Factors regulating the function and assembly of the sarcoglycan complex in brain

A thesis submitted for the Degree of Doctor of Philosophy at Cardiff University School of Medicine

Francesca Carlisle

2016

Supervised by Professor Derek Blake and Professor Anthony Isles

Thesis summary

Myoclonus dystonia (MD) is a neurogenic movement disorder that can be caused by mutations in the *SGCE* gene encoding ε-sarcoglycan. ε-sarcoglycan belongs to the sarcoglycan family of cell surface-localised, single-pass transmembrane proteins originally identified in muscle where they form a heterotetrameric subcomplex of the dystrophin-associated glycoprotein complex (DGC). Mutations in the *SGCA*, *SGCB*, *SGCG* and *SGCD* genes encoding α-, β-, γ- and δ-sarcoglycan cause limb-girdle muscular dystrophy (LGMD). There is no phenotypic overlap between MD and LGMD. LGMD-associated sarcoglycan mutations impair trafficking of the entire sarcoglycan complex to the cell surface and destabilise the DGC in muscle, while MD-associated mutations typically result in loss of ε-sarcoglycan from the cell surface. This suggests cell surface ε-sarcoglycan but not other sarcoglycans is required for normal brain function. To gain insight into ε-sarcoglycan's function(s) in the brain, immunoaffinity purification was used to identify ε-sarcoglycan-interacting proteins. Ubiquitous and brain-specific ε-sarcoglycan isoforms co-purified with three other sarcoglycans including ζ-sarcoglycan (encoded by *SGCZ*) from the brain. Incorporation of an LGMD-associated β-sarcoglycan mutant into the brain sarcoglycan complex impaired the formation of the βδ-sarcoglycan core but failed to abrogate the association and trafficking of ε- and ζ-sarcoglycan in heterologous cells. Both ε-sarcoglycan isoforms also co-purified with β-dystroglycan, indicating inclusion in DGC-like complexes. Additionally, the brain-specific ε-sarcoglycan isoform co-purified with the perineuronal net component tenascin-R, potentially suggesting a unique function for this isoform in modulating synapses. In common with *SGCE*, transcripts from the genes encoding α-, β-, δ-, γ- and ζ-sarcoglycans were found to undergo extensive alternative splicing, in some cases producing novel isoforms that affected assembly and trafficking of the sarcoglycan complex. In summary, data presented herein show that alternatively spliced sarcoglycan isoforms are part of the DGC in brain. These data contribute to our understanding of MD pathophysiology and the role of the sarcoglycan protein family.

Acknowledgements

Firstly, I would like to thank my supervisors Professor Derek Blake and Professor Anthony Isles for giving me the opportunity to work on this project. It has been a fascinating and challenging experience.

I am immensely grateful to all those who have helped me out in the lab with training, advice, cheer, commiserations and desperate begging for just a tiny bit of this one reagent. In particular, I would like to thank Adrian for his endless advice, support, and willingness to share his ridiculously impressive knowledge of biochemistry. Other stars of the last three years include Alis, Matt, Taniesha, Anna, the other Anna, Lesley, Alex, Magda and Kira. Their help got me through many a trying day.

Eloise, Jenny and Sarah – thank you from the bottom of my heart for the coffee dates, bake off hysteria, science jokes and everything else. You helped me keep it together when things got dark, and celebrated with me when things went well. Thank you also to my many friends at Ceroc South Wales, for the silliness and fun we had at each class. Without your good cheer and friendship, I would probably be irredeemably insane by now. To my fellow Aslam House princesses, Alison, Hannah, Stacy and Laura: we made it!!!

Last but definitely not least, to my family: Mum, Dad, Emily, Charlie, Grandma, Sally, and my innumerable extended family – I cannot properly express how grateful I am to you for the support, encouragement and love. Even when I thought I couldn't do it, you did. Thank you.

Abbreviations

dATP	2’-deoxyadenosine 5’-triphosphate
dCTP	2’-deoxycytidine 5’-triphosphate
dGTP	2’-deoxyguanosine 5’-triphosphate
dNTP	2’-deoxynucleoside 5’-triphosphate
dTTP	2’-deoxythymidine 5’-triphosphate
ADom	autosomal dominant
AR	autosomal recessive
ATP	adenosine triphosphate
AP	affinity purification
α-SG	alpha sarcoglycan
aa	amino acid
N-terminus	amino-terminus
APS	ammonium persulfate
Amp	ampicillin
bp	base pair
BLAST	basic local alignment search tool
β-SG	beta sarcoglycan
BLAT	BLAST-like alignment tool
BSA	bovine serum albumin
CIP	calf intestinal alkaline phosphatase
CO_2	carbon dioxide
C-terminus	carboxyl-terminus
CNS	central nervous system
cDNA	complementary DNA
°C	degrees Celsius
δ-SG	delta sarcoglycan
DNA	deoxyribonucleic acid
DMP	dimethylpimelimidate dihydrochloride
DMSO	dimethylsulphoxide
DRD	dopamine-responsive dystonia
DTT	dithiothreitol
DMD	Duchenne muscular dystrophy
DMEM	Dulbeccos modified Eagle medium
DGC	dystrophin-associated glycoprotein complex
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum-associated degradation
ε-SG	epsilon sarcoglycan
EDTA	ethylenediaminetetraacetic acid
EGTA	ethyleneglycoltetraacetic acid
ECACC	European Collection of Authenticated Cell Cultures
EST	expressed sequence tag
ECM	extracellular matrix
FBS	foetal bovine serum
fMRI	functional magnetic resonance imaging
γ-SG	gamma sarcoglycan
GST	glutathione-S-transferase
g	gram

HA	human influenza haemagglutinin
Tris-HCl	HCl-buffered tris(hydroxymethyl)aminomethane
h	hour
HEK	human embryonic kidney
IAP	immunoaffinity purification
IgG	immunoglobulin G
IP	immunoprecipitation
kb	kilobase
kDa	kilodalton
LGMD	limb-girdle muscular dystrophy
l	litre
LB	Luria-Bertani media
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
miRNA	micro ribonucleic acid
µg	microgram
µl	microliter
µM	micromolar
ml	millilitre
mM	millimolar
min	minute
M	molar
MD	myoclonus dystonia
TEMED	N,N,N',N'-tetramethylethylenediamine
ng	nanogram
nM	nanomolar
NCBI	National Centre for Biotechnology Information
NE	nuclear envelope
nt	nucleotide
OD	optical density
PNN	perineuronal net
PBS	phosphate-buffered saline
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PET	positron emission tomography
PTC	premature termination codon
pre-mRNA	pre-messenger ribonucleic acid
RIPA	radioimmunoprecipitation assay
RT-PCR	reverse transcriptase polymerase chain reaction
rpm	revolutions per minute
RNA	ribonucleic acid
SG	sarcoglycan
SGC	sarcoglycan complex
s	second
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	standard error of the mean
Tet	tetracycline
Thx	thioredoxin
TMS	transcranial magnetic stimulation

Tris	tris(hydroxymethyl)aminomethane
TBST	tris-buffered saline-Tween-20
U	units
UTR	untranslated region
v/v	volume to volume ratio
w/v	weight to volume ratio
Zeo	zeocin
ζ-SG	zeta sarcoglycan
β-me	β-mercaptoethanol

Table of contents

Thesis summary i
Declarations ii
Acknowledgements iii
Abbreviations iv
Table of contents vii
List of figures and tables xi
Chapter 1: General introduction 1
1.1. Introduction 1
1.2. Overview of dystonia 1
1.2.1. Classification of dystonia 2
1.2.1.1. Axis I: clinical features 3
1.2.1.2. Axis II: aetiology 3
1.2.2. Neuroanatomy 4
1.2.3. Neurophysiological abnormalities in dystonia 5
1.2.3.1. Loss of inhibition 5
1.2.3.2. Impaired sensorimotor integration 6
1.2.3.3. Maladaptive neural plasticity 6
1.2.4. The monogenic dystonias: insights into the molecular aetiology of dystonia 7
1.2.4.1. Isolated dystonias 7
1.2.4.2. Combined dystonias 14
1.2.4.3. Insights from the monogenic dystonias 17
1.3. Myoclonus-dystonia (MD) 19
1.3.1. Overview of MD 19
1.3.2. Mutations in *SGCE* cause MD 19
1.3.2.1. *SGCE* mutation-positive MD 19
1.3.2.2. *SGCE* encodes ε-sarcoglycan 21
1.3.2.3. *SGCE* mutations associated with MD 23
1.3.2.4. Mouse models of *SGCE* mutation-positive MD 23
1.3.3. Neuroanatomy and physiology 24
1.3.4. MD: a genetically heterogeneous syndrome? 25
1.3.4.1. *DRD2* and *TOR1A* 25
1.3.4.2. New genes associated with MD and MD-like phenotypes: *RELN, CACNA1B,* and *KCTD17* 25
1.4. The sarcoglycans: a family of transmembrane glycoproteins 27
1.4.1. Overview of the sarcoglycans 27
1.4.2. Synthesis and assembly of sarcoglycan complexes 29
1.4.3. Sarcoglycans as part of the dystrophin-associated glycoprotein complex 30
1.4.4. The sarcoglycans in human disease 34
1.4.5. Animal models of sarcoglycan mutations 36
1.4.6. Function(s) of the sarcoglycans 37
1.5. Thesis aims 39
Chapter 2: Materials and methods 40
2.1. Molecular Biology 40
2.1.1. First strand synthesis from total RNAs 40
2.1.2. Oligonucleotides (primers) 40
2.1.3. Polymerase chain reaction 41
2.1.3.1. RT-PCR from first strand cDNA 41

2.1.3.2. General PCR....43
2.1.3.3. Colony PCR....43
2.1.3.4. Site-directed mutagenesis....44
2.1.4. Agarose gel electrophoresis....45
2.1.5. PCR product purification....45
2.1.6. Nucleic acid quantitation....45
2.1.7. Cloning into plasmid vectors....46
2.1.7.1. Vectors used and constructs produced....46
2.1.7.2. Restriction digest of DNA....47
2.1.7.3. Ligation of cohesive DNA fragment termini....48
2.2. Plasmid amplification and protein expression in *E.coli*....48
2.2.1. Preparation of chemically competent *E.coli* XL1-Blue....48
2.2.2. Transformation of XL1-Blue cells using the heat-shock method....49
2.2.3. Transformation of XL10-Gold cells using the heat-shock method....49
2.2.4. Transformation of MAX Efficiency Stbl2 Competent cells using the heat-shock method 50
2.2.5. Preparation of plasmid DNA from *E.coli*....50
2.2.6. Expression of recombinant murine Thx-ε-SG C-terminus fusion protein in *E.coli* using the pET-32 system....51
2.2.7. Purification of recombinant murine Thx-ε-SG isoform 1 C-terminus fusion protein from *E.coli*....51
2.3. Antibody preparation....52
2.3.1. Generation of ε-SG isoform 2-specific antibodies....52
2.3.2. Preparation of peptide immunogen affinity chromatography column....53
2.3.3. Preparation of denatured fusion protein affinity chromatography column....53
2.3.4. Antibody purification....54
2.3.4.1. Purification of polyclonal antibodies from terminal bleed serum....54
2.3.4.2. Purification of monoclonal antibodies from hybridoma culture media....55
2.3.5. Cross-linking of antibodies to Protein A-agarose or Protein G-sepharose beads..56
2.3.6. Antibodies used in this study....58
2.4. Tissue culture....60
2.4.1. Mammalian cell line....60
2.4.2. Standard cell culture conditions....60
2.4.3. Transfection of mammalian cell cultures....61
2.4.4. Treatment of cells with bortezomib to inhibit proteasome activity....61
2.5. Protein analysis....61
2.5.1. General sample preparation for SDS-PAGE and Western blots....61
2.5.1.1. Sample preparation from tissue....61
2.5.1.2. Sample preparation from cultured cells....62
2.5.2. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) under denaturing conditions....62
2.5.3. Coomassie staining of SDS-PAGE gels....63
2.5.4. Western blotting....64
2.5.5. Quantitation of Western blots....65
2.6. Proteomics....66
2.6.1. PNGase F treatment of recombinant protein....66
2.6.2. Immunoprecipitation from cell lines....66
2.6.3. Cell surface biotinylation....67
2.6.4. Immunoaffinity purification (IAP) from mouse tissue....68
2.6.5. IAP sample preparation for mass spectrometry....70

2.6.6. Protein identification by mass spectrometry 70
2.7. General bioinformatics 71
Chapter 3: ε-SG protein interactions in the brain 72
3.1. Introduction 72
3.2. Results 73
3.2.1. Generation and validation of the esg2-1358 antibody specific to ε-SG isoform 2 73
3.2.2. Identification of a brain sarcoglycan complex containing ε-SG isoform 2 75
3.2.2.1. ε-SG isoform 2 co-purifies with β-SG, δ-SG and ζ-SG in mouse brain 75
3.2.2.2. Alternative splicing of ε-SG does not affect formation and localisation of an εβδζ tetramer in heterologous cells 85
3.2.3. Dp71 and β-dystroglycan co-purify with ε-SG under mild lysis conditions 91
3.2.4. The Shaker-related voltage-gated potassium channels cross-react with the esg2-1358 antibody 94
3.2.5. Identifying other components of ε-SG-containing complexes in the brain 95
3.2.5.1. IAP MS data were filtered to remove non-specifically interacting and cross-reacting proteins 97
3.2.5.2. Identification of high-confidence ε-SG interactions in multiple mouse strains 103
3.2.5.3. Evaluation of an *Sgce* gene trap mouse line to identify esg2-1358 cross-reacting proteins 106
3.2.6. ε-SG isoform 2 co-purifies with Tenascin-R from mouse brain. 109
3.3. Discussion 112
Chapter 4: Alternative splicing of sarcoglycan pre-mRNA 119
4.1. Introduction 119
4.2. Results 120
4.2.1. Few alternatively spliced *SGCB* transcripts were identified in cerebellum and skeletal muscle 124
4.2.2. Numerous alternatively spliced *SGCD* transcripts were identified in cerebellum and skeletal muscle 129
4.2.3. Alternatively spliced *SGCZ* transcripts in the cerebellum were varied 138
4.2.4. *SGCA* pre-mRNA undergoes extensive alternative splicing in skeletal muscle .143
4.2.5. Few alternatively spliced *SGCG* transcripts were identified 149
4.3. Discussion 154
Chapter 5: Characterisation of alternatively spliced sarcoglycan isoforms 159
5.1. Introduction 159
5.2. Results 159
5.2.1. Alternatively spliced β-SG isoforms 161
5.2.1.1. β-SG isoform 2 is degraded by the proteasome. 161
5.2.1.2. Lack of β-SG alters sarcoglycan complex assembly 162
5.2.2. Alternatively spliced δ-SG isoforms 165
5.2.2.1. Some alternatively spliced *SGCD* transcripts were translated to produce glycosylated δ-SG isoforms 165
5.2.2.2. Alternatively spliced δ-sarcoglycan isoforms do not affect assembly or trafficking of sarcoglycan εβδζ heterotetramers 167
5.2.3. Alternatively spliced ζ-SG isoforms 169
5.2.3.1. ζ-SG isoforms are translated and N-glycosylated 169
5.2.3.2. ζ-SG isoforms 2 and 3 interfere with the assembly and trafficking of sarcoglycan heterotetramers 169
5.2.4. Alternatively spliced γ-SG isoforms 171

5.2.4.1. Alternatively spliced *SGCG* transcripts encode glycosylated γ-SG isoforms. 171
5.2.4.2. γ-SG isoforms alter sarcoglycan complex assembly and trafficking 171
5.2.5. Alternatively spliced α-SG isoforms 173
5.2.5.1. Alternative α-SG isoforms were N-glycosylated, and stabilised by the proteasome 173
5.2.5.2. Alternatively spliced α-SG isoforms had different effects on the assembly and trafficking of the SGC.....175
5.3. Discussion 175
Chapter 6: General discussion 184
6.1. Introduction 184
6.2. ε-SG protein interactions in the brain 185
6.3. Alternative splicing of *SGCA*, *SGCB*, *SGCD*, *SGCG* and *SGCZ* pre-mRNA 189
6.4. Future directions 191
Bibliography 192
Appendix I: Oligonucleotide primers 227
Appendix II: Pre-immune immunoglobulin IAP, CD-1 mice 231
Appendix III: Pre-immune immunoglobulin IAP, C57BL/6J mice.....247
Appendix IV: Papers and presentations arising from this thesis 252

List of figures and tables

Figures

Figure 1.1 Overview of cell pathways implicated in dystonia by the monogenic dystonias, after (Lohmann and Klein, 2013). The proteins encoded by dystonia-associated genes have been implicated in a number of cellular pathways. These are summarised here, but as research into the functions of these genes continues this list will likely require amendment. 18

Figure 1.2 Genomic architecture, alternative splicing and isoforms of human SGCE encoding ε-SG. A) The genomic architecture of the human SGCE gene, not to scale. After Ritz et al., 2011. B) The splicing pattern for each of the three main SGCE transcripts is depicted against the basic genomic architecture. Lines connect exons included in the mature transcript. The polypeptide encoded by each transcript is illustrated below with major domains marked. Red shading marks the signal peptide, which is cleaved off to produce the mature protein; yellow marks the transmembrane domain; blue marks the peptide encoded by exon 8; purple or pink represent isoform-specific sequences; and the red N represents the N-linked glycosylation site. 21

Figure 1.3 Comparison of human sarcoglycan proteins. For each sarcoglycan protein, the full-length/canonical isoform 1 is depicted. Transmembrane domains are marked in yellow, signal peptides marked in red, N marks an N-linked glycosylation site, and green lines represent disulphide bonds between cysteine residues. The intracellular and extracellular regions are marked against the polypeptide. Numbers at the C-terminus of each protein give the polypeptide length before processing. Polypeptides are approximately to scale. ε-SG and α-SG signal peptides are as reported in the Leiden Open Variation Database (Fokkema et al., 2011). 28

Figure 1.4 Putative molecular organisation of the DGC in muscle and DGC-like complexes in the brain. Adapted from (Waite et al., 2009). Illustration of the basic organisation for the core DGC and DGC-like complexes in A) muscle, B) CNS neurons, and C) CNS glia. The muscle DGC has been determined through direct biochemical purification of the intact complex from tissue, whereas CNS DGC-like complexes have been inferred from protein interactions identified using a variety of protocols including yeast two-hybrid and co-immunoprecipitation from cultured cells and/or tissue. Only the basic structure and a subset of associated proteins are illustrated for each complex, and full details can be obtained in recent reviews (Allen et al., 2016, Waite et al., 2012, Waite et al., 2009). Arrows with question marks indicate putative but not confirmed protein interactions. Dystrobrevin and dystrophin are not to scale. Abbreviations: DGC, dystrophin-associated glycoprotein complex; DG, dystroglycan; SG, sarcoglycan; SGC, sarcoglycan complex; SSPN, sarcospan; nNOS, neuronal-type nitric oxide synthase; SAST, syntrophin-associated serine/threonine kinase; ABD, actin binding domain; PH, plecstrin homology domain; PDZ, PSD-95 (postsynaptic density protein 95), discs_large and zonula occludens-1 domain; SU, syntrophin-unique region; Kir 4.1, inwardly rectifying potassium channel 4.1. 32

Tables

Table 1.1 Classification of dystonia according to (Albanese et al., 2013). Under the most recent consensus update on dystonia classification, dystonia can be divided along two axes: clinical characteristics and aetiology. This table provides the complete proposed system for classifying dystonia. 2
Table 1.2 The monogenic dystonias. After Klein 2014. A complete list of the current genetically defined inherited dystonias. These are split by associated features in isolated dystonias (dystonia and tremor only) and combined dystonias (dystonia plus other signs). Disorders are marked as confirmed if they have been independently described by 2 different groups. Novel disorders were identified too recently for confirmation. COL6A3 is marked as dubious due to recent evidence that mutations in this gene may be common in healthy controls. Abbreviations: ADom, autosomal dominant; AR, autosomal recessive, XR, X-linked recessive. 9
Table 1.3 The sarcoglycan protein family with associated diseases. Complete list of human sarcoglycan proteins with gene name, genomic locus, polypeptide size in amino acids, apparent molecular weight on SDS-PAGE, number of N-linked glycans, number of disulphide bonds, and the associated disease. Protein information obtained from UniProt database. Abbreviations: SG, sarcoglycan; LGMD, limb-girdle muscular dystrophy; MD, myoclonus dystonia. 27

Chapter 1: General introduction

1.1.Introduction

Dystonia is a movement disorder characterised by involuntary, abnormal muscle contractions leading to abnormal postures and/or movements. It was first described in 1911 by Hermann Oppenheim as "dystonia musculorum deformans" (Albanese et al., 2013). Though for many years regarded as a psychogenic condition, it is now commonly accepted that dystonia arises from neurological dysfunction (Frucht, 2013). It can occur in a number of contexts including neurological trauma, infection, metabolic disease and as part of a genetic disorder (Geyer and Bressman, 2006).

This chapter will provide an overview of dystonia with particular emphasis on the monogenic dystonias. The focus of this thesis is on Myoclonus Dystonia (MD), 20-80% of cases of which are caused by mutations in *SGCE* encoding ε-sarcoglycan (Asmus et al., 2009, Gerrits et al., 2006, Nardocci et al., 2008, Peall et al., 2013, Ritz et al., 2009, Tezenas du Montcel et al., 2006). Therefore, MD and the sarcoglycans will be discussed in detail. Thesis aims will be outlined at the end of the chapter.

1.2.Overview of dystonia

The abnormal movements and postures seen in patients with dystonia are generated by simultaneous activation of both agonist and antagonist muscles. Dystonic movements tend to be patterned, can display motor overflow into adjacent muscles, and involve the same group of muscles each time. Voluntary movements may exacerbate or trigger dystonic movements (Albanese et al., 2013, Geyer and Bressman, 2006, Phukan et al., 2011, Sitburana et al., 2009). Some patients find that a sensory trick ("geste antagoniste") such as a light touch to

Axis I. Clinical characteristics		
Age at onset	Infancy (birth-2yrs) Childhood (3-12yrs) Adolescence (13-20yrs) Early adulthood (21-40yrs) Late adulthood (>40yrs)	
Body distribution	Focal Segmental Multifocal Generalised Hemidystonia	
Temporal pattern	Disease course	Static Progressive
	Day-to-day	Persistent Diurnal Action-specific Paroxysmal
Associated features	Isolated (dystonia and tremor only) Combined with other movement disorder(s) Additional neurological/systemic features	
Axis II. Aetiology		
Origin	Inherited	Autosomal dominant Autosomal recessive X-linked recessive Mitochondrial
	Acquired	Brain injury Vascular Neoplastic Perinatal brain injury Toxin Drug Infection Psychogenic
	Idiopathic	Sporadic Familial
Neuropathology	Neurodegeneration Structural lesions Neither	

Table 1.1 Classification of dystonia according to (Albanese et al., 2013). Under the most recent consensus update on dystonia classification, dystonia can be divided along two axes: clinical characteristics and aetiology. This table provides the complete proposed system for classifying dystonia.

the affected area can temporarily ameliorate dystonia (Kagi et al., 2013, Poisson et al., 2012, Ramos et al., 2014). Most cases of dystonia are not associated with gross structural abnormalities on cerebral imaging, and therefore dystonia is typically regarded as a disorder of neurofunction (Breakefield et al., 2008, Geyer and Bressman, 2006).

1.2.1. Classification of dystonia

The most recent consensus on dystonia phenomenology and classification subdivides dystonia along two principal axes: clinical features and aetiology, as summarised in table 1.1 (Albanese et al., 2013).

1.2.1.1. Axis I: clinical features

A major clinical classification is by distribution of dystonia-affected body parts. Focal dystonias affect a single region of the body, and include cervical dystonia, blepharospasm, writer's cramp or hand dystonia, and spasmodic dysphonia or laryngeal dystonia. If dystonia affects multiple contiguous muscle groups it is referred to as segmental, in contrast to multifocal dystonia where multiple non-contiguous muscle groups are affected. Hemidystonia involves multiple muscle groups on one side of the body only, while generalised dystonia involves the trunk and at least two other body regions.

Age at onset is grouped into five classes: infancy (0-2 years at onset), childhood (3-12 years), adolescence (13-20 years), early adulthood (21-40 years) or late adulthood (>40 years) (Albanese et al., 2013). Early onset typically correlates with a more severe phenotype that is more likely to progress with time (Albanese et al., 2013). However, a substantial delay between onset and diagnosis can complicate determining age at onset (Phukan et al., 2011).

Dystonia may also be subdivided by whether it occurs in isolation (only dystonia is present) or in combination with other movement disorder(s) (combined dystonia) (Albanese et al., 2013, Balint and Bhatia, 2014). The isolated dystonias include syndromes featuring dystonia and tremor only, whereas combined dystonias can feature other movement disorders such as myoclonus, parkinsonism or spasticity (Albanese et al., 2013). In combined dystonias, the non-dystonic movement disorder(s) may be more prominent and/or the presenting feature (Albanese et al., 2013, Balint and Bhatia, 2014, Fung et al., 2013).

1.2.1.2. Axis II: aetiology

Within the aetiology axis, the principle division is between inherited, acquired and idiopathic dystonias (Albanese et al., 2013). Acquired dystonias stem from known insults to the central nervous system (CNS), while inherited dystonias have a defined genetic cause. Idiopathic

dystonias currently have no known cause, but may be reassigned to one of the other two categories with further study.

1.2.2. Neuroanatomy

Multiple brain regions have been implicated in the pathogenesis of dystonia. Originally, the basal ganglia were regarded as the principal brain region involved in dystonia as basal ganglia lesions were frequently observed in association with dystonia in both humans and animal models (Bhatia and Marsden, 1994, Lehéricy et al., 2013, Neychev et al., 2011, Wilson and Hess, 2013). However, lesional studies have also implicated other regions including the cerebellum, cortex and brainstem in dystonia (LeDoux and Brady, 2003, Matsumura et al., 1992, Matsumura et al., 1991, Neychev et al., 2011, Prudente et al., 2013). Notably, subtle cerebellar alterations such as reduced Purkinje cell numbers have been identified in some patients with dystonia, while specific disruption of cerebellar function can produce dystonia in model animals (LeDoux and Brady, 2003, Ma et al., 2012, Pizoli et al., 2002, Prudente et al., 2013, Raike et al., 2013).

Neuroimaging studies have also implicated multiple regions in dystonia. Case-control analyses using voxel-based morphometry, positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) have consistently demonstrated differences in the basal ganglia, cortex, cerebellum and thalamus of patients with dystonia compared to healthy controls (Carbon et al., 2009, Carbon et al., 2004, Draganski et al., 2009, Egger et al., 2007, Lehéricy et al., 2013, Suzuki et al., 2007). Diffusion tensor imaging (DTI) examines water movement in the brain, which is influenced by the white matter microstructure; DTI studies of patients with dystonia have found altered connectivity between motor control regions including the cerebellum and basal ganglia (Argyelan et al., 2009, Carbon et al., 2008, Carbon et al., 2004, Cheng et al., 2012).

Animal models also implicate multiple brain regions in dystonia. Lesions and administration of stimulants or inhibitors to the basal ganglia, cerebellum, pre-motor cortex, motor cortex, and thalamus can all induce dystonia in animal models (Neychev et al., 2011). Cell type-specific knockout of genes including *Atp1a3*, *Itpr1*, and *Sgce* (section 1.2.4) in the cerebellum or striatum can also produce dystonic phenotypes (Fremont et al., 2015, Hisatsune et al., 2013, Yokoi et al., 2012a, Yokoi et al., 2012b). Additionally, dystonia in the dt rat (*Atcay* mutation) and Tottering mouse (*Cacna1a* mutation) can be prevented by ablation of the cerebellum, cerebellar outputs, or the cerebellar Purkinje cells alone (Campbell et al., 1999, LeDoux, 2011, LeDoux et al., 1993, LeDoux et al., 1995, Neychev et al., 2008).

1.2.3. Neurophysiological abnormalities in dystonia

The general neurophysiological abnormalities observed in dystonia can be grouped into three broad themes: loss of inhibition, sensory dysfunction, and abnormal neural plasticity (Quartarone and Hallett, 2013).

1.2.3.1. Loss of inhibition

Normal movements are thought to involve both activation of the required muscle(s) and surround inhibition of other muscles to ensure movement specificity. Failure of surround inhibition would cause movement overflow into unintended muscles, as seen in dystonia (Hallett, 2011, Quartarone and Hallett, 2013). Reduced surround inhibition has been demonstrated for several dystonias including focal hand, cervical, blepharospasm, segmental and generalised dystonias (Berardelli et al., 1985, Nakashima et al., 1989, Quartarone et al., 2005, Ridding et al., 1995, Rona et al., 1998, Tisch et al., 2006a, Tisch et al., 2006b). GABAergic signalling may be involved in surround inhibition, since injecting the GABA-A antagonist bicuculline into the motor or pre-motor cortex of monkeys results in increased neuronal activity and co-contraction of antagonist muscles during movement (Matsumura et

al., 1992, Matsumura et al., 1991). Subsequent PET, magnetic resonance spectroscopy (MRS) and MRI studies have provided mixed support for GABA abnormalities in the basal ganglia and cortex of dystonia patients compared to healthy controls (Garibotto et al., 2011, Herath et al., 2010, Levy and Hallett, 2002).

1.2.3.2. Impaired sensorimotor integration

Although dystonia can appear a "pure" motor disorder, sensory symptoms have been reported in many forms of dystonia: pain is associated with cervical dystonia, while blepharospasm patients often report photosensitivity and other ocular symptoms (Patel et al., 2014). Sensory input can modify dystonia as demonstrated by symptom reduction with sensory tricks, reduced dystonic movements after topical anaesthetic application, and worsening of dystonia by vibration (Kaji et al., 1995). Patients with dystonia also have difficulty in discriminating between temporally or spatially separated stimuli (Bara-Jimenez et al., 2000a, Bara-Jimenez et al., 2000b, Molloy et al., 2003, Scontrini et al., 2009). These deficits might relate to the disorganisation of somatotopic representation reported in both humans with dystonia and primate models of dystonia (Patel et al., 2014, Quartarone and Hallett, 2013). This suggests a defect of integration between sensory input and motor control in dystonia.

1.2.3.3. Maladaptive neural plasticity

Neural plasticity allows the brain to remodel connections in response to changes in input and desired output, but abnormal plasticity is thought to erode motor control and so contribute to dystonia (Neychev et al., 2011, Quartarone and Hallett, 2013, Sadnicka et al., 2014). Several studies have used transcranial magnetic stimulation (TMS) to identify enhanced plasticity of multiple regions in patients with different types of focal dystonia (Meunier et al., 2012, Quartarone et al., 2003, Quartarone et al., 2008, Weise et al., 2011, Weise et al., 2006). A similar study of individuals with dystonia-associated *TOR1A* mutations (section 1.2.4.1)

identified enhanced plasticity in individuals with dystonia but reduced plasticity in non-manifesting carriers, suggesting a protective effect of reduced plasticity (Edwards et al., 2006). This supports abnormal plasticity as a driver of dystonia rather than a consequence.

1.2.4. The monogenic dystonias: insights into the molecular aetiology of dystonia

Since the late 1990s, substantial progress has been made in identifying the genes responsible for inherited dystonias. There are approximately 21 genetically defined dystonic disorders at the time of writing, although evidence for some of these is still preliminary (table 2.1).

1.2.4.1. Isolated dystonias

TOR1A: DYT1 early-onset generalised dystonia

The most extensively researched heritable form of isolated dystonia is DYT1 early-onset generalised dystonia caused by autosomal dominant mutations in *TOR1A*. This severe dystonic disorder typically begins in early life (5-28 years) with focal limb dystonia that then spreads, usually sparing the craniocervical region (Bressman et al., 2000). It has a penetrance of 30-40%, in part due to a missense SNP that complements pathogenic mutations (Bressman et al., 1989, Risch et al., 1990, Risch et al., 2007). The most common *TOR1A* mutation is deletion of a GAG trinucleotide (ΔGAG), which removes a glutamate residue from the encoded torsinA protein to produce torsinA ΔE (Ozelius et al., 1997, Vulinovic et al., 2014).

TOR1 is broadly expressed throughout the adult CNS with localisation of torsinA protein to the endoplasmic reticulum (ER) lumen, nuclear envelope (NE) and synaptic vesicles in neurons (Augood et al., 1999, Augood et al., 1998, Konakova et al., 2001, Konakova and Pulst, 2001, Naismith et al., 2004, Shashidharan et al., 2000). In heterotetrameric complexes with either NE-localised lamina-associated polypeptide 1 (LAP1) or ER-localised luminal domain like LAP1 (LULL1), torsinA becomes a functional AAA+ ATPase

DYT designation	Disorder	Gene	Encoded protein	Inheritance	Status
Isolated dystonia					
DYT1	Early-onset generalised dystonia	*TOR1A*	TorsinA	ADom	Confirmed
DYT6	Adolescent-onset dystonia of mixed type	*THAP1*	Thanatos-associated protein domain-containing apoptosis-associated protein 1	ADom	Confirmed
DYT25	Adult onset cranial-cervical dystonia	*GNAL*	Gαolf	ADom	Confirmed
DYT4	Spasmodic/"whispering" dysphonia	*TUBB4A*	β-tubulin 4A	ADom	Confirmed
DYT23	Adult onset cranial-cervical dystonia	*CIZ1*	Cdkn1A-interacting zinc finger protein 1	ADom	Unconfirmed
DYT24	Adult onset cranial-cervical dystonia	*ANO3*	Anoctamin-3/TMEM16C	ADom	Unconfirmed
-	Primary isolated dystonia	*HPCA*	Hippocalcin	AR	Novel
DYT27	Isolated dystonia	*COL6A3*	Collagen alpha-3 (VI) chain	AR	Dubious
Combined dystonia					
DYT3	X-linked dystonia-parkinsonism, "lubag"	*TAF1*	Transcription initiation factor TFIID subunit 1	XR	Confirmed
DYT5	Dopa-responsive dystonia, Segawa syndrome	*GCH1*	GTP cyclohydrolase 1	ADom	Confirmed
	Dopa-responsive dystonia, Segawa syndrome	*TH*	Tyrosine hydroxylase	AR	Confirmed
	Dopa-responsive dystonia	*SPR*	Sepiapterin reductase	AR	Confirmed
DYT8	Paroxysmal nonkinesigenic dyskinesia 1	*MR1*	Myofibrillogenesis regulator 1	ADom	Confirmed
DYT9/ DYT18	Paroxysmal choreoathetosis with episodic ataxia and spasticity/Paroxysmal exertion-induced dyskinesia	*SLC2A1*	Glucose transporter 1	ADom	Confirmed
DYT10	Paroxysmal kinesigenic dyskinesia and infantile convulsions	*PRRT2*	Proline-rich transmembrane protein 2	ADom	Confirmed
DYT11	Myoclonus dystonia	*SGCE*	ε-sarcoglycan	ADom (imprinted)	Confirmed
DYT12	Rapid-onset dystonia-parkinsonism	*ATP1A3*	Sodium/potassium-transporting ATPase subunit alpha-3	ADom	Confirmed
DYT16	Young-onset dystonia-parkinsonism	*PRKRA*	Interferon-inducible double-stranded RNA-dependent protein kinase activator A	AR	Confirmed

-	Myoclonus dystonia-like	*CACNA1B*	Voltage-dependent N-type calcium channel subunit alpha-1B	ADom	Novel
-	Myoclonus dystonia	*RELN*	Reelin	ADom	Novel
-	Myoclonus dystonia	*KCTD17*	BTB/POZ domain-containing protein KCTD17	ADom	Novel

Table 1.2 The monogenic dystonias. After Klein 2014. A complete list of the current genetically defined inherited dystonias. These are split by associated features in isolated dystonias (dystonia and tremor only) and combined dystonias (dystonia plus other signs). Disorders are marked as confirmed if they have been independently described by 2 different groups. Novel disorders were identified too recently for confirmation. *COL6A3* is marked as dubious due to recent evidence that mutations in this gene may be common in healthy controls. Abbreviations: ADom, autosomal dominant; AR, autosomal recessive, XR, X-linked recessive.

