. doi: 10.1111/febs.12267.

Pasch, E. *et al.* (2015) 'The LINC complex component Sun4 plays a crucial role in sperm head formation and fertility', *Biology Open*, 4(12), pp. 1792–1802. doi: 10.1242/bio.015768.

Pegoraro, G. *et al.* (2009) 'Ageing-related chromatin defects through loss of the NURD complex', *Nature Cell Biology*, 11(10), pp. 1261–1267. doi: 10.1038/ncb1971.

Pekovic, V. *et al.* (2007) 'Nucleoplasmic LAP2α-lamin A complexes are required to maintain a proliferative state in human fibroblasts', *Journal of Cell Biology*, 176(2), pp. 163–172. doi: 10.1083/jcb.200606139.

Percival, J. M. (2011) 'nNOS regulation of skeletal muscle fatigue and exercise performance', *Biophysical Reviews*, 3(4), pp. 209–217. doi: 10.1007/s12551-011-0060-9.

Percival, J. M. *et al.* (2012) 'Sildenafil reduces respiratory muscle weakness and fibrosis in the mdx mouse model of Duchenne muscular dystrophy', *Journal of Pathology*, 228(1), pp. 77–87. doi: 10.1002/path.4054.

Pereboev, A. V. *et al.* (2001) 'Epitopes in the interacting regions of β-dystroglycan (PPxY motif) and dystrophin (WW domain)', *Biochimica et Biophysica Acta - General Subjects*, 1527(1–2), pp. 54–60. doi: 10.1016/S0304-4165(01)00147-

7.

Peretto, G. *et al.* (2018) 'Updated clinical overview on cardiac laminopathies: An electrical and mechanical disease', *Nucleus*, 9(1), pp. 380–391. doi: 10.1080/19491034.2018.1489195.

Perovanovic, J. *et al.* (2016) 'Laminopathies disrupt epigenomic developmental programs and cell fate', *Science Translational Medicine*, 8(335). doi: 10.1126/scitranslmed.aad4991.

Peter, A. K., Marshall, J. L. and Crosbie, R. H. (2008) 'Sarcospan reduces dystrophic pathology: Stabilization of the utrophin-glycoprotein complex', *Journal of Cell Biology*, 183(3), pp. 419–427. doi: 10.1083/jcb.200808027.

Peter, A. K., Miller, G. and Crosbie, R. H. (2007) 'Disrupted mechanical stability of the dystrophin-glycoprotein complex causes severe muscular dystrophy in sarcospan transgenic mice', *Journal of Cell Science*, 120(6), pp. 996–1008. doi: 10.1242/jcs.03360.

Petrof, B. J. *et al.* (1993) 'Dystrophin protects the sarcolemma from stresses developed during muscle contraction', *Proceedings of the National Academy of Sciences of the United States of America*, 90(8), pp. 3710–3714. doi: 10.1073/pnas.90.8.3710.

Pham, D. H. *et al.* (2008) 'Attenuation of leakiness in doxycycline-inducible expression via incorporation of 3' AU-rich mRNA destabilizing elements', *BioTechniques*, 45(2), pp. 155–162. doi: 10.2144/000112896.

Piehler, J. (2005) 'New methodologies for measuring protein interactions in vivo and in vitro', *Current Opinion in Structural Biology*, 15(1 SPEC. ISS.), pp. 4–14. doi: 10.1016/j.sbi.2005.01.008.

Piekarowicz, K. *et al.* (2019) 'A Muscle Hybrid Promoter as a Novel Tool for Gene Therapy', *Molecular Therapy - Methods and Clinical Development*, 15(December), pp. 157–169. doi: 10.1016/j.omtm.2019.09.001.

Piggott, R. (2014) 'The Regulation of Beta- Dystroglycan Internalization', (July).

Postel, R. *et al.* (2011) 'Nesprin-3 augments peripheral nuclear localization of intermediate filaments in zebrafish', *Journal of Cell Science*, 124(5), pp. 755–764. doi: 10.1242/jcs.081174.

Powell, L. and Burke, B. (1990) 'Internuclear exchange of an inner nuclear membrane protein (p55) in heterokaryons: In vivo evidence for the interaction of p55 with the nuclear lamina', *Journal of Cell Biology*, 111(6 I), pp. 2225–2234. doi: 10.1083/jcb.111.6.2225.

Prelich, G. (2012) 'Gene overexpression: Uses, mechanisms, and interpretation', *Genetics*, 190(3), pp. 841–854. doi: 10.1534/genetics.111.136911.

Prins, K. W. *et al.* (2009) 'Dystrophin is a microtubule-associated protein', *Journal of Cell Biology*, 186(3), pp. 363–369. doi: 10.1083/jcb.200905048.

Pronobis, M. I., Deuitch, N. and Peifer, M. (2016) 'The Miraprep: A protocol that uses a Miniprep kit and provides Maxiprep yields', *PLoS ONE*, 11(8), pp. 1–12.

doi: 10.1371/journal.pone.0160509.

Puente, X. S. *et al.* (2011) 'Exome sequencing and functional analysis identifies BANF1 mutation as the cause of a hereditary progeroid syndrome', *American Journal of Human Genetics*, 88(5), pp. 650–656. doi: 10.1016/j.ajhg.2011.04.010.

Puri, T. *et al.* (2007) 'Dodecameric Structure and ATPase Activity of the Human TIP48/TIP49 Complex', *Journal of Molecular Biology*, 366(1), pp. 179–192. doi: 10.1016/j.jmb.2006.11.030.

Putyrski, M. and Schultz, C. (2012) 'Protein translocation as a tool: The current rapamycin story', *FEBS Letters*, 586(15), pp. 2097–2105. doi: 10.1016/j.febslet.2012.04.061.

Raab, M. *et al.* (2016) 'ESCRT III repairs nuclear envelope ruptures during cell migration to limit DNA damage and cell death', *Science*, 352(6283), pp. 359–362. doi: 10.1126/science.aad7611.