(Brown et al., 2014, Shashidharan et al., 2000, Sosa et al., 2014, Vander Heyden et al., 2009, Zhao et al., 2013a). NE TorsinA-LAP1 complexes rapidly induce structural reorganisation through LINC complex component displacement, implicating torsinA in nuclear envelope structure (Nery et al., 2008, Vander Heyden et al., 2009). In the ER, torsinA co-localises with known ER-associated protein degradation (ERAD) proteins and its absence inhibits degradation of specific ERAD targets (Nery et al., 2011). Consequently, torsinA is thought to contribute to the degradation of misfolded proteins and therefore inhibit stress responses. In keeping with this, overexpression of a *TOR1A* paralogue in *C.elegans* inhibits ER stress response onset while mouse fibroblasts deficient in torsinA have an increased stress response; torsinA expression can also reduce protein aggregates in cell and *C.elegans* models (Caldwell et al., 2003, Chen et al., 2010, McLean et al., 2002). Finally, torsinA interacts directly with the vesicle exocytosis protein snapin and motor complex component kinesin light chain 1 (KLC1), and promotes endocytosis of polytopic membrane proteins (Granata et al., 2008, Kamm et al., 2004, Torres et al., 2004). TorsinA also inhibits neurite outgrowth, possibly through interactions with the cytoskeletal proteins vimentin and tubulin (Ferrari-Toninelli et al., 2004, Hewett et al., 2006).

The ΔE mutation disrupts many of these interactions. TorsinA ΔE demonstrates enhanced translocation to the NE, but even in complex with LAP1 or LULL1 cannot hydrolyse ATP and therefore only slowly induces remodelling of the NE (Gonzalez-Alegre and Paulson, 2004, Goodchild and Dauer, 2004, Naismith et al., 2009, Vander Heyden et al., 2009, Zhao et al., 2013a). In addition, torsinA ΔE cannot suppress protein aggregation and reduces ERAD efficiency compared to the wild-type protein (Caldwell et al., 2003, Nery et al., 2011). This is hypothesized to lead to toxic accumulation in neurons that compromises their function, potentially resulting in dystonia alongside disrupted NE structure (Nery et al., 2011).

THAP1: DYT6 adolescent-onset dystonia of mixed type

The second dystonia gene identified was *THAP1*, partially penetrant dominant mutations in which cause an adolescent-onset dystonia that is typically craniocervical or upper limb at first with subsequent progression to segmental or generalised dystonia (Blanchard et al., 2011, Fuchs et al., 2009, Houlden et al., 2010, Xiromerisiou et al., 2012). *THAP1* is expressed throughout the brain with temporal and regional regulation, and encodes Thanatos-associated protein domain-containing, apoptosis-associated protein 1 (THAP1) (Cayrol et al., 2007, Zhao et al., 2013b). Dimeric THAP1 binds specific 11-nucleotide DNA sequences and is thought to function primarily as a transcription factor regulating cell cycle progression (Bessiere et al., 2008, Blanchard et al., 2011, Cayrol et al., 2007, Clouaire et al., 2005, Roussigne et al., 2003, Sabogal et al., 2010, Sengel et al., 2011). THAP1 also binds the *TOR1A* promoter, although evidence for regulation of *TOR1A* expression by THAP1 is conflicting (Gavarini et al., 2010, Kaiser et al., 2010). Dystonia-associated *THAP1* mutations reduce THAP1 stability and/or functionality, implicating perturbed neuronal gene expression in dystonia (Blanchard et al., 2011, Campagne et al., 2012, Sengel et al., 2011).

GNAL: DYT25 adult-onset cranial-cervical dystonia

Incompletely penetrant dominant mutations in *GNAL* cause craniocervical dystonia with a mean age at onset of 26 years (Fuchs et al., 2013, Vemula et al., 2013). Dystonia typically begins in the cervical region, subsequently progressing to the cranial musculature and occasionally other regions (Dauer, 2014, Fuchs et al., 2013, Vemula et al., 2013). *GNAL* encodes $G\alpha_{olf}$, a G protein α subunit originally identified in olfactory neurons and paralogous to $G\alpha_s$ but specifically enriched in the cerebellar Purkinje cells and striatum where $G\alpha_s$ expression is low (Drinnan et al., 1991, Herve et al., 1993, Zhuang et al., 2000). Striatal D1 dopamine receptors and adenosine A_{2A} receptors therefore rely on $G\alpha_{olf}$ alone for intracellular signalling (Corvol et al., 2001, Herve et al., 2001, Kull et al., 2000, Zhuang et al., 2000).

Most dystonia-associated *GNAL* mutations are loss-of-function, so abnormal striatal response to dopaminergic input may underlie *GNAL* mutation-positive dystonia (Fuchs et al., 2013).

CIZ1: DYT23 Adult-onset cervical dystonia

Several *CIZ1* mutations have been identified in patients showing cervical dystonia and tremor with a mean age at onset of 35 years, although the pathogenicity of some mutations has not yet been confirmed (Dufke et al., 2015, Uitti and Maraganore, 1993, Xiao et al., 2012). *CIZ1* is widely expressed in many tissues and encodes Cdkn1A-interacting zinc finger protein 1 (Mitsui et al., 1999, Warder and Keherly, 2003). CIZ1 interacts with cell cycle regulatory proteins and nuclear matrix proteins through its N- and C-terminal regions respectively, and coordinates replication initiation at DNA replication origins (Ainscough et al., 2007, Copeland et al., 2010, Coverley et al., 2005, Liu et al., 2016). Correspondingly, CIZ1 hyperphosphorylation inhibits the initiation of DNA replication (Copeland et al., 2015). How this might contribute to dystonia pathogenesis is currently unknown.

ANO3: DYT24 craniocervical dystonia

Mutations in *ANO3* cause adult-onset tremor and focal craniocervical dystonia progressing to segmental (Charlesworth et al., 2012, Ma et al., 2015, Miltgen et al., 2016, Stamelou et al., 2014, Zech et al., 2014b). *ANO3* is predominantly expressed in the nervous system, and encodes a member of the anoctamin/TMEM16 calcium-activated multi-pass transmembrane protein family (Huang et al., 2013, Kunzelmann et al., 2016, Picollo et al., 2015). Although ANO1 and ANO2 are calcium-activated chloride channels, so far no channel activity has been described for ANO3 (Kunzelmann et al., 2016, Picollo et al., 2015, Suzuki et al., 2013). Although its function remains unknown, recent studies found that ANO3 modulates KCNT1 sodium-activated potassium channel kinetics, and alters cell surface phospholipid distribution in heterologous cells (Huang et al., 2013, Picollo et al., 2015, Suzuki et al., 2013).

TUBB4A: DYT4 spasmodic "whispering" dysphonia

Specific *TUBB4A* mutations cause spasmodic "whispering" dysphonia with a mean age at onset of 30 years that progresses to generalised dystonia (Hersheson et al., 2013, Lohmann et al., 2013, Wilcox et al., 2011). In addition to dystonia, patients often have a distinctive "hobby horse" gait with a thin facial appearance and body habitus (Lohmann et al., 2013, Wilcox et al., 2011). *TUBB4A* expression is mostly restricted to the brain, and other mutations cause hypomyelinating leukodystrophy with microcephaly (Romaniello et al., 2015). Dystonia-associated *TUBB4A* mutations predominantly affect the autoregulatory region of the encoded β-tubulin 4A protein and appear to reduce total β-tubulin 4A levels (Hersheson et al., 2013, Lohmann et al., 2013). How this causes dystonia is unknown.

HPCA: autosomal recessive early-onset dystonia

Autosomal recessive *HPCA* mutations have been identified in two families with childhood-onset focal lower limb dystonia progressing to segmental or generalised dystonia with tremor (Charlesworth et al., 2015, Khan et al., 2003). To date, no additional families with *HPCA*-associated dystonia have been identified (Dobricic et al., 2016). *HPCA* encodes the neuronal calcium sensor hippocalcin and is expressed primarily in hippocampus principal cells, cerebellar Purkinje cells and the striatum (Amici et al., 2009, Paterlini et al., 2000, Saitoh et al., 1993). In its calcium-bound state hippocalcin activates a number of pathways including clathrin adaptor AP2-dependent endocytosis of NMDA and AMPA glutamate receptors (Jo et al., 2010, Kobayashi et al., 2005, Oh et al., 2006, Palmer et al., 2005). Hippocalcin also activates KCNQ potassium channels, contributing to slow afterhyperpolarisation (Kim et al., 2012, Tzingounis et al., 2007). Hippocalcin is thus thought to be important for synaptic plasticity, which is defective in dystonia (section 1.2.3.3).

1.2.4.2. Combined dystonias

***GCH1*, *TH*, *SPR*: DYT5, Dopa-responsive dystonia**

The dopa-responsive dystonias (DRDs) are a heterogeneous group of dystonias characterised by a distinct response of symptoms to treatment with levodopa. Onset is typically in childhood with lower limb dystonia progressing to generalised dystonia and demonstrating diurnal variation (Wijemanne and Jankovic, 2015). DRD can be caused by loss of function or dominant negative mutations in three genes: autosomal dominant (ADom) or recessive (AR) mutations in *GCH1* encoding GTP cyclohydrolase 1, AR mutations in *TH* encoding tyrosine hydroxylase, and AR mutations in *SPR* encoding sepiapterin reductase (Wijemanne and Jankovic, 2015). All three genes encode proteins essential for dopamine synthesis, so mutations in these genes potentially disrupt basal ganglia dopaminergic neurotransmission (Furukawa et al., 2002, Furukawa et al., 1999, Wijemanne and Jankovic, 2015).

***ATP1A3*: DYT12 rapid-onset dystonia parkinsonism**

Rapid-onset dystonia-parkinsonism (RDP) is one of a spectrum of disorders caused by highly penetrant mutations in *ATP1A3*, which encodes the α3 subunit of the Na^+, K^+ ATPase (de Carvalho Aguiar et al., 2004, Sweney et al., 2015). RDP is characterised by sudden onset of dystonia in a rostral-caudal (face, arms, then legs) pattern after a significant stressor such as fever or physical trauma during adolescence or adulthood. Parkinsonian features are also prominent (Brashear et al., 2007, Sweney et al., 2015). The Na^+, K^+ ATPase comprises α, β, and regulatory γ subunits, and helps maintain the plasma membrane electrochemical gradient; there are multiple isomers of each subunit (Bottger et al., 2011, DeAndrade et al., 2011). Mutant α3 has a reduced affinity for Na^+, impairing ion transport (Blanco-Arias et al., 2009, Einholm et al., 2010, Rodacker et al., 2006, Toustrup-Jensen et al., 2014). Distinct neuronal subtypes express different α-subunit combinations, so hypomorphic α3 may preferentially

affect certain brain regions (McGrail et al., 1991). Indeed, in mice loss of α3 activity specifically in the cerebellum causes dystonia with abnormal Purkinje cell firing (DeAndrade et al., 2011, Fremont et al., 2014, Fremont et al., 2015, Ikeda et al., 2013, McGrail et al., 1991, Sugimoto et al., 2014). Therefore cerebellar dysfunction may be central to RDP.

PRKRA: DYT16 young onset dopa-responsive dystonia and parkinsonism

Dopa-responsive dystonia and parkinsonism caused by *PRKRA* mutations is characterised by adolescence-onset dystonia progressing to generalised dystonia, with evidence of mild parkinsonism in some patients (Camargos et al., 2008, Quadri et al., 2016, Zech et al., 2014a). Ubiquitously expressed *PRKRA* encodes interferon-inducible double-stranded RNA-dependent protein kinase activator A or PACT, which is phosphorylated under stress conditions. It activates the latent kinase PKR to modulate gene expression and stop cell cycle progression (García et al., 2007, Patel et al., 2000, Peters et al., 2006, Singh et al., 2011). PACT also affects small RNA molecule biogenesis through interactions with Dicer and Argonaute-2 (Lee et al., 2006). Mutant PACT alters PKR activation in response to stress, implicating a defective stress response in dystonia (Vaughn et al., 2015).

X-linked dystonia parkinsonism DYT3, "Lubag"

X-linked recessive dystonia parkinsonism (XDP) or lubag is most common amongst Panay Islanders with a mean age at onset of 39 years; it presents with focal dystonia that usually generalises within 6 years, although a few patients present with parkinsonism (Domingo et al., 2015, Lee et al., 2002, Lee et al., 2011). Over time the parkinsonian features typically become more prominent and progressive atrophy of the basal ganglia can be seen on cerebral imaging (Goto et al., 2005, Lee et al., 2002, Lee et al., 2011, Rosales, 2010, Waters et al., 1993). Although linkage analysis and disequilibrium studies have refined the XDP locus to 427kb on Xq13.1, no mutations in protein-coding regions have been identified (Domingo et

al., 2015). However, five disease-specific single nucleotide changes, a 48bp deletion and a SVA retrotransposon insertion within the critical XDP region are inherited together as a single haploblock co-segregating with the disease (Deng et al., 2008, Domingo et al., 2015, Makino et al., 2007). These variants are in the vicinity of *TAF1* encoding transcription initiation factor TFIID subunit 1 (Domingo et al., 2015, Herzfeld et al., 2007). The XDP haplotype is hypothesised to reduce expression of specific neuronal TAF1 isoforms and/or affect splicing patterns (Makino et al., 2007, Muller et al., 2007, Sako et al., 2011).

MR-1: DYT8 paroxysmal nonkinesigenic dyskinesia

Paroxysmal nonkinesigenic dyskinesia (PNKD) is a childhood-onset disorder characterised by episodes of dystonia, chorea and ballism that are precipitated by factors including alcohol, caffeine, exercise, heat, hunger, stress and fatigue (Bruno et al., 2007). Highly penetrant autosomal dominant mutations in *MR-1* encoding the myofibrillogenesis regulator 1 protein have been identified in multiple families with PNKD (Ghezzi et al., 2015, Lee et al., 2004, Rainier et al., 2004). MR-1 may function as a detoxifying enzyme, though a more recent study showed that it also interacts with synaptic RIM proteins to regulate neurotransmitter exocytosis (Shen et al., 2015, Shen et al., 2011b). *MR-1* mutations typically affect cleavage of the N-terminus and alter protein stability (Ghezzi et al., 2009, Shen et al., 2011b). Mice carrying an *Mr1* knock-in mutation recapitulate the human PNKD phenotype and show abnormal basal ganglia dopaminergic signalling, suggesting MR-1 may be involved in monoaminergic neurotransmission (Lee et al., 2012b).

SLC2A1: DYT9 and DYT18 paroxysmal exercise induced dystonia and episodic ataxia

Paroxysmal exercise-induced dyskinesia (PED) is a rare childhood-onset disorder characterised by dyskinesia with dystonia, chorea, ballism and athetosis after exercise (Weber et al., 2008). Reduced penetrance mutations in *SLC2A1,* encoding the glucose transporter

GLUT1, have been identified in several families with PED, placing PED within the broad clinical spectrum of GLUT1 deficiency syndrome (Schneider et al., 2009, Suls et al., 2008, Weber et al., 2008). GLUT1 is the principle transporter for glucose uptake into neurons, astrocytes and erythrocytes (Pearson et al., 2013). PED-associated *SLC2A1* mutations reduce GLUT1 capacity for glucose transport, compromising CNS metabolism during periods of high energy demand (Suls et al., 2008, Weber et al., 2011, Weber et al., 2008).

PRRT2: DYT10 and DYT19 paroxysmal kinesigenic choreoathetosis

Patients with paroxysmal kinesigenic dyskinesia (PKD) have childhood or early adolescent onset of brief, frequent attacks combining dystonia and chorea, typically triggered by sudden movement (Silveira-Moriyama et al., 2013, Tan et al., 2014). Highly penetrant, autosomal dominant mutations in *PRRT2* have been identified in some PKD patients (Bhatia and Schneider, 2012, Wang et al., 2011). *PRRT2* is widely expressed in the brain, and most PKD-associated mutations prevent synthesis of the encoded transmembrane protein (Heron et al., 2012, Lee et al., 2012a, Valente et al., 2016, Wu et al., 2014). PRRT2 protein localises to the axonal plasma membrane where it interacts with neurotransmitter release SNARE proteins and the fast calcium sensor synaptotagmin to regulate calcium-stimulated neurotransmitter release (Lee et al., 2012a, Li et al., 2015, Valente et al., 2016, Wu et al., 2014).

Myoclonus-dystonia

This will be discussed in section 1.3.

1.2.4.3. Insights from the monogenic dystonias

Genes implicated in monogenic dystonias are involved in a variety of biological pathways, as shown in Figure 1.1. However, there are several points of convergence. Dopaminergic signalling to the striatum is repeatedly implicated, firstly by the dopamine-responsive dystonias arising from deficiencies in dopamine synthesis. $G\alpha_{olf}$ is required for downstream

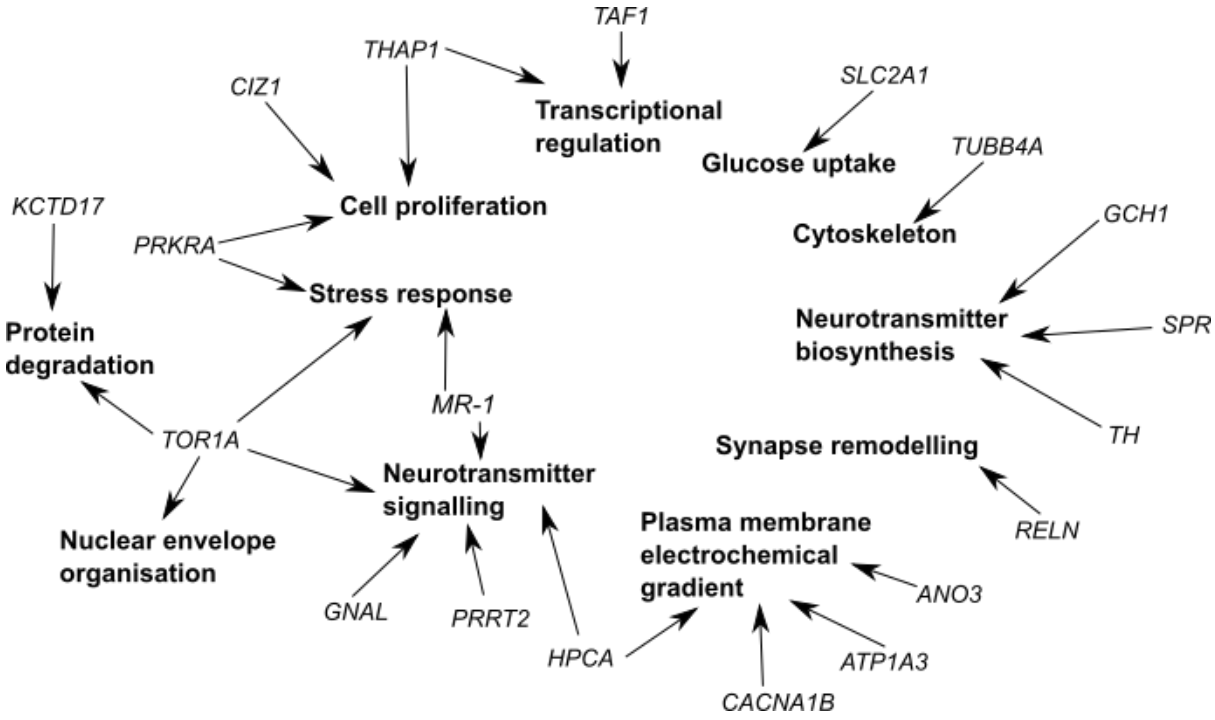


Figure 1.1 Overview of cell pathways implicated in dystonia by the monogenic dystonias, after (Lohmann and Klein, 2013). The proteins encoded by dystonia-associated genes have been implicated in a number of cellular pathways. These are summarised here, but as research into the functions of these genes continues this list will likely require amendment.

signalling from striatal D1 dopamine receptors, linking *GNAL* mutation-positive dystonia to dopaminergic signalling (Corvol et al., 2001, Herve et al., 2001, Kull et al., 2000, Zhuang et al., 2000). Additionally, patients with *TOR1A* and *THAP1* mutations have reduced striatal dopamine receptor availability on cerebral imaging; therefore these genes may also be involved in basal ganglia dopaminergic signalling (Carbon et al., 2009). Other monogenic dystonias implicate intracellular calcium signalling. *ANO3* and *PRRT2* both encode calcium-sensitive proteins, while hippocalcin (*HPCA*) is itself a calcium sensor. Abnormal calcium homeostasis and/or response could contribute to neuronal dysfunction resulting in dystonia. Overall, however, the monogenic dystonias also provide strong support for the view of dystonia as a network disorder. While the above genes appear to predominantly affect the basal ganglia, *ATP1A3* mutations specifically affect cerebellar Purkinje cells and *PRRT2* has been linked to multiple sensorimotor brain regions. This supports involvement of multiple brain regions in dystonia.

1.3. Myoclonus-dystonia (MD)

1.3.1. Overview of MD

The focus of this thesis is specifically on the rare combined dystonia Myoclonus Dystonia (MD). MD typically begins in the first two decades of life and is characterised by predominantly upper body myoclonic jerks with or without dystonia affecting the neck and/or arms (Asmus et al., 2009, Kinugawa et al., 2009, Nardocci et al., 2008, Peall et al., 2014, Raymond et al., 2008, Ritz et al., 2009, Roze et al., 2008). The myoclonus tends to worsen with activity or posture maintenance, and improve with alcohol (Kojovic et al., 2011, Li et al., 2008, Marelli et al., 2008, Quinn, 1996, Roze et al., 2008). Psychiatric symptomology is also prominent, particularly in *SGCE* mutation-positive MD patients (section 1.3.2.1) (Carecchio et al., 2013, Peall et al., 2016, Peall et al., 2011).

1.3.2. Mutations in *SGCE* cause MD

1.3.2.1. *SGCE* mutation-positive MD

Linkage analysis of several large MD families in the 1990s linked MD to a 3.2Mb region at 7q21 (Asmus et al., 2001, Klein et al., 2000b, Nygaard et al., 1999, Vidailhet et al., 2001, Zimprich et al., 2001). Sequencing of the 15 genes within this region revealed mutations in *SGCE* in affected individuals from six different MD families (Zimprich et al., 2001). Subsequent studies have detected pathogenic *SGCE* mutations or deletions in 20-80% of MD patients (Asmus et al., 2009, Gerrits et al., 2006, Grunewald et al., 2008, Han et al., 2003, Nardocci et al., 2008, Peall et al., 2013, Ritz et al., 2009, Schule et al., 2004, Tezenas du Montcel et al., 2006). The *SGCE* locus is maternally imprinted, and therefore typically only paternally inherited mutations cause MD. Studies of mice and humans demonstrated exclusive expression of *SGCE* from the paternal allele (Grabowski et al., 2003, Muller et al., 2002, Piras et al., 2000). Additionally, the human *SGCE* promoter CpG island is almost

entirely methylated on the maternal chromosome but unmethylated on the paternal chromosome (Grabowski et al., 2003, Muller et al., 2002). A few cases of apparently maternally inherited *SGCE* mutation-positive MD have been reported, although a loss of imprinting was not clearly demonstrated in any of these (Asmus et al., 2002, Beukers et al., 2011, Grabowski et al., 2003).

SGCE mutation-positive MD is distinguished by onset of the movement disorder within the first decade, and a higher than expected incidence of psychiatric disorders (Carecchio et al., 2013, Gerrits et al., 2006, Nardocci et al., 2008, Peall et al., 2014, Tezenas du Montcel et al., 2006). Approximately two-thirds of patients are symptomatic for at least one psychiatric disorder, a rate greater than that seen in either the general population or cohorts with other chronic, disabling movement disorders (Carecchio et al., 2013, Nardocci et al., 2008, Peall et al., 2016, Peall et al., 2014, Peall et al., 2013, van Tricht et al., 2012, Weissbach et al., 2013). Therefore psychiatric disorders are unlikely to be a secondary effect of the movement disorder. *SGCE* mutation-positive MD is associated with a distinct psychiatric phenotype characterised by primarily compulsive obsessive-compulsive disorder (OCD), and anxiety (Carecchio et al., 2013, Hess et al., 2007, Nardocci et al., 2008, Peall et al., 2016, Peall et al., 2014, Peall et al., 2013, Saunders-Pullman et al., 2002, van Tricht et al., 2012, Weissbach et al., 2013). Social phobia, panic disorder, generalised anxiety disorder, agoraphobia, attention deficit/hyperactivity disorder, psychosis and anorexia nervosa have also been described in patients with MD (Dale et al., 2011, Nardocci et al., 2008, Peall et al., 2016, Peall et al., 2013, Peall et al., 2011, van Tricht et al., 2012). Alcohol misuse is also common. Though initially considered secondary to alcohol's therapeutic effect on the movement disorder, recent studies suggest alcohol misuse may actually be a core phenotype of *SGCE* mutation (Hess et al., 2007, Peall et al., 2016, Peall et al., 2013, Saunders-Pullman et al., 2002, Weissbach et al., 2013).

1.3.2.2. *SGCE* encodes ε-sarcoglycan

SGCE was originally identified as a widely-expressed paralogue of the muscle-restricted *SGCA* gene (Ettinger et al., 1997, McNally et al., 1998). Expression of *SGCE* is highest in the brain, with particular enrichment in midbrain monoaminergic neurons, cerebellar Purkinje cells and the olfactory bulb (Chan et al., 2005, Nishiyama et al., 2004, Xiao and LeDoux, 2003). *SGCE* encodes the transmembrane glycoprotein ε-sarcoglycan (SG), a member of the sarcoglycan family of proteins (section 1.4). Similar to its paralogue α-SG, ε-SG is a type I

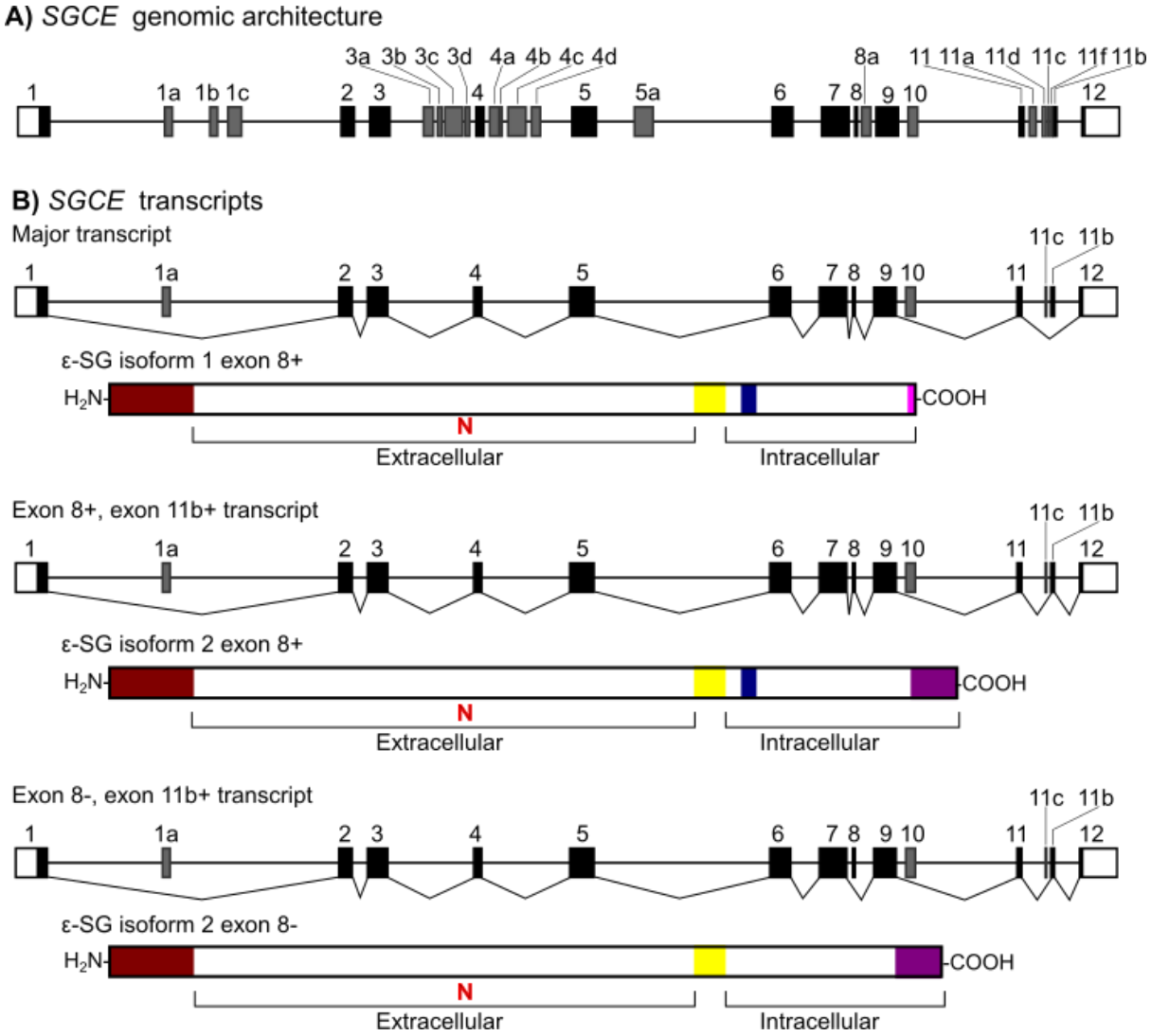


Figure 1.2 Genomic architecture, alternative splicing and isoforms of human *SGCE* encoding ε-SG. A) The genomic architecture of the human *SGCE* gene, not to scale. After Ritz et al., 2011. B) The splicing pattern for each of the three main *SGCE* transcripts is depicted against the basic genomic architecture. Lines connect exons included in the mature transcript. The polypeptide encoded by each transcript is illustrated below with major domains marked. Red shading marks the signal peptide, which is cleaved off to produce the mature protein; yellow marks the transmembrane domain; blue marks the peptide encoded by exon 8; purple or pink represent isoform-specific sequences; and the red N represents the N-linked glycosylation site.

transmembrane glycoprotein consisting of a posttranslationally cleaved N-terminal signal peptide, extracellular region with a single N-linked glycan, transmembrane domain, and short intracellular region as shown in Figure 1.2 (Esapa et al., 2007, Ettinger et al., 1997).

SGCE pre-mRNA undergoes general and tissue-specific alternative splicing. The Alu element-derived cryptic exon 10 is rarely included in the mature transcript, as are several other exons identified through RNA sequencing (Figure 1.2a) (Nishiyama et al., 2004, Ritz et al., 2011). By contrast, exon 11b is incorporated into approximately 30% of human *SGCE* transcripts in the brain only; exon 11b+ transcripts encode ε-SG isoform 2 with an extended intracellular C-terminal tail compared to the more widespread ε-SG isoform 1 (Figure 1.2b) (Nishiyama et al., 2004, Ritz et al., 2011, Yokoi et al., 2005). Also, exon 8 is spliced out of some *SGCE* brain transcripts through the activity of RNA binding protein SAM68 (Chawla et al., 2009, Nishiyama et al., 2004, Paronetto et al., 2011, Ritz et al., 2011, Yokoi et al., 2005). Intriguingly, ε-SG isoforms 1 and 2 may differ in cell type and subcellular localisation in the brain. Fractionation analysis of mouse brain lysates has suggested that only isoform 1 is present in capillary endothelial cells and astrocytes, while isoforms 1 and 2 have differential synapse localisation in neurons (Chan et al., 2005, Nishiyama et al., 2004). The differences in intracellular C-terminal tail sequence and localisation between the two main ε-SG isoforms suggest that brain-specific ε-SG isoform 2 could have an as-yet unknown unique brain-specific function. No effects of exon 8 inclusion/exclusion have been identified so far.

While α-SG requires stabilisation by other sarcoglycan proteins to be retained at the plasma membrane, ε-SG stably localises to the plasma membrane with or without other sarcoglycans (Draviam et al., 2006b, Esapa et al., 2007). In smooth muscle, peripheral nerve, adipose tissue and cardiac muscle, ε-SG is a component of prototypical sarcoglycan complexes that are part of dystrophin-associated glycoprotein complexes (DGCs) and DGC-like complexes (section 1.4.2-3) (Cai et al., 2007, Durbeej and Campbell, 1999, Groh et al., 2009, Imamura

et al., 2000, Lancioni et al., 2011, Straub et al., 1999). However, ε-SG's function at the plasma membrane with or without other sarcoglycans is unclear.

1.3.2.3. *SGCE* mutations associated with MD

Cell surface ε-SG appears to be essential for normal neural function. *SGCE* sequence changes that cause MD include missense and nonsense point mutations, small insertions/deletions, whole gene deletions, and larger genomic deletions covering multiple genes (Asmus et al., 2007, Dale et al., 2011, DeBerardinis et al., 2003, Grunewald et al., 2008, Peall et al., 2014). The same *SGCE* mutation can lead to a variety of MD phenotypes even within the same family (Foncke et al., 2006, Nardocci et al., 2008, Tezenas du Montcel et al., 2006). Most pathogenic *SGCE* mutations result in total ε-SG deficiency, but even those retaining some ε-SG expression result in absence of ε-SG from the cell surface (Esapa et al., 2007, Waite et al., 2011). When expressed in heterologous cells, missense ε-SG mutants are usually localised to the ER, polyubiquitinated and degraded by the proteasome (Esapa et al., 2007, Waite et al., 2011). This contrasts with the predominantly plasma membrane and Golgi localisation of wild-type ε-SG in those same cells (Esapa et al., 2007, Waite et al., 2011). Most mutant ε-SG not degraded by the proteasome is still retained in the ER, while a cell surface trafficking-competent mutant was degraded by the endosomal-lysosomal system (Waite et al., 2011). Therefore, *SGCE* mutations causing MD lead to absence of ε-SG from the plasma membrane/cell surface.

1.3.2.4. Mouse models of *SGCE* mutation-positive MD

Several mouse *Sgce* mutants have been generated. The first knockout allele involved deletion of exon 4, and resulted in complete loss of *Sgce* expression when paternally inherited or homozygous (Yokoi et al., 2006, Yokoi et al., 2005). Knockout mice exhibited spontaneous myoclonus and motor abnormalities, plus increased striatal dopamine levels associated with

reduced striatal dopamine receptor 2 and abnormal dopaminergic signalling (Yokoi et al., 2006, Zhang et al., 2012). These mice also had a behavioural phenotype suggestive of psychiatric abnormalities, including significantly higher scores on anxiety measures and a trend toward higher scores on measures of depression-like behaviour (Yokoi et al., 2006). If *Sgce* exon 4 loss is restricted to the striatum, mice had no overt myoclonus or dystonia but displayed deficits in motor learning, coordination and balance deficits (Yokoi et al., 2012b). By contrast, specific *Sgce* knockout in the cerebellar Purkinje cells resulted in mice with a slight motor learning deficit but no additional phenotypes (Yokoi et al., 2012a). This supports involvement of multiple brain regions in the *Sgce* deficiency phenotype. A second *Sgce* knockout allele replacing exons 6-9 with a cassette encoding neo and β-galactosidase has been produced, but not characterised beyond absence of ε-SG protein and normal muscle sarcoglycan complexes (Lancioni et al., 2011).

1.3.3. Neuroanatomy and physiology

Cerebral imaging studies of MD patients, particularly those with *SGCE* mutations, implicated several brain regions in MD pathogenesis. Several cerebral imaging techniques highlighted cerebellar dysfunction, including PET and fMRI demonstrating hypermetabolism and hyperactivity respectively (Beukers et al., 2010, Carbon et al., 2013, van der Salm et al., 2013). In addition, cerebellum-controlled eye blink conditioning and saccadic eye movements are abnormal in *SGCE* mutation-positive MD patients (Hubsch et al., 2011, Popa et al., 2014). The cortex has been implicated by fMRI studies demonstrating hyperresponsiveness of specific sensorimotor regions but overall reduced metabolism in MD patients compared to healthy controls (Beukers et al., 2010, Beukers et al., 2011, Carbon et al., 2013, Meunier et al., 2008, van der Salm et al., 2013, van der Salm et al., 2009). PET studies showed thalamus hypermetabolism also, in both manifesting and non-manifesting *SGCE* carriers though to a lesser degree in the latter (Carbon et al., 2013). White matter abnormalities have also been

described in MD patients, particularly in the region connecting the basal ganglia to the cerebellum (van der Meer et al., 2012). This was postulated to be a consequence of increased plasticity, resulting in changes to regional connections (van der Meer et al., 2012).