Radu, A., Moore, M. S. and Blobel, G. (1995) 'The peptide repeat domain of nucleoporin Nup98 functions as a docking site in transport across the nuclear pore complex', *Cell*, 81(2), pp. 215–222. doi: 10.1016/0092-8674(95)90331-3.

Raharjo, W. H. *et al.* (2001) 'Nuclear envelope defects associated with LMNA mutations cause dilated cardiomyopathy and Emery-Dreifuss muscular dystrophy', *Journal of Cell Science*, 114(24), pp. 4447–4457. doi: 10.1007/s00412-012-0388-3.

Ran, F. A., Hsu, P. D., Lin, C. Y., *et al.* (2013) 'Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity', *Cell*, 154(6), pp. 1380–1389. doi: 10.1016/j.cell.2013.08.021.

Ran, F. A., Hsu, P. D., Wright, J., *et al.* (2013) 'Genome engineering using the CRISPR-Cas9 system', *Nature Protocols*, 8(11), pp. 2281–2308. doi: 10.1038/nprot.2013.143.

Randles, K. N. *et al.* (2010) 'Nesprins, but not sun proteins, switch isoforms at the nuclear envelope during muscle development', *Developmental Dynamics*, 239(3), pp. 998–1009. doi: 10.1002/dvdy.22229.

Razafsky, D. *et al.* (2016) 'Lamin B1 and lamin B2 are long-lived proteins with distinct functions in retinal development', *Molecular Biology of the Cell*, 27(12), pp. 1928–1937. doi: 10.1091/mbc.E16-03-0143.

Reddy, K. L. *et al.* (2008) 'Transcriptional repression mediated by repositioning of genes to the nuclear lamina', *Nature*, 452(7184), pp. 243–247. doi: 10.1038/nature06727.

Reichelt, R. *et al.* (1990) 'Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components', *Journal of Cell Biology*, 110(4), pp. 883–894. doi: 10.1083/jcb.110.4.883.

Reimann, J., Irintchev, A. and Wernig, A. (2000) 'Regenerative capacity and the number of satellite cells in soleus muscles of normal and mdx mice', *Neuromuscular Disorders*, 10(4–5), pp. 276–282. doi: 10.1016/S0960-8966(99)00118-2.

Rentschler, S. *et al.* (1999) 'The WW domain of dystrophin requires EF-hands

region to interact with ??-dystroglycan', *Biological Chemistry*, 380(4), pp. 431–442. doi: 10.1515/BC.1999.057.

Ribbeck, K. and Görlich, D. (2002) 'The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion', *EMBO Journal*, 21(11), pp. 2664–2671.

Richards, S. A., Carey, K. L. and Macara, I. G. (1997) 'Requirement of guanosine triphosphate-bound ran for signal-mediated nuclear protein export.', *Science*, 276(5320), pp. 1842–1844. doi: 10.1126/science.276.5320.1842.

Riemersma, M. *et al.* (2015) 'Absence of alpha-and beta-dystroglycan is associated with Walker-Warburg syndrome', *Neurology*, 84(21), pp. 2177–2182. doi: 10.1212/WNL.0000000000001615.

Rivier, François *et al.* (1999) 'Dystrophin and utrophin complexed with different associated proteins in cardiac Purkinje fibres', *Histochemical Journal*, 31(7), pp. 425–432. doi: 10.1023/A:1003805905456.

Rivier, FRANÇOIS *et al.* (1999) 'Utrophin and dystrophin-associated glycoproteins in normal and dystrophin deficient cardiac muscle', *Journal of Muscle Research and Cell Motility*, 20(3), pp. 305–314. doi: 10.1023/A:1005426920070.

Robbins, J. *et al.* (1991) 'Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: Identification of a class of bipartite nuclear targeting sequence', *Cell*, 64(3), pp. 615–623. doi: 10.1016/0092-8674(91)90245-T.

De Rosa, M. C. *et al.* (2011) 'A second Ig-like domain identified in dystroglycan by molecular modelling and dynamics', *Journal of Molecular Graphics and Modelling*, 29(8), pp. 1015–1024. doi: 10.1016/j.jmgm.2011.04.008.

Van Rosmalen, M., Krom, M. and Merkx, M. (2017) 'Tuning the Flexibility of Glycine-Serine Linkers to Allow Rational Design of Multidomain Proteins', *Biochemistry*, 56(50), pp. 6565–6574. doi: 10.1021/acs.biochem.7b00902.

Roux, K. J. *et al.* (2009) 'Nesprin 4 is an outer nuclear membrane protein that can induce kinesin-mediated cell polarization', *Proceedings of the National Academy of Sciences of the United States of America*, 106(7), pp. 2194–2199. doi: 10.1073/pnas.0808602106.

Roux, K. J. *et al.* (2012) 'A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells', *Journal of Cell Biology*, 196(6), pp. 801–810. doi: 10.1083/jcb.201112098.

Roux, K. J., Kim, D. I. and Burke, B. (2013) 'BioID : A Screen for Protein-Protein Interactions', *Current Protocols in Protein Science*, 19.23(November), pp. 1–14. doi: 10.1002/0471140864.ps1923s74.

Rusiñol, A. E. and Sinensky, M. S. (2006) 'Farnesylated lamins, progeroid syndromes and farnesyl transferase inhibitors', *Journal of Cell Science*, 119(16), pp. 3265–3272. doi: 10.1242/jcs.03156.

Russo, K. *et al.* (2000) 'Characterization of the β-dystroglycan-growth factor receptor 2 (Grb2) interaction', *Biochemical and Biophysical Research Communications*, 274(1), pp. 93–98. doi: 10.1006/bbrc.2000.3103.

Sabatelli, P. *et al.* (2003) 'Extracellular matrix and nuclear abnormalities in skeletal muscle of a patient with Walker-Warburg syndrome caused by POMT1 mutation', *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1638(1), pp. 57–62. doi: 10.1016/S0925-4439(03)00040-1.

Sadaie, M. *et al.* (2015) 'Cell-based screen for altered nuclear phenotypes reveals senescence progression in polyploid cells after Aurora kinase B inhibition', *Molecular Biology of the Cell*, 26(17), pp. 2971–2985. doi: 10.1091/mbc.E15-01-0003.

Sadoulet-Puccio, H. M., Rajala, M. and Kunkel, L. M. (1997) 'Dystrobrevin and dystrophin: An interaction through coiled-coil motifs', *Proceedings of the National Academy of Sciences of the United States of America*, 94(23), pp. 12413–12418. doi: 10.1073/pnas.94.23.12413.

Saito, F. *et al.* (2003) 'Unique role of dystroglycan in peripheral nerve myelination, nodal structure, and sodium channel stabilization', *Neuron*, 38(5), pp. 747–758. doi: 10.1016/S0896-6273(03)00301-5.

Salazar-Mendiguchía, J. *et al.* (2020) 'The p.(Cys150Tyr) variant in CSRP3 is associated with late-onset hypertrophic cardiomyopathy in heterozygous individuals', *European Journal of Medical Genetics*, 63(12), p. 104079. doi: 10.1016/j.ejmg.2020.104079.

Salpingidou, G. *et al.* (2007) 'A novel role for the nuclear membrane protein emerin in association of the centrosome to the outer nuclear membrane', *Journal of Cell Biology*, 178(6), pp. 897–904. doi: 10.1083/jcb.200702026.

Sandonà, D. and Betto, R. (2009) 'Sarcoglycanopathies: Molecular pathogenesis and therapeutic prospects', *Expert Reviews in Molecular Medicine*, 11(September), pp. 1–27. doi: 10.1017/S1462399409001203.

De Sandre-Giovannoli, A. *et al.* (2002) 'Homozygous defects in LMNA, encoding lamin A/C nuclear-envelope proteins, cause autosomal recessive axonal neuropathy in human (Charcot-Marie-Tooth disorder type 2) and mouse', *American Journal of Human Genetics*, 70(3), pp. 726–736. doi: 10.1086/339274.

Sandre-Giovannoli, A. De *et al.* (2003) 'Lamin A Truncation in Hutchinson-Gilford Progeria', *Science*, 300, p. 2055. doi: 10.1126/science.1084125.

Sasse, B. *et al.* (1997) 'In vitro assembly of Drosophila lamin Dm0. Lamin polymerization properties are conserved', *European Journal of Biochemistry*, 250(1), pp. 30–38. doi: 10.1111/j.1432-1033.1997.t01-1-00030.x.

Sato, A. *et al.* (2009) 'Cytoskeletal Forces Span the Nuclear Envelope to Coordinate Meiotic Chromosome Pairing and Synapsis', *Cell*, 139(5), pp. 907–919. doi: 10.1016/j.cell.2009.10.039.

Satz, J. S. *et al.* (2008) 'Brain and Eye Malformations Resembling Walker-Warburg Syndrome Are Recapitulated in Mice by Dystroglycan Deletion in the Epiblast', *Journal of Neuroscience*, 28(42), pp. 10567–10575. doi: 10.1523/JNEUROSCI.2457-08.2008.

Sauer, B. and Henderson, N. (1988) 'Site-specific DNA recombination in

mammalian cells by the Cre recombinase of bacteriophage P1.', *Proceedings of the National Academy of Sciences of the United States of America*, 85(14), pp. 5166–5170. doi: 10.1073/pnas.85.14.5166.

Sawilowsky, S. S. (2009) 'New Effect Size Rules of Thumb', *Journal of Modern Applied Statistical Methods*, 8(2), pp. 597–599. doi: 10.22237/jmasm/1257035100.

Sbardella, D. *et al.* (2012) 'Enzymatic processing by MMP-2 and MMP-9 of wild-type and mutated mouse β-dystroglycan', *IUBMB Life*, 64(12), pp. 988–994. doi: 10.1002/iub.1095.

Schirmer, E. C. *et al.* (2003) 'Nuclear membrane proteins with potential disease links found by subtractive proteomics', *Science*, 301(5638), pp. 1380–1382. doi: 10.1126/science.1088176.

Schirmer, E. C. and Gerace, L. (2004) 'The stability of the nuclear lamina polymer changes with the composition of lamin subtypes according to their individual binding strengths', *Journal of Biological Chemistry*, 279(41), pp. 42811–42817. doi: 10.1074/jbc.M407705200.

Schirmer, E. C. and Gerace, L. (2005) 'The nuclear membrane proteome: Extending the envelope', *Trends in Biochemical Sciences*, 30(10), pp. 551–558. doi: 10.1016/j.tibs.2005.08.003.

Schmidt, H. B. and Görlich, D. (2016) 'Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles', *Trends in Biochemical Sciences*, 41(1), pp. 46–61. doi: 10.1016/j.tibs.2015.11.001.

Schmierer, B. and Hill, C. S. (2007) 'TGFβ-SMAD signal transduction: Molecular specificity and functional flexibility', *Nature Reviews Molecular Cell Biology*, 8(12), pp. 970–982. doi: 10.1038/nrm2297.

Schmitt, J. *et al.* (2007) 'Transmembrane protein Sun2 is involved in tethering mammalian meiotic telomeres to the nuclear envelope', *Proceedings of the National Academy of Sciences of the United States of America*, 104(18), pp. 7426–7431. doi: 10.1073/pnas.0609198104.

Schneider, M. *et al.* (2011) 'Molecular mechanisms of centrosome and cytoskeleton anchorage at the nuclear envelope', *Cellular and Molecular Life Sciences*, 68(9), pp. 1593–1610. doi: 10.1007/s00018-010-0535-z.

Schopp, I. M. *et al.* (2017) 'Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes', *Nature Communications*, 8. doi: 10.1038/ncomms15690.