1.3.4. MD: a genetically heterogeneous syndrome?

Although *SGCE* remains the only gene definitively implicated in hereditary MD, mutations in a number of other genes have been linked to MD.

1.3.4.1. *DRD2* and *TOR1A*

The first mutation identified as segregating with MD in a family was in a highly conserved region of *DRD2* encoding the D2 dopamine receptor (Klein et al., 1999). However, examination of other MD families did not reveal further *DRD2* mutations, and subsequent reassessment of the original *DRD2*-linked MD family identified a pathogenic *SGCE* mutation segregating with the MD phenotype (Durr et al., 2000, Klein et al., 2000a, Klein et al., 2002). Similarly, in another family MD appeared to co-segregate with a *TOR1A* mutation, but subsequent investigation identified a pathogenic *SGCE* mutation in all affected individuals (Doheny et al., 2002, Klein et al., 2002, Leung et al., 2001). Therefore, *TOR1A* and *DRD2* are not believed to contribute to MD pathogenesis, though multigenic inheritance remains a possibility for some cases of *SGCE* mutation-negative MD.

1.3.4.2. New genes associated with MD and MD-like phenotypes: *RELN, CACNA1B,* and *KCTD17*

In the last two years, three new genes have been linked to MD or MD-like phenotypes in patients lacking *SGCE* mutations. A total of four distinct *RELN* mutations have been identified in three MD families and two sporadic MD patients with a classic MD phenotype involving upper body myoclonus, dystonia and anxiety- or depression-related disorders (Groen et al., 2015b). Age at onset was highly variable however, with adult onset in two

families and the sporadic cases but onset at 3-4 years in the final family (Groen et al., 2015b). *RELN* encodes the large glycoprotein reelin, an extracellular matrix (ECM) protein that binds to cell surface receptors including VLDLR, APOER2 and potentially α3β1 integrin (Folsom and Fatemi, 2013, Howell et al., 1999, Howell et al., 2000). This initiates an intracellular signalling cascade that regulates neuronal positioning and migration during development (Folsom and Fatemi, 2013). Correspondingly, reeler mice deficient in *Reln* and human patients with rare recessive *RELN* mutations or chromosomal rearrangements display disorganised brain structures or lissencephaly (Chang et al., 2007a, Folsom and Fatemi, 2013, Hong et al., 2000, Tissir and Goffinet, 2003, Zaki et al., 2007). Reelin also regulates neurotransmission, and contributes to formation and plasticity of the synapse (Folsom and Fatemi, 2013, Iafrati et al., 2014, Lakatosova and Ostatnikova, 2012).

A family with an unusual MD-like syndrome linked to a point mutation in *CACNA1B* has also been identified (Groen et al., 2011, Groen et al., 2015a). Affected individuals had dystonia, progressive upper and lower limb myoclonus affecting gait and stability, painful cramps and cardiac arrhythmias (Groen et al., 2011, Groen et al., 2015a). As with *SGCE* mutation-positive MD, alcohol intake alleviates myoclonus in these patients (Groen et al., 2011). *CACNA1B* encodes a pore-forming subunit of voltage-gated calcium channels, and in heterologous cells the R1389H mutation identified in the MD-like patients affected ion flow through the channel (Ament et al., 2015, Groen et al., 2015a). No additional *CACNA1B* mutations in MD patients have been identified to date, and the complex phenotype has raised questions regarding whether this clinical entity represents MD or an MD-like disorder (Groen et al., 2015a, Mencacci et al., 2015a).

Finally, an autosomal dominant missense mutation in *KCTD17* has been identified in two families with MD (Mencacci et al., 2015b). Affected individuals had onset of jerks or jerky tremor in the upper limbs between 5 and 20 years of age, with subsequent development of

predominantly craniocervical and upper limb dystonia (Mencacci et al., 2015b). *KCTD17* encodes potassium channel tetramerisation domain-containing protein 17, a cytosolic BTB/POZ domain-containing protein (Mencacci et al., 2015b). Studies of MD patient fibroblasts implicated KCTD17 in calcium signalling, but there is also evidence the protein may be an ubiquitin ligase (Kasahara et al., 2014, Mencacci et al., 2015b).

1.4. The sarcoglycans: a family of transmembrane glycoproteins

ε-SG, encoded by *SGCE* and involved in the molecular pathogenesis of MD (section 1.3.2), belongs to the sarcoglycan protein family. This family of six transmembrane glycoproteins was originally identified in skeletal muscle as part of the dystrophin-associated glycoprotein complex (DGC), but has subsequently been detected in other tissues (Yoshida et al., 1994).

1.4.1. Overview of the sarcoglycans

SG	Gene	Genomic locus	Amino acids	Molecular weight	N-linked glycans	Disulphide bonds	Associated disease
α	*SGCA*	17q21	387	50	2	0	LGMD 2D
β	*SGCB*	4q12	318	43	3	2	LGMD 2E
δ	*SGCD*	5q33	290	35	3	2	LGMD 2F
γ	*SGCG*	13q12	291	35	1	2	LGMD 2C
ε	*SGCE*	7q21	437	45	1	0	MD
ζ	*SGCZ*	8p22	312	35	2	1	none

Table 1.3 The sarcoglycan protein family with associated diseases. Complete list of human sarcoglycan proteins with gene name, genomic locus, polypeptide size in amino acids, apparent molecular weight on SDS-PAGE, number of N-linked glycans, number of disulphide bonds, and the associated disease. Protein information obtained from UniProt database. Abbreviations: SG, sarcoglycan; LGMD, limb-girdle muscular dystrophy; MD, myoclonus dystonia.

Sarcoglycans are single-pass transmembrane glycoproteins with a large glycosylated extracellular region, transmembrane domain, and short intracellular tail. The six sarcoglycans are α-SG encoded by *SGCA*, β-SG encoded by *SGCB*, δ-SG encoded by *SGCD*, γ-SG encoded by *SGCG*, ζ-SG encoded by *SGCZ*, and the previously discussed ε-SG encoded by *SGCE* (table 1.3). ε-SG and α-SG are paralogous type I transmembrane proteins with highly similar exon/intron borders, approximately 62% similarity and 45% identity at the amino acid level, and an extracellular cadherin/immunoglobulin superfamily domain similar to those

present in α-dystroglycan (Figure 1.3) (Bozic et al., 1998, De Rosa et al., 2011, Dickens et al., 2002, Ettinger et al., 1997, McNally et al., 1998). α-SG expression is restricted to muscle, whereas ε-SG is broadly expressed with particular enrichment in the brain (section 1.3.2.2).

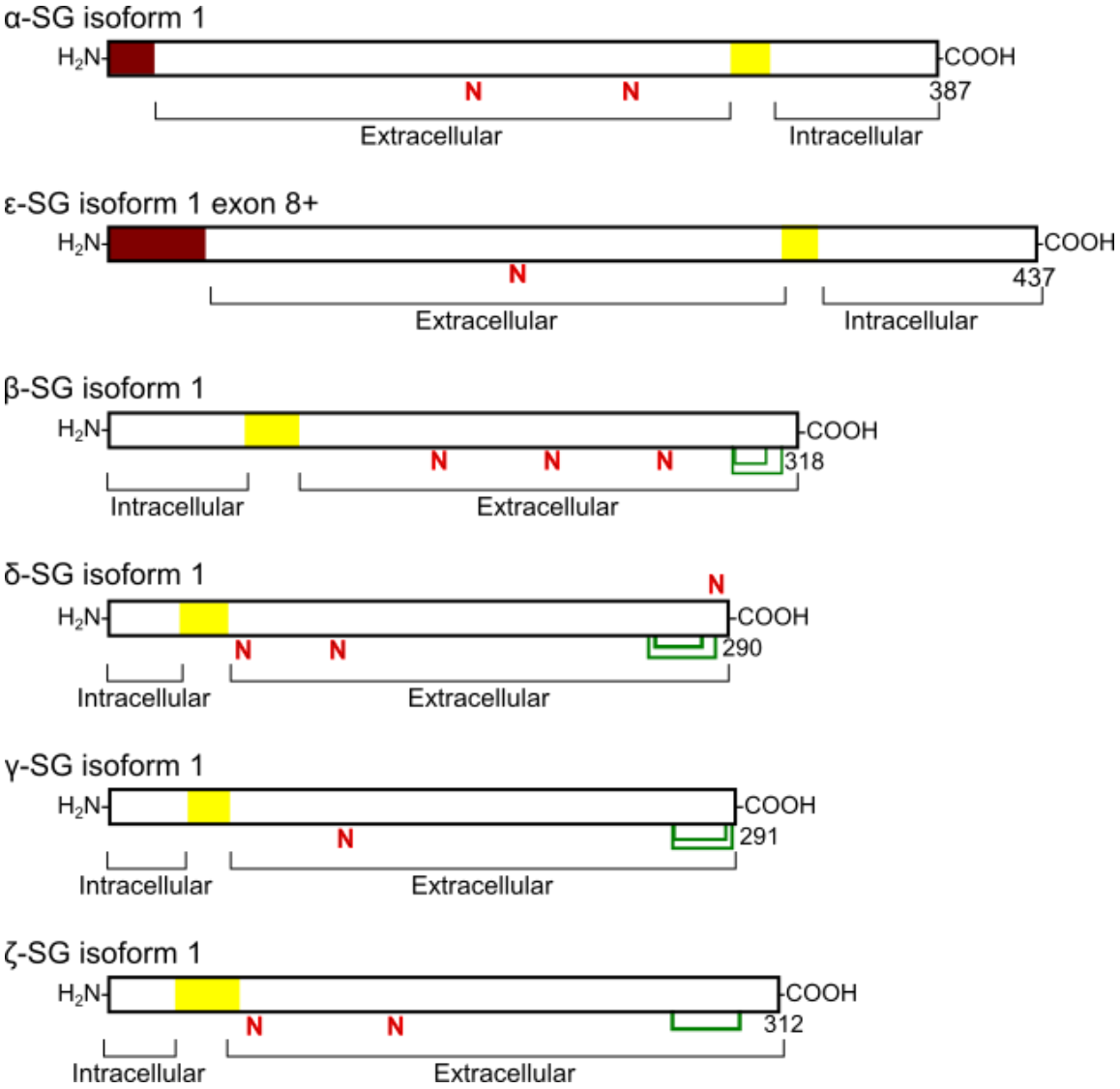


Figure 1.3 Comparison of human sarcoglycan proteins. For each sarcoglycan protein, the full-length/canonical isoform 1 is depicted. Transmembrane domains are marked in yellow, signal peptides marked in red, N marks an N-linked glycosylation site, and green lines represent disulphide bonds between cysteine residues. The intracellular and extracellular regions are marked against the polypeptide. Numbers at the C-terminus of each protein give the polypeptide length before processing. Polypeptides are approximately to scale. ε-SG and α-SG signal peptides are as reported in the Leiden Open Variation Database (Fokkema et al., 2011).

δ-SG, γ-SG and ζ-SG form another group of paralogues, with 74% similarity at the amino acid level but differential N-glycosylation (Jung et al., 1996, Shiga et al., 2006, Wheeler et al., 2002). These three sarcoglycans together with β-SG are type II transmembrane proteins with intracellular N-terminal regions and extracellular C-terminal regions containing 1-2

conserved disulphide bonds (Figure 1.3) (Chan et al., 1998, Shi et al., 2004, Wheeler et al., 2002). γ-SG expression is muscle-specific, whereas ζ-SG is expressed more broadly with highest expression in the brain and lowest in skeletal muscle (Barresi et al., 2000b, Shiga et al., 2006, Yamamoto et al., 1994). Expression of β-SG and δ-SG is also widespread, though considerably greater in muscle compared to other tissues (Cheng et al., 2006, Jung et al., 1996, Lim et al., 1995, Nigro et al., 1996).

1.4.2. Synthesis and assembly of sarcoglycan complexes

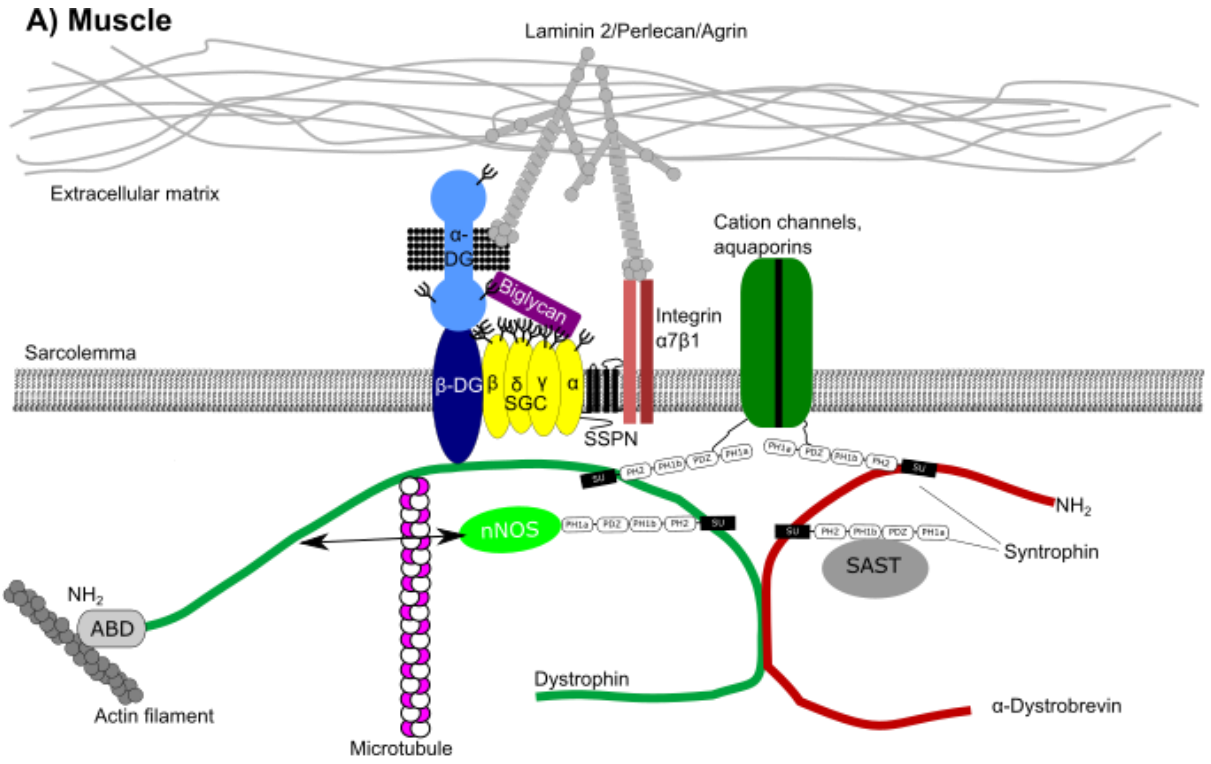
Immediately after synthesis, the sarcoglycans interact to form complexes containing equimolar quantities of each sarcoglycan and thought to be predominantly heterotetrameric (Holt and Campbell, 1998, Jung et al., 1996, Noguchi et al., 2000). With the exception of α- and ε-SG, sarcoglycans require heterotetramer assembly for cell surface trafficking (Noguchi et al., 2000). Paralogues α-SG and ε-SG are thought to be interchangeable in sarcoglycan complex assembly, as are γ-SG and ζ-SG (Cai et al., 2007, Liu and Engvall, 1999). Sarcoglycan heterotetramers composed of αβδγ-SG predominate in muscle, but εβδγ-SG heterotetramers have also been described in cardiac and smooth muscle (Duclos et al., 1998, Liu and Engvall, 1999, Straub et al., 1999, Yoshida et al., 1994). By contrast, Schwann cells and adipose tissue contain εβδζ-SG heterotetramers (Cai et al., 2007, Groh et al., 2009, Imamura et al., 2000). So far no functional differences between sarcoglycan complexes have been identified.

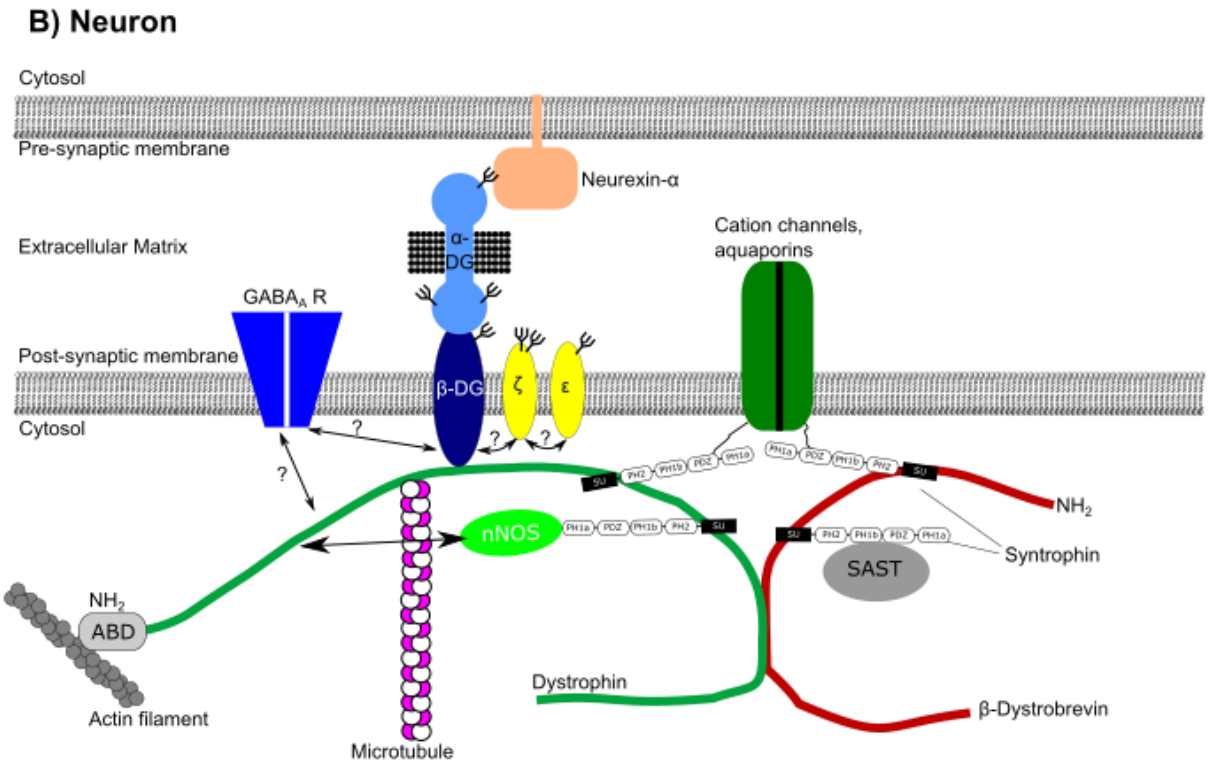
Studies of sarcoglycan complex assembly in cultured myotubes, transiently transfected heterologous cells, and animal models of sarcoglycan deficiency (section 1.4.5) have highlighted the importance of β-SG, which appears to initiate sarcoglycan heterotetramer assembly by interacting strongly with δ-SG; this forms a βδ-SG “core” required for heterotetramer assembly and trafficking to the plasma membrane (Chan et al., 1998, Draviam et al., 2006a, Shi et al., 2004). Sequential deletion of the extracellular domain has narrowed

down the interaction between β-SG and δ-SG to the δ-SG extracellular juxtamembrane region between residues 57 and 92 (Chen et al., 2006). Co-immunoprecipitation (IP) experiments indicate that γ-SG/ζ-SG interacts strongly with the βδ-SG core via the more distal residues 94-194 of γ-SG (Chan et al., 1998, Chen et al., 2006, Draviam et al., 2006a, Noguchi et al., 2000, Shi et al., 2004, Yoshida et al., 1997, Yoshida et al., 1994). Based on further co-IP experiments, γ-SG/ζ-SG has been proposed to mediate a weaker interaction between α-SG/ε-SG and the remainder of the complex, although in heterologous cells α-SG can interact directly with β-SG and δ-SG (Chan et al., 1998, Draviam et al., 2006a, Hack et al., 2000, Noguchi et al., 2000, Shi et al., 2004). Trafficking to the plasma membrane is thought to require normal N-linked glycosylation and disulphide bond formation in sarcoglycans plus the core interaction between β-SG and δ-SG (Chan et al., 1998, Chen et al., 2006, Draviam et al., 2006a, Shi et al., 2004).

1.4.3. Sarcoglycans as part of the dystrophin-associated glycoprotein complex

In muscle, the sarcoglycan complex forms a subcomplex of the dystrophin-associated glycoprotein complex (DGC) based on the cytosolic protein dystrophin (Figure 1.4a). This large multimeric complex can be biochemically subdivided into three subcomplexes: the sarcoglycan subcomplex containing the sarcoglycan complex (section 1.4.2) plus sarcospan; the dystroglycan subcomplex; and the cytosolic subcomplex of dystrophin, dystrobrevins and syntrophins (Yoshida et al., 1994)(Figure 1.4a). The dystroglycan subcomplex comprises single-pass transmembrane β-dystroglycan and heavily glycosylated extracellular α-dystroglycan, both produced through proteolysis of a single precursor protein (Esapa et al., 2003, Holt et al., 2000, Ibraghimov-Beskrovnaya et al., 1992). This complex connects ECM proteins that interact with α-dystroglycan (e.g. laminin) to dystrophin which interacts with the cytosolic region of β-dystroglycan and cytoskeletal structures (Allen et al., 2016, Chung and Campanelli, 1999, Prins et al., 2009, Way et al., 1992). Therefore the DGC connects the


A) Muscle
Laminin 2/Perlecan/Agrin
Extracellular matrix
α-DG
Biglycan
Cation channels, aquaporins
Integrin α7β1
Sarcolemma
β-DG
β δ γ α
SGC
SSPN
nNOS
NH2
SAST
Syntrophin
NH2
ABD
Actin filament
Dystrophin
α-Dystrobrevin
Microtubule


B) Neuron
Cytosol
Pre-synaptic membrane
Neurexin-α
Extracellular Matrix
α-DG
Cation channels, aquaporins
GABAA R
Post-synaptic membrane
β-DG
ζ
ε
Cytosol
?
nNOS
NH2
SAST
Syntrophin
NH2
ABD
Actin filament
Dystrophin
β-Dystrobrevin
Microtubule

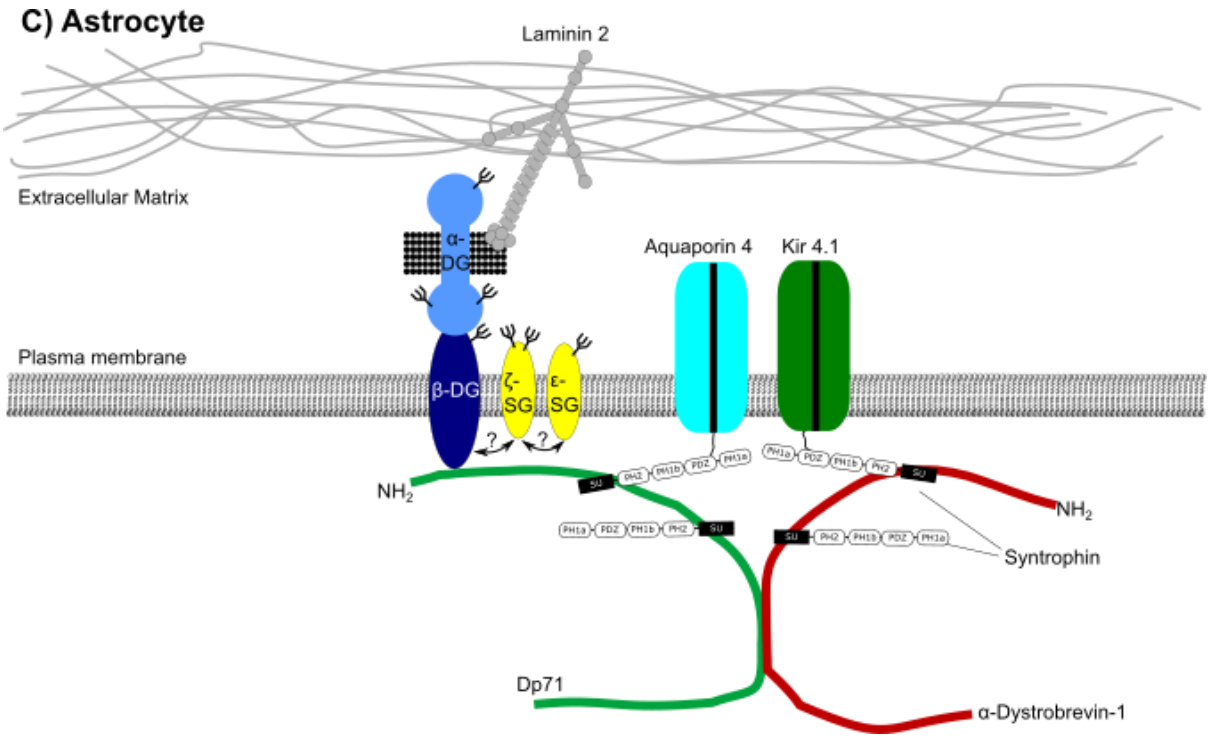


Figure 1.4 Putative molecular organisation of the DGC in muscle and DGC-like complexes in the brain. Adapted from (Waite et al., 2009). Illustration of the basic organisation for the core DGC and DGC-like complexes in A) muscle, B) CNS neurons, and C) CNS glia. The muscle DGC has been determined through direct biochemical purification of the intact complex from tissue, whereas CNS DGC-like complexes have been inferred from protein interactions identified using a variety of protocols including yeast two-hybrid and co-immunoprecipitation from cultured cells and/or tissue. Only the basic structure and a subset of associated proteins are illustrated for each complex, and full details can be obtained in recent reviews (Allen et al., 2016, Waite et al., 2012, Waite et al., 2009). Arrows with question marks indicate putative but not confirmed protein interactions. Dystrobrevin and dystrophin are not to scale. Abbreviations: DGC, dystrophin-associated glycoprotein complex; DG, dystroglycan; SG, sarcoglycan; SGC, sarcoglycan complex; SSPN, sarcospan; nNOS, neuronal-type nitric oxide synthase; SAST, syntrophin-associated serine/threonine kinase; ABD, actin binding domain; PH, plecstrin homology domain; PDZ, PSD-95 (postsynaptic density protein 95), discs_large and zonula occludens-1 domain; SU, syntrophin-unique region; Kir 4.1, inwardly rectifying potassium channel 4.1.

cell's cytoskeleton to the extracellular matrix, and is thought to protect the sarcolemma from mechanical stress (Allen et al., 2016, Petrof et al., 1993, Straub et al., 1997). Dystrophin is essential for DGC assembly; its absence due to *DMD* mutation in Duchenne and Becker muscular dystrophies results in secondary absence of all other DGC components from the muscle sarcolemma (Allen et al., 2016).

In addition to its structural function, the DGC is thought to influence signal transduction and calcium homeostasis by acting as a scaffold for recruitment of other proteins. Intracellular calcium levels are typically elevated with DGC dysfunction, which may be caused by dysregulation of calcium homeostasis or increased plasma membrane permeability (Allen et al., 2016). The cytosolic subcomplex comprising dystrophin, dystrobrevins and syntrophins

interacts with a number of signalling proteins, most notably neuronal-type nitric oxide synthase (nNOS) which is localised to the sarcolemma through its interactions with syntrophin and dystrophin (Adams et al., 2001, Allen et al., 2016, Brenman et al., 1996, Lai et al., 2013). Nitrous oxide (NO) produced by nNOS is a major signalling molecule in muscle that helps regulate functions such as glucose uptake and calcium release from intracellular stores (Allen et al., 2016). Syntrophin also helps recruit a number of cation channels and aquaporin to the vicinity of the DGC (Adams et al., 2001, Allen et al., 2016). In addition, environment- or activity-induced phosphorylation of β-dystroglycan and dystrophin regulates their interaction and interactions with other proteins (Ilsley et al., 2001, Sotgia et al., 2001, Swiderski et al., 2014). Similar phosphorylation has been described for other DGC components (Allen et al., 2016). In the ECM, the chondroitin sulphate proteoglycan biglycan interacts with α-dystroglycan to modulate expression and plasma membrane association of other DGC components (Bowe et al., 2000, Mercado et al., 2006, Rafii et al., 2006). Finally, integrin α7β1 interacts directly with sarcospan and is thought to contribute to both mechanical and signalling functions of the DGC (Allen et al., 2016, Marshall et al., 2012).

The sarcoglycan heterotetramer is integrated into the DGC through several protein interactions, and may contribute both to DGC stability in the plasma membrane and to signal transduction. The βδ-SG core interacts with dystroglycan, while the extracellular domain of γ-SG interacts with and stabilises the tetraspanin protein sarcospan in the plasma membrane (Chan et al., 1998, Coral-Vazquez et al., 1999, Crosbie et al., 1997, Crosbie et al., 1999, Crosbie et al., 2000, Hayashi et al., 2006). The sarcoglycan heterotetramer may also interact with the N-terminal domain of dystrobrevin, although this is thought to be stronger with CNS-expressed β-dystrobrevin compared to muscle-expressed α-dystrobrevin (Blake et al., 1998, Yoshida et al., 2000). Consequently, this interaction could be of greater importance in the CNS than in muscle (Blake et al., 1998). *In vitro* experiments identified an additional

interaction between the βδ-SG core and the carboxyl terminus of dystrophin, though this has not been replicated *in vivo* (Chen et al., 2006). Finally, biglycan interacts with and modulates expression of α-SG and γ-SG (Rafii et al., 2006). Thus, in muscle the sarcoglycans form a heterotetrameric complex integrated into the wider DGC by multiple interactions.

A number of DGC-like complexes have been described in the CNS (Figure 1.4b-c) (Blake et al., 1999). These can differ in dystrophin isoform (for example, Dp71 in glia), inclusion of α-dystrobrevin-1 versus β-dystrobrevin in glia versus neurons, and the syntrophins incorporated (Blake et al., 1999, Blake et al., 1992, Blake et al., 1998, Tadayoni et al., 2012, Waite et al., 2012). Brain DGC-like complexes have been inferred from protein interactions rather than through isolation of intact complexes, so the exact organisation of these complexes remains unknown. However, DGC-like complexes are clearly essential for normal neurological development and function. Dystroglycan glycosylation defects can result in severe brain structure abnormalities, while Duchenne/Becker muscular dystrophy patients commonly have cognitive impairment and neuropsychiatric disorders (Anderson et al., 2002, Snow et al., 2013, Waite et al., 2012). Dystrophin, dystroglycan and dystrobrevin are all involved in the clustering of GABA-A and other neurotransmitter receptors clustering at the synapse, with disorganised postsynaptic regions in their absence although the mechanism behind this remains unknown (Brünig et al., 2002, Knuesel et al., 1999, Krasowska et al., 2014, Waite et al., 2012). In astrocytes, DGC-like complexes containing Dp71 are involved in localisation of aquaporin-4 and the Kir 4.1 inwardly rectifying potassium channel which help maintain the blood-brain barrier (Connors et al., 2004, Tadayoni et al., 2012, Waite et al., 2012). However, contributions of the sarcoglycans to these DGC-like complexes in brain remain unknown.

1.4.4. The sarcoglycans in human disease

Mutations in *SGCA, SGCB, SGCD* and *SGCG* cause autosomal recessive limb-girdle muscular dystrophy (LGMD), while mutations in *SGCE* cause MD (section 1.3.2). LGMD is

characterised by progressive degeneration of skeletal muscle, predominantly affecting the muscles around the pelvis and scapulae (Sandona and Betto, 2009). Patients with mutations in *SGCB*, *SGCD* or *SGCG* resulting in LGMD 2E, 2F and 2C respectively often develop cardiomyopathy in addition to skeletal muscle degeneration, though this is less common in patients with *SGCA* mutations, probably due to compensation for α-SG deficiency by ε-SG (Barresi et al., 2000a, Fayssoil, 2010, Lancioni et al., 2011, Politano et al., 2001, Sandona and Betto, 2009). LGMD can range from mild to severe even with the same sarcoglycan mutation, and is associated with partial or complete loss of the entire sarcoglycan complex from the sarcolemma in addition to deficiency of the mutated sarcoglycan (Draviam et al., 2001, McNally et al., 1996, Sandona and Betto, 2009). In LGMD 2E or 2F, the entire sarcoglycan complex is typically reduced or lost from the sarcolemma with β-dystroglycan and dystrophin levels also reduced (Draviam et al., 2001, Klinge et al., 2008, Vainzof et al., 1996). By contrast, *SGCA* or *SGCG* mutations typically result in a variable presence of residual sarcoglycans at the sarcolemma and no changes to β-dystroglycan or dystrophin (Klinge et al., 2008, Sandona and Betto, 2009). However, these are trends rather than firm differences and so cannot be used to differentiate between genotypes (Klinge et al., 2008).

As with MD-associated *SGCE* mutations (section 1.3.2.3), most LGMD-associated sarcoglycan mutations result in sarcoglycan deficiency from the plasma membrane. Several LGMD-associated *SGCA* missense mutations have been shown to result in mutant α-SG proteins that are ubiquitinated and degraded by the proteasome (Gastaldello et al., 2008, Sandona and Betto, 2009). If the ER quality control pathway is inhibited, some of these mutant proteins can interact normally with other sarcoglycans and traffic to the plasma membrane, suggesting they retain at least partial functionality (Bartoli et al., 2008, Bianchini et al., 2014, Gastaldello et al., 2008, Soheili et al., 2012). Studies in heterologous cells have demonstrated that several LGMD-associated β-SG, δ-SG and γ-SG mutants impair assembly

of the sarcoglycan complex, while others interact with other sarcoglycans but reduce trafficking to the plasma membrane (Chen et al., 2006, Shi et al., 2004). Interestingly, one *SGCD* mutation was found to result in mislocalisation of the sarcoglycans to the NE in transgenic mice in addition to their absence from the sarcolemma (Heydemann et al., 2007).

1.4.5. Animal models of sarcoglycan mutations

In addition to the *Sgce* mutant mouse lines (section 1.3.2.4), rodent models of *SGCA, SGCB, SGCD* and *SGCG* deficiency have been generated. The first was the BIO14.6 strain of Syrian hamsters, which carries a spontaneous *Sgcd* exon 1 deletion (Nigro et al., 1997). BIO14.6 hamsters develop muscular dystrophy plus cardiomyopathy associated with complete loss of the SGC and reduction of α-dystroglycan at the sarcolemma in skeletal and cardiac muscle (Ikeda et al., 2002, Roberds et al., 1993b, Straub et al., 1998). There are also several transgenic mouse lines with knock-in mutations or null alleles of *Sgca, Sgcb, Sgcd* and *Sgcg*. Mice homozygous for *Sgca, Sgcb* or *Sgcd* null alleles exhibit complete loss of the entire SGC from the skeletal muscle sarcolemma with severe muscular dystrophy (Araishi et al., 1999, Coral-Vazquez et al., 1999, Duclos et al., 1998, Durbeej et al., 2000). By contrast, *Sgcg*-deficient mice have severe muscular dystrophy with loss of β-SG and δ-SG but residual α-SG at the sarcolemma (Hack et al., 1998). *Sgcb, Sgcd* and *Sgcg* null mice all develop cardiomyopathy but *Sgca* null mice do not, partially due to *Sgce* expression in cardiac muscle (Araishi et al., 1999, Coral-Vazquez et al., 1999, Duclos et al., 1998, Durbeej et al., 2000, Hack et al., 1998, Lancioni et al., 2011). Intriguingly, transgenic mice substantially overexpressing *Sgcg* also develop severe muscular dystrophy, with intracellular retention of γ-SG aggregates and upregulation of α- and β-SG (Zhu et al., 2001). This highlights the importance of appropriate levels of each sarcoglycan for normal muscle function.