Schröder, J. E. *et al.* (2007) 'Dystroglycan regulates structure, proliferation and differentiation of neuroepithelial cells in the developing vertebrate CNS', *Developmental Biology*, 307(1), pp. 62–78. doi: 10.1016/j.ydbio.2007.04.020.

Schütz, W. *et al.* (2005) 'Nuclear envelope remodeling during mouse spermiogenesis: Postmeiotic expression and redistribution of germline lamin B3', *Experimental Cell Research*, 307(2), pp. 285–291. doi: 10.1016/j.yexcr.2005.03.023.

Sciandra, F. *et al.* (2009) 'Mutagenesis at the α-β Interface impairs the cleavage of the dystroglycan precursor', *FEBS Journal*, 276(17), pp. 4933–4945. doi:

10.1111/j.1742-4658.2009.07196.x.

Sciandra, F. *et al.* (2012) 'Dystroglycan is associated to the disulfide isomerase ERp57', *Experimental Cell Research*, 318(19), pp. 2460–2469. doi: 10.1016/j.yexcr.2012.07.006.

Senior, A. and Gerace, L. (1988) 'Integral membrane proteins specific to the inner nuclear membrane and associated with the nuclear lamina', *Journal of Cell Biology*, 107(6 I), pp. 2029–2036. doi: 10.1083/jcb.107.6.2029.

Sewduth, R. N., Baietti, M. F. and Sablina, A. A. (2020) 'Cracking the monoubiquitin code of genetic diseases', *International Journal of Molecular Sciences*, 21(9), pp. 1–18. doi: 10.3390/ijms21093036.

Sgambato, A. *et al.* (2004) 'Increased expression of dystroglycan inhibits the growth and tumorigenicity of human mammary epithelial cells', *Cancer Biology and Therapy*, 3(10), pp. 967–975. doi: 10.4161/cbt.3.10.1132.

Shabek, N. *et al.* (2012) 'The Size of the Proteasomal Substrate Determines Whether Its Degradation Will Be Mediated by Mono- or Polyubiquitylation', *Molecular Cell*, 48(1), pp. 87–97. doi: 10.1016/j.molcel.2012.07.011.

Shackleton, S. *et al.* (2000) 'LMNA, encoding lamin A/C, is mutated in partial lipodystrophy', *Nature Genetics*, 24(2), pp. 153–156. doi: 10.1038/72807.

Shaikh, S. and Nicholson, L. F. B. (2006) 'Optimization of the Tet-on system for inducible expression of RAGE', *Journal of Biomolecular Techniques*, 17(4), pp. 283–292.

Sheikh, M. O., Halmo, S. M. and Wells, L. (2017) 'Recent advancements in understanding mammalian O-mannosylation', *Glycobiology*, 27(9), pp. 806–819. doi: 10.1093/glycob/cwx062.

Shen, X. *et al.* (2000) 'A chromatin remodelling complex involved in transcription and DNA processing', *Nature*, 406, pp. 541–544.

Shimi, T. *et al.* (2008) 'The A- and B-type nuclear lamin networks: Microdomains involved in chromatin organization and transcription', *Genes and Development*, 22(24), pp. 3409–3421. doi: 10.1101/gad.1735208.

Shimi, T. *et al.* (2011) 'The role of nuclear lamin B1 in cell proliferation and senescence', *Genes and Development*, 25(24), pp. 2579–2593. doi: 10.1101/gad.179515.111.

Siebel, C. and Lendahl, U. (2017) 'Notch signaling in development, tissue homeostasis, and disease', *Physiological Reviews*, 97(4), pp. 1235–1294. doi: 10.1152/physrev.00005.2017.

Signorino, G. *et al.* (2018) 'A dystroglycan mutation (p.Cys667Phe) associated to muscle-eye-brain disease with multicystic leucodystrophy results in ER-retention of the mutant protein', *Human Mutation*, 39(2), pp. 266–280. doi: 10.1002/humu.23370.

Simha, V. and Garg, A. (2002) 'Body fat distribution and metabolic derangements in patients with familial partial lipodystrophy associated with

mandibuloacral dysplasia', *Journal of Clinical Endocrinology and Metabolism*, 87(2), pp. 776–785. doi: 10.1210/jcem.87.2.8258.

Sinensky, M. *et al.* (1994) 'The processing pathway of prelamin A', *Journal of Cell Science*, 107(1), pp. 61–67.

Singh, J. *et al.* (2004) 'Proteolytic enzymes and altered glycosylation modulate dystroglycan function in carcinoma cells', *Cancer Research*, 64(17), pp. 6152–6159. doi: 10.1158/0008-5472.CAN-04-1638.

Smith, A. and Benavente, R. (1992) 'Identification of a short nuclear lamin protein selectively expressed during meiotic stages of rat spermatogenesis', *Differentiation*, 52(1), pp. 55–60. doi: 10.1111/j.1432-0436.1992.tb00499.x.

Smith, L. R. *et al.* (2011) 'Hamstring contractures in children with spastic cerebral palsy result from a stiffer extracellular matrix and increased in vivo sarcomere length', *Journal of Physiology*, 589(10), pp. 2625–2639. doi: 10.1113/jphysiol.2010.203364.

Snapp, E. L. and Hegde, R. S. (2006) 'Rational Design and Evaluation of FRET Experiments to Measure Protein Proximities in Cells', *Curr Protoc Cell Biol.*, 23(1), pp. 1–7. doi: 10.1002/0471143030.cb1709s32.

Solovei, I. *et al.* (2013) 'LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation', *Cell*, 152(3), pp. 584–598. doi: 10.1016/j.cell.2013.01.009.

Soriano, P. (1999) 'Generalized lacZ expression with the ROSA26 Cre reporter strain', *Nature genetics*, 21, pp. 70–71.

Sosa, B. A. *et al.* (2012) 'LINC complexes form by binding of three KASH peptides to domain interfaces of trimeric SUN proteins', *Cell*, 149(5), pp. 1035–1047. doi: 10.1016/j.cell.2012.03.046.