1.4.6. Function(s) of the sarcoglycans

Despite ongoing research, the precise function(s) of the sarcoglycan complex in muscle and in brain are unknown. In the muscle, the sarcoglycan complex is thought to help stabilise the wider DGC in the plasma membrane and thus support sarcolemma mechanical stability. However, there is increasing evidence for direct involvement of the sarcoglycans in calcium homeostasis and signal transduction. *Sgcb* and *Sgcd* null mice plus *Sgcd* null hamsters have abnormal calcium homeostasis in muscle despite residual DGCs at the sarcolemma (Fraysse et al., 2010, Iwata et al., 2003, Nakamura et al., 2001, Solares-Perez et al., 2010a, Solares-Perez et al., 2010b). Genetically or chemically modulating calcium signalling in these rodents can improve the dystrophic phenotype (Iwata et al., 2005, Parsons et al., 2007). In *Sgcb* null mice, calcium leaking from intracellular storage through RyR1 channels in skeletal muscle was observed and may contribute to abnormal calcium homeostasis (Andersson et al., 2012). Increased calcium levels alone can cause muscular dystrophy in mice, possibly through activation of proteases that break down muscle (Iwata et al., 2005, Millay et al., 2009). Aberrant skeletal muscle calcium handling has also been described in an LGMD 2C patient, indicating this is not restricted to rodent LGMD models (Hassoni and Cullen, 1999).

Several protein interactions implicate sarcoglycans in signal transduction. As previously described (section 1.4.3), the sarcoglycans interact with signalling scaffold protein dystrobrevin (Yoshida et al., 2000). Correspondingly, loss of the SGC in sarcoglycan-deficient mice results in loss of nNOS from the sarcolemma with downstream effects on NO signalling (Crosbie et al., 2002, Heydemann et al., 2004). γ-SG and δ-SG specifically interact with two further proteins implicated in signal transduction: γ-filamin and the 16 kDa subunit of the vacuolar proton ATPase (16K) (Chen et al., 2007, Thompson et al., 2000). Filamins are involved in cytoskeletal remodelling after mechanotransduction through interaction with actin, and therefore may implicate sarcoglycans in mechanotransduction (Guyon et al., 2003,

Thompson et al., 2000). The contribution of the 16K protein is less clear, as in addition to its canonical role in the vacuolar ATPase it has been implicated in other functions including neurotransmitter exocytosis (Chen et al., 2007, Morel and Poea-Guyon, 2015). Both 16K and γ-filamin also interact with integrins, suggesting similar roles and/or bidirectional signalling between the integrins and sarcoglycans (Chen et al., 2007, Thompson et al., 2000).

Several sarcoglycans also undergo tyrosine phosphorylation. In cultured myocytes, specific intracellular tyrosine residues in α-SG and γ-SG are phosphorylated during the formation of focal adhesions; these sarcoglycans also co-purify with adhesion proteins (Yoshida et al., 1998). γ-SG is also phosphorylated in mouse skeletal muscle after contraction, and this phosphorylation is required for normal mechanotransduction (Barton, 2006, Barton, 2010). *Sgcg* null mice exhibit increased protein phosphorylation in skeletal muscle after contraction compared to wild-type mice (Barton, 2006). This also occurs in LGMD 2C patients and may relate to γ-SG's interaction with archvillin, a stimulus-dependent scaffold for ERK1/2 and other signal transduction proteins (Spinazzola et al., 2015). In mouse muscle, the interaction between archvillin and phosphorylated ERK1/2 requires γ-SG (Spinazzola et al., 2015).

The sarcoglycans could also contribute to signal transduction through interactions and activity in the extracellular space. α-SG and γ-SG interact with and are regulated by the ECM proteoglycan biglycan (section 1.4.3) (Rafii et al., 2006). In muscle biglycan modulates sarcolemmal localisation of the DGC, but it is also involved in numerous other processes including stabilising neuromuscular junction synapses (Amenta et al., 2012, Bowe et al., 2000, Nastase et al., 2012, Rafii et al., 2006). Furthermore, α-SG's extracellular region is an ecto-ATPase that can modulate extracellular ATP concentrations (Betto et al., 1999, Sandona et al., 2004). Therefore α-SG may modulate signalling via manipulation of extracellular ATP. However, all of the above protein interactions and possible functions for the sarcoglycans derive from studies performed on muscle, cultured myocytes, or heterologous cells. Several

sarcoglycans are known to be expressed in other tissues including the CNS, as previously discussed; however, their functions in those tissues remain largely unknown.

1.5. Thesis aims

Despite the well-established involvement of *SGCE* mutations in MD, at the start of this study very little was known about ε-SG and other sarcoglycans in the brain. Clues to a protein's function can be obtained from the other proteins with which it interacts. Therefore, the overall aim of this study was to gain insight into the function(s) of ε-SG in the brain through identification of proteins with which it interacts.

In pursuit of this aim, the study had three major objectives:

1. To determine whether ε-SG interacts with other sarcoglycans and DGC components in the brain. A positive result would implicate DGC-like complexes in the molecular pathogenesis of MD (Chapter 3).
2. To identify additional, non-DGC protein(s) with which ε-SG interacts in the brain (Chapter 3).
3. To examine alternative splicing of *SGCB*, *SGCD*, *SGCZ*, *SGCA* and *SGCG*. Alternatively spliced transcripts encoding novel isoforms could contribute to heterogeneity of sarcoglycan complexes between tissues (Chapters 4 and 5).

Finally, data presented in this thesis will be summarised and discussed in Chapter 6.

Chapter 2: Materials and methods

All experiments used analytical grade chemicals from either Thermo Fisher Scientific or Sigma Aldrich unless otherwise indicated.

2.1. Molecular Biology

2.1.1. First strand synthesis from total RNAs

Pre-extracted mouse whole brain total RNA was supplied by Dr Adrian Waite and used as a template for first strand cDNA synthesis. Pre-extracted control human cerebellum RNA from a study of neurological disorders in Wales (REC for Wales 09/MRE09/35) and total RNA from human skeletal muscle obtained from Clontech (item 636534) were used as templates for first strand cDNA synthesis. First strand cDNA from 1µg of RNA was synthesised with the Protoscript II First Strand cDNA Synthesis Kit (New England Biolabs) according to the manufacturer's instructions and using the supplied poly-d$(T)_{23}$VN primer to enrich for mature transcripts. First strand cDNA was stored at -20°C.

2.1.2. Oligonucleotides (primers)

All oligonucleotides (primers) are listed in appendix I. Primers were designed using a combination of Primer3 (Rozen and Skaletsky, 1999, Untergasser et al., 2012) and NCBI Primer-BLAST (Ye et al., 2012) using standard PCR design guidelines (Green and Sambrook, 2012). Primers were ordered from Sigma-Aldrich, re-suspended in molecular-grade water at a concentration of 100µM and stored at -20°C. For each pair of primers, the optimal annealing temperature for polymerase chain reactions (PCRs) was determined by assessing PCR yield and specificity at a range of annealing temperatures from 56-63°C using the general PCR protocol with RedTaq polymerase (section 2.1.3.2). The highest annealing temperature with decent yield and the fewest amplified products visible on agarose gel

electrophoresis (section 2.1.4) was used for all subsequent PCRs with that primer pair (Green and Sambrook, 2012).

2.1.3. Polymerase chain reaction

All polymerase chain reactions (PCRs) were carried out using C1000/S1000 Thermal Cyclers (BioRad). For each PCR, the final reaction mixture contained 1X DNA polymerase buffer (1.5-2mM Mg^{2+} depending on enzyme), 0.2-0.5µM each forward and reverse primers, 0.16-0.2mM deoxyribonucleotide mixture (0.16-0.2mM each dATP, dTTP, dGTP and dCTP), DNA polymerase, sterile molecular biology-grade water, and template DNA. Final PCR volume ranged from 10µl to 25µl depending on the DNA polymerase used and purpose of the reaction. For sequences with a high GC content, DMSO was added at a final concentration of 5% to facilitate amplification (Green and Sambrook, 2012). The following DNA polymerase enzymes were used:

- EasyA Hi-Fi Cloning enzyme (Agilent) for reverse transcription polymerase chain reaction (RT-PCR) from first strand cDNA
- Q5 High-Fidelity DNA polymerase (New England Biolabs) for RT-PCR from first strand cDNA
- RedTaq DNA polymerase (Sigma-Aldrich) for colony and routine PCRs
- PfuUltra II Fusion DNA polymerase (Agilent) for site-directed mutagenesis (SDM)

2.1.3.1. RT-PCR from first strand cDNA

To amplify specific transcripts from first strand cDNA, either the EasyA Hi-Fi cloning enzyme (Agilent) or Q5 High-Fidelity DNA polymerase (New England Biolabs) was used. Using the EasyA Hi-Fi cloning enzyme, the typical PCR mixture was as follows:

Reagent	Volume (µl)	Final concentration
10x EasyA reaction buffer	2.5	1x
Deoxyribonucleotide mix (10mM)	0.5	200µM
Forward primer (10µM)	0.5	0.2µM
Reverse primer (10µM)	0.5	0.2µM
EasyA DNA polymerase	0.25	1.25 U/25µl
Template	0.5-2 depending on application	Depends on application
Molecular grade water	To 25	

Reaction mixtures using the Q5 High-Fidelity DNA Polymerase enzyme for RT-PCR were as follows:

Reagent	Volume (µl)	Final concentration
Q5 reaction buffer (5x)	5	1x
Deoxyribonucleotide mix (10mM)	0.5	200µM
Forward primer (10µM)	1.25	0.5µM
Reverse primer (10µM)	1.25	0.5µM
Q5 DNA polymerase	0.5	1 U/25µl
Template	0.5-2 depending on application	Depends on application
Molecular grade water	To 25	

For both cloning DNA polymerases, typical PCR cycling parameters were as follows:

1. 95°C for 4min
2. 95°C for 30s
3. Optimised (section 2.1.2) annealing temperature for primers (usually 55-60°C), 30s
4. 72°C for 1min/kb of product length
5. Go to step 2, 20-30 times depending on the abundance of the target sequence and desired output quantity
6. 72°C, 7min
7. 8°C, hold.

PCR products were visualised by electrophoresis as described in section 2.1.4.

2.1.3.2. General PCR

For general PCRs using the RedTaq DNA polymerase system, the typical reaction mixture was as follows:

Reagent	Volume (µl)	Final concentration
10x RedTaq buffer	2	1x
Deoxyribonucleotide mix (10mM)	0.32	160µM
Forward primer (10µM)	0.8	0.4µM
Reverse primer (10µM)	0.8	0.4µM
RedTaq DNA polymerase	0.6	0.6 U/20µl
Template	0.5-2 depending on application	Depends on application
Molecular grade water	To 20	

The typical PCR cycling parameters and post-PCR analysis were as for the RT-PCRs.

2.1.3.3. Colony PCR

To screen bacterial colonies for plasmid inserts, individual bacterial colonies were picked using radiation-sterilised inoculation needles (Sarstedt/Fisher Scientific) and resuspended in 30-40µl of sterile Lennox formulation (5g/l NaCl) Luria-Bertani (LB) broth or molecular biology grade water. A 2µl aliquot of the resuspended colony was used as the template in the following PCR reaction mixture with appropriate oligonucleotide primers:

Reagent	Volume (µl)	Final concentration
10x RedTaq buffer	2	1x
Deoxyribonucleotide mix (10mM)	0.32	160µM
Forward primer (10µM)	0.8	0.4µM
Reverse primer (10µM)	0.8	0.4µM
RedTaq DNA polymerase	0.6	0.6 U/20µl
Template	2µl colony resuspension	-
Molecular grade water	To 20	

Typical thermal cycler parameters were as in section 2.1.3.1. PCR products were resolved by agarose gel electrophoresis as described in section 2.1.4.

2.1.3.4. Site-directed mutagenesis

Site-directed mutagenesis (SDM) was used to correct PCR errors and introduce mutations in cDNA plasmid constructs. The QuikChange (Aligent) mutatgenesis protocol using the PfuUltra high-fidelity DNA polymerase (Agilent) was employed for this as per the manufacturer's protocol. The typical PCR mixture for SDM PCRs was as follows:

Reagent	Volume (µL)	Final concentration
10x PfuUltra reaction buffer	1	1x
Deoxyribonucleotide mix (10mM)	0.2	200µM
Forward primer (10µM)	0.2	0.2µM
Reverse primer (10µM)	0.2	0.2µM
PfuUltra DNA polymerase	0.2	1 U/10µl
Plasmid template	1	-
Molecular grade water	To 10	-

The PCR cycling protocol was as follows:

1. 95°C for 30s
2. 95°C for 30s
3. 55°C for 1min
4. 68°C for 15min
5. Go to step 2 a total of 18 times
6. 12°C hold

1µl of the PCR product was analysed by agarose gel electrophoresis (section 2.1.4) to check that the reaction had been successful. DNA passaged through *E.coli* is methylated by deoxyadenosine methylase, so template DNA was removed from the SDM PCR product by adding 1µl of the methylated DNA-specific *Dpn*I enzyme (New England Biolabs) and incubating this mixture at 37°C for 2h. The resultant product was used to transform

chemically competent XL1-Blue cells by heat shock (section 2.2.2); only *E.coli* transformed with intact (PCR product) plasmids would survive antibiotic selection. Plasmids from the resultant colonies were sequenced to verify that mutagenesis was successful.

2.1.4. Agarose gel electrophoresis

Agarose gels for resolving DNA fragments were made using 0.5-3% (w/v) agarose (ATGC) in 0.5x TBE (Fisher; 45mM Tris-borate, 1mM EDTA, pH 8.3). Agarose concentration was determined by the sizes of the DNA fragments to be resolved. Ethidium bromide was added to the gels at a final concentration of 100ng/ml to visualise DNA. DNA samples to be resolved were mixed with a suitable volume of 5x DNA loading buffer (30% (v/v) glycerol, 20mM EDTA pH 8.0, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol in molecular grade water) before being loaded onto the gel. Samples were resolved by electrophoresis at 100-120V for 20-90min alongside a DNA fragment size standard, either the 1Kb Plus DNA Ladder (Life Technologies), or the Quick-Load 2-Log DNA Ladder (New England Biolabs). UV transillumination of the gel was performed to visualise resolved, ethidium bromide-stained DNA (BioRad Gel Doc XR+ Imaging System).

2.1.5. PCR product purification

PCR products were purified using either the QIAGEN QIAquick Gel Extraction kit or the QIAGEN QIAquick PCR Purification kit as per the manufacturer's instructions. Samples were eluted in 30-50µl of supplied elution buffer (EB) or molecular biology grade water depending on intended downstream use.

2.1.6. Nucleic acid quantitation

The concentration and purity of DNA and RNA samples were analysed using a NanoDrop 8000 spectrophotometer (Thermo Scientific) according to the manufacturer's instructions with the inclusion of an appropriate blank.

2.1.7. Cloning into plasmid vectors

2.1.7.1. Vectors used and constructs produced

Four different plasmid vectors were employed, depending on the qualities of the cDNA insert. All four plasmids were mammalian expression vectors, to facilitate expression of the inserted cDNA in cultured mammalian cells. These plasmids were:

- pCI-neo (Promega) – standard expression vector with no included epitope tag; ampicillin selection in *E.coli*
- pCMV-myc (Clontech) – expression of cDNA with an N-terminal c-Myc epitope tag; ampicillin selection in *E.coli*
- pCMV-HA (Clontech) - expression of cDNA with an N-terminal human influenza haemagglutinin (HA) epitope tag; ampicillin selection in *E.coli*
- pFUSE-hlgG2-F2 (Invivogen) – expression of cDNA with an N-terminal IL2 signal sequence for secretion; Zeocin selection in *E.coli*.

The constructs produced are summarised in Table 2.1. c-Myc, HA and FLAG epitope tags were introduced into pCI-neo-based constructs by PCR. The c-Myc N-terminal tag was introduced into pFUSE-hlgG2-F2 constructs by PCR.

Encoded protein	N-term epitope tag	C-term epitope tag	Species	Plasmid vector	5' RE site	3' RE site
ε-SG isoform 1 exon 8+	c-Myc	-	Human	pFUSE-hIgG2-F2	*Nco*I	*Nco*I
ε-SG isoform 1 exon 8-	c-Myc	-	Human	pFUSE-hIgG2-F2	*Nco*I	*Nco*I
ε-SG isoform 2 exon 8+	c-Myc	-	Human	pFUSE-hIgG2-F2	*Nco*I	*Nco*I
ε-SG isoform 2 exon 8-	c-Myc	-	Human	pFUSE-hIgG2-F2	*Nco*I	*Nco*I
α-SG isoform 1	-	c-Myc	Human	pCI-neo	*Sal*I	*Not*I
α-SG isoform 3 (exon 6 truncation)	-	c-Myc	Human	pCI-neo	*Sal*I	*Not*I
α-SG isoform 4 (Δ exon 9b)	-	c-Myc	Human	pCI-neo	*Sal*I	*Not*I
β-SG isoform 1	-	FLAG	Human	pCI-neo	*Sal*I	*Not*I
β-SG isoform 2	-	FLAG	Human	pCI-neo	*Sal*I	*Not*I
β-SG T182A	-	FLAG	Human	pCI-neo	*Sal*I	*Not*I
δ-SG isoform 1 (exon 9 termination)	-	HA	Human	pCI-neo	*Sal*I	*Not*I
δ-SG isoform 2 (exon 8b termination)	-	HA	Human	pCI-neo	*Sal*I	*Not*I
δ-SG isoform 3 (Δ exon 7)	-	HA	Human	pCI-neo	*Sal*I	*Not*I
δ-SG isoform 1 exon 3 truncation	-	HA	Human	pCI-neo	*Sal*I	*Not*I
δ-SG isoform 1 Δ exon 6	-	HA	Human	pCI-neo	*Sal*I	*Not*I
ζ-SG isoform 1	c-Myc	-	Human	pCMV-myc	*Sal*I	*Not*I
ζ-SG isoform 2 (Δ exon 3)	c-Myc	-	Human	pCMV-myc	*Sal*I	*Not*I
ζ-SG isoform 3 (Δ exon 5)	c-Myc	-	Human	pCMV-myc	*Sal*I	*Not*I
γ-SG isoform 1	c-Myc	-	Human	pCMV-myc	*Sal*I	*Not*I
γ-SG isoform 2 (Δ exon 5)	c-Myc	-	Human	pCMV-myc	*Sal*I	*Not*I
γ-SG isoform 3 (exon 6b+)	c-Myc	-	Human	pCMV-myc	*Sal*I	*Not*I
Kcna3	c-Myc	-	Mouse	pCMV-myc	*Sal*I	*Not*I
Tenascin-R (full-length isoform)	-	FLAG	Mouse	pCI-neo	*Sal*I	*Not*I

Table 2.1 Plasmid cDNA constructs for expression in mammalian cells. For each construct, the encoded protein, N-terminal or C-terminal tag as appropriate, species of origin, vector and restriction enzyme sites used to produce the construct are provided. For novel sarcoglycan isoforms, the alternative splicing event responsible for that isoform is provided in parentheses. Abbreviations: 5' RE site, restriction enzyme site at the 5' end of the cDNA insert; 3' RE site, restriction enzyme site at the 3' end of the cDNA insert; C term, carboxyl terminus; N term, amine terminus.

2.1.7.2. Restriction digest of DNA

Plasmid vectors and PCR products were digested with restriction enzymes purchased from New England Biolabs (NEB). For single digests, the recommended NEB buffer was used; for double digests, either the recommended double digest buffer was used or sequential digests were performed. For a typical digest, 42.5µl of purified PCR product or 0.5-1µg of DNA was combined with 1x buffer, 1x bovine serum albumin (BSA), 5 units of each enzyme, and molecular grade water in a total volume of 50µl. This was incubated at 37°C for 3h to ensure

complete digestion of DNA. To prevent re-circularisation, digested plasmids were further incubated with 20 units of calf intestinal alkaline phosphatase (CIP; New England Biolabs) at 37°C for an additional hour. This removed 5' phosphate groups from the cut ends of the DNA. Digests were purified as described in section 2.1.5.

2.1.7.3. Ligation of cohesive DNA fragment termini

The termini of digested DNA fragments were ligated using the T4 DNA ligase (Promega). A typical reaction contained 1µl 10x ligase buffer, 1µl digested and purified plasmid vector (30-100ng), 1µl (2U) of T4 ligase, 4µl of purified, digested PCR product, and molecular grade water to a total volume of 10µl. This reaction was thoroughly mixed and incubated at 4°C overnight to allow the ligation reaction to proceed to completion. For each preparation of linearised vector, a control reaction was also performed in which the insert DNA was omitted; this identified whether the vector had been fully linearised and dephosphorylated. Approximately 2-5µl of ligation reaction was used to transform cells of an appropriate *E.coli* strain (sections 2.2.2-4).

2.2. Plasmid amplification and protein expression in *E.coli*

2.2.1. Preparation of chemically competent *E.coli* XL1-Blue

Chemically competent *E.coli* XL1-Blue cells (Agilent; genotype: *rec*A1 *end*A1 *gyr*A96 *thi-1 hsd*R17 *supE44 relA1 lac* [F' *proAB lacI*^{q}*ZΔM15* Tn*10* (Tet^{r})]) were prepared in-house using a slightly modified version of the calcium chloride protocol described in (Alexander, 1987). A single colony of XL1-Blue *E.coli* (Agilent) was used to inoculate 10ml of LB media containing 10µg/ml tetracycline to select for growth of only the tetracycline-resistant XL1-Blue bacteria. This starter culture was incubated overnight at 37°C, 200rpm. The following morning, 5ml of the starter culture was used to inoculate 500ml of fresh, sterilised LB media and this culture was incubated at 30°C with shaking at 195rpm until the culture's optical

density at 600nm (OD_{600}) reached 0.45-0.55. The culture was then chilled on ice for 2h, and centrifuged at 2000rpm, 4°C for 20min. Harvested cells were resuspended gently in 24ml of ice-cold filter-sterilised salt solution (100mM $CaCl_2$, 70mM $MnCl_2$, 40mM sodium acetate (NaOAc), pH 5.5). After incubation on ice for another 45min, the cells were pelleted by centrifugation at 3500rpm, 4°C, for 10min. Cells were finally resuspended in 50ml of ice-cold salt solution containing 15% (v/v) glycerol, aliquoted into sterile tubes at 200µl per tube, and stored at -80°C until use. Transformation efficiency was calculated by performing a set of standard heat-shock transformations (section 2.2.2) with known quantities of pUC18 plasmid: 100pg, 50pg, 25pg, and 10pg. The number of bacterial colonies from these transformations were counted to determine the number of colony-forming units per microgram of DNA.

2.2.2. Transformation of XL1-Blue cells using the heat-shock method

In-house stocks of chemically competent XL1-Blue cells were transformed using a standard heat-shock method. For each transformation, 100µl of cells were combined with an appropriate volume of plasmid or ligation reaction in a thin-walled Falcon 2059 polypropylene tube and incubated on ice for 20-30min. Cells were heat shocked at 42°C for 45s, then placed on ice for 2min. To recover cells, 750µl of LB was added and the cells were incubated at 37°C for 45-60min. Cells were then plated on LB agar containing an appropriate antibiotic, and incubated at 37°C overnight before being checked for bacterial colony growth.

2.2.3. Transformation of XL10-Gold cells using the heat-shock method

XL10-Gold ultracompetent *E.coli* (Agilent; genotype: Tetr∆(*mcrA*)*183* ∆(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F´ *proAB lacIqZDM15* Tn*10* (Tet^r) Amy Cam^r]) were used for situations requiring higher transformation efficiency, such as the generation of cDNA mini-libraries to be screened for alternatively spliced cDNA inserts.

Transformations were carried out using a scaled-down version of the manufacturer's protocol: 2µl of the supplied 2-mercaptoethanol and 50µl of cells were mixed in a Falcon 2059 tube, and incubated on ice for 10min with mixing approximately every 2min. To this mixture 2.5µl of ligation reaction or 1µl of a 1:50 plasmid dilution was added, and incubated on ice for 30min. Cells were heat shocked at 42°C for 30s and placed on ice for 2min. To recover cells, 750µl of LB broth was added and the cells were incubated at 37°C for 1h with shaking at 200rpm. Cells were spun down and re-suspended in a suitable volume of LB broth for plating of the entire transformation mixture on LB agar containing an appropriate antibiotic. Plates were incubated at 37°C overnight and checked for bacterial colony growth.

2.2.4. Transformation of MAX Efficiency Stbl2 Competent cells using the heat-shock method

MAX Efficiency Stbl2 competent cells (Life Technologies; genotype: F- *mcrA Δ(mcr*BC-hsdRMS-*mrr) rec*A1 *end*A1*lon gyr*A96 *thi sup*E44 *rel*A1 λ- Δ(*lac-pro*AB)) were used for transformations of plasmids unstable in other *E.coli* strains. Transformations were carried out as per a scaled-down version of the manufacturer's protocol: 3µl of plasmid plus 50µl cells were incubated on ice in a Falcon 2059 tube for 30min, then heat shocked at 42°C for 25s. After 2min on ice, cells were recovered by adding 750µl of the supplied SOC medium and incubated at 30°C for 60-90min with shaking at 225rpm. Cells were then spun down and resuspended in an appropriate volume of LB broth for plating on LB agar containing the appropriate antibiotic. Plates were incubated at 37°C overnight and then checked for bacterial colony growth.

2.2.5. Preparation of plasmid DNA from *E.coli*

Plasmid DNA was isolated from small (5-10ml), medium (50-100ml) or large-scale (150-250ml) *E.coli* cultures. QIAGEN Spin Miniprep, Spin or Filter Midiprep, or Spin Maxiprep

kits were used to isolate DNA from these cultures respectively as per the manufacturer's protocol.

2.2.6. Expression of recombinant murine Thx-ε-SG C-terminus fusion protein in *E.coli* using the pET-32 system

Recombinant His-tagged Thx-ε-SG isoform 1 C-terminus fusion protein was expressed in *E.coli* strain BL21 (DE3) (Agilent; genotype: *E.coli* B F^- *dcm ompT hsdS*(r_B^- m_B^-) *gal* λ(DE3)). The fusion protein consisted of thioredoxin (Thx) plus the last 98 amino acids of mouse ε-SG isoform 1 (Esapa et al., 2007, Waite et al., 2011). The BL21 (DE3) strain is deficient in the *ompT* and *lon* proteases, which can reduce recombinant protein stability. A glycerol stock of BL21 (DE3) strain *E.coli* transformed with the fusion protein expression plasmid (supplied by A. Waite) was used to inoculate 10ml of LB containing an appropriate antibiotic, and this was incubated overnight at 37°C, 200rpm. A 2ml aliquot of the overnight culture was used to inoculate 250ml of LB containing an appropriate antibiotic and incubated at 37°C, 200rpm for about 2h until an OD_{600} of 0.6 was reached. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the culture at a final concentration of 1mM to induce expression of the recombinant protein. The culture was incubated at 37°C, 200rpm for 3h to permit protein synthesis. Cells were harvested from the culture by centrifugation at 8000rpm, 4°C for 20 minutes. Samples (1ml) of bacterial culture were taken immediately before and after induction then analysed by SDS-PAGE and Coomassie blue staining (sections 2.5.1-2.5.3) to determine the level of protein induction.

2.2.7. Purification of recombinant murine Thx-ε-SG isoform 1 C-terminus fusion protein from *E.coli*

During purification of the recombinant murine Thx-ε-SG isoform 1 C-terminus fusion protein, buffers containing 2M urea were used to obtain denatured protein. First, the cell

pellet from 2.2.6 was resuspended in 15ml of sonication buffer (20mM Tris pH 8.0, 100mM NaCL, 2M urea) and sonicated thoroughly using a Vibra-Cell Ultrasonic Processor (Sonics) in 30s pulses at 50W with chilling on ice between sonication pulses. The sample was then centrifuged again at 13000rpm, 4°C for 20min to pellet cell debris and insoluble material. His-tagged Thx-ε-SG isoform 1 C-terminus fusion protein is soluble, so was purified from the supernatant (Esapa et al., 2007, Waite et al., 2011). In order to purify the fusion protein from the supernatant, an Econo-Pac® disposable chromatography column (Bio-Rad) was packed with 2ml of TALON® Resin (Clontech). The resin was washed once with molecular biology-grade water, and then equilibrated in 20ml sonication buffer containing 2M urea. The supernatant containing fusion protein was added to the capped column, and incubated at room temperature for 30min with rotation. The column was then drained of supernatant and unbound protein, and washed three times in sonication buffer. The column was washed once more, with 10ml sonication buffer containing 5mM imidazole. Recombinant protein was eluted from the column in 3 elutions of sonication buffer containing 100mM imidazole to displace His-tagged protein from the TALON resin: elution 1 was 1.8ml, elution 2 was 4ml, and elution 3 was 2ml. Aliquots from the resuspended cell pellet, supernatant, and elutions were processed for SDS-PAGE and stained with Coomassie blue (sections 2.5.1-2.5.3) to assess protein purification. Most of the purified protein was present in the second elution.

2.3. Antibody preparation

2.3.1. Generation of ε-SG isoform 2-specific antibodies

To obtain antibodies specific to isoform 2 of ε-SG, rabbit polyclonal antibodies against the unique intracellular C-terminus of human ε-SG isoform 2 were generated. Two New Zealand white rabbits were immunised with a keyhole limpet haemocyanin-fused synthetic peptide corresponding to the human ε-SG isoform 2 C-terminus unique peptide NH_2-C-

QRFEVNGIPEERKLTEAMSL-COOH by CovaLab according to their standard immunisation protocols. Polyclonal antibodies against this peptide were purified from terminal bleed antiserum as described in section 2.3.2.2.

2.3.2. Preparation of peptide immunogen affinity chromatography column

The polyclonal rabbit antibodies against the ε-SG isoform 2-specific peptide immunogen were purified from terminal bleed antiserum by affinity chromatography with the immunogen peptide. Peptide immunogen affinity chromatography columns were produced by first filling an Econo-Pac® disposable chromatography column (Bio-Rad) with 2ml packed SulfoLink coupling resin (Thermo Scientific). This was equilibrated in 20ml coupling buffer without urea (50mM Tris pH 8.5, 5mM EDTA). The column was then capped before adding 2mg peptide immunogen (NH_2-C-QRFEVNGIPEERKLTEAMSL-COOH) in 5ml coupling buffer; this was incubated at room temperature with rotation for 15min. The column was removed from the tube rotator and allowed to stand at room temperature for a further 30min. After the column was washed with 10ml coupling buffer, free iodoacetyl groups on the Sulfolink resin were blocked by incubating the column with 3ml filter-sterilised 50mM L-cysteine in coupling buffer for 15min at room temperature with rotation. The column was then allowed to stand at room temperature for a further 30 minutes before being washed with 20ml 1M NaCl. After the column was washed with phosphate-buffered saline (PBS; 0.1M phosphate buffer, 0.0027M potassium chloride, and 0.137M sodium chloride, pH 7.4), a frit was added and the column was ready for use. Columns were stored at 4°C with the resin equilibrated in PBS containing 0.025% (w/v) sodium azide.

2.3.3. Preparation of denatured fusion protein affinity chromatography column

Generating a fusion protein affinity chromatography column using purified denatured Thx-ε-SG isoform 1 C-terminus fusion protein (section 2.2.7) required reduced protein. To

accomplish this, the protein was first desalted from sonication buffer into reduction buffer containing urea (0.1M sodium phosphate, 5mM EDTA, 2M urea) by running 3ml of protein purification elution 2 (section 2.2.7) over an Econo-Pac® 10DG desalting column (Bio-Rad) that had previously been equilibrated in reduction buffer. Protein was eluted in 4ml reduction buffer, and dithiothreitol (DTT, Sigma) was added to a final concentration of 25mM. This was incubated at 37°C for 90min in the dark to reduce the fusion protein. To the reduced protein, 2ml of coupling buffer (50mM Tris pH 8.5, 5mM EDTA, 2M urea) was added. This was passed over a fresh Econo-Pac® 10DG desalting column equilibrated in coupling buffer in two aliquots of 3ml each, with protein eluted from the column after each aliquot using 4ml coupling buffer. Once desalted into coupling buffer, the reduced fusion protein was incubated with 2ml packed coupling buffer equilibrated SulfoLink coupling resin (Thermo Scientific) for 15min at room temperature with rotation in an Econo-Pac® disposable chromatograph column (Bio-Rad). The column was then allowed to stand for 30min at room temperature before being drained and washed once with coupling buffer. To block remaining free iodoacetyl groups on the SulfoLink resin (Thermo Scientific), the column was incubated with 2ml of 50mM L-cysteine at room temperature for 15min with rotation, and then allowed to stand for 30min at room temperature. The column was drained and washed once with 1M NaCl, then once with PBS. A frit was added, and the column was stored at 4°C with the resin equilibrated in PBS containing 0.025% (w/v) sodium azide.

2.3.4. Antibody purification

2.3.4.1. Purification of polyclonal antibodies from terminal bleed serum

Antibodies were purified from rabbit terminal bleed antiserum using previously prepared affinity chromatography columns with the immunogen used to produce each antibody. For the esg3790 antibody the denatured Thx-ε-SG isoform 1 carboxyl terminus column (section 2.3.3) was used, while for esg2-1355 and esg2-1358 the ε-SG isoform 2-specific peptide

column (section 2.3.2) was used for affinity chromatography. All other antibodies were purified using immunogen affinity chromatography columns previously prepared by A. Waite (Table 2.2, section 2.3.6). Approximately 10ml of terminal bleed serum diluted to 20ml with PBS was repeatedly passed over a pre-clearing column (Table 2.2) equilibrated in PBS if required for 1h at room temperature. The diluted, pre-cleared serum was then passed over the appropriate PBS-equilibrated immunogen affinity chromatography column for 1h to bind antibodies. The column was then washed in PBS before antibodies were eluted using low-pH IgG elution buffer (Thermo Scientific). Eluted antibody was neutralised using 50µl 1M Tris pH 9 per 1ml IgG elution buffer. Columns were regenerated with a wash of IgG elution buffer and a wash of PBS. Columns were stored at 4°C with PBS containing 0.025% (w/v) sodium azide. Antibody elutions were quantified using the enhanced BCA assay kit (Piece) as per the manufacturer's instructions, and stored at 4°C short term or -20°C for longer term.

2.3.4.2. Purification of monoclonal antibodies from hybridoma culture media

Mouse hybridoma culture supernatant containing monoclonal antibodies was supplied by A. Waite for the 9E10 antibody, or purchased from the Developmental Studies Hybridoma Bank (DSHB) for the MANDAG2 and MANDRA1 antibodies. MANDAG2 and MANDRA1 were both produced and deposited to the DSHB by G.E. Morris (Helliwell et al., 1994, Nguyen et al., 1992). Antibodies were purified from the culture supernatant by passing the supernatant over a PBS-equilibrated 1ml bed of packed PBS-equilibrated Protein G-sepharose (GE Healthcare Life Sciences) in an Econo-Pac® disposable chromatography column (Bio-Rad) for 2h. The column was then washed once with PBS before antibodies were eluted using IgG elution buffer (Thermo Scientific). Eluted antibody was neutralised using 50µl 1M Tris pH 9 per 1ml antibody elution. Antibodies were stored at 4°C short-term or -20°C long-term. The Protein G-sepharose columns were regenerated by washing once with IgG elution buffer and once with PBS. They were stored in PBS containing 0.025% (w/v) sodium azide at 4°C.

2.3.5. Cross-linking of antibodies to Protein A-agarose or Protein G-sepharose beads

Mouse monoclonal antibodies 9E10, MANDAG2 and MANDRA1 (section 2.3.6, Table 2.2) were coupled to Protein G-sepharose beads (GE Healthcare Life Sciences) for immunoprecipitation (IP) and immunoaffinity purification (IAP) experiments (sections 2.6.2 and 2.6.4), while rabbit anti-ε-SG and anti-Tenascin-R polyclonal antibodies (section 2.3.6, Table 2.2) were coupled to Protein A-agarose beads (Life Technologies). 2mg of PBS-diluted antibody per 1ml packed PBS-equilibrated beads was incubated with the beads either overnight at 4°C or for 2 hrs at room temperature with rotation. 10µl samples of the diluted antibody before and after incubation with the beads, and a 5µl (packed) sample of the beads after incubation with the antibody were taken. Beads were then spun down at 2000rpm for 5min, washed once with PBS, and washed once with 0.1M borate buffer (0.2M di-sodium tetraborate, 0.2M boric acid, pH 9.0). To cross-link the antibodies to the beads, the beads were incubated for 30min at room temperature with 5ml 20mM dimethyl pimelimidate dihydrochloride (DMP) in 0.1M borate buffer. Beads were spun down at 2000rpm for 5min, and washed once in 5ml 0.1M ethanolamine pH 8.0. Remaining DMP was quenched by incubating the beads in 5ml 0.1M ethanolamine for 1h at room temperature with rotation. Beads were then centrifuged at 2000rpm for 5min and washed once in PBS, once in 5ml IgG elution buffer (Thermo Scientific) to remove unbound antibody, and 2-4 times in PBS. A sample of cross-linked antibody beads was taken. The beads were stored as a 50% slurry in PBS containing 0.025% (w/v) sodium azide at 4°C.