Sotgia, F. *et al.* (2001) 'Tyrosine phosphorylation of β-dystroglycan at its WW domain binding motif, PPxY, recruits SH2 domain containing proteins', *Biochemistry*, 40(48), pp. 14585–14592. doi: 10.1021/bi011247r.

Sotgia, F. *et al.* (2003) 'Localization of phospho-beta-dystroglycan (pY892) to an intracellular vesicular compartment in cultured cells and skeletal muscle fibers in vivo', *Biochemistry*, 42(23), pp. 7110–7123. doi: 10.1021/bi0271289.

Soullam, B. and Worman, H. J. (1995) 'Signals and structural features involved in integral membrane protein targeting to the inner nuclear membrane', *Journal of Cell Biology*, 130(1), pp. 15–27. doi: 10.1083/jcb.130.1.15.

Spence, H. J., Dhillon, A. S., *et al.* (2004) 'Dystroglycan, a scaffold for the ERK-MAP kinase cascade', *EMBO Reports*, 5(5), pp. 484–489. doi: 10.1038/sj.embor.7400140.

Spence, H. J., Chen, Y.-J., *et al.* (2004) 'Ezrin-dependent regulation of the actin cytoskeleton by beta-dystroglycan.', *Human molecular genetics*, 13(15), pp. 1657–1668. doi: 10.1093/hmg/ddh170.

Srsen, V., Nadia, K. and Schirmer, E. C. (2011) 'Nuclear envelope influences on

cell-cycle progression', *Biochemical Society Transactions*, 39(6), pp. 1742–1746. doi: 10.1042/BST20110656.

Starr, D. A. and Han, M. (2002) 'Role of ANC-1 in tethering nuclei to the actin cytoskeleton', *Science*, 298(5592), pp. 406–409. doi: 10.1126/science.1075119.

Sternberg, N. and Hamilton, D. (1981) 'Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites', *Journal of Molecular Biology*, 150(4), pp. 467–486. doi: 10.1016/0022-2836(81)90375-2.

Stewart, R. M., Rodriguez, E. C. and King, M. C. (2019) 'Ablation of SUN2-containing LINC complexes drives cardiac hypertrophy without interstitial fibrosis', *Molecular Biology of the Cell*, 30(14), pp. 1664–1675. doi: 10.1091/mbc.E18-07-0438.

Stone, M. R. *et al.* (2005) 'Specific Interaction of the Actin-binding Domain of Dystrophin with Intermediate Filaments Containing Keratin 19', *Molecular Biology of the Cell*, 16(November), pp. 4280–4293. doi: 10.1091/mbc.E05.

Straub, V., Murphy, A. and Udd, B. (2018) '229th ENMC international workshop: Limb girdle muscular dystrophies – Nomenclature and reformed classification Naarden, the Netherlands, 17–19 March 2017', *Neuromuscular Disorders*, 28(8), pp. 702–710. doi: 10.1016/j.nmd.2018.05.007.

Stuurman, N., Heins, S. and Aebi, U. (1998) 'Nuclear lamins: Their structure, assembly, and interactions', *Journal of Structural Biology*, 122(1–2), pp. 42–66. doi: 10.1006/jsbi.1998.3987.

Sullivan, T. *et al.* (1999) 'Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy', *Journal of Cell Biology*, 147(5), pp. 913–919. doi: 10.1083/jcb.147.5.913.

Sun, L. *et al.* (2020) 'CRISPR/Cas9 mediated establishment of a human CSRP3 compound heterozygous knockout hESC line to model cardiomyopathy and heart failure', *Stem Cell Research*, 49(2). doi: 10.1016/j.scr.2020.102077.

Suzuki, A., Yoshida, M. and Ozawa, E. (1995) 'Mammalian α1- and β1-syntrophin bind to the alternative splice-prone region of the dystrophin COOH terminus', *Journal of Cell Biology*, 128(3), pp. 373–381. doi: 10.1083/jcb.128.3.373.

Swartz, R. K., Rodriguez, E. C. and King, M. C. (2014) 'A role for nuclear envelope-bridging complexes in homology-directed repair', *Molecular Biology of the Cell*, 25(16), pp. 2461–2471. doi: 10.1091/mbc.E13-10-0569.

Tajik, A. *et al.* (2016) 'Transcription upregulation via force-induced direct stretching of chromatin', *Nature Materials*, pp. 1–20. doi: 10.1038/nmat4729.

Talamas, J. A. and Capelson, M. (2015) 'Nuclear envelope and genome interactions in cell fate', *Frontiers in Genetics*, 5(FEB), pp. 1–16. doi: 10.3389/fgene.2015.00095.

Talamas, J. A. and Hetzer, M. W. (2011) 'POM121 and Sun1 play a role in early steps of interphase NPC assembly', *Journal of Cell Biology*, 194(1), pp. 27–37. doi: 10.1083/jcb.201012154.

Taniguchi-Ikeda, M. *et al.* (2016) 'Mechanistic aspects of the formation of α-dystroglycan and therapeutic research for the treatment of α-dystroglycanopathy: A review', *Molecular Aspects of Medicine*, 51, pp. 115–124. doi: 10.1016/j.mam.2016.07.003.

Taranum, S. *et al.* (2012) 'LINC complex alterations in DMD and EDMD/CMT fibroblasts', *European Journal of Cell Biology*, 91(8), pp. 614–628. doi: 10.1016/j.ejcb.2012.03.003.

Taylor, M. R. G. *et al.* (2005) 'Thymopoietin (lamina-associated polypeptide 2) gene mutation associated with dilated cardiomyopathy', *Human Mutation*, 26(6), pp. 566–574. doi: 10.1002/humu.20250.

Tazir, M. *et al.* (2004) 'Phenotypic variability in autosomal recessive axonal Charcot-Marie-Tooth disease due to the R298C mutation in lamin A/C', *Brain*, 127(1), pp. 154–163. doi: 10.1093/brain/awh021.

Tazir, M. *et al.* (2013) 'Autosomal recessive Charcot-Marie-Tooth disease: From genes to phenotypes', *Journal of the Peripheral Nervous System*, 18(2), pp. 113–129. doi: 10.1111/jns5.12026.

Tedesco, F. S. *et al.* (2010) 'Repairing skeletal muscle: regenerative potential of skeletal muscle stem cells', *The Journal of Clinical Investigation*, 120(1), pp. 11–19. doi: 10.1172/JCI40373.and.

Termini, C. M. and Gillette, J. M. (2017) 'Tetraspanins function as regulators of cellular signaling', *Frontiers in Cell and Developmental Biology*, 5(APR), pp. 1–14. doi: 10.3389/fcell.2017.00034.

Thakar, K. *et al.* (2017) 'Opposing roles for distinct LINC complexes in regulation of the small GTPase RhoA', *Molecular Biology of the Cell*, 28(1), pp. 182–191. doi: 10.1091/mbc.E16-06-0467.

Thomas, G. D. (2013) 'Functional muscle ischemia in Duchenne and Becker muscular dystrophy', *Frontiers in Physiology*, 4 DEC(December), pp. 1–6. doi: 10.3389/fphys.2013.00381.

Thompson, O. *et al.* (2008) 'Dystroglycan, Tks5 and Src mediated assembly of podosomes in myoblasts', *PLoS ONE*, 3(11). doi: 10.1371/journal.pone.0003638.

Thompson, O. *et al.* (2010) 'Modulation of cell spreading and cell-substrate adhesion dynamics by dystroglycan', *Journal of Cell Science*, 123(1), pp. 118–127. doi: 10.1242/jcs.047902.

Tinsley, J. *et al.* (1998) 'Expression of full-length utrophin prevents muscular dystrophy in mdx mice', *Nature Medicine*, 4(12), pp. 1441–1444. doi: 10.1038/4033.

Tinsley, J. M. *et al.* (1992) 'Primary structure of dystrophin-related protein', *Nature*, 360(December), pp. 591–593. doi: 10.1038/360591a0.

Tsai, L. H. and Gleeson, J. G. (2005) 'Nucleokinesis in neuronal migration', *Neuron*, 46(3), pp. 383–388. doi: 10.1016/j.neuron.2005.04.013.

Turgay, Y. *et al.* (2017) 'The molecular architecture of lamins in somatic cells', *Nature*, 543(7644), pp. 261–264. doi: 10.1038/nature21382.

Tytgat, H. L. P. *et al.* (2015) 'Endogenous biotin-binding proteins: An overlooked factor causing false positives in streptavidin-based protein detection', *Microbial Biotechnology*, 8(1), pp. 164–168. doi: 10.1111/1751-7915.12150.

Vadrot, N. *et al.* (2014) 'The p.R482W substitution in a-type lamins deregulates SREBP1 activity in dunnigan-type familial partial lipodystrophy', *Human Molecular Genetics*, 24(7), pp. 2096–2109. doi: 10.1093/hmg/ddu728.

Varas, J. *et al.* (2015) 'Absence of SUN1 and SUN2 proteins in Arabidopsis thaliana leads to a delay in meiotic progression and defects in synapsis and recombination', *Plant Journal*, 81(2), pp. 329–346. doi: 10.1111/tpj.12730.

Vásquez-Limeta, A. *et al.* (2014) 'Nuclear import of β-dystroglycan is facilitated by ezrin-mediated cytoskeleton reorganization', *PLoS ONE*, 9(3). doi: 10.1371/journal.pone.0090629.

Vélez-Aguilera, G. *et al.* (2018) 'Control of nuclear β-dystroglycan content is crucial for the maintenance of nuclear envelope integrity and function', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1865(2), pp. 406–420. doi: 10.1016/j.bbamcr.2017.11.013.

Vergnes, L. *et al.* (2004) 'Lamin B1 is required for mouse development and nuclear integrity', *Proceedings of the National Academy of Sciences*, 101(28), pp. 10428–10433. doi: 10.1073/pnas.0401424101.

Verstraeten, V. L. R. M. *et al.* (2008) 'Increased mechanosensitivity and nuclear stiffness in Hutchinson-Gilford progeria cells: Effects of farnesyltransferase inhibitors', *Aging Cell*, 7(3), pp. 383–393. doi: 10.1111/j.1474-9726.2008.00382.x.

De vos, W. H. *et al.* (2011) 'Repetitive disruptions of the nuclear envelope invoke temporary loss of cellular compartmentalization in laminopathies', *Human Molecular Genetics*, 20(21), pp. 4175–4186. doi: 10.1093/hmg/ddr344.

Wang, A. S. and Dreesen, O. (2018) 'Biomarkers of cellular senescence and skin aging', *Frontiers in Genetics*, 9(AUG), pp. 1–14. doi: 10.3389/fgene.2018.00247.

Wang, J.-Y. *et al.* (2015) 'Sun1 deficiency leads to cerebellar ataxia in mice.', *Disease Models & Mechanisms*, 8, pp. 957–967. doi: 10.1242/dmm.019240.

Wang, Q. *et al.* (2006) 'Characterization of the structures involved in localization of the SUN proteins to the nuclear envelope and the centrosome.', *DNA and cell biology*, 25(10), pp. 554–62. doi: 10.1089/dna.2006.25.554.

Wang, R. and Brattain, M. G. (2007) 'The maximal size of protein to diffuse through the nuclear pore is larger than 60 kDa', *FEBS Letters*, 581(17), pp. 3164–3170. doi: 10.1016/j.febslet.2007.05.082.

Watanabe, N. *et al.* (2007) 'Cys669-Cys713 disulfide bridge formation is a key to dystroglycan cleavage and subunit association', *Genes to Cells*, 12(1), pp. 75–88. doi: 10.1111/j.1365-2443.2006.01033.x.

Way, M. *et al.* (1992) 'Expression of the N-terminal domain of dystrophin in E. coli and demonstration of binding to F-actin', *FEBS Letters*, 301(3), pp. 243–245. doi: 10.1016/0014-5793(92)80249-G.