To assess whether cross-linking of the beads was successful, samples collected during the protocol (diluted antibody, depleted antibody after incubation with the beads, beads after incubation with the antibody, and cross-linked antibody beads) were analysed via SDS-PAGE and Coomassie blue staining (sections 2.5.2-2.5.3). If antibody was adsorbed to the beads, then antibody would be reduced in the post-incubation antibody sample compared to the

original diluted antibody sample. If cross-linking was successful, less protein would be detected in the cross-linked antibody beads compared to the first antibody beads sample.

2.3.6. Antibodies used in this study

Antibody	Immunogen	Pre-clear column	Specificity	Western blot dil.	IP?	Source
Rabbit polyclonal						
esg3788	Thx-ε-SG isoform 1 C-term	Thx	ε-SG C-term, all isoforms	1:250	Yes	(Waite, 2009)
esg3790	Thx-ε-SG isoform 1 C-term denatured in 2M urea	Thx	ε-SG C-term, all isoforms	1:350	No	(Waite, 2009)
Biotinylated esg3790	Thx-ε-SG isoform 1 C-term denatured in 2M urea	Thx	ε-SG C-term, all isoforms	1:250	No	Gift from A.Waite
esg4990	GST-ε-SG isoform 2 C-term	GST	ε-SG C-term, all isoforms	1:350	Yes	(Esapa et al., 2007)
esg2-1358	ε-SG isoform 2 peptide	None	ε-SG isoform 2 C-term	-	Yes	Described herein (Waite et al., 2016)
esg2-1355	ε-SG isoform 2 peptide	None	ε-SG isoform 2 C-term	-	Yes	Described herein (Waite et al., 2016)
anti-α-tubulin (loading control)	Human α-tubulin aa1-100	-	α-tubulin	1:2000	No	Abcam
Mouse monoclonal						
9E10	Human c-Myc aa408-439	-	c-Myc	1:350	Yes	DSHB
MANDAG2 (clone 7A10)	Last 16 aa of β-dystroglycan	-	β-dystroglycan	1:250	Yes	DSHB (Helliwell et al., 1994)
MANDRA1 (clone 7D11)	Human dystrophin exons 45-50	-	Dystrophin C-term	1:200	Yes	DSHB (Nguyen et al., 1992)
M2 anti-FLAG	FLAG epitope tag	-	FLAG epitope tag	1:2000	No	Sigma Aldrich
HA.11	HA epitope tag	-	HA epitope tag	1:1000	No	BioLegend
Tenascin-R clone 619	Purified tenascin-R	-	Tenascin-R	1:200	Yes	R&D Systems
Secondary antibodies						
Alexa Fluor 680-conjugated donkey anti-rabbit IgG	rabbit IgG	-	rabbit IgG	1:10,000	No	Life Technologies

IRDye800CW-conjugated donkey anti-mouse IgG	mouse IgG	-	mouse IgG	1:10,000	No	Rockland Immunocyto-chemicals
Alexa Fluor 680-conjugate streptavidin	-	-	biotin	1:5000	No	Life Technologies

Table 2.2 Antibodies used in this study. This table provides details for each antibody used in this thesis, including both commercially purchased and in-house antibodies. For each antibody, the immunogen against which it was raised and the antigen to which it binds are given. Also provided are the Western blot dilution at which the antibody was used, whether it was used for immunoprecipitation (IP) experiments, and the source of the antibody. For antibodies developed in-house, the affinity chromatography column used to pre-clear non-specific antibodies from the antiserum is also given. Abbreviations: dil., dilution; IP, immunoprecipitation; Thx, thioredoxin; GST, glutathione S-transferase; C-term, carboxyl terminus; aa, amino acids; IgG, immunoglobulin G; DSHB, Developmental Studies Hybridoma Bank.

2.4. Tissue culture

2.4.1. Mammalian cell line

For the biochemistry experiments described herein, HEK293T cells were employed. These are a human embryonic kidney cell line, although the gene expression profile suggests potentially neuronal or adrenal origin (DuBridge et al., 1987, Lin et al., 2014). The cell line was originally obtained from the European Collection of Authenticated Cell Cultures (ECACC).

2.4.2. Standard cell culture conditions

HEK293T cells were cultured in high-glucose Dulbecco's modified Eagle medium containing GlutaMAX supplement (DMEM, Life Technologies), with 10% foetal bovine serum (FBS, Life Technologies) and 1% w/v penicillin/streptomycin (Life Technologies) added. They were maintained as semi-adherent cultures in standard tissue culture flasks with appropriate volumes of media, in humidified incubators at 5% CO_2. At 60-90% confluence cultures were passaged to maintain their viability. To passage cells, the media was removed and the cells washed twice in Ca^{2+}- and Mg^{2+}-free Hanks balanced salt solution (HBSS, Sigma). Cells were then detached from the flask by incubation in an appropriate volume of 1x trypsin-EDTA (Sigma) solution at 37°C for 3min. Trypsinisation was terminated by adding the standard culture growth DMEM, and cells were collected through centrifugation at 1000rpm for 3 minutes at room temperature. The cell pellet was resuspended in an appropriate volume of DMEM, and reseeded in culture flasks at a 1:5 to 1:20 dilution depending on original confluence. For cryopreservation of cell aliquots, the cell pellet was resuspended in DMEM containing 10% (v/v) DMSO for a density of 1-3x10^6 cells/ml, aliquoted into cryovials, and frozen at -80°C using a Mr. Frosty (Life Technologies) to preserve cells.

For experiments, cells were seeded into dishes or multi-well plates as follows:

- 10cm dish: $2.5x10^6$ cells in 10ml media per dish
- 6cm dishes: $8\text{-}8.5x10^5$ cells in 4ml media per dish
- 6 well dish: $3.5x10^5$ cells in 2ml media per well
- 12 well dish: $2x10^5$ cells in 1ml media per well
- 24 well dish: $1x10^5$ cells in 0.5ml media per well

2.4.3. Transfection of mammalian cell cultures

Transfection of HEK293T cells at 40-50% confluency was carried out using the FuGENE 6 transfection reagent (Promega) as per the manufacturer's protocol. For each 1µg of DNA to be transfected, 3µL of FuGENE was used. For each 6cm dish, no more than 6µg total DNA was transfected; for other dish/well sizes, this was scaled up or down as appropriate.

2.4.4. Treatment of cells with bortezomib to inhibit proteasome activity

A 1mM stock of the proteasome inhibitor bortezomib (Santa Cruz Biotech) in DMSO was diluted to 1µM with standard cell culture media. The culture media on 100% confluent cells was replaced with bortezomib-containing media. To control for the effects of the diluent DMSO on the cells, a second set of 100% confluent cells were treated with an equal volume of DMSO diluted in standard cell culture media. Cells were allowed to grow overnight (15-18hrs) after addition of bortezomib or DMSO before collection and analysis by SDS-PAGE and Western blot (sections 2.5.1-2.5.4).

2.5. Protein analysis

2.5.1. General sample preparation for SDS-PAGE and Western blots

2.5.1.1. Sample preparation from tissue

Mouse tissue was homogenised in treatment buffer (75mM Tris pH 6.8, 3.8% (w/v) SDS, 4M urea, 20% (v/v) glycerol) using a Polytron PT 3100 bench-top homogeniser (Kinematica).

Cell debris was removed from the homogenate by centrifugation at 4000rpm for 10 minutes at 4°C. Protein concentration in the homogenate was determined using the BCA assay (section 2.4.3), and the homogenate was diluted to 10mg/ml using treatment buffer containing 0.001% (w/v) bromophenol blue and a final concentration of 5% (v/v) 2-mercaptoethanol. Samples were stored at -20°C. For SDS-PAGE, 50-100µg was diluted to an appropriate volume in treatment buffer containing bromophenol blue but no 2-mercaptoethanol. 2-mercaptoethanol was then added to a final concentration of 5% (v/v), the sample was boiled at 95°C for 5min and cooled on ice before being loaded onto the polyacrylamide gel (section 2.5.2).

2.5.1.2. Sample preparation from cultured cells

Cells at 100% confluency were lysed in an appropriate volume of 2x Laemmli sample buffer (0.125M Tris pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 0.001% (w/v) bromophenol blue), and scraped off the bottom of the plate or well. These lysates were transferred to microcentrifuge tubes, and sonicated using a Vibra-Cell Ultrasonic Processor (Sonics), with 10s pulses at 50W for 2 cycles. Lysates were stored at -20°C. Prior to SDS-PAGE, 2-mercaptoethanol was added to the lysate at a final concentration of 5% (v/v) and the sample was boiled at 95°C for 5min then cooled on ice before being loaded onto the polyacrylamide gel.

2.5.2. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) under denaturing conditions

Depending on molecular weight and the protein size differences to be resolved, proteins were visualised on 8-12% linear polyacrylamide gels by SDS-PAGE using the Mini Protean III gel system (Bio-Rad). For each gel, a resolving gel containing 380mM Tris-HCl pH 8.9, 8-12% (v/v) acrylamide (ProtoGel 30% (w/v) acrylamide:0.8% (w/v) bis-acrylamide 37.5:1 National

Diagnostics), 0.1% (w/v) sodium dodecyl sulphate (SDS), 0.1% (w/v) ammonium persulphate (APS) and 0.08% (v/v) N,N,N',N'-Tetramethylethylenediamine (TEMED) was poured in either an 0.75mm or a 1.5mm spacer plate with short plate to complete the mould. During polymerisation, isopropanol was layered on the top of the resolving gel to ensure a level interface between that and the stacking gel. Once the gel was polymerised, the isopropanol was rinsed off the gel with distilled water and the gel was dried. Subsequently, a stacking gel containing 125mM Tris-HCl pH 6.8, 5% (v/v) acrylamide (30% (w/v) acrylamide:0.8% (w/v) bis-acrylamide 37.5:1), 0.1% (w/v) SDS, 0.1% (w/v) APS and 0.1% (v/v) TEMED was layered on top of the resolving gel. A lane-forming comb was inserted into this stacking gel, and it was allowed to fully polymerise at room temperature. Gels were then immobilised in the Mini-Protean clamp system (Bio-Rad) and submerged in 1x SDS-PAGE running buffer (25mM Tris-base, 192mM glycine, 1% (w/v) SDS). Samples were loaded into the gel lanes with volume loaded dictated by sample concentration and gel thickness: 1.5mm thickness gels accommodated up to about 45µl of sample per lane, whereas 0.75mm thickness gels accommodated up to 25µl of sample. An identical volume was loaded into each well on a gel to ensure even running of the gel. To determine approximate protein sizes, the pre-stained Protein Marker, Broad Range 6-175kDa (New England Biolabs) or its replacement the Blue Prestained Protein Standard, Broad Range 11-190kDa (New England Biolabs) was run alongside the samples. Proteins were separated by electrophoresis at 150V for 60-90min or until the protein markers or standards were appropriately resolved.

2.5.3. Coomassie staining of SDS-PAGE gels

To directly visualise proteins in an SDS-PAGE gel, the gel was stained using Coomassie blue stain (0.1% (w/v) Coomassie brilliant blue, 40% (v/v) methanol, 10% (v/v) acetic acid). The gel was incubated in the stain for 2h at room temperature with agitation. The stain was then drained off, and the gel was destained in 40% (v/v) methanol, 10% (v/v) acetic acid at room

temperature for 2-6h at room temperature or overnight at 4°C with agitation. After rinsing in distilled water, the stained gel was scanned using the Odyssey® Infrared Imaging System (LI-COR Biosciences) in the 700nm channel.

2.5.4. Western blotting

After electrophoresis, proteins from unstained SDS-PAGE gels were transferred to 0.2µm pore size nitrocellulose (Amersham™ Protran®). This was accomplished using the Mini Trans-blot electrophoretic transfer cell (Bio-Rad). Transfers were carried out with the gel and nitrocellulose completely submerged in transfer buffer (25mM Tris-base, 192mM glycine, 1% (w/v) SDS, 20% (v/v) methanol) for 1h at 75V if a single gel or 85V if proteins from two gels were being transferred. After protein transfer, non-specific binding sites on the nitrocellulose membrane were blocked by incubating the membrane in TBST (1.5M NaCl, 500mM Tris pH 7.5, 1% (v/v) Tween) containing 5% (w/v) dried skimmed milk powder either for 1h at room temperature or overnight at 4°C with agitation. The blocking solution was removed, and replaced with the primary antibody or antibodies diluted appropriately (section 2.3.6, table 2.2) in 5% milk-TBST solution. The membrane was incubated in the primary antibody solution for 1 hr at room temperature with agitation, then washed three times in 1x TBST at room temperature for 5min per wash with agitation. The Alexa Fluor 680 (Life Technologies) or IRDye 800 (Rockland Immunocytochemicals) conjugated secondary antibodies were diluted in TBST as appropriate (section 2.3.6, table 2.2) and added to the membrane; this was incubated in the dark for 30min at room temperature with agitation. The membrane was washed 3 further times in TBST as before. The membrane was then imaged using the two-channel Odyssey Infrared Imaging System (LI-COR Biosciences).

2.5.5. Quantitation of Western blots

Western blot signal intensity was quantified using the Odyssey Infrared Imaging System (LI-COR Biosciences) with Image Studio Lite Version 5.2 software (LI-COR Biosciences). To quantify signal intensity in a Western blot protein band, the Image Studio software was used to draw a rectangle around the entire band on the digital image of the blot. The intensity of immunofluorescence signal per channel was determined for the region inside the rectangle and the upper and lower boundaries of the rectangle. The background signal was calculated from the upper and lower boundaries of the rectangle, and subtracted from the signal within the rectangle to give the signal intensity of the protein band, corrected for background signal.

2.6. Proteomics

2.6.1. PNGase F treatment of recombinant protein

PNGase F (New England Biolabs) is an amidase cleaving between the innermost N-acetylglucosamine of an oligosaccharide and the asparagine to which it is linked. Transfected HEK293T cells in 6-well plates were collected 24-48h after transfection and washed once in PBS. Cells were then lysed in 100µl per well of 1x glycoprotein denaturing buffer (New England Biolabs). The lysate was scraped up from the well bottom, transferred to a microcentrifuge tube, and briefly sonicated. It was then boiled at 95°C for 10min and chilled on ice. For each sample, two reactions were set up as follows:

Reagent	**PNGase F**	**Undigested**
Lysate	35µl	35 µl
GlycoBuffer 2 (New England Biolabs)	5 µl	5 µl
PNGaseF (New England Biolabs)	1 µl	-
NP-40 (New England Biolabs)	5 µl	5 µl
PBS	4 µl	5 µl

The reaction mixtures were incubated at 37°C for 2h, and then 2x Laemmli sample buffer plus 2-mercaptoethanol to a final concentration of 5% (v/v) was added. Samples were analysed by SDS-PAGE and Western blot, as described in sections 2.5.1.2, 2.5.2 and 2.5.4 respectively.

2.6.2. Immunoprecipitation from cell lines

Transfected or wild-type HEK293T cells were collected 24-48h after transfection, at 100% confluency. Cells were washed in ice-cold PBS, and then lysed in either ice-cold RIPA buffer (150mM NaCl, 50mM Tris pH 8.0, 1% (v/v) Triton X-100, 0.5% (w/v) sodium deoxycholate, and 1mM EGTA) or NP-40 buffer (50mM Tris pH 7.4, 150mM NaCl, 1% (v/v) NP-40) for 30min on ice. Cells were scraped off the dish or well bottom, and triturated by repeated passage through a 20-23 gauge needle. Cell debris was removed by centrifugation at

13,000rpm for 10min at 4°C. An aliquot of the lysate was taken, and the remainder was pre-cleared for 1h at room temperature under rotation with 10µl of packed Protein A-agarose or Protein G-sepharose equilibrated in the lysis buffer used. The pre-cleared lysate was then incubated overnight at 4°C under rotation with either cross-linked antibody beads or antibody pre-absorbed onto Protein A-agarose/Protein G-sepharose beads. The antibody beads with bound protein were then washed three times in the appropriate lysis buffer, and bound proteins were eluted in 20-60µl of 2x Laemmli sample buffer containing 5% 2-mercaptoethanol by boiling at 95°C, 5min. Samples were analysed by SDS-PAGE and Western blot (sections 2.5.2 and 2.5.4 respectively).

2.6.3. Cell surface biotinylation

Biotinylation of cell surface proteins was performed using a modified version of the protocol in (Esapa et al., 2007). Transfected HEK293T cells at approximately 100% confluency and a minimum of 24 hrs post-transfection were washed once in ice-cold PBS. To biotinylated cell surface proteins, cells were then incubated with an appropriate volume of freshly-prepared 0.5-1mg/ml membrane-impermeable EZ-Link sulfo-NHS-LC biotin (Pierce) in PBS at 4°C for 30min with gentle agitation. For a 6cm dish, 1ml of biotin solution was used; this volume was scaled up or down as appropriate for the surface area of the culture dish or well. To quench the biotinylation reaction, the biotin solution was removed and cells were incubated in 50mM Tris pH8.0 for 10 minutes at room temperature with gentle agitation. Cells were then rinsed twice in PBS; the PBS used was centrifuged to retrieve any cells that detached during the rinsing process. Cells were lysed in either ice-cold NP-40 lysis buffer (50mM Tris pH 7.4, 150mM NaCl, 1% (v/v) NP-40) or ice-cold RIPA buffer (150mM NaCl, 50mM Tris pH 8.0, 1% (v/v) Triton X-100, 0.5% (w/v) sodium deoxycholate, and 1mM EGTA) for 30min on ice (500µl per 6cm dish). Lysed cells were scraped off the dish, triturated by passage through a 20-23G needle, and centrifuged at 13,000rpm for 10 minutes to remove

cell debris. Samples of the lysate were taken for subsequent analysis of protein content using Western blot. Proteins were then immunoprecipitated from the lysate as described in section 2.6.2. To enrich for total biotinylated protein, High-Capacity NeutrAvidin Agarose Resin (Pierce) was used for affinity purification (AP). Immunoprecipitation, affinity purification and lysate samples were analysed by SDS-PAGE and Western blot (sections 2.5.2 and 2.5.4).

2.6.4. Immunoaffinity purification (IAP) from mouse tissue

Mouse tissue (brain or lung) was homogenised in 10ml of either ice-cold RIPA buffer (150mM NaCl, 50mM Tris pH8.0, 1% (v/v) Triton X-100, 0.5% (w/v) sodium deoxycholate, and 1mM EGTA) or ice-cold digitonin buffer (150mM NaCl, 50mM Tris pH8.0, 1% (w/v) digitonin (Merck Chemicals), and 1mM EGTA) containing EDTA-free protease inhibitors (Roche) and 1mM sodium orthovanadate. RIPA buffer extractions were performed in 10ml of buffer per 2 mouse brains, while digitonin buffer extractions were performed in 10ml of buffer per 1 mouse brain. Homogenisation was carried out using a Polytron PT 3100 bench-top homogeniser (Kinematica). The homogenates were incubated on ice for 30min, volume adjusted to 12ml with either RIPA buffer or dilution buffer (150mM NaCl, 50mM Tris pH8.0, and 1mM EGTA) for RIPA-lysed and digitonin buffer-lysed samples respectively, and transferred to 13.2ml Thinwall polypropylene tubes (Beckman Coulter). To remove cell debris and clarify the samples, homogenates were centrifuged for 30min at 25000rpm (100000g) at 4°C in an ultracentrifuge using a SW41Ti rotor (Beckman Coulter). Samples of clarified homogenates were taken, and the remainder pre-cleared for 3 hrs at 4°C with rotation on Protein A-agarose or Protein G-Sepharose beads equilibrated in RIPA buffer/dilution buffer, depending on the antibody to be used for the IAP. Approximately 200µl of packed beads per brain or lung were used for pre-clearing. The pre-cleared homogenate was then divided into individual samples and incubated with 60µl dilution buffer-equilibrated packed antibody beads per brain or lung overnight at 4°C under rotation.

A sample of the unbound fraction was taken, and the beads were then washed three times in either RIPA buffer or dilution buffer depending on the original lysis buffer used. Beads were then transferred to 1.5ml screwcap microcentrifuge tubes, and washed once more.

To elute bound proteins, two sequential elutions in 60µl each of 2x lithium dodecyl sulphate (LDS) sample buffer (Life Technologies) or 2x Laemmli sample buffer were performed:

1. Beads incubated in sample buffer without reducing agent at 65°C for 30min, then boiled at 95°C for 5min before being cooled on ice. The supernatant containing eluted proteins was removed to a sterile microcentrifuge tube, and DTT was added to a final concentration of 50mM. The elution with added DTT was then boiled again at 95°C for 5min. This was elution 1.
2. Beads boiled in sample buffer containing 50mM DTT at 95°C for 5min. This was elution 2.

For one immunoaffinity purification (IAP) experiment, proteins were eluted sequentially in 60µl Rapigest SF elution buffer (Waters) as follows:

1. Beads incubated with 60µl Rapigest at 65°C for 35min, then boiled at 95°C for 5min. Supernatant removed to a sterile tube as elution 1.
2. Beads boiled in 60µl Rapigest at 95°C for 5min. Supernatant removed to a sterile tube as elution 2.
3. Beads boiled in 60µl 2x Laemmli sample buffer with 50mM DTT at 95°C for 5min. This was elution 3.

For each IAP elution, an aliquot was analysed by SDS-PAGE and Western blot as described in sections 2.5.2 and 2.5.4.

2.6.5. IAP sample preparation for mass spectrometry

If an IAP sample was to be analysed by mass spectrometry (MS), a 40µl aliquot of elution 1 (section 2.6.4) was resolved by electrophoresis on a 4-12% gradient NuPAGE® Novex Bis-Tris gel (Life Technologies). Gels were run at 150V in 1x MOPS Running Buffer (Life Technologies), using the XCell *Surelock*™ Mini-Cell system (Life Technologies). If the entire IAP sample was to be analysed by MS, the gel was electrophoresed for 12-15min. If specific protein bands were to be excised and analysed by MS, the gel was run for 90min. After electrophoresis, gels were fixed by incubation at room temperature for 15min in 50% (v/v) methanol and 7% (v/v) acetic acid diluted with molecular biology-grade water. Gels were then washed twice for 5min per wash in molecular biology-grade water. To stain for proteins, gels were incubated at room temperature in colloidal Coomassie blue stain, either Gelcode Blue Stain (Pierce) for 2h or InstantBlue (Expedeon) for 1h. If Gelcode Blue Stain was used, the gel was then destained in molecular biology-grade water. The stained gels were imaged with the Odyssey® Infrared Imaging System (LI-COR Biosciences) in the 700nm channel. Bands were then excised with clean scalpel blades and placed in sterile 1.5ml screwcap microcentrifuge tubes to be sent for MS analysis.

2.6.6. Protein identification by mass spectrometry

Excised gel plugs were sent to The Functional Genomics and Proteomics Laboratories in the School of Biosciences at the University of Birmingham for processing. There, the samples were first digested using automated in-gel trypsin digest with a Qiagen 3000 robot. The resultant peptides were extracted into 1% formic acid, and size separated through a 2-dimensional nano-liquid chromatography system (Dionex). Peptides were then ionised and sprayed into a Velos Orbitrap mass spectrometer (Thermo Fisher) using the Triversa Nanomate (Advion) chip-based electrospray system. The Velos Orbitrap is a hybrid mass spectrometer containing a linear ion trap and Orbitrap analyser (Eliuk and Makarov, 2015).

Output mass spectra were analysed using the SEQUEST algorithm as implemented in the Proteome Discoverer software suite (Thermo Fisher) to correlate mass spectra with protein identities (Eng et al., 1994, Yates et al., 1995). Excel pivot tables of assigned spectra that were exported from Proteome Discoverer were used in further cross-sample comparison and filtration as described in Chapter 3.

2.7. General bioinformatics

Gene and protein sequences were obtained from four databases: NCBI GenBank and RefSeq, Ensembl, UniProtKB/Swiss-Prot, and the Leiden Open Variation Database (LOVD) (Brown et al., 2015, Flicek et al., 2014, Fokkema et al., 2011, Harrow et al., 2012). General DNA and protein sequence similarity searches were carried out using the BLAST (Basic Local Alignment Search Tool) suite of programs (Altschul et al., 1990, Altschul et al., 1997, Camacho et al., 2009). To compare sequences between species, the BLAT (BLAST-Like Alignment Tool) program was used (Kent, 2002). Sequences were aligned using Clustal Omega (Sievers et al., 2011). The RepeatMasker program web server with the cross_match search engine was used to identify sequences derived from transposable elements (Smit et al., 2013-2015). The Swiss Institute of Bioinformatics ExPASy (Expert Protein Analysis System) proteomics portal was used for DNA to protein translation (Artimo et al., 2012). To identify translation initiation codons, ATGpr and NetStart servers were used (Nishikawa et al., 2000, Pedersen and Nielsen, 1997, Salamov et al., 1998). The NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/) was used to predict N-linked glycosylation sites, while signal sequences were predicted using SignalP 4.1 (Petersen et al., 2011). Venn diagrams were produced using the Venn diagram tool from the Bioinformatics and Systems Biology server at Ghent University was used (http://bioinformatics.psb.ugent.be/webtools/Venn/).

Chapter 3: ε-SG protein interactions in the brain

3.1. Introduction

While the role of *SGCE* mutations in the genetic aetiology of MD is well-established, little is known about the function of ε-SG in the brain. As discussed in Chapter 1 (section 1.4.2), complexes containing ε-SG have been described in cardiac and smooth muscle, adipose tissue and peripheral nerve Schwann cells (Cai et al., 2007, Durbeej and Campbell, 1999, Groh et al., 2009, Imamura et al., 2000, Lancioni et al., 2011, Liu and Engvall, 1999, Straub et al., 1999). In these tissues, ε-SG is part of prototypical εβδγ or εβδζ heterotetramers that associate with other DGC components (Cai et al., 2007, Durbeej and Campbell, 1999, Groh et al., 2009, Imamura et al., 2000, Lancioni et al., 2011, Liu and Engvall, 1999, Straub et al., 1999). Despite ε-SG's broad expression and integration into muscle sarcoglycan complexes, *SGCE* mutation-positive MD patients have no evident muscle pathology (Asmus et al., 2002, Durbeej and Campbell, 1999, Hjermind et al., 2008, Straub et al., 1999). By contrast, no MD-like features have been described in LGMD patients with sarcoglycan mutations (Hjermind et al., 2008). Several explanations have been proposed for this apparent paradox. For example, the early onset and severe muscle phenotype of LGMD could mask any movement disorder features in patients with sarcoglycan mutations. Alternatively, ε-SG in the brain may have a function independent of the DGC. Alternative splicing of *SGCE* pre-mRNA results in brain-specific ε-SG isoforms, and a putative unique function of the brain-specific ε-SG isoform 2 has been suggested to underlie the purely neurological phenotype of MD (Nishiyama et al., 2004, Ritz et al., 2011, Yokoi et al., 2005). To gain further insight into the role of ε-SG in the brain and thus in the pathogenesis of MD, I used immunoaffinity purification (IAP) followed by mass spectrometry (MS) to identify proteins interacting with ubiquitous isoform 1 and brain-specific isoform 2 of ε-SG in the brain.

3.2. Results

IAP was used to identify proteins interacting with ε-SG in the brain because it enabled unbiased identification of protein interactions from native protein complexes in physiologically relevant tissues. Other approaches to defining protein-protein interactions rely on prior knowledge of probable interactions (co-immunoprecipitation, fluorescence resonance energy transfer), use non-native environments to identify interactions (yeast 2-hybrid), or require protein overexpression that can lead to false positives through excessive protein abundance (tandem affinity purification) (Carneiro et al., 2016). By contrast, IAP using validated anti-ε-SG antibodies allowed isolation of native ε-SG-containing complexes directly from brain tissue. MS analysis of the purified ε-SG-containing complexes was then used to identify the proteins in those complexes. In order to determine whether the brain-specific ε-SG isoform 2 was part of the same protein complexes as the ubiquitous isoform 1, an antibody that binds only the brain-specific ε-SG isoform 2 was used for IAP along with antibodies against all ε-SG isoforms.

3.2.1. Generation and validation of the esg2-1358 antibody specific to ε-SG isoform 2

To specifically detect ε-SG isoform 2, polyclonal antibodies against the unique C-terminal peptide of human ε-SG isoform 2 were raised in rabbits (Chapter 2 section 2.3.1 and 2.3.4). Briefly, two rabbits were immunised with a peptide antigen (Figure 3.1), and polyclonal antibodies were isolated from terminal bleed serum by peptide affinity chromatography. Initial evaluation of these antibodies indicated that neither was useful for Western blotting, while the esg2-1358 antibody performed better for immunoprecipitation (IP) and was therefore used in preference to the esg2-1355 antibody in subsequent experiments (data not shown).

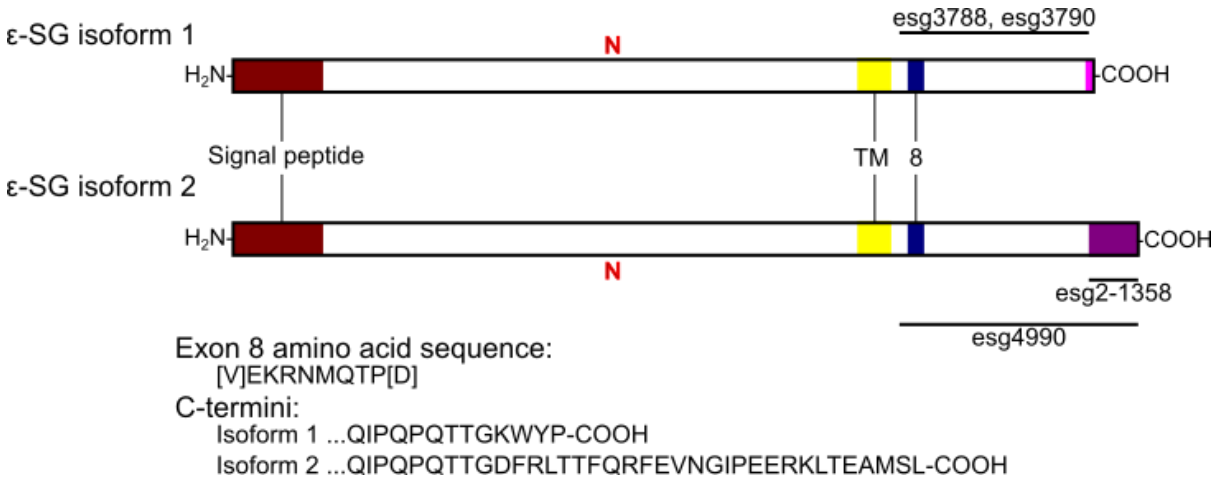


Figure 3.1 Major features of ε-SG isoforms 1 and 2, and antibody antigens. The polypeptides for ε-SG isoforms 1 and 2 are shown, with the signal sequence marked in red, transmembrane domain (TM) marked in yellow, region encoded by exon 8 marked in blue (8), N-linked glycosylation sites marked with N and the unique C-terminus sequences of isoforms 1 and 2 marked in pink and purple respectively. The amino acid sequences of the exon 8-encoded region and the C-termini for the isoforms are provided below. Brackets enclose residues encoded at splice sites. Mouse protein immunogens used to raise antibodies used in this study are indicated. Abbreviations: TM, transmembrane domain; 8, peptide encoded by exon 8.

To determine the specificity of esg2-1358 in immunoprecipitation, this antibody was used to perform an IAP of ε-SG from RIPA-solubilised mouse whole brain tissue in parallel with the previously characterised esg4990 antibody (Chapter 2 section 2.6.4, Table 2.2) (Waite et al., 2011). Esg4990 was raised against the entire C-terminus of ε-SG isoform 2 and detects all ε-SG isoforms (Figure 3.1). As predicted, the esg4990 antibody enriched ε-SG isoform 1, ε-SG isoform 2 exon 8+ and ε-SG isoform 2 exon 8- (Figure 3.2). By contrast, esg2-1358 enriched only ε-SG isoform 2 exon 8+ and ε-SG isoform 2 exon 8- (Figure 3.2). This confirmed specificity of esg2-1358 to ε-SG isoform 2 and validated its use for IAP.

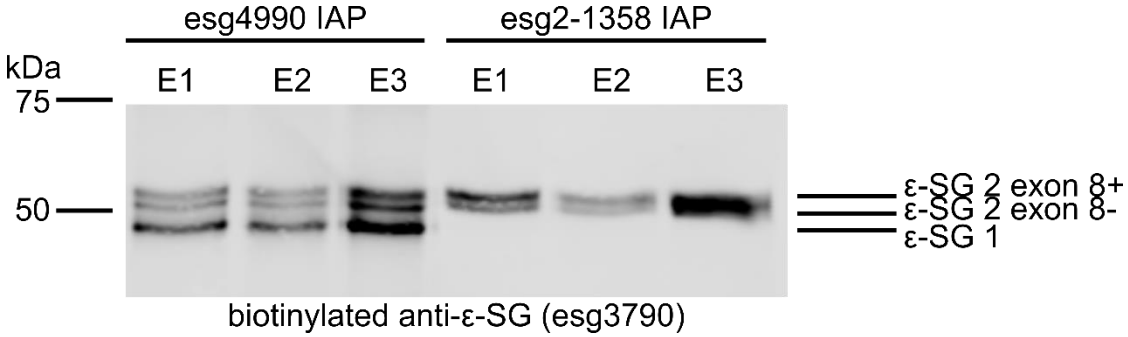


Figure 3.2 IAP of endogenous ε-SG from mouse whole brain using two distinct anti-ε-SG antibodies. The pan-ε-SG antibody esg4990 binding all major isoforms of ε-SG and the ε-SG isoform 2-specific esg2-1358 antibody were used to enrich ε-SG from RIPA-lysed mouse whole brain through IAP. Three sequential elutions were performed after IAP to ensure elution of all ε-SG from the protein A-agarose-conjugated antibodies. Immunoblot of IAP elutions using the biotinylated pan-ε-SG antibody esg3790 (see Figure 3.1) demonstrated that three distinct ε-SG proteins were present in the esg4990 IAP; these were ε-SG isoform 1, ε-SG isoform 2 exon 8-, and ε-SG isoform 2 exon 8+ from smallest to largest. By contrast just two ε-SG bands were detected in the esg2-1358 IAP, corresponding to ε-SG isoform 2 exon 8- and ε-SG isoform 2 exon 8+ only. No ε-SG isoform 1 could be detected in any esg2-1358 IAP elution, indicating that the esg2-1358 antibody specifically isolated ε-SG isoform 2. Abbreviations: kDa, kilodaltons; E1, IAP elution 1; E2, IAP elution 2; E3, IAP elution 3; ε-SG 2, ε-SG isoform 2; ε-SG 1, ε-SG isoform 1.

3.2.2. Identification of a brain sarcoglycan complex containing ε-SG isoform 2

3.2.2.1. ε-SG isoform 2 co-purifies with β-SG, δ-SG and ζ-SG in mouse brain

The esg2-1358 and esg2-1355 antibodies were conjugated to Protein A-agarose beads and used for IAP of ε-SG isoform 2-containing complexes from RIPA-solubilised wild-type mouse whole brain as previously described (Chapter 2 section 2.6.4) (Waite et al., 2016). A control IAP using non-specific immunoglobulin derived from an immunogen-naïve rabbit (pre-immune immunoglobulin) conjugated to protein A-agarose beads was also performed in parallel to identify non-specific interactions with immunoglobulin and the Protein A-agarose beads. Western blot confirmed enrichment of ε-SG isoform 2 only in IAPs from mouse brain tissue using the esg2-1355 and esg2-1358 antibodies but not in the pre-immune immunoglobulin IAP (Figure 3.3a). ε-SG isoform 1 running at approximately 48kDa (Figure 3.2) was not detected in any of these IAPs. Importantly, these data also indicated that ε-SG isoforms 1 and 2 do not co-purify in mouse brain and therefore form separate protein complexes. Subsequent IAPs using the esg2-1358 antibody (sections 3.2.4, 3.2.5) confirmed this result.

Having shown enrichment of ε-SG isoform 2 in the esg2-1358 IAPs, samples were analysed by MS to identify proteins co-purifying with ε-SG isoform 2. To prepare samples for MS analysis, the IAPs were resolved by polyacrylamide gel electrophoresis, fixed and stained with colloidal Coomassie blue (Chapter 2 section 2.6.5). Eight protein bands covering a range of sizes including the 25-55kDa region predicted to contain all sarcoglycans were excised from the esg2-1358 IAP as indicated (Figure 3.3b). These were analysed by MS at the University of Birmingham Functional Genomics and Proteomics Laboratories (Chapter 2 section 2.6.6). After filtering out common contaminants such as keratins and immunoglobulins, the list of protein identities in each band were ranked by Proteome Discoverer protein identity score as shown in Table 3.1 (Trinkle-Mulcahy et al., 2008).

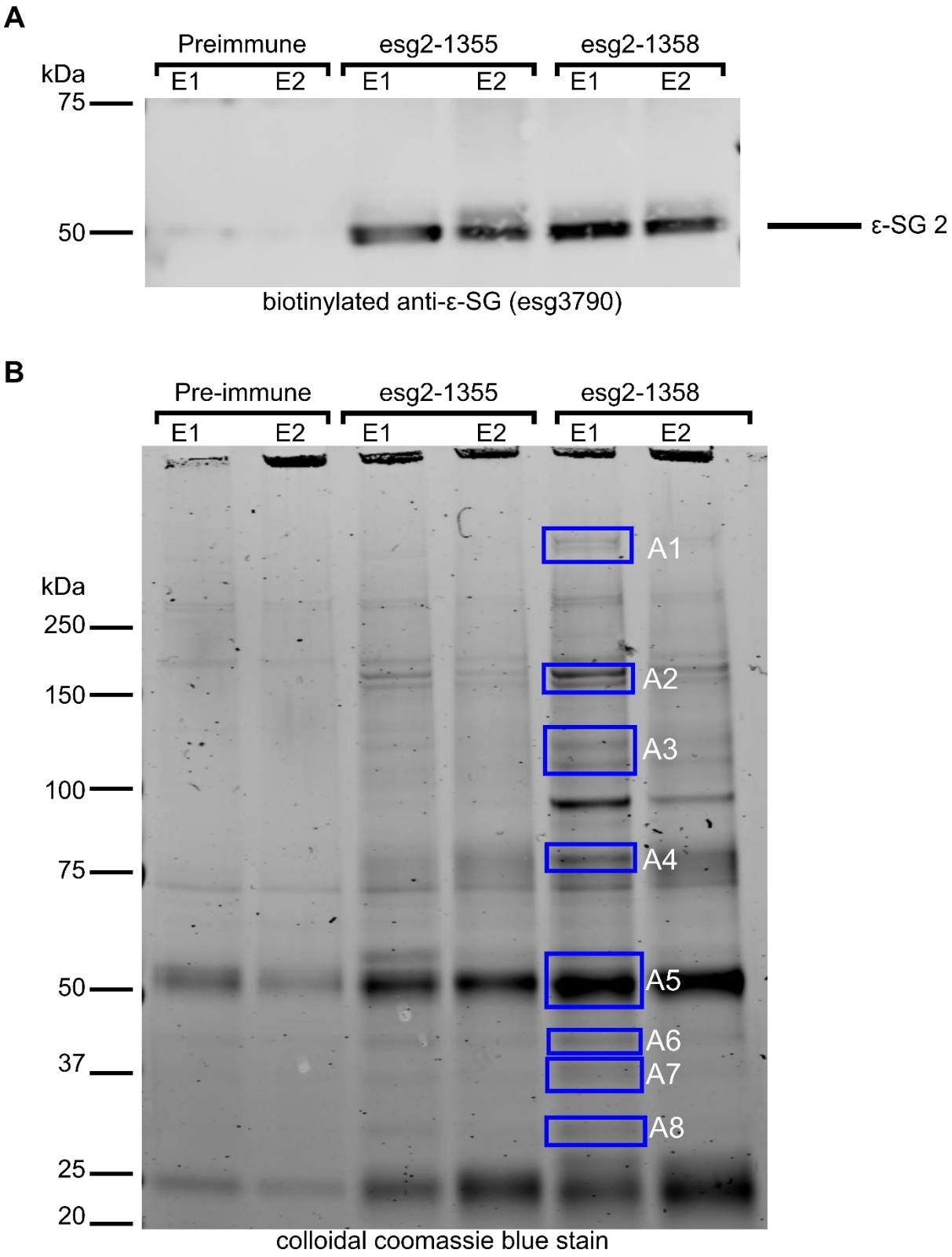


Figure 3.3 IAP of complexes containing ε-SG isoform 2 from whole mouse brain lysed in RIPA buffer. Protein complexes were enriched from mouse whole brain lysed in RIPA buffer using pre-immune IgG as a control, or the ε-SG isoform 2-specific antibodies esg2-1355 and esg2-1358 to enrich for complexes containing ε-SG isoform 2. A) Immunoblot analysis of the IAP elutions using a biotinylated pan-ε-SG antibody confirmed enrichment of ε-SG isoform 2 in the esg2-1355 and esg2-1358 IAPs but not in the pre-immune IgG control IAP. B) IAP elutions were resolved via PAGE, fixed, and stained for proteins using colloidal Coomassie blue. Bands A1 through A8 (blue rectangles) were excised and analysed by MS. Abbreviations: kDa, kilodaltons; Pre-immune, IAP using pre-immune IgG; esg2-1355, IAP using esg2-1355; esg2-1358, IAP using esg2-1358; E1, IAP elution 1; E2, IAP elution 2.

Band	Gene	Description	Protein identity score	Number of unique peptides	Spectral count
A1	Prss1	protease, serine 1	124.34	2	21
	Jup	Junction plakoglobin	74.73	2	5
A2	Map3k5	mitogen activated protein kinase kinase kinase 5	2924.17	48	232
	Rab11fip5	rab11 family-interacting protein 5 isoform 1	845.64	25	50
	Uhrf1bp1l	UHRF1-binding protein 1-like	690.12	22	44
	Map3k15	mitogen-activated protein kinase kinase kinase 15	369.35	11	48
	Tnr	Tenascin R	339.79	12	16
	Clasp2	CLIP-associating protein 2	300.99	10	14
	Tjp1	Tight junction protein ZO-1	187.13	10	14
	Nefh	Neurofilament heavy polypeptide	184.4	7	9
	Kiaa1109	Uncharacterized protein KIAA1109	174.73	6	11
	Kcna3	Potassium voltage-gated channel subfamily A member 3	167.48	3	15
	Dgkh	Diacylglycerol kinase eta	160.09	5	10
	Nsf	Vesicle-fusing ATPase	151.5	6	13
	Map3k6	Mitogen-activated protein kinase kinase kinase 6	139.38	9	21
	Atp1a3	Sodium/potassium-transporting ATPase subunit alpha-3	122.8	1	4
	Ampd2	AMP deaminase 2	113.16	5	12
	Kcna6	Potassium voltage-gated channel subfamily A member 6	105.93	2	9
	Atp1a1	Sodium/potassium-transporting ATPase subunit alpha-1	103.09	1	3
	Kcna2	Potassium voltage-gated channel subfamily A member 2	99.98	1	9
	Cadps	Calcium-dependent secretion activator 1	97.49	4	5
	Ccdc136	coiled-coil domain-containing protein 136 isoform 2	83.49	2	2
	Prss1	protease, serine 1	80.49	1	10
	Fmnl1	Formin-like protein 1	80.04	1	7
	Tjp2	Tight junction protein ZO-2	74.33	2	9
	Slc1a2	Excitatory amino acid transporter 2	72.01	3	7
	Wdr7	WD repeat-containing protein 7	71.74	3	4

A3	Tpd52l2	tumor protein D52-like 2	746.33	3	95
	C1orf43	Uncharacterized protein C1orf43 homolog	292.74	8	28
	Rab34	Ras-related protein Rab-34	249.92	7	17
	Fam26e	Protein FAM26E	141.07	3	15
	Atp6v1d	V-type proton ATPase subunit D	140.89	3	4
	Tppp	Tubulin polymerization-promoting protein	129.51	3	4
	Vapa	Vesicle-associated membrane protein-associated protein A	118.11	3	8
	Hsd17b8	Estradiol 17-beta-dehydrogenase 8	114.97	3	3
	Impa1	Inositol monophosphatase 1	113.23	4	5
	Ppp3ca	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform	111.93	4	8
	Prss1	protease, serine 1	109.92	1	24
	Tpd52	Tumor protein D52	108.4	2	3
	Hsd17b12	Very-long-chain 3-oxoacyl-CoA reductase	106.57	2	3
	Rbm15	Putative RNA-binding protein 15	104.98	4	8
	Atp6v1e1	V-type proton ATPase subunit E 1	97.59	4	6
	Fhl3	Four and a half LIM domains protein 3	87.97	1	2
	Vapb	Vesicle-associated membrane protein-associated protein B	85.89	2	6
	Ampd2	AMP deaminase 2	76.87	3	7
A5	Flot1	Flotillin-1	1170.32	22	47
	Mb21d2	Protein MB21D2	920.69	16	54
	Scrn1	Secernin-1	524.14	13	29
	Trim21	E3 ubiquitin-protein ligase TRIM21	387.62	11	22
	Eno1	Alpha-enolase	328.32	6	16
	Kcna6	Potassium voltage-gated channel subfamily A member 6	297.08	8	19
	Kcna3	Potassium voltage-gated channel subfamily A member 3	280.61	7	18
	Ap2m1	AP-2 complex subunit mu	258.02	11	17
	Camk2a	Calcium/calmodulin-dependent protein kinase type II subunit alpha	254.79	7	12
	Sept7	Septin-7	240.56	6	12
	Ccdc136	Coiled-coil domain-containing protein 136	238.14	8	12

	Blmh	Bleomycin hydrolase	195.47	5	8
	Eno2	Gamma-enolase	175.73	2	8
	Sept6	Septin-6	156.53	4	9
	Sept11	Septin-11	145.54	2	8
	Btbd17	BTB/POZ domain-containing protein 17	121.32	2	3
	Clcc1	Chloride channel CLIC-like protein 1	117.78	2	3
	Ppp2r2a	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform	114.83	2	3
	Endod1	Endonuclease domain-containing 1 protein	113.64	5	6
	Slc1a3	Excitatory amino acid transporter 1	104.79	3	6
	Ckb	Creatine kinase B-type	98.4	2	2
	Gda	Guanine deaminase	90.75	2	3
	Ina	Alpha-internexin	88.49	3	7
	Ampd2	AMP deaminase 2	88.12	6	8
	Dctn2	Dynactin subunit 2	87.36	2	3
	Atp1b1	Sodium/potassium-transporting ATPase subunit beta-1	87.3	4	6
	Rufy3	Protein RUFY3	78.11	2	4
	Map3k5	mitogen activated protein kinase kinase kinase 5	73.13	3	4
	Samm50	Sorting and assembly machinery component 50 homolog	69.28	3	6
	Rbm15	Putative RNA-binding protein 15	63.06	2	4
	Slc1a2	Excitatory amino acid transporter 2	54.73	2	3
	Atp6v1h	V-type proton ATPase subunit H	54.71	3	4
	Nsf	Vesicle-fusing ATPase	54.12	2	2
	Ctbp1	C-terminal binding protein 1	51.3	3	4
	Psmc3	26S protease regulatory subunit 6A	49.32	2	2
	Fam26e	Protein FAM26E	46.01	2	3
	Sh3glb2	Endophilin-B2	45.92	3	5
	Fscn1	Fascin	45.58	2	2
	Mapt	Microtubule-associated protein tau	45.47	3	4

	Serbp1	Plasminogen activator inhibitor 1 RNA-binding protein	44.17	2	3
	Sgce	Epsilon sarcoglycan, isoform 2	302.64	8	24
	Sgce	Epsilon sarcoglycan	276.02	10	25
	Prss1	protease, serine 1	141.38	1	16
	Atp6v1b2	V-type proton ATPase subunit B, brain isoform	70.58	1	1
	Pdia6	Protein disulfide-isomerase A6	70.03	1	1
A6	Glul	Glutamine synthetase	401.75	14	33
	Ckb	Creatine kinase B-type	318.23	9	21
	Kcnab1	Voltage-gated potassium channel subunit beta-1	307.56	8	23
	Mbip	MAP3K12-binding inhibitory protein 1	301.03	10	22
	Map2k1	Dual specificity mitogen-activated protein kinase kinase 1	282.89	7	23
	Jup	Junction plakoglobin	260.28	8	14
	Kcnab3	Voltage-gated potassium channel subunit beta-3	252.87	4	8
	Sept5	Septin-5	221.05	7	17
	Erlin2	Erlin-2	189.13	6	8
	Aldoa	Fructose-bisphosphate aldolase A	178.69	8	17
	Actr1b	Beta-centractin	158.81	2	8
	Actr1a	Alpha-centractin	157.55	3	12
	Aldoc	Fructose-bisphosphate aldolase C	150.56	2	6
	Pdcd6	Programmed cell death protein 6	137.33	4	8
	Serpinb5	Serpin B5	132.15	4	6
	Sept2	Septin-2	119.81	4	5
	Eppk1	Epiplakin	116.33	2	14
	Prss1	protease, serine 1	102.95	1	19
	Pcmtd1	protein-L-isoaspartate O-methyltransferase domain-containing protein 1	101.11	2	5
	Actr2	Actin-related protein 2	99.82	3	7
	Sh3gl2	Endophilin-A1	98.08	2	4
	Vcp	Transitional endoplasmic reticulum ATPase	97.54	4	5
	Eno1	Alpha-enolase	93.79	4	6

	Evpl	Envoplakin	92.09	3	9
	Got1	Aspartate aminotransferase, cytoplasmic	89.08	3	4
	Kcnab2	Voltage-gated potassium channel subunit beta-2	83.51	2	7
	Pkp1	Plakophilin-1	82.82	1	2
	Sept3	Neuronal-specific septin-3	80.23	2	6
	Ampd2	adenosine monophosphate deaminase 2	77.11	3	8
	Clcc1	chloride channel CLIC-like protein 1	76.21	2	3
	Sgcb	Beta sarcoglycan	76.08	4	6
	Atp1b1	Sodium/potassium-transporting ATPase subunit beta-1	74.66	2	2
	Nccrp1	F-box only protein 50	74.64	1	1
	Pdia6	Protein disulfide-isomerase A6	73.74	1	1
	Pkm	Pyruvate kinase PKM	71.06	2	3
A7	Kcnab2	Voltage-gated potassium channel subunit beta-2	707.42	15	92
	Gnao1	Guanine nucleotide-binding protein G(o) subunit alpha	370.96	9	23
	Kcnab1	Voltage-gated potassium channel subunit beta-1	271.04	4	48
	Aldoa	Fructose-bisphosphate aldolase A	202.2	3	7
	Clcc1	chloride channel CLIC-like protein 1	172.36	8	13
	Gnai2	Guanine nucleotide-binding protein G(i) subunit alpha-2	170.71	3	11
	Aldoc	Aldolase C, fructose-bisphosphate	156.58	5	9
	Capza2	F-actin-capping protein subunit alpha-2	152.41	7	15
	Acot7	Cytosolic acyl coenzyme A thioester hydrolase	146.95	4	6
	Mapk1	mitogen-activated protein kinase 1	134.84	9	15
	Bub3	Mitotic checkpoint protein BUB3	130.83	6	11
	Tpd52l2	tumor protein D52-like 2	130.42	4	7
	Rbm15	Putative RNA-binding protein 15	125.42	4	13
	Strap	Serine-threonine kinase receptor-associated protein	123.28	5	9
	Sirt2	NAD-dependent protein deacetylase sirtuin-2	123.23	6	7
	Phyhip	Phytanoyl-CoA hydroxylase-interacting protein	119.09	5	7
	Psat1	Phosphoserine aminotransferase	112.24	4	6

	Glul	Glutamine synthetase	111.24	4	8
	Ppp1ca	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	111.01	1	6
	Rpsa	40S ribosomal protein SA	110.73	3	6
	Gnai1	guanine nucleotide binding protein (G protein), alpha inhibiting 1	105.72	2	9
	Ppp1cc	Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	105.14	2	7
	Taldo1	transaldolase	104.9	4	9
	Ppp1cb	Serine/threonine-protein phosphatase PP1-beta catalytic subunit	103.02	1	5
	Atp6v0d1	V-type proton ATPase subunit d 1	101.72	6	11
	Syp	Synaptophysin	101.26	3	8
	Mdh1	Malate dehydrogenase, cytoplasmic	100.6	4	4
	Sgcd	Delta sarcoglycan	96.54	4	5
	Dcaf7	DDB1- and CUL4-associated factor 7	92.72	4	4
	Dnajb2	DnaJ homolog subfamily B member 2	91.46	4	4
	Fbxo2	F-box only protein 2	91.14	1	1
	Prss1	protease, serine 1	90.88	1	14
	Sgcz	Zeta sarcoglycan	88.3	1	1
	Blvra	Biliverdin reductase A	87.15	3	6
	Lancl1	LanC-like protein 1	85.95	4	10
	Clu	Clusterin	83.97	4	5
	Kcna3	Potassium voltage-gated channel subfamily A member 3	82.6	3	6
	Gpd1	Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic	82.26	4	6
	Gnb5	Guanine nucleotide-binding protein subunit beta-5	81.22	3	4
	Gnb1	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	78.41	4	6
	Gpd1l	Glycerol-3-phosphate dehydrogenase 1-like protein	77.59	2	8
	PPP2CB	Serine/threonine-protein phosphatase 2A catalytic subunit beta isoform	75.73	2	3
	Pcmtd2	Protein-L-isoaspartate O-methyltransferase domain-containing protein 2	75.59	4	5
	Ampd2	Adenosine monophosphate deaminase 2	74.92	5	10
	Zadh2	Zinc-binding alcohol dehydrogenase domain-containing protein 2	73.2	2	3
	Fam26e	Protein FAM26E	72.3	2	6

	Napg	Gamma-soluble NSF attachment protein	70.74	2	6
A8	Tpd52l2	tumor protein D52-like 2	395.85	3	26
	Phb	Prohibitin	192.65	7	13
	Fam26e	Protein FAM26E	179.44	6	24
	Vdac2	Voltage-dependent anion-selective channel protein 2	132.15	7	14
	Tppp	Tubulin polymerization-promoting protein	115.34	3	5
	Map1b	Microtubule-associated protein 1B	110.15	3	6
	Napa	Alpha-soluble NSF attachment protein	102.36	3	8
	Cops7a	COP9 signalosome complex subunit 7a	96.18	2	2
	Vdac3	Voltage-dependent anion-selective channel protein 3	91.67	5	14
	Clta	Clathrin light chain A	91.17	4	6
	Rack1	Receptor of activated protein C kinase 1	91.03	6	7
	Atp6v1e1	V-type proton ATPase subunit E 1	82.03	4	5
	Vdac1	Voltage-dependent anion-selective channel protein 1	77.65	3	6
	Vapb	Vesicle-associated membrane protein-associated protein B	76.28	1	4
	Fhl2	Four and a half LIM domains protein 2	75.42	4	11
	Rogdi	Protein rogdi homolog	73.49	2	5

Table 3.1 Proteins identified by MS in the esg2-1358 IAP from RIPA-lysed mouse whole brain. The esg2-1358 antibody was used for IAP of ε-SG isoform 2 from RIPA-lysed mouse whole brain tissue. This table summarises the proteins identified in the IAP via mass spectrometry, ordered by the gel band in which the protein was identified and the protein identity score. Protein identity score is calculated by the SEQUEST algorithm implemented in Proteome Discover, and indicates the likelihood that a given protein was present in the sample. Also provided are the gene symbol, number of unique peptides corresponding to that particular protein and accession number detected in the sample, and the spectral count associated with that protein. Spectral count represents the number of mass spectra detected in the sample that corresponded to that particular protein, and can reflect abundance. Band A4 contained only immunoglobulins and contaminants. The sarcoglycans are highlighted in green, TnR is highlighted in blue, and the Shaker-related voltage-gated potassium channel subunits are highlighted in orange.

Gene	Uniprot acc.	Gel band	Unique peptide sequences	Identity score	Total peptides	Spectral count	Protein coverage
SGCE	O70258	A5	NVYPSAGVLFVHVLER QVSTYQEVVR GEGILPDGGEYKPPSDSLK NMNVEEMLASEVLGDFLGAVK THFHIDWCK **FEVNGIPEER** VPLPINDMK TPYSDGVLYGSPTAENVGKPTIIEITAYNRR EVENPQNQLR **KLTEAMSL** NVWQPER DMQTPDIQLVHHSSIQK	302.64	23	49	35.13%
SGCB	P82349	A6	AAAAAAAAATEQQGSNGPVKK RNENLVITGNNQPIVFQQGTTK THNILFSTDYETHEFHLPSGVK TSITSDIGMQFFDPR	76.08	4	6	25%
SGCD	P82347	A7	LLFSADDSEVVVGAER LEGDSEFLQPLYAK VLTQLVTGPK VLGAEGTVFPK SIETPNVR	96.54	5	5	20.42%
SGCZ	Q8BX51	A7	VLFSADEDEITIGAEK	88.3	1	1	6.87%

Table 3.2 Sarcoglycan peptides identified in the esg2-1358 IAP of ε-SG from whole CD-1 mouse brain lysed in RIPA buffer. For each sarcoglycan identified in the esg2-1358 IAP from RIPA-lysed mouse whole brain, all unique peptide sequences detected in the sample and the band in which they were present are given. The protein identity score provides the confidence in that protein's presence in the sample analysed by MS. The protein's Uniprot accession code, the number of total peptides detected, spectral count, and coverage by identified peptides (Protein coverage) are all also provided. Peptides in bold are specific to ε-SG isoform 2.

Only immunoglobulins and keratins were detected in protein band A4, but all other bands contained multiple protein matches. Peptides corresponding to ε-SG, β-SG, δ-SG and ζ-SG were all identified with high confidence in the IAP, including peptides derived from the unique ε-SG isoform 2 C-terminus (Table 3.2). No γ-SG peptides were identified, which was consistent with the reported low expression of γ-SG in the brain (Durbeej and Campbell, 1999, Noguchi et al., 1995, Noguchi et al., 2001, Waite et al., 2016). Co-purification of β-SG, δ-SG and ζ-SG with ε-SG isoform 2 suggested that ε-SG isoform 2 may be incorporated into a prototypical heterotetramer comprising εβδζ in the brain.

In addition to the sarcoglycans, Shaker-related voltage-gated K^+ channel subunits were identified in bands A2, A5, A6 and A7 (Table 3.1). These were later examined for a potential cross-reaction with the esg2-1358 antibody (section 3.2.4). Additionally, the extracellular matrix protein proteoglycan tenascin-R was identified in band A2 (Table 3.1). Subsequent IAP-MS experiments suggested this protein did indeed co-purify with ε-SG (section 3.2.5).

3.2.2.2. Alternative splicing of ε-SG does not affect formation and localisation of an εβδζ tetramer in heterologous cells

The IAP-MS experiment described above supported the existence of prototypical εβδζ sarcoglycan heterotetramers in the brain. Moreover, ε-SG isoforms 1 and 2 must be part of separate tetramers since isoform 1 did not co-purify with isoform 2 in the esg2-1358 IAP. Though both isoforms were part of prototypical heterotetramers, alternatively spliced ε-SG isoforms as described in Chapter 1 (section 1.3.2.2) could affect complex assembly and trafficking. Therefore, assembly and trafficking of the brain sarcoglycan complex with different ε-SG isoforms was examined in HEK293T cells co-transfected with different combinations of plasmids encoding epitope-tagged sarcoglycans. For these experiments, plasmids encoding human sarcoglycans were used since they had already been generated as

part of the study described in Chapter 4 and have high similarity to their mouse orthologues.

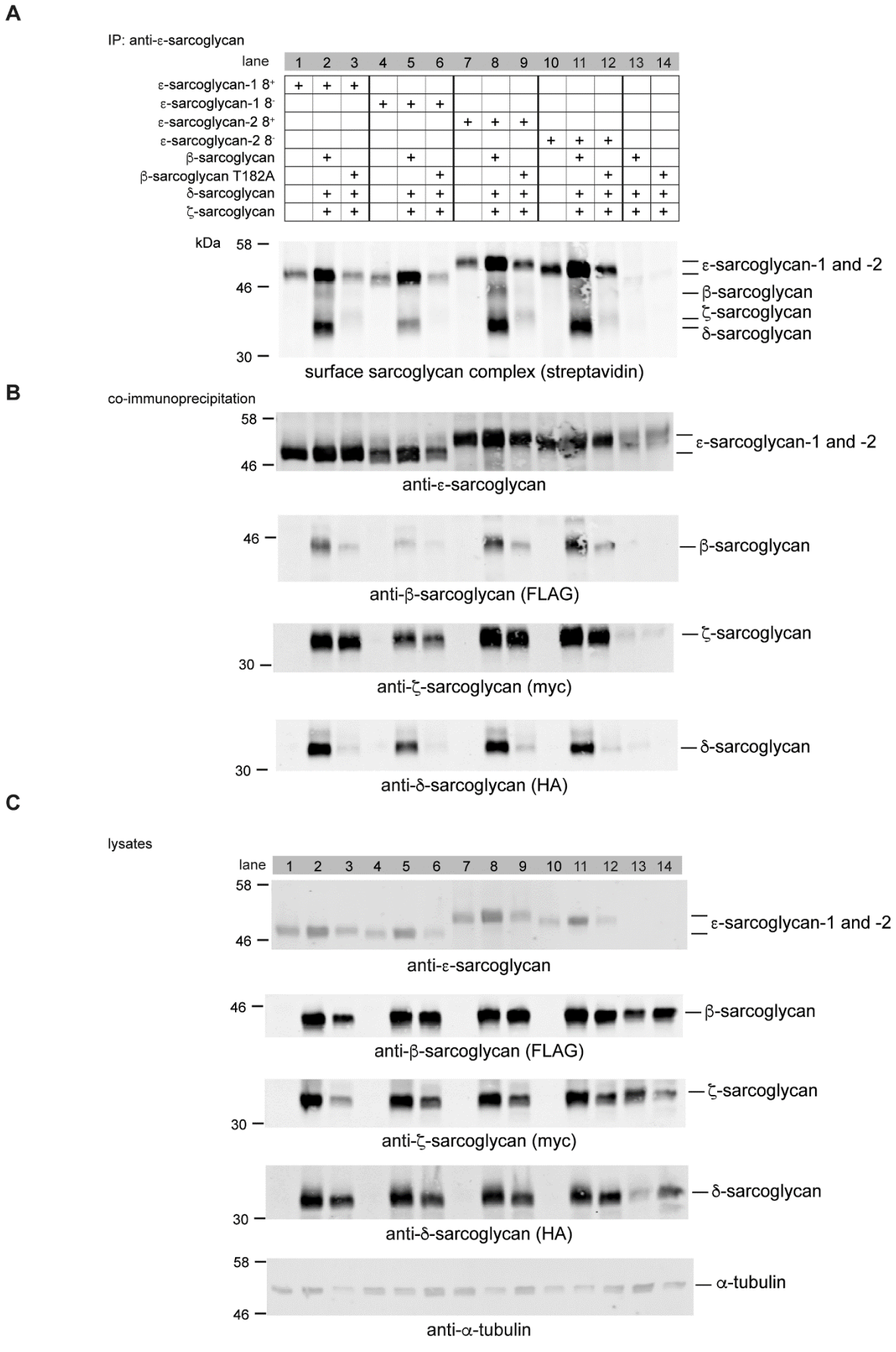


Figure 3.4 Membrane trafficking of the brain sarcoglycan complex *in vitro* with different ε-SG isoforms. HEK293T cells were transfected with different combinations of plasmids encoding epitope-tagged sarcoglycans as indicated, and cell surface proteins were biotinylated. ε-SG and associated proteins were immunoprecipitated using esg4990 antibody beads. For comparative purposes, data are presented in groups of three (lanes 1-12) corresponding to the four major brain ε-SG isoforms. Each group follows the same scheme; ε-SG isoform alone, ε-SG isoform in an εβδζ tetramer with wild-type β-SG, and ε-SG isoform in a tetramer with mutant β-SG. **A**) Biotinylated surface proteins immunoprecipitated by esg4990 were identified using streptavidin-Alexa Fluor 680. Although ε-SG isoforms trafficked to the cell surface in the absence of the other sarcoglycans (panel A, lanes 1, 4, 7 and 10), the highest levels of sarcoglycans at the membrane were observed in cells co-expressing a tetramer with wild-type β-SG (lanes 2, 5, 8 and 11). By contrast, levels of δ-SG and to a lesser extent ε- and ζ-SG were drastically reduced at the cell surface in cells expressing the LGMD2E-associated β-SG T182A mutant (lanes 3, 6, 9 and 12). **B**) Complex formation was demonstrated by co-immunoprecipitation for all four ε-SG isoforms. Robust co-immunoprecipitation of each sarcoglycan was observed for all heterotetramers that contain wild-type β-SG (lanes 2, 5, 8 and 11). By contrast, βδ-SG core formation was severely disrupted in cells expressing mutant β-SG compared to wild-type; however, the association of ε- and ζ-SG was apparently unaffected by the mutant (lanes 3, 6, 9 and 12). Lanes 13 and 14 shows that trimers formed in the absence of transfected ε-SG do not traffic to the cell surface. **C)** Whole cell RIPA lysates are shown for comparative purposes while α-tubulin was used as a loading control. Abbreviations: kDa, kilodaltons.

Transfected cells were treated with a membrane-impermeable biotin reagent to label cell surface sarcoglycan complexes only, and proteins were either immunoprecipitated using esg4990-Protein A agarose beads to enrich for ε-SG-containing complexes or affinity purified using NeutrAvidin resin to enrich for biotinylated proteins (Chapter 2 section 2.6.3). Immunoprecipitated proteins were analysed by Western blot. All experiments were repeated three times as previously described (Waite et al., 2016).

Western blotting with streptavidin-Alexa Fluor 680 to detect biotinylated proteins showed that all ε-SG isoforms trafficked to the cell surface both independently and with other sarcoglycans, though co-expression with other sarcoglycans increased cell surface ε-SG (Figures 3.4 and 3.5, lanes 2, 5, 8, 11; Figure 3.6). By contrast, β-SG and δ-SG were robustly detected at the cell surface only when co-expressed with ε-SG and ζ-SG (Figure 3.4a, b lanes 2, 5, 8 and 11). The low levels of endogenous expressed ε-SG in HEK293T cells were insufficient for membrane trafficking of β-, δ- and ζ-SG without exogenous ε-SG (Figure 3.4a lanes 13 and 14). These data recapitulate previous studies describing assembly of ε-SG into a membrane-associated sarcoglycan complex in heterologous cells, and highlight the importance of the βδ-SG interaction for trafficking of the sarcoglycan heterotetramer to the plasma membrane, despite ε-SG's ability to traffic independently (Chen et al., 2006, Draviam et al., 2006a, Esapa et al., 2007, Hack et al., 2000, Shi et al., 2004, Shiga et al., 2006). All four ε-SG isoforms tested co-immunuoprecipitated with the other sarcoglycans, indicating

that they were all capable of assembling into prototypical sarcoglycan heterotetramers (Figure 3.4b).

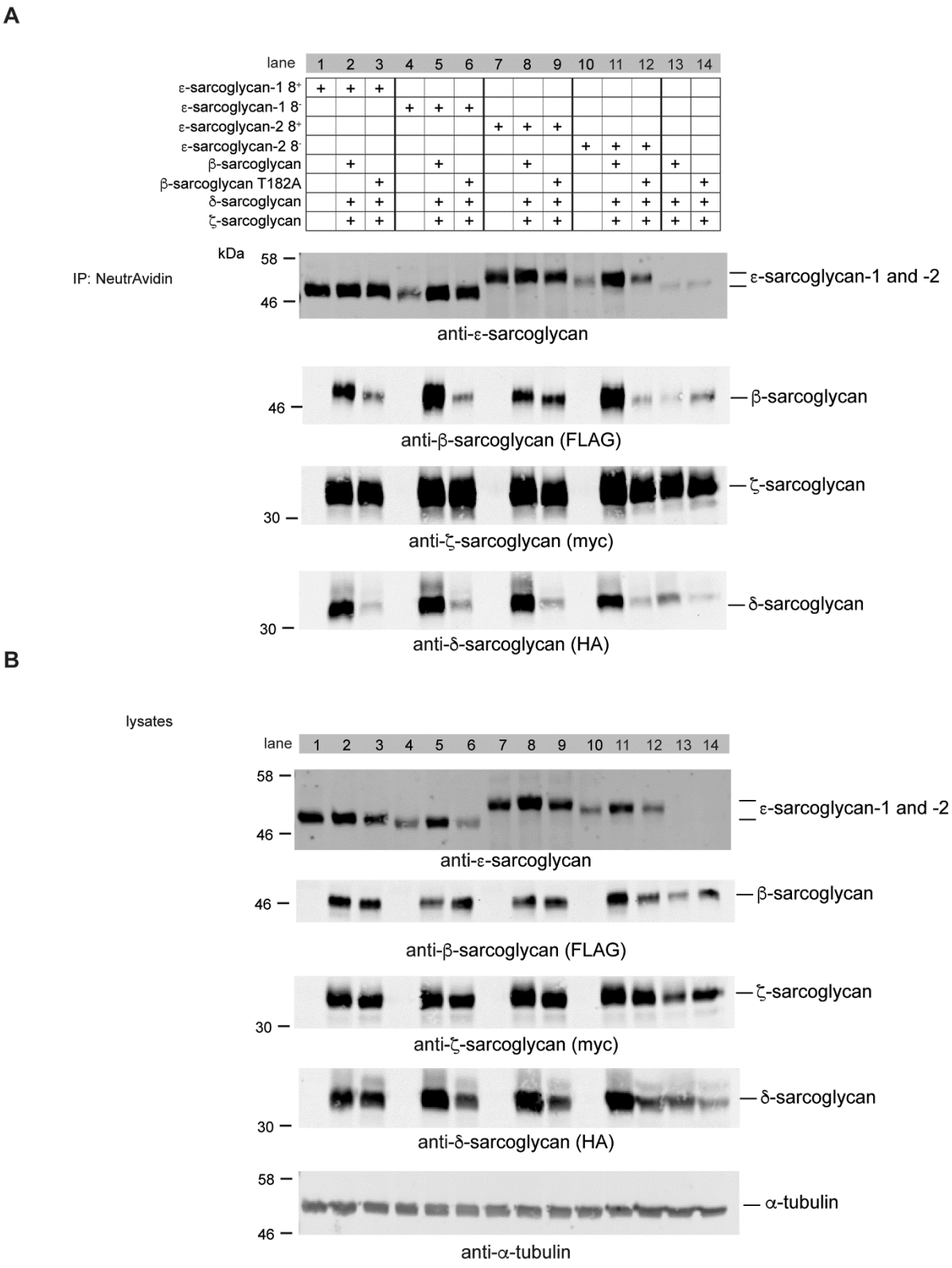


Figure 3.5 Assessing membrane trafficking of the brain SGC using NeutrAvidin capture. The figure is arranged following the scheme detailed for figure 3. HEK293T cells were transfected with different combinations of plasmids encoding epitope-tagged sarcoglycans as indicated. After surface biotinylation, NeutrAvidin agarose beads were used to affinity purify biotinylated membrane proteins. **A)** NeutrAvidin capture of cell surface localised sarcoglycans in heterologous cells. Although ε- and ζ-SG were always detected at the membrane, co-expression of all four sarcoglycans promoted robust trafficking of the wild-type sarcoglycan tetramer (lanes 2, 5, 8 and 11). By contrast, in cells expressing the LGMD2E-associated β-sarcoglycan T182A mutant, the levels of βδ-SG core were severely reduced at the cell surface (lanes 3, 6, 9 and 12). Importantly, membrane trafficking of ε- and ζ-sarcoglycan was apparently unaffected by the T182A mutant. **B)** Whole cell lysates are shown for comparative purposes while α-tubulin was used as a loading control. Abbreviations: kDa, kilodaltons.