Weber, K., Plessmann, U. and Traub, P. (1989) 'Maturation of nuclear lamin A involves a specific carboxy-terminal trimming, which removes the polyisoprenylation site from the precursor; implications for the structure of the nuclear lamina', *FEBS Letters*, 257(2), pp. 411–414. doi: 10.1016/0014-5793(89)81584-4.

Wei, M. G. *et al.* (1996) 'Assembly of lamins in vitro', *Cell Research*, 6(1), pp. 11–22. doi: 10.1038/cr.1996.2.

Weis, K., Ryder, U. and Lamond, A. I. (1996) 'The conserved amino-terminal domain of hSRP1 alpha is essential for nuclear protein import.', *The EMBO Journal*, 15(8), pp. 1818–1825. doi: 10.1002/j.1460-2075.1996.tb00531.x.

Van Westering, T. L. E., Betts, C. A. and Wood, M. J. A. (2015) 'Current understanding of molecular pathology and treatment of cardiomyopathy in duchenne muscular dystrophy', *Molecules*, 20(5), pp. 8823–8855. doi: 10.3390/molecules20058823.

Wilhelmsen, K. *et al.* (2005) 'Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin', *Journal of Cell Biology*, 171(5), pp. 799–810. doi: 10.1083/jcb.200506083.

Wilkie, G. S. *et al.* (2011) 'Several novel nuclear envelope transmembrane proteins identified in skeletal muscle have cytoskeletal associations', *Molecular and Cellular Proteomics*, 10(1), pp. 1–16. doi: 10.1074/mcp.M110.003129.

Willer, M. K. and Carroll, C. W. (2017) 'Substrate stiffness-dependent regulation of the SRF–Mkl1 co-activator complex requires the inner nuclear membrane protein Emerin', *Journal of Cell Science*, 130(13), pp. 2111–2118. doi: 10.1242/jcs.197517.

Williamson, R. A. *et al.* (1997) 'Dystroglycan is essential for early embryonic development: Disruption of Reichert's membrane in Dag1-null mice', *Human Molecular Genetics*, 6(6), pp. 831–841. doi: 10.1093/hmg/6.6.831.

Wilson, I. A. *et al.* (1984) 'The structure of an antigenic determinant in a protein', *Cell*, 37(3), pp. 767–778. doi: 10.1016/0092-8674(84)90412-4.

Wilson, K. *et al.* (2017) 'Duchenne and Becker Muscular Dystrophies: A Review of Animal Models, Clinical End Points, and Biomarker Quantification', *Toxicologic Pathology*, 45(7), pp. 961–976. doi: 10.1177/0192623317734823.

Wilson, K. L. and Foisner, R. (2010) 'Lamin-binding Proteins', *Cold Spring Harbor Perspectives in Biology*, 2(a000554), pp. 1–18. doi: 10.1101/cshperspect.a000554.

Wilson, M. H. and Holzbaur, E. L. F. (2012) 'Opposing microtubule motors drive robust nuclear dynamics in developing muscle cells', pp. 4158–4169. doi: 10.1242/jcs.108688.

Wilson, M. H. and Holzbaur, E. L. F. (2015) 'Nesprins anchor kinesin-1 motors to the nucleus to drive nuclear distribution in muscle cells', *Development*, 142(1), pp. 218–228. doi: 10.1242/dev.114769.

Witting, N. *et al.* (2014) 'Effect of sildenafil on skeletal and cardiac muscle in Becker muscular dystrophy', *Annals of Neurology*, 76(4), pp. 550–557. doi:

10.1002/ana.24216.

Wong, X., Luperchio, T. R. and Reddy, K. L. (2014) 'NET gains and losses: The role of changing nuclear envelope proteomes in genome regulation', *Current Opinion in Cell Biology*, 28(1), pp. 105–120. doi: 10.1016/j.ceb.2014.04.005.

Wong, X. and Stewart, C. L. (2020) 'The Laminopathies and the Insights They Provide into the Structural and Functional Organization of the Nucleus', *Annual Review of Genomics and Human Genetics*, 21(23), pp. 1–26.

Worman, H. J. (2012) 'Nuclear lamins and laminopathies', *Journal of Pathology*, pp. 316–325. doi: 10.1002/path.2999.

Worman, H. J. and Bonne, G. (2007) '"Laminopathies": A wide spectrum of human diseases', *Experimental Cell Research*, 313(10), pp. 2121–2133. doi: 10.1016/j.yexcr.2007.03.028.

Wright, W. E., Binder, M. and Funk, W. (1991) 'Cyclic amplification and selection of targets (CASTing) for the myogenin consensus binding site.', *Molecular and Cellular Biology*, 11(8), pp. 4104–4110. doi: 10.1128/mcb.11.8.4104.

Xie, W. *et al.* (2016) 'A-type Lamins Form Distinct Filamentous Networks with Differential Nuclear Pore Complex Associations', *Current Biology*, 26(19), pp. 2651–2658. doi: 10.1016/j.cub.2016.07.049.

Xie, Z. *et al.* (2019) 'Clinical and genetic spectrum of sarcoglycanopathies in a large cohort of Chinese patients', *Orphanet Journal of Rare Diseases*, 14(1), pp. 1–13. doi: 10.1186/s13023-019-1021-9.

Xing, S. *et al.* (2016) 'Techniques for the analysis of protein-protein interactions in vivo', *Plant Physiology*, 171(2), pp. 727–758. doi: 10.1104/pp.16.00470.

Yamada, H. *et al.* (2001) 'Processing of β-dystroglycan by matrix metalloproteinase disrupts the link between the extracellular matrix and cell membrane via the dystroglycan complex', *Human Molecular Genetics*, 10(15), pp. 1563–1569. doi: 10.1093/hmg/10.15.1563.

Yang, B. *et al.* (1995) 'SH3 Domain-mediated Interaction of Dystroglycan and Grb2', *J Biol Chem*, 270, pp. 11711–11714.