Previous studies in cell models showed that sarcoglycan mutants impair assembly and cell surface trafficking of the sarcoglycan complex, recapitulating the cell surface sarcoglycan complex deficiencies observed in sarcoglycanopathies (Chen et al., 2006, Draviam et al., 2006a, Duggan et al., 1997a, Duggan and Hoffman, 1996, Fanin et al., 1997, Holt and Campbell, 1998). The *SGCB* c.544A>G mutation, encoding T182A mutant β-SG protein, causes a severe muscular dystrophy phenotype and cardiomyopathy with complete loss of the SGC from the sarcolemma in skeletal muscle, and has been shown to disrupt the βδ-SG core in heterologous cells (Chen et al., 2006, Draviam et al., 2006a, Duggan et al., 1997a, Duggan and Hoffman, 1996, Fanin et al., 1997, Holt and Campbell, 1998). Different ε-SG isoforms might react differently to the presence of mutant β-SG in the sarcoglycan heterotetramer. Therefore, each ε-SG isoform was tested in a complex containing the T182A mutant β-SG, wild-type δ-SG and wild-type ζ-SG as above. As expected, only residual amounts of T182A mutant β-SG were detected at the cell surface using streptavidin-Alexa Fluor 680 (Figure 3.4a) or in NeutrAvidin affinity purifications (Figure 3.5a). Similarly, cell surface levels of δ-SG were substantially reduced. However, ε-SG and ζ-SG still co-purified, and some was biotinylated indicating membrane trafficking (Figures 3.4 and 3.5, lanes 3, 6, 9 and 12). These findings were consistent for all ε-SG isoforms. Therefore the T182A mutant β-SG did not disrupt the physical interaction between and membrane trafficking of ε-SG and ζ-SG. These data correspond with a recent observation of reduced but still present (residual) cell surface ε-SG and ζ-SG in the brain of δ-SG-deficient BIO14.6 hamster (Waite et al., 2016).

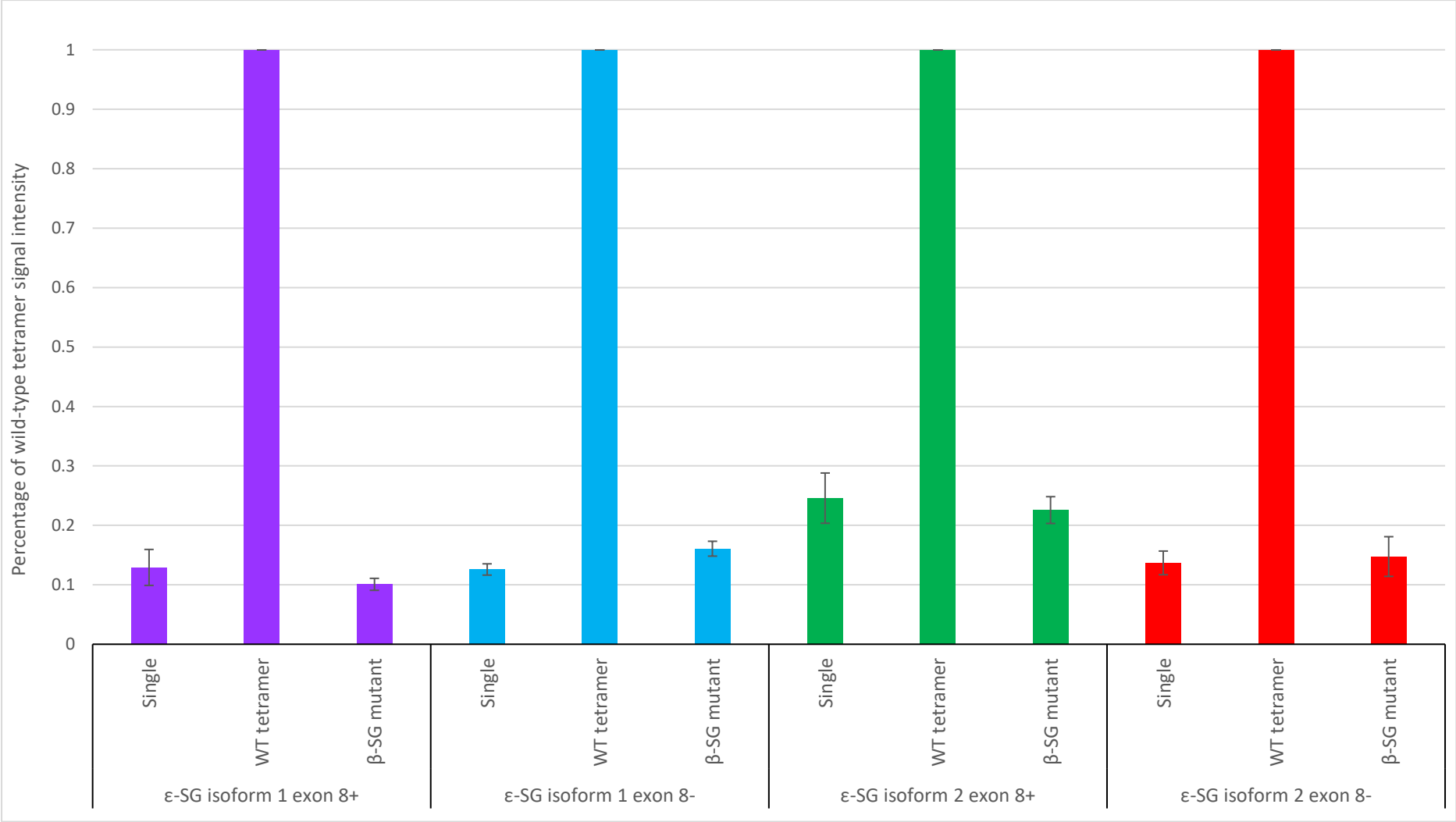


Figure 3.6 Levels of ε-SG at the cell surface do not differ significantly between single transfections and transfections in the presence of mutant β-SG. The relative amount of biotinylated (cell surface) ε-SG in IP of ε-SG from cells transfected with ε-SG alone, in combination with wild-type sarcoglycans, or in a tetrameric context with mutant β-SG was estimated by quantifying immunofluorescence on the immunoblot, and averaging over three experimental replicates. The level of biotinylated ε-SG in cells transfected with the wild-type εβδζ tetramer were set as 100%, and levels in the single or β-SG mutant transfections were expressed as a percentage of that wild-type tetramer level within each ε-SG isoform set. No significant difference between surface ε-SG in single transfections and with mutant β-SG was found using Student's T-test assuming unequal variances. Error bars represent standard error of the mean.

3.2.3. Dp71 and β-dystroglycan co-purify with ε-SG under mild lysis conditions

MS analysis of the esg2-1358 antibody IAP using RIPA-solubilised brain tissue described above included protein sizes expected to include a range of DGC proteins. However, only peptides corresponding to the sarcoglycans and no other DGC proteins were identified. Buffer stringency could have disrupted interactions between these other DGC proteins, despite apparent stability of the sarcoglycan complex in RIPA. Therefore, a buffer containing the milder, non-ionic detergent digitonin (1% (w/v)) was used to solubilise proteins from mouse brain and lung tissue, and IAP was performed using the esg4990 antibody against all ε-SG isoforms. In lung, ε-SG isoform 1 is known to be part of a prototypical sarcoglycan

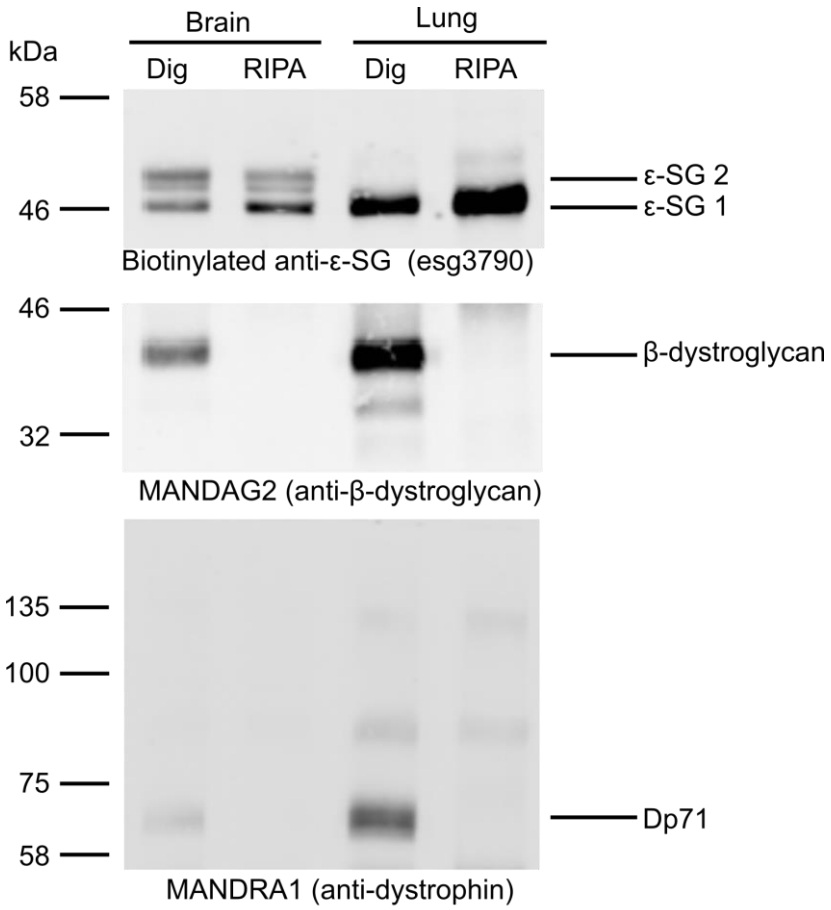


Figure 3.7 β-dystroglycan and Dp71 co-purify with ε-SG from digitonin-solubilised but not RIPA-solubilised brain and lung. Mouse brain and lung tissue was solubilised in either digitonin (Dig) or RIPA buffer, and an antibody (esg3790) against all ε-SG isoforms was used to enrich for ε-SG. Western blot with a biotinylated anti-ε-SG antibody demonstrated that ε-SG isoforms 1 and 2 can be purified from brain, but only ε-SG isoform 1 can be purified from lung. β-dystroglycan and dystrophin isoform Dp71 co-purified with ε-SG from both brain and lung, but only in digitonin buffer. However, ε-SG solubilisation was greater in RIPA compared to digitonin buffer. Abbreviations: kDa, kilodaltons; Dig, digitonin buffer-lysed tissue; RIPA, RIPA buffer-lysed tissue; ε-SG 2, ε-SG isoform 2; ε-SG 1, ε-SG isoform 1.

complex that is integrated into the DGC (Durbeej and Campbell, 1999, Liu and Engvall, 1999). Western blotting of immunoaffinity purified proteins from brain or lung demonstrated co-purification of β-dystroglycan and Dp71 with ε-SG from digitonin-solubilised but not RIPA-solubilised tissue (Figure 3.7). This demonstrated that in brain as well as in lung, ε-SG is associated with other DGC proteins.

The IAPs were repeated using esg2-1358 (detects ε-SG isoform 2 only) and esg3788 (detects all ε-SG isoforms) to determine whether both the widely-expressed ε-SG isoform 1 and the brain-specific isoform 2 co-purified with β-dystroglycan and Dp71 (Figure 3.8). As expected, β-dystroglycan and Dp71 co-purified with ε-SG using the esg3788 antibody from digitonin-solubilised but not RIPA-solubilised tissue. However, only β-dystroglycan and not Dp71

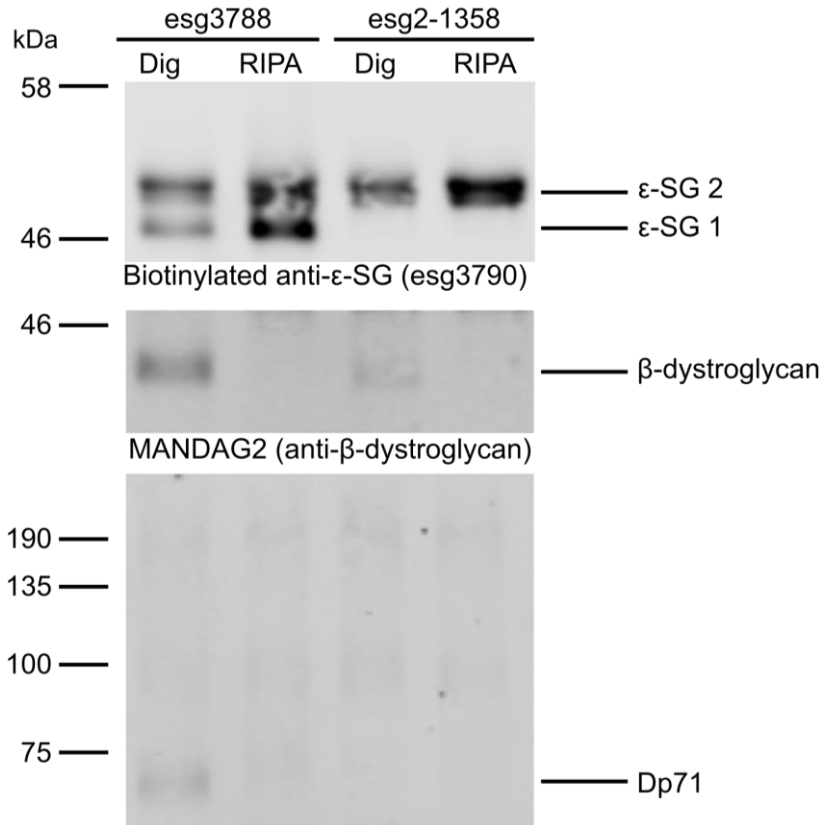


Figure 3.8 β-dystroglycan and Dp71 predominantly co-purified with ε-SG isoform 1 from whole brain tissue. Complexes containing ε-SG were isolated from mouse brain tissue solubilised in either digitonin or RIPA buffer using either the pan-ε-SG antibody esg3788 or the ε-SG isoform 2-specific antibody esg2-1358. Western blot of immunoaffinity purified proteins using biotinylated anti-ε-SG antibody showed enrichment of ε-SG isoforms 1 and 2 with esg3788, but only ε-SG isoform 2 with esg2-1358. β-dystroglycan and Dp71 co-purified with ε-SG in the esg3788 IAP, while β-dystroglycan co-purified in the esg2-1358 IAP as well; in both cases, this occurred in digitonin buffer but not RIPA buffer. Thus, both major isoforms of ε-SG in the brain interacted with β-dystroglycan, and ε-SG isoform 1 interacted with Dp71. Abbreviations: kDa, kilodaltons; Dig, digitonin-lysed tissue; RIPA, RIPA-lysed tissue; ε-SG 2, ε-SG isoform 2; ε-SG 1, ε-SG isoform 1.

could be detected in esg2-1358 immunoaffinity-purified proteins (Figure 3.8). Additionally, ε-SG and β-dystroglycan signal intensity on Western blot was weaker in the esg2-1358 IAP compared to the esg3788 IAP. These results were most likely due to the smaller total amount of ε-SG purified using the esg2-1358 antibody compared to the esg3788 antibody (Figure 3.8). Subsequent IAPs carried out on wild-type C57BL/6J and *Sgce*$^{m+/pGT}$ mice as described later in this chapter (sections 3.2.5.3-4) replicated these results.

To confirm that β-dystroglycan and dystrophin were genuinely co-purifying with ε-SG, reciprocal IAPs from digitonin-solubilised tissue using antibodies against β-dystroglycan and

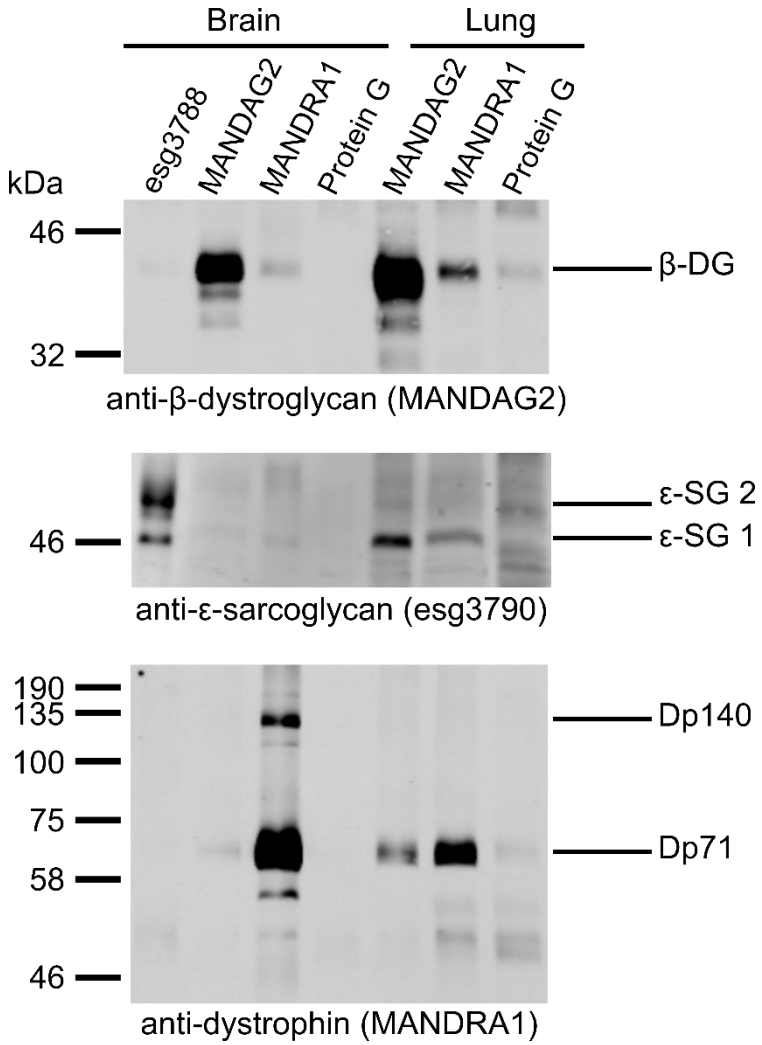


Figure 3.9 ε-SG co-purified with β-dystroglycan and Dp71 from both brain and lung. IAPs using the MANDAG2 and MANDRA1 antibodies to enrich for β-dystroglycan and dystrophin respectively were carried out with digitonin-solubilised brain and lung tissue. Parallel IAPs using esg3788 to enrich for all ε-SG isoforms and Protein G-sepharose only to isolate non-specific interactions were also included. In both tissues, Dp71 and β-dystroglycan weakly co-purified. ε-SG was also detected in the MANDAG2, MANDRA1 and esg3788 IAPs but not the Protein G-sepharose only IAP. Dp140 was also identified in the MANDRA1 IAP from brain tissue only. Abbreviations: kDa, kilodaltons; ε-SG 2, ε-SG isoform 2; ε-SG 1, ε-SG isoform 1.

dystrophin were performed. The antibodies used were MANDAG2 against β-dystroglycan and MANDRA1 against the dystrophin C-terminus, both of which were developed and validated by Professor Glenn Morris, and obtained through the Developmental Studies Hybridoma Bank (Helliwell et al., 1994, Nguyen et al., 1992). Western blot analysis of immunoaffinity purified protein demonstrated that ε-SG was present in both the MANDAG2 and the MANDRA1 IAPs, from brain and lung tissue (Figure 3.9). However, only isoform 1 was clearly present; ε-SG isoform 2 might have been present in the IAPs from brain tissue, but could not be clearly detected on the Western blot. Dp71 and β-dystroglycan also weakly co-purified with each other from both tissues (Figure 3.9). These data supported genuine co-purification of β-dystroglycan, Dp71 and ε-SG from brain tissue.

3.2.4. The Shaker-related voltage-gated potassium channels cross-react with the esg2-1358 antibody

Alongside the sarcoglycans, several Shaker-related voltage-gated potassium channel subunits were identified with high confidence by MS in the esg2-1358 IAP from RIPA-solubilised mouse brain tissue (Table 3.1). While these proteins may genuinely interact with ε-SG isoform 2, alternatively their presence in the IAP could be caused by cross-reaction with the esg2-1358 antibody. Kcna3 was identified at high confidence in multiple bands from the esg2-1358 IAP, and was therefore tested for cross-reaction with the esg2-1358 antibody. A c-Myc-tagged Kcna3 construct was overexpressed in HEK293T cells, and protein was immunoprecipitated from cell lysates using Protein A-agarose-conjugated esg2-1358, esg4990 or pre-immune immunoglobulin to check for non-specific interactions (Figure 3.10). Western blot of immunoprecipitated proteins showed that c-Myc-Kcna3 co-purified with the esg2-1358 antibody but not with the esg4990 antibody or pre-immune immunoglobulin (Figure 3.10). Therefore, Kcna3 cross-reacted with the esg2-1358 antibody. Since the Shaker-related voltage-gated potassium channels are tightly associated, purification of one subunit

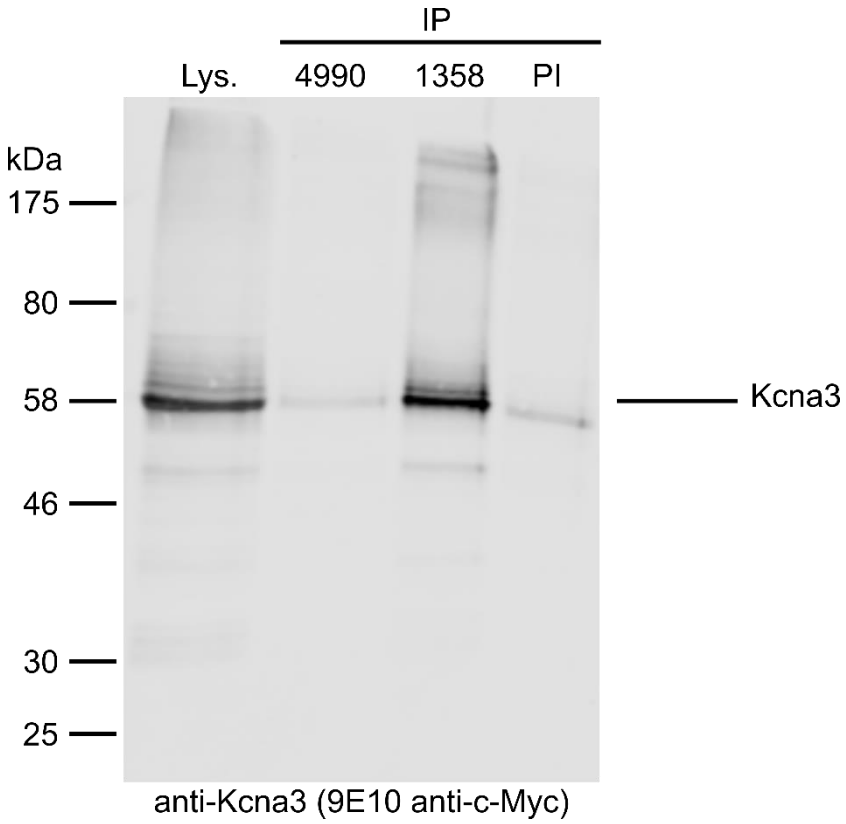


Figure 3.10 Kcna3 cross-reacts with the anti-ε-SG isoform 2 antibody esg2-1358. HEK293T cells were transfected with a plasmid encoding c-Myc-Kcna3. After transfection, proteins were immunoprecipitated from NP-40 buffer cell lysates using the esg4990 antibody, the esg2-1358 antibody, or pre-immune immunoglobulin. Immunoblot of the IPs for c-Myc-Kcna3 showed that Kcna3 purified with esg2-1358 but not esg4990 or pre-immune IgG. Therefore, Kcna3 cross-reacted with the esg2-1358 antibody. Abbreviations: kDa, kilodaltons; lys, cell lysate; IP, immunoprecipitation; 4990, esg4990 antibody; 1358, esg2-1358 antibody; PI, pre-immune immunoglobulin

typically results in co-purification of other subunits (Vacher et al., 2008). Consequently, enrichment of all Shaker-related voltage-gated potassium channels in the esg2-1358 IAP was most likely due to the cross-reaction between Kcna3 and the esg2-1358 antibody.

3.2.5. Identifying other components of ε-SG-containing complexes in the brain

The interactions between ε-SG and DGC proteins described above were all identified by looking for specific proteins in the ε-SG IAPs, either through MS analysis or Western blot. To identify other proteins that co-purify and therefore might interact with ε-SG in brain tissue, a series of hypothesis-free IAP-MS experiments using anti-ε-SG antibodies were carried out. IAPs were performed using two different anti-ε-SG antibodies: esg2-1358 which detects ε-SG isoform 2 only, and esg3788 which detects all ε-SG isoforms. These antibodies were raised against non-overlapping immunogens, and therefore a protein that co-purified

with ε-SG using both antibodies would have a greater likelihood of genuinely interacting with ε-SG rather than cross-reacting with an anti-ε-SG antibody (Figure 3.1). Digitonin-solubilised mouse brain tissue was used for these IAPs, since this buffer appeared to preserve protein interactions better than RIPA buffer as demonstrated by co-purification of Dp71 and β-dystroglycan with ε-SG (section 3.2.3). Additionally, rather than selecting specific protein bands from the IAP for MS analysis, each IAP was divided into two samples that were both analysed by MS (Figure 3.11). This approach allowed identification of all protein sizes, but also increased MS sample complexity and therefore reduced protein coverage compared to the individual protein bands approach used in the initial esg2-1358 IAP (section 3.2.2.1).

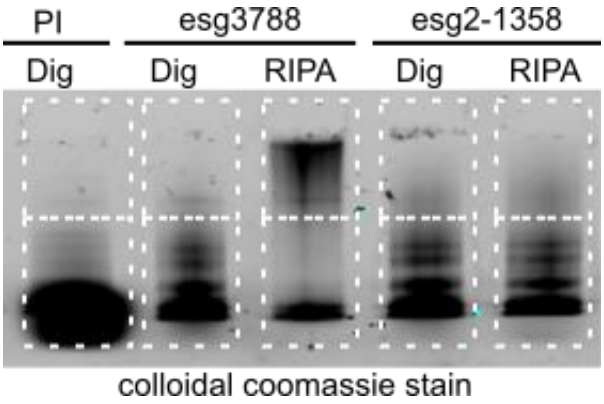


Figure 3.11 Polyacrylamide gel showing immunoaffinity purified protein complexes prepared for MS analysis. Two-thirds of each IAP described in figure 3.8 – esg3788 and esg2-1358 IAPs from digitonin (Dig)-solubilised or RIPA-solubilised brain tissue, pre-immune immunoglobulin (PI) IAP from digitonin-solubilised brain tissue – were briefly resolved by PAGE. Proteins were visualised by staining the gel with a colloidal Coomassie stain. Each IAP/lane shown here was divided into two samples across the horizontal axis as shown, and both samples were processed for MS. Abbreviations: PI, IAP using pre-immune IgG; esg3788, IAP using the esg3788 antibody; esg2-1358, IAP using the esg2-1358 antibody; dig, IAP from tissue lysed in digitonin buffer; RIPA, IAP from tissue lysed in RIPA buffer.

The proteins identified in each IAP-MS experiment were ranked using several parameters generated by the Proteome Discoverer SEQUEST algorithm implementation used to assign peptide identities to spectra (Chapter 2 section 2.6.6). The SEQUEST protein identity score gave the likelihood that a given protein was genuinely present in the sample; a higher score indicated greater confidence. The number of unique peptides provided the number of distinct peptide sequences assigned to that particular protein/accession number detected in the sample. Finally, spectral count is the number of spectra corresponding to that

protein/accession number detected in the sample, and was used as a semi-quantitative representation of protein abundance in the IAP (Bantscheff et al., 2007).

3.2.5.1. IAP MS data were filtered to remove non-specifically interacting and cross-reacting proteins

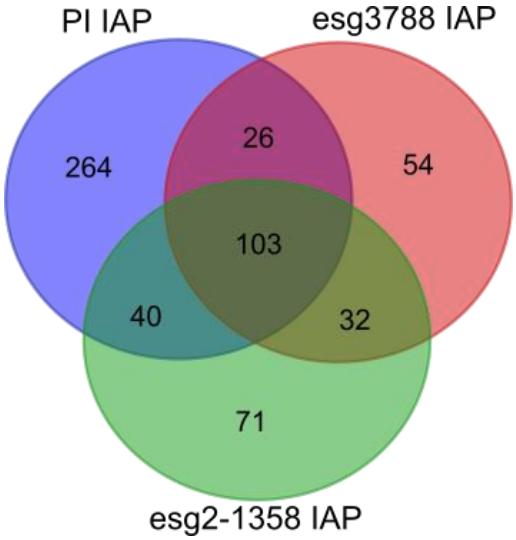


Figure 3.12 Comparison of proteins identified in the esg3788, esg2-1358, and pre-immune IAPs from digitonin-solubilised mouse brain. Venn diagram of overlap between proteins identified by MS in the pre-immune immunoglobulin (PI), esg3788 IAP, and esg2-1358 IAPs performed in parallel on digitonin-solubilised mouse brain tissue. Overlap was based on gene symbols. Proteins present in the PI IAP were likely to be enriched through non-specific interactions with immunoglobulin or Protein A-agarose, and were eliminated from the esg3788 and esg2-1358 IAP datasets. This removed 143 proteins from the esg2-1358 IAP and 129 proteins from the esg3788 IAP. The remaining 103 proteins in the esg2-1358 IAP dataset and 86 proteins in the esg3788 IAP were expected to be enriched through interaction with ε-SG or cross-reaction with the antibody.

In addition to proteins co-purifying with ε-SG itself, IAPs using anti-ε-SG antibodies were expected to enrich proteins interacting non-specifically with immunoglobulin or Protein A-agarose, or cross-reacting with the anti-ε-SG antibody. To identify and filter out non-specifically interacting proteins, an IAP using pre-immune immunoglobulin conjugated to protein A-agarose was performed in parallel with other IAPs and analysed by MS (Appendix II). Only proteins interacting non-specifically with immunoglobulin or Protein A-agarose – the bead proteome - would be present in this IAP. Therefore, all proteins detected in the pre-immune IAP were filtered out of IAPs using anti-ε-SG antibodies as shown in Figure 3.12.

Proteins cross-reacting with the anti-ε-SG antibodies were identified through two different approaches. As described above, use of two distinct antibodies could help identify cross-

reacting proteins. However, esg3788 and esg2-1358 have only partially overlapping reactivity: esg3788 binds all ε-SG isoforms, whereas esg2-1358 binds only ε-SG isoform 2. Therefore a protein genuinely interacting with ε-SG isoform 1 but not isoform 2 would be enriched in an IAP using esg3788, but not in an IAP using esg2-1358. To further identify cross-reacting proteins, the proteins enriched by IAP from digitonin-solubilised tissue were compared to those enriched by IAP from RIPA-solubilised tissue. Proteins that cross-react with an antibody should always be enriched through IAP unless antibody-antigen interactions are disrupted. By contrast, proteins enriched through interaction with ε-SG will only be enriched in the IAP if lysis conditions are mild enough to maintain the interaction. Therefore, cross-reacting proteins were expected to be more abundant in IAPs from RIPA-solubilised tissue, whereas proteins co-purified through interaction with ε-SG were expected to be more abundant in IAPs from digitonin-solubilised tissue; ε-SG itself should be enriched equally under both conditions.

The parallel esg3788 and esg2-1358 IAPs from digitonin-solubilised and RIPA-solubilised tissue described in section 3.2.3 (Figure 3.8) were therefore analysed by MS and the resultant datasets were compared. As predicted, spectral counts (a semi-quantitative measure of abundance) for β-SG, δ-SG and ζ-SG which are known to genuinely interact with ε-SG were generally increased in IAPs from digitonin-solubilised tissue compared to IAPs from RIPA-solubilised tissue (Table 3.3). Conversely, the Shaker-related voltage-gated potassium channels specifically cross-react with the esg2-1358 antibody and had increased spectral counts in IAPs from RIPA-solubilised tissue compared to IAPs from digitonin-solubilised tissue (Table 3.3). ε-SG spectral counts were comparable between the two tissue solubilisation conditions. Dp71 and β-dystroglycan could not be detected in any IAP by MS, despite their detection in digitonin-solubilised IAPs using Western blot. Based on these results, additional proteins directly binding the antibodies were identified by their greater or

Gene	Digitonin lysis, esg3788 antibody		RIPA lysis, esg3788 antibody		Digitonin lysis, esg2-1358 antibody		RIPA lysis, esg2-1358 antibody		Change in abundance, RIPA to digitonin
	Spectral count	Unique peptides	Spectral count	Unique peptides	Spectral count	Unique peptides	Spectral count	Unique peptides	
Sgce	28	4	15	3	14	3	14	4	+
Sgcz	8	6	1	1	6	5	0	0	+
Sgcd	6	5	2	2	1	1	0	0	+
Sgcb	2	2	1	1	2	2	0	0	+
Kcna3	0	0	0	0	11	8	50	9	-
Kcna2	0	0	0	0	8	1	30	7	-
Kcnab2	0	0	0	0	11	6	18	9	-

Table 3.3 Semi-quantitative MS data for sarcoglycans and Shaker-related voltage-gated potassium channel subunits in ε-SG IAPs from RIPA- and digitonin-solubilised mouse brain tissue. For IAPs using the same antibody but tissue solubilised under different conditions, the spectral counts and number of unique peptides for sarcoglycans and Shaker-related voltage-gated potassium channel subunits were compared. Spectral count refers to the number of spectra corresponding to the protein detected in the sample, while the number of unique peptides is how many distinct peptides in the sample could be assigned to the protein identity. Shaker-related voltage-gated potassium channel subunits cross-react with the esg2-1358 antibody; their spectral counts and numbers of unique peptides were increased in IAPs from RIPA-solubilised tissue compared to digitonin-solubilised tissue. By contrast, β-, δ- and ζ-SG are known to interact with ε-SG from other tissues. Their spectral counts and numbers of unique peptides were increased in IAPs from digitonin-solubilised tissue compared to RIPA-solubilised tissue.

equal spectral counts in IAPs from RIPA-solubilised tissue compared to digitonin-solubilised tissue (Figure 3.13). The remaining proteins in the IAPs from digitonin-solubilised tissue were ranked by protein identity score (Figure 3.13). The lowest sarcoglycan protein identity score was used as the minimum threshold to define proteins with a high probability of genuinely interacting with ε-SG. This process generated a list of 26 high-confidence proteins from the esg3788 IAP and 23 proteins from the esg2-1358 IAP; 17 proteins including the sarcoglycans were shared between the two lists (Tables 3.4 and 3.5).

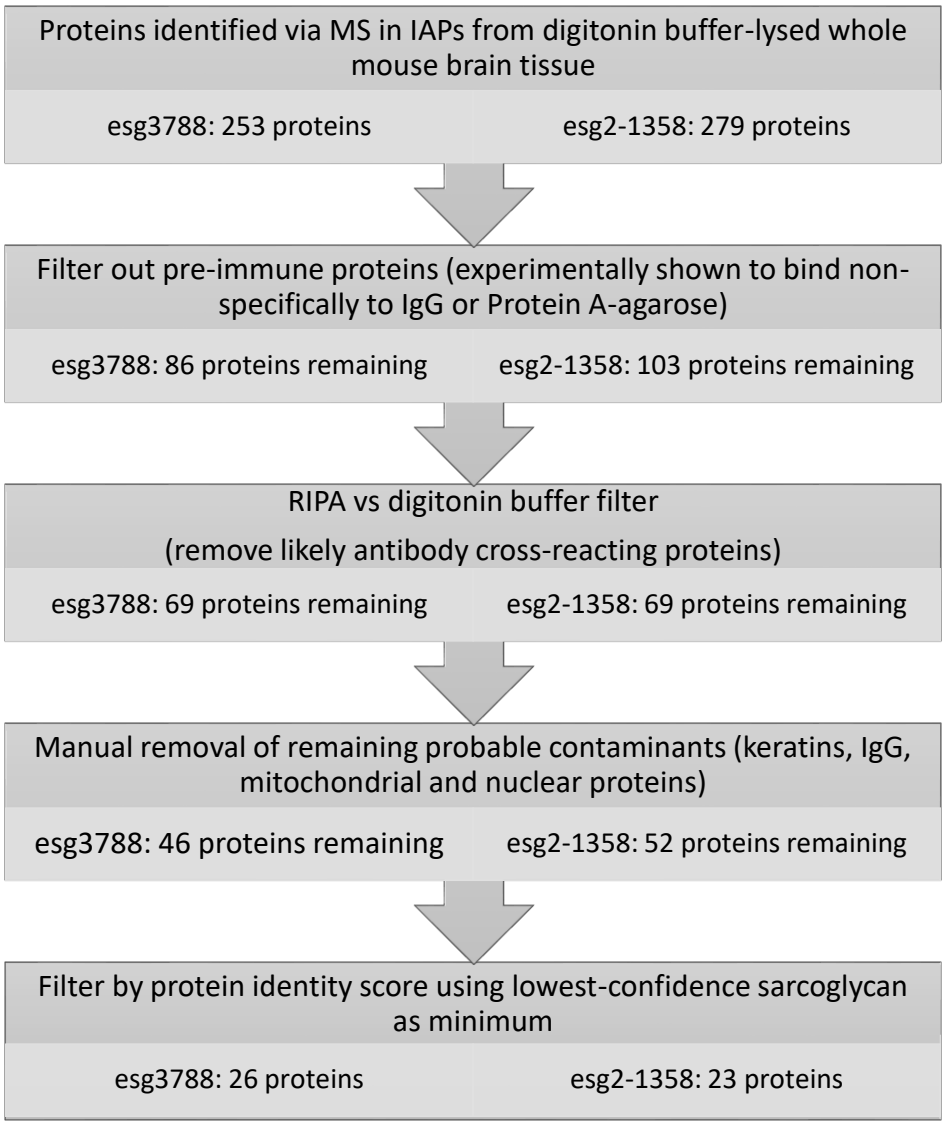


Figure 3.13 Overview of the IAP MS analysis filtering process to remove non-specific interacting proteins, cross-reacting proteins, and contaminants. Starting with MS data from IAPs performed with digitonin buffer-lysed whole mouse brains, several filtering steps were applied to remove potential contaminants and likely cross-reacting proteins. Each step is summarised here, with the number of proteins remaining after the filtering step below. Probable contaminants removed in the penultimate step were identified primarily by subcellular localisation; predominantly cell surface-localised ε-SG is unlikely to interact with proteins localised to the mitochondrial matrix or nucleus.

Gene symbol	Protein	Protein identity score	Unique peptides (digitonin)	Spectral count (digitonin)	Spectral count (RIPA)
Map2	Microtubule-associated protein 2	948.82	85	284	105
Dync1h1	Cytoplasmic dynein 1 heavy chain 1	205.15	47	61	0
Sptan1	Spectrin alpha chain, non-erythrocytic 1	171.39	45	53	3
Plec	Plectin	154.97	46	50	0
Sptbn1	Spectrin beta chain, non-erythrocytic 1	91.06	20	28	0
Rnf213	E3 ubiquitin-protein ligase RNF213	81.76	23	26	1
Sgce	Epsilon sarcoglycan	81.3	4	28	15
Map1a	Microtubule-associated protein 1A	64.98	20	25	0
Dmxl2	DmX-like protein 2	63.27	13	19	0
Sgcz	Zeta sarcoglycan	27.22	6	8	1
Myh10	Myosin-10	22.14	6	6	0
Itpr1	Inositol 1,4,5-trisphosphate receptor type 1	21.56	8	8	0
Usp9x	Ubiquitin carboxyl-terminal hydrolase	19.16	4	6	0
Sgcd	Delta sarcoglycan	18.51	5	6	2
Sptbn2	Spectrin beta chain, non-erythrocytic 2	18.25	4	6	0
Sbf1	Myotubularin-related protein 5	17.51	6	6	0
Slc12a5	Solute carrier family 12 member 5	16.92	3	8	0
Tnr	Tenascin-R	15.65	6	6	0
Tln2	Talin-2	15.4	2	4	0
Fasn	Fatty acid synthase	14.24	3	4	0
Myh9	Myosin-9	13.69	2	5	0
Flg2	Filaggrin-2	9.16	3	3	0
Nf1	Neurofibromin	9.08	2	3	0
Ryr2	Ryanodine receptor 2	8.94	2	3	0
Arfgef2	Brefeldin A-inhibited guanine nucleotide-exchange protein 2	7.56	2	2	0
Sgcb	Beta sarcoglycan	6	2	2	1

Table 3.4 High-confidence proteins identified in the esg3788 IAP after filtering against RIPA IAP results to remove cross-reacting proteins. After filtering against the RIPA IAP MS results, proteins identified in the esg3788 IAP were ranked by protein identity score and number of high-quality unique peptides found to produce a list of top priority proteins. Protein identity score gives the probability that a protein was genuinely present in the sample analysed. Unique peptides refers to the number of distinct peptides from a given protein identified in the sample; each unique peptide may be detected multiple times in the sample. Spectral count refers to the number of spectra corresponding to that protein detected in the sample. Sarcoglycans are highlighted in green; other proteins shared with the esg2-1358 IAP are highlighted in orange.

Gene	Protein	Protein identity score	Unique peptides (digitonin)	Spectral count (digitonin)	Spectral count (RIPA)
Sptan1	Spectrin alpha chain, non-erythrocytic 1	220.86	51	70	37
Dync1h1	Cytoplasmic dynein 1 heavy chain 1	138.77	32	41	6
Itpr1	Inositol 1,4,5-trisphosphate receptor type 1	119.8	32	42	1
Ryr2	Ryanodine receptor 2	92.75	21	29	20
Dmxl2	DmX-like protein 2	91.4	23	29	1
Sptbn1	Spectrin beta chain, non-erythrocytic 1	79.24	24	26	18
Map1a	Microtubule-associated protein 1A	77.47	21	25	7
Clcc1	Chloride channel CLIC-like protein 1	61.73	10	17	13
Sgce	Epsilon sarcoglycan	55.19	3	14	14
Sptbn2	Spectrin beta chain, non-erythrocytic 2	45.42	11	13	3
Ehd2	EH domain-containing protein 2	20.17	5	6	0
Plec	Plectin	19.84	8	9	0
Tenm4	Teneurin-4	19.33	6	6	5
Sgcz	Zeta sarcoglycan	18.82	5	6	0
Tnr	Tenascin-R	10.66	4	4	3
Cul9	Cullin-9	9.35	3	3	1
Myh10	Myosin-10	8.58	2	2	0
4933434 E20Rik	Uncharacterized protein C1orf43 homolog	8.23	3	3	2
Sgce	Epsilon sarcoglycan isoform 2	7.96	1	2	0
Sgcb	Beta sarcoglycan	7.69	2	2	0
Nf1	Neurofibromin	6.15	2	2	0
Sbf1	Myotubularin-related protein 5	5	2	2	0
Sgcd	Delta sarcoglycan	2.86	1	1	0

Table 3.5 High-confidence proteins identified in the esg2-1358 IAP after filtering against RIPA IAP results to remove cross-reacting proteins. After filtering against the RIPA IAP MS results, proteins identified in the esg2-1358 IAP were ranked by protein identity score and number of high-quality unique peptides found to produce a list of top priority proteins likely to genuinely interact with ε-SG in the brain. Protein identity score gives the probability that a protein was genuinely present in the sample analysed. Unique peptides refers to the number of distinct peptides from a given protein identified in the sample; each unique peptide may be detected multiple times in the sample. Spectral count refers to the number of mass spectra from the protein detected in the sample. Sarcoglycans are highlighted in green; other proteins shared with the esg3788 IAP are highlighted in orange.

3.2.5.2. Identification of high-confidence ε-SG interactions in multiple mouse strains

Up to this point, IAP-MS experiments had been carried out exclusively on brain tissue from CD-1 mice, which were obtained from an inbred colony at Cardiff University based on the outbred CD-1 stock (Chia et al., 2005). CD-1 mice were assumed to have generally similar brain biochemistry to other *Mus musculus* strains. However, multiple transcriptomic studies have demonstrated that gene expression varies between mouse strains to produce detectable differences in protein levels, including those for several synaptic proteins (Loos et al., 2016, Nadler et al., 2006, Sandberg et al., 2000, Turk et al., 2004). Differences in protein abundance between mouse strains could result in a different pattern of contaminants (non-specifically interacting proteins and cross-reacting proteins) in IAPs using the same antibody from these different mouse strains. By contrast, proteins that genuinely interact with ε-SG should be consistently enriched in ε-SG IAPs regardless of mouse strain. Therefore, a protein consistently enriched in ε-SG IAPs from brain tissue of different mouse strains would be more likely to genuinely interact with ε-SG.

Consequently, the ε-SG IAPs were repeated on frozen brain tissue from C57BL/6J mice, a gift from Mark LeDoux and Jianfeng Xiao. These mice were wild-type littermates of the *Sgce*$^{m+/pGt}$ mice described below (section 3.2.5.3); IAPs on wild-type C57BL/6J and *Sgce*$^{m+/pGt}$ tissue were performed in parallel. Alongside the esg3788 and esg2-1358 IAPs, a pre-immune immunoglobulin control IAP was performed from digitonin-solubilised C57BL/6J brain tissue to provide strain-specific contaminant data. MS data for the two anti-ε-SG antibody IAPs were filtered against the C57BL/6J pre-immune IAP, and then compared to equivalently filtered data from IAPs using CD-1 mouse brains. In addition to use of different strains, these two sets of experiments differed in person performing the dissection,

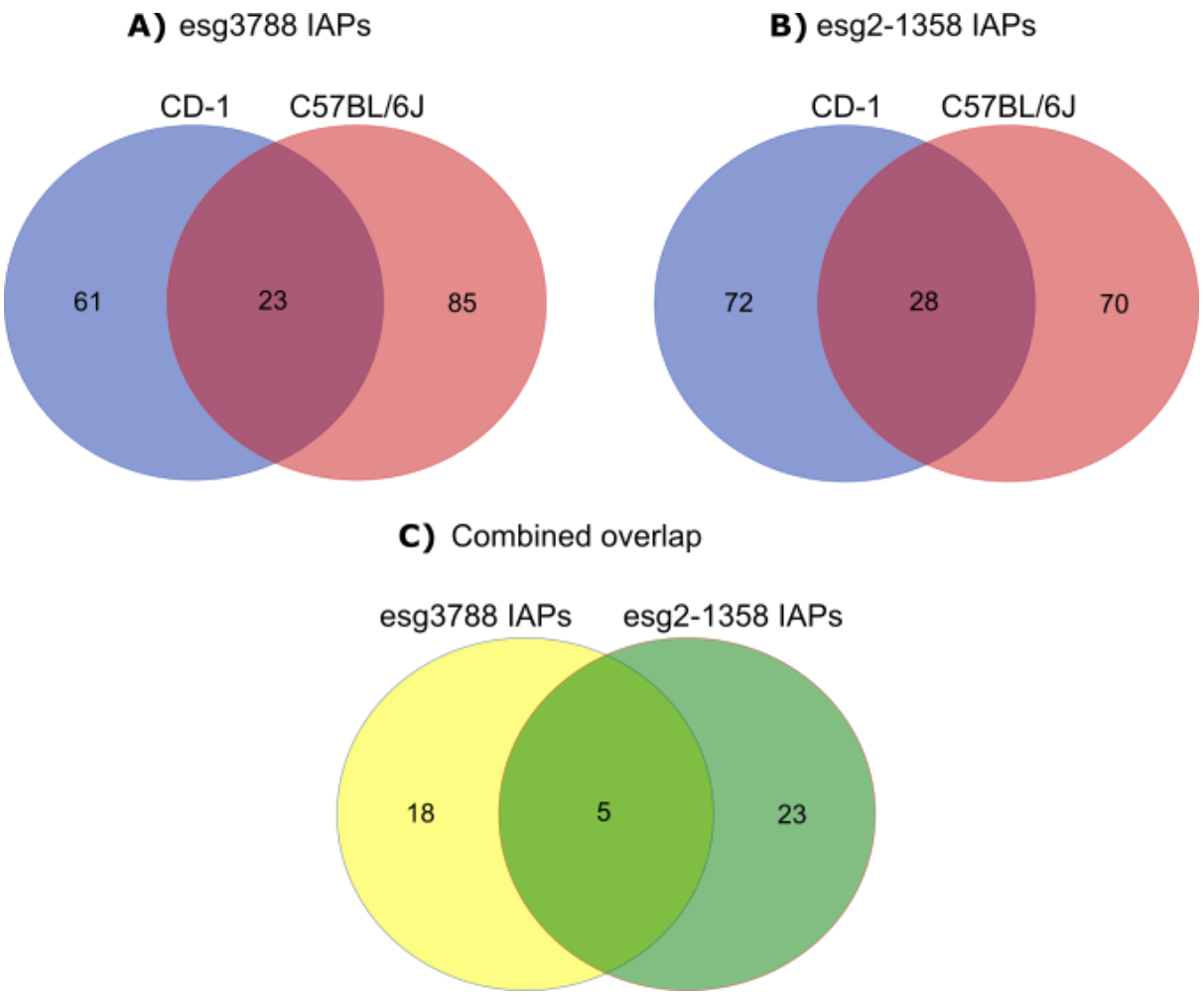


Figure 3.14 Overlap of proteins identified in ε-SG IAPs from CD-1 and C57BL/6J whole mouse brains. After filtering against pre-immune IAPs, the esg3788 and esg2-1358 IAPs from digitonin-solubilised CD-1 and C57BL/6J mouse brains were compared to identify shared proteins that would have the highest probability of being genuine ε-SG interactors. A) Overlap of IAPs using the esg3788 antibody between CD-1 and C57BL/6J whole mouse brains. B) Overlap of IAPs using the esg2-1358 antibody between CD-1 and C57BL/6J. C) Overlap between proteins identified in both mouse strain esg3788 IAPs and proteins identified in both esg2-1358 IAPs. The five proteins shared between all four IAPs were the four sarcoglycans plus Glul, a common contaminant typically filtered out.

tissue processing (the CD-1 brains were used fresh while C57BL/6J brains were snap-frozen and stored at -80°C before use), and time elapsed between dissection and experiment.

The protein overlap between IAP-MS datasets from the two mouse strains was smaller than expected (Figure 3.14). Sarcoglycan complexes were identified by MS in IAPs from both mouse strains using both anti-ε-SG antibodies, and β-dystroglycan was detected by Western blot (Figures 3.8 and 3.15). β- and α-dystroglycan were also identified by MS analysis in the esg3788 IAP from C57BL/6J mouse brains (Table 3.6). This was the first identification of dystroglycan by MS in IAPs, though it had been consistently identified by Western blot

analysis in all previous IAPs using digitonin-solubilised tissue. Aside from the sarcoglycans, only glutamine synthase (*Glul*) was identified in IAPs from both mouse strains with both antibodies; this protein is a known contaminant (Trinkle-Mulcahy et al., 2008). Several proteins were identified in IAPs from both strains but not with both antibodies. While most of these were predicted to be cross-reacting with antibodies based on IAPs from RIPA-solubilised tissue, two were identified as probably co-purifying with ε-SG itself: inositol 1,4,5-triphosphate receptor 1 (*Itpr1*) and tenascin-R (*Tnr*).

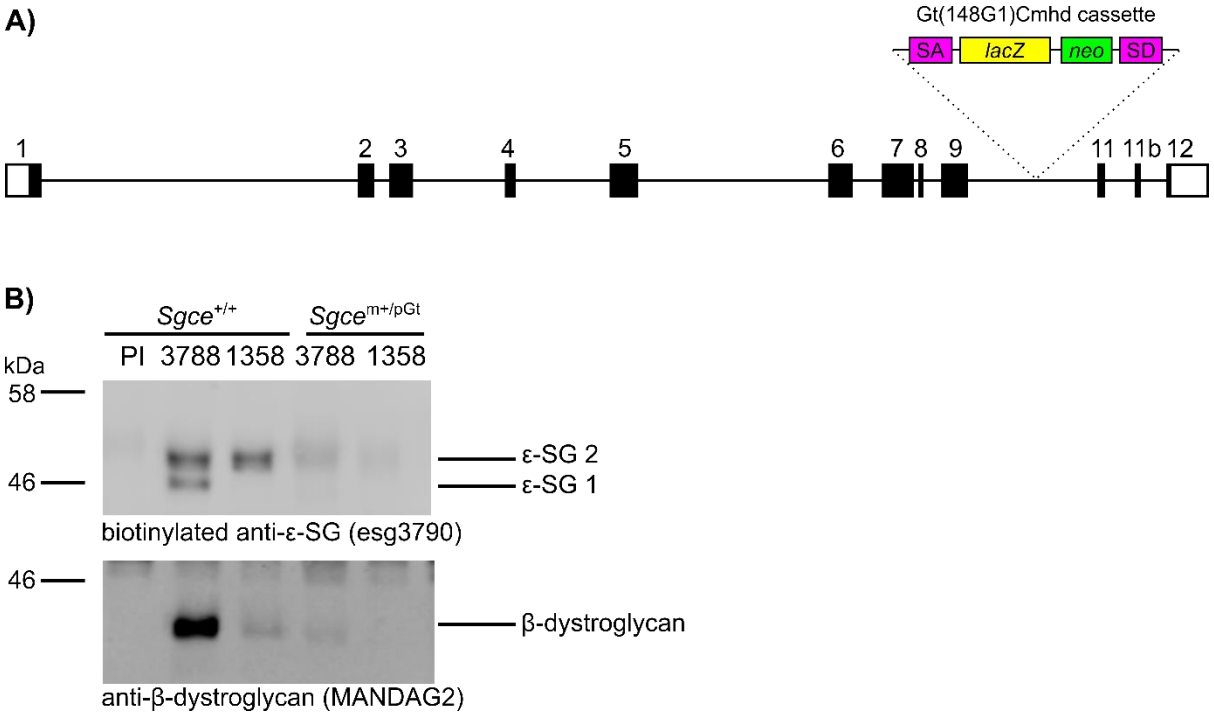


Figure 3.15 The *Sgce* intron 11 gene trap mouse has greatly reduced levels of ε-SG. A) Schematic of the *Sgce* gene-trap allele. The Gt(148G1)Cmhd gene-trap cassette including the *lacZ* and *neo*^{r} coding sequences is inserted into intron 9. Note: mouse *Sgce* does not have an exon 10. B) Complexes containing ε-SG were isolated from digitonin-solubilised brains of *Sgce*$^{m+/pGt}$ and wild-type (*Sgce*$^{+/+}$) littermates, using both the esg3788 antibody that detects all ε-SG isoforms and the esg2-1358 antibody that detects only ε-SG isoform 2. A negative control IAP using pre-immune IgG was also conducted on wild-type tissue. Immunoblot of the IAPs using a separate biotinylated anti-ε-SG antibody identified both ε-SG isoforms 1 and 2 in wild-type mice but no clear signal in *Sgce*$^{m+/pGt}$ mice. A smeared band observed around 50kDa in both the pre-immune IAP and the *Sgce*$^{m+/pGt}$ ε-SG IAPs may represent an IgG-related species. Overall, substantially less ε-SG was detected in IAPs from the *Sgce*$^{m+/pGt}$ mice compared to wild-type littermates. Additionally, β-dystroglycan which interacts with ε-SG could be detected in both ε-SG IAPs from wild-type mice but at a reduced level in the esg3788 IAP only from *Sgce*$^{m+/pGt}$ mice. Abbreviations: kDa, kilodaltons; PI, pre-immune IgG IAP; 3788, esg3788 IAP; 1358, esg2-1358 IAP; ε-SG 1, ε-SG isoform 1; ε-SG 2, ε-SG isoform 2.

3.2.5.3. Evaluation of an *Sgce* gene trap mouse line to identify esg2-1358 cross-reacting proteins

In parallel with the above IAPs on brain tissue from wild-type C57BL/6J mice, a novel *Sgce* mutant mouse line on the C57BL/6J background was investigated as an alternative method to identify proteins cross-reacting with the esg2-1358 antibody. The *Sgce*$^{Gt(148G1)Cmhd}$ gene trap, henceforth abbreviated as *Sgce*^{Gt}, was generated by the Centre for Modelling Human Diseases and consists of a Gep-SD5 *LacZ-neo* construct inserted into intron 9 of the mouse *Sgce* locus (Figure 3.15a) (Xiao et al., 2017). Homozygous (*Sgce*$^{Gt/Gt}$) and paternal heterozygous (*Sgce*$^{m+/pGt}$) mutant mice exhibited a transient movement disorder from postnatal days 14-16 and a mild anxiety-like behavioural phenotype, but were otherwise normal (Xiao et al., 2017). Quantitative RT-PCR indicated overall reduction of *Sgce* transcript levels by 60-70% compared to wild-type littermates in the brains of *Sgce*$^{m+/pGt}$ mice, with no detectable exon 11b+ transcripts (Xiao et al., 2017). The absence of *Sgce* exon 11b-containing transcripts suggested this model might represent a specific knockout of ε-SG isoform 2. In that case, any proteins enriched from *Sgce*$^{m+/pGt}$ mouse brains using the esg2-1358 antibody would be non-specific interactors or cross-reacting proteins.

IAPs using the esg3788 and esg2-1358 antibodies were performed on digitonin-solubilised *Sgce*$^{m+/pGt}$ mouse brain tissue to identify which ε-SG isoform(s) were present and identify proteins purified by the esg2-1358 antibody from these mice. This was done in parallel with the IAPs on wild-type C57BL/6J littermates as described above (section 3.2.5.2). Snap-frozen *Sgce*$^{m+/pGt}$ mouse brain and lung tissue was a gift from Mark LeDoux and Jianfeng Xiao. Immunoaffinity-purified proteins from the *Sgce*$^{m+/pGt}$ mice were compared to proteins purified from wild-type littermates in parallel (section 3.2.5.2) by Western blot analysis. IAP using the esg3788 antibody showed reduced ε-SG from the *Sgce*$^{m+/pGt}$ mice compared to

		C57BL/6J esg3788 IAP				C57BL/6J esg21358 IAP			
Gene	Uniprot acc.	Wild-type		Gene trap		Wild-type		Gene trap	
		Unique peptides	Spectral count	Unique peptides	Spectral count	Unique peptides	Spectral count	Unique peptides	Spectral count
Sgce	O70258	TPYSDGVLYGSPTAENVGKPTIIEITAYNRR QVSTYQEVVR EVENPQNQLR **FEVNGIPEER** **KLTEAMSL**	9	QVSTYQEVVR	1	QVSTYQEVVR EVENPQNQLR	2	0	0
Sgcb	P82349	RNENLVITGNNQPIVFQQGTTK LPSSSSGDQSGSGDWVR THNILFSTDYETHEFHLPSGVK TSITSDIGMQFFDPR LCMCADGTLFK GNEGVFIMGK	17	RNENLVITGNNQPIVFQQGTTK TSITSDIGMQFFDPR	2	RNENLVITGNNQPIVFQQGTTK LPSSSSGDQSGSGDWVR TSITSDIGmQFFDPR LCMCADGTLFK GNEGVFIMGK	8	0	0
Sgcd	P82347	GVEINAEAGNMEAICR LEGDSEFLQPLYAK LLFSADDSEVVVGAER VLGAEGTVFPK VFEVCVCANGR VLTQLVTGPK SRPGNALYFK SLVMEAPK	13	LEGDSEFLQPLYAK LLFSADDSEVVVGAER VLGAEGTVFPK VLTQLVTGPK	5	GVEINAEAGNMEAICR LEGDSEFLQPLYAK LLFSADDSEVVVGAER VLGAEGTVFPK VFEVCVCANGR VLTQLVTGPK	8	0	0
Sgcz	Q8BX51	ELHLQSTEGEIFLNADSIR VLFSADEDEITIGAEK LEGISEFLLPLYVK LGNLPIGSFSSSTSSSNSR QTVYELCVCPNGK GVQVSAAAGDFK	8	LEGISEFLLPLYVK GVQVSAAAGDFK	2	ELHLQSTEGEIFLNADSIR VLFSADEDEITIGAEK LEGISEFLLPLYVK LGNLPIGSFSSSTSSSNSR QTVYELCVCPNGK GVQVSAAAGDFK STDLDIQELK VTGTEGAVFGHSVETPHIR	10	0	0
Dag1	Q62165	IPSDTFYDNEDTTTDKLK ALSIAVTGSGScR	2	0	0	0	0	0	0

Table 3.6 High-confidence sarcoglycan and dystroglycan peptides detected in ε-SG IAPs from wild-type C57BL/6J and *Sgce* intron 11 gene-trap homozygote mice. IAPs were carried out on wild-type C57BL/6J and *Sgce* intron 11 gene-trap homozygote mice using the esg3788 and esg2-1358 antibodies. The proteins present in each IAP were identified through MS analysis

of the sample. The sarcoglycan peptides present in each IAP are listed here, along with the spectral count for that protein in that IAP; the table is arranged so that identical peptides corresponding to a protein are shown in the same row. Peptides derived from the ε-SG isoform 2 unique C-terminal tail are in bold. No sarcoglycans at all were detected in the esg2-1358 IAP from gene-trap mouse brain, suggesting that ε-SG isoform 2 may be severely reduced or absent from these mice.

		CD-1 wild-type		C57BL/6J wild-type		*Sgce* gene trap	
Gene	**Protein**	**esg3788 spectral count**	**esg2-1358 spectral count**	**esg3788 spectral count**	**esg2-1358 spectral count**	**esg3788 spectral count**	**esg2-1358 spectral count**
Tnr	Tenascin-R	6	4	3	0	0	0
Sgcz	Zeta sarcoglycan	8	6	8	10	2	0
Sgce	Epsilon sarcoglycan (both isoforms)	28	16	9	2	1	0
Sgcd	Delta sarcoglycan	6	1	2	8	5	0
Sgcb	Beta sarcoglycan	2	2	7	8	2	0
Itpr1	Inositol 1,4,5-trisphosphate receptor type 1	8	42	1	0	0	1

Table 3.7 Comparison of spectral counts for proteins most likely to be genuine ε-SG interactors between CD-1 wild-type mice, C57BL/6J wild-type mice, and *Sgce*$^{m+/pGt}$ mice. High probability proteins present in IAPs from both CD-1 and C57BL/6J mouse strains were identified, and the spectral counts in each IAP were compared: esg3788 and esg2-1358 IAPs from CD-1 wild-type mice, C57BL/6 wild-type mice, and *Sgce*$^{m+/pGt}$ mice. Abbreviations: gene trap, *Sgce*$^{m+/pGt}$ mice.

wild-type littermates, and correspondingly lower levels of co-purified β-dystroglycan (Figure 3.15b). ε-SG isoform 2 was clearly detected at approximately 50kDa in the wild-type tissue, but not in the *Sgce*$^{m+/pGt}$ mice; faint signal just above ε-SG isoform 2 of approximately 50kDa was detected in *Sgce*$^{m+/pGt}$ mice, but this was also present in the pre-immune IAP and therefore was most likely immunoglobulin-related (Figure 3.15, Table 3.6). MS analysis of the esg3788 immunoaffinity-purified protein confirmed reduced sarcoglycan abundance in the *Sgce*$^{m+/pGt}$ mice based on lower spectral counts (Table 3.6). No sarcoglycans were detected in the esg2-1358 IAP from *Sgce*$^{m+/pGt}$ mice, and combined with the previously described RT-PCR data this suggested absence of ε-SG isoform 2 from the *Sgce*$^{m+/pGt}$ mice (Table 3.6). Therefore, the *Sgce*$^{m+/pGt}$ mice could be used to identify proteins cross-reacting with the ε-SG isoform 2-specific esg2-1358 antibody. Of the two high-likelihood proteins identified in ε-SG IAPs above (section 3.2.5.3), inositol 1,4,5-triphosphate receptor 1 (*Itpr1*) was detected by MS analysis in the *Sgce*$^{m+/pGt}$ mice (Table 3.7). Therefore this protein may cross-react with the esg2-1358 antibody rather than genuinely co-purifying with ε-SG.

3.2.6. ε-SG isoform 2 co-purifies with Tenascin-R from mouse brain.

Tenascin-R was consistently identified in ε-SG IAPs from whole mouse brain tissue using both esg3788 and esg2-1358 antibodies (Table 3.8). To determine whether tenascin-R was genuinely co-purifiying with ε-SG rather than cross-reacting with anti-ε-SG antibodies or interacting non-specifically, the mouse monoclonal anti-TnR antibody clone 619 (Chapter 2 section 2.3.12, Table 2.2) was used for IAP to enrich tenascin-R-containing complexes from mouse brain. The 619 antibody binds the fibrinogen-like domain of tenascin-R and has been well-validated in several previous studies (Morawski et al., 2014, Morganti et al., 1990, Weber et al., 1999, Xiao et al., 1996). For comparison, parallel IAPs using esg3788 or Protein G-sepharose without antibody were also performed. As expected, the anti-tenascin-R antibody IAP successfully enriched tenascin-R as demonstrated by Western blot (Figure

Antibody	Lysis buffer	Mouse strain	Year	Protein identity score	Distinct tenascin-R peptides	Total TnR peptides	TnR spectral count	TnR protein coverage
esg2-1358	RIPA	CD-1	2013	339.79	DGQEAVFAYYDK DKEEDMLEVLLDATKR ITFTPSSGISSEVTVPR LDSSVVPNTVTEFAITR LILNYSPR LYPATEYEISLNSVR SPPTSASVSTVIDGPTQILVR SSLTSTVFTTGGR VATHLSTPQGLQFK VVYSTLAGEQYHEVLVPK YEVSISAVR YGLVGGEGGK	12	15	15.40%
esg3788	Digitonin	CD-1	2015	15.65	DKEEDMLEVLLDATKR LEGLSENTDYTVLLQAAQEATR SSLTSTVFTTGGR VATHLSTPQGLQFK VGFGNLEDEFWLGLDNIHR YEVSISAVR	6	6	15.65%
esg2-1358	Digitonin	CD-1	2015	10.66	SSLTSTVFTTGGR VATHLSTPQGLQFK VGFGNLEDEFWLGLDNIHR YEVSISAVR	4	4	4.05%
esg2-1358	RIPA	CD-1	2015	8.3	VGFGNLEDEFWLGLDNIHR VATHLSTPQGLQFK YEVSISAVR	3	3	3.09%
esg3788	Digitonin	C57BL/6J	2015	7.37	AAIENYVLTYK LDSSVVPNTVTEFAITR YEVSISAVR	3	3	2.72%

Table 3.8 Tenascin-R peptides identified in IAPs from whole mouse brain using anti-ε-SG antibodies. Details of each ε-SG IAP in which tenascin-R (TnR) was identified are listed. The TnR peptides detected in the IAPs are also given, alongside the protein identity score (likelihood that TnR was actually present in the sample analysed by MS), total number of TnR peptides detected, TnR spectral count, and the percentage of the TnR protein covered by the peptides detected. Abbreviations: TnR, tenascin-R.

3.16). In addition, ε-SG isoform 2 but not isoform 1 could be detected in the tenascin-R IAP (Figure 3.16). Therefore, ε-SG isoform 2 and tenascin co-purify using antibodies against both proteins as assessed by immunoblot and MS analysis of IAPs. However, tenascin-R could not be detected in the esg3788 IAP, despite its consistent detection in ε-SG IAPs by MS analysis (Table 3.8, Figure 3.16). Spectral counts for tenascin-R in ε-SG IAPs were comparatively low, so the inability of the anti-tenascin-R antibody to detect tenascin-R in the esg3788 IAP could reflect low antibody sensitivity on Western blot.

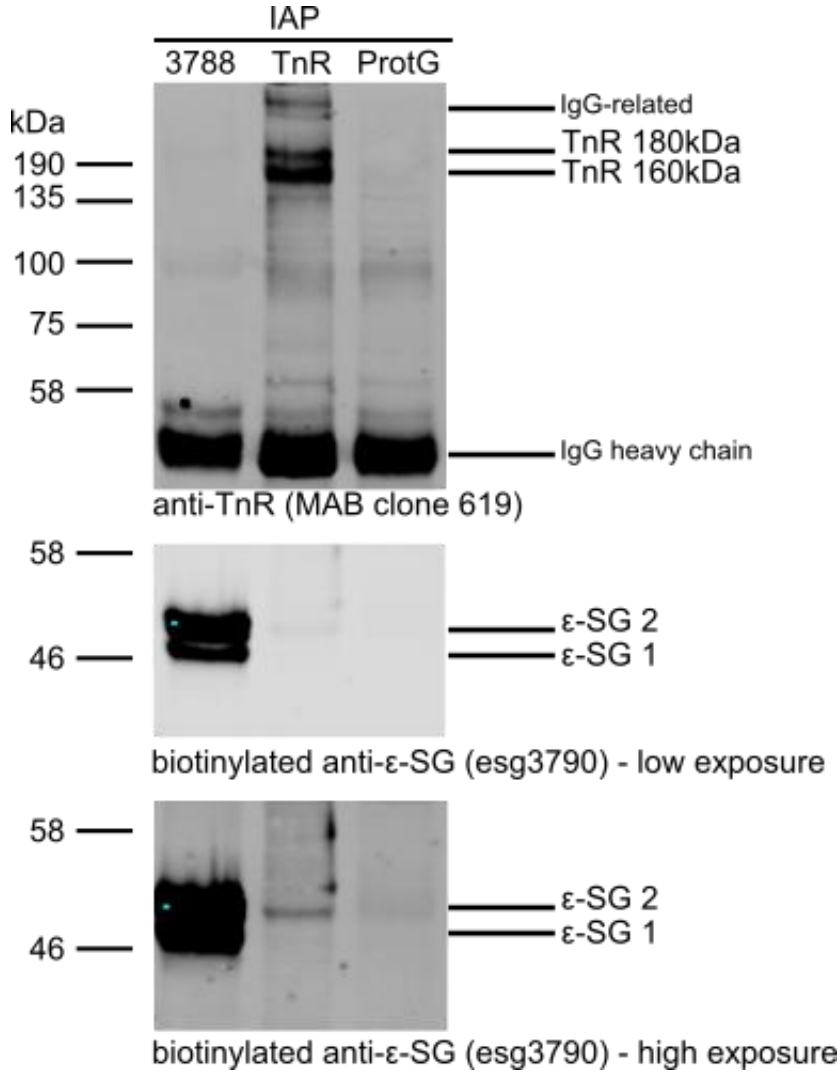


Figure 3.16 ε-SG isoform 2 but not isoform 1 co-purifies with TnR from digitonin-solubilised mouse brain. An anti-TnR mouse monoclonal antibody was used to isolate TnR-containing complexes from mouse whole brain solubilised with digitonin lysis buffer. The esg3788 antibody was used to enrich for ε-SG, and a Protein G-sepharose only capture with no antibody was included to evaluate for non-specific protein binding. Immunoblot using the anti-TnR antibody indicated detectable enrichment in the anti-TnR IAP but not in the ε-SG IAP. This may be due to inability of the anti-TnR antibody to detect low abundance TnR such as that observed in ε-SG IAPs, as TnR has consistently been detected in ε-SG IAPs via MS. Immunoblot using a biontinylated pan-ε-SG antibody demonstrated that both isoforms of ε-SG were present in the esg3788 IAP, while only ε-SG isoform 2 co-purified with TnR. Thus, TnR appeared to specifically co-purify with ε-SG isoform 2. No ε-SG or TnR was present in the Protein G-sepharose control. Abbreviations: kDa, kilodaltons; 3788, IAP using the esg3788 antibody; TnR, IAP using the anti-TnR antibody; ProtG, IAP using Protein G-sepharose with no antibody.

3.3. Discussion

The proteins with which ε-SG interacts in the brain could provide insight into its function and the molecular pathogenesis of MD. Therefore, in this study I aimed to identify proteins interacting with the widely-expressed ε-SG isoform 1 and the brain-specific ε-SG isoform 2 in the mouse brain. MS and Western blot analysis of immunoaffinity-purified proteins from brain demonstrated that both ε-SG isoforms consistently co-purified with β-, δ- and ζ-SGs, demonstrating the existence of prototypical εβδγ sarcoglycan heterotetramers in the brain. Furthermore, under mild solubilisation conditions ε-SG co-purified with the other DGC proteins Dp71 and β-dystroglycan. This co-purification indicated that ε-SG may be part of DGC-like complexes in the brain. ε-SG isoform 2 was also found to co-purify with the extracellular matrix protein tenascin-R in the brain.

Sgce is predominantly expressed in neuronal tissue, with some expression in vascular tissue and glial cells including astrocytes (Boulay et al., 2015, Chan et al., 2005, Nishiyama et al., 2004). However, a recent study identified transcripts from all six sarcoglycans in blood vessels isolated from mouse brain tissue (Boulay