Yang, L., Guan, T. and Gerace, L. (1997) 'Lamin-binding fragment of LAP2 inhibits increase in nuclear volume during the cell cycle and progression into S phase', *Journal of Cell Biology*, 139(5), pp. 1077–1087. doi: 10.1083/jcb.139.5.1077.

Yang, S. H. *et al.* (2006) 'A farnesyltransferase inhibitor improves disease phenotypes in mice with a Hutchinson-Gilford progeria syndrome mutation', *Journal of Clinical Investigation*, 116(8), pp. 2115–2121. doi: 10.1016/j.bbalip.2007.11.003.

Yao, J. *et al.* (2011) 'Subnuclear segregation of genes and core promoter factors in myogenesis', *Genes and Development*, 25(6), pp. 569–580. doi: 10.1101/gad.2021411.

Yassine, S. *et al.* (2015) 'Dynamics of Sun5 localization during spermatogenesis in wild type and Dpy19l2 knock-out mice indicates that Sun5 is not involved in acrosome attachment to the nuclear envelope', *PLoS ONE*, 10(3), pp. 1–20. doi:

10.1371/journal.pone.0118698.

Yoshida-Moriguchi, T. *et al.* (2010) 'O-Mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding', *Science*, 327(5961), pp. 88–92. doi: 10.1126/science.1180512.

Yoshida-Moriguchi, T. *et al.* (2013) 'SGK196 is a glycosylation-specific O-mannose kinase required for dystroglycan function', *Science*, 341(6148), pp. 896–899. doi: 10.1126/science.1239951.

Yoshida, M. *et al.* (2000) 'Biochemical evidence for association of dystrobrevin with the sarcoglycan-sarcospan complex as a basis for understanding sarcoglycanopathy', *Human Molecular Genetics*, 9(7), pp. 1033–1040. doi: 10.1093/hmg/9.7.1033.

YOSHIDA, M. *et al.* (1994) 'Dissociation of the complex of dystrophin and its associated proteins into several unique groups by n-octyl beta-d-glucoside', *European Journal of Biochemistry*, 222(3), pp. 1055–1061. doi: 10.1111/j.1432-1033.1994.tb18958.x.

Yu, J. *et al.* (2011) 'KASH protein Syne-2/Nesprin-2 and SUN proteins SUN1/2 mediate nuclear migration during mammalian retinal development', *Human Molecular Genetics*, 20(6), pp. 1061–1073. doi: 10.1093/hmg/ddq549.

Yuh, M. C. and Blobel, G. (2001) 'Karyopherins and nuclear import', *Current Opinion in Structural Biology*, 11(6), pp. 703–715. doi: 10.1016/S0959-440X(01)00264-0.

Zaccaria, M. L. *et al.* (2001) 'Dystroglycan distribution in adult mouse brain: A light and electron microscopy study', *Neuroscience*, 104(2), pp. 311–324. doi: 10.1016/S0306-4522(01)00092-6.

Zhang, G. *et al.* (2019) 'Dystroglycan is involved in the activation of ERK pathway inducing the change of AQP4 expression in scratch-injured astrocytes', *Brain Research*, 1721(June), p. 146347. doi: 10.1016/j.brainres.2019.146347.

Zhang, Q. *et al.* (2001) 'Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues.', *Journal of cell science*, 114(Pt 24), pp. 4485–4498. doi: 10.1074/jbc.274.50.35375.

Zhang, Q. *et al.* (2007) 'Nesprin-1 and -2 are involved in the pathogenesis of Emery - Dreifuss muscular dystrophy and are critical for nuclear envelope integrity', *Human Molecular Genetics*, 16(23), pp. 2816–2833. doi: 10.1093/hmg/ddm238.

Zhang, X. *et al.* (2009) 'SUN1/2 and Syne/Nesprin-1/2 Complexes Connect Centrosome to the Nucleus during Neurogenesis and Neuronal Migration in Mice', *Neuron*, 64(2), pp. 173–187. doi: 10.1016/j.neuron.2009.08.018.

Zhao, X. L. *et al.* (2010) 'MMP-mediated cleavage of β-dystroglycan in myelin sheath is involved in autoimmune neuritis', *Biochemical and Biophysical Research Communications*, 392(4), pp. 551–556. doi: 10.1016/j.bbrc.2010.01.062.

Zhen, Y. *et al.* (2002) 'NUANCE, a giant protein connecting the nucleus and actin cytoskeleton', *Journal of Cell Science*, 115(2), pp. 3207–3222.

Zhong, D. *et al.* (2006) 'Characterization of the protease activity that cleaves the extracellular domain of β-dystroglycan', *Biochemical and Biophysical Research Communications*, 345(2), pp. 867–871. doi: 10.1016/j.bbrc.2006.05.004.

Zhou, C. *et al.* (2017) 'Novel nesprin-1 mutations associated with dilated cardiomyopathy cause nuclear envelope disruption and defects in myogenesis', *Human molecular genetics*, 26(12), pp. 2258–2276. doi: 10.1093/hmg/ddx116.

Zhou, Z. *et al.* (2012) 'Structure of Sad1-UNC84 homology (SUN) domain defines features of molecular bridge in nuclear envelope', *Journal of Biological Chemistry*, 287(8), pp. 5317–5326. doi: 10.1074/jbc.M111.304543.

Zhu, R., Antoku, S. and Gundersen, G. G. (2017) 'Centrifugal Displacement of Nuclei Reveals Multiple LINC Complex Mechanisms for Homeostatic Nuclear Positioning', *Current Biology*, 27(20), pp. 3097-3110.e5. doi: 10.1016/j.cub.2017.08.073.

Zuleger, N. *et al.* (2013) 'Specific nuclear envelope transmembrane proteins can promote the location of chromosomes to and from the nuclear periphery', *Genome Biology*, 14(2). doi: 10.1186/gb-2013-14-2-r14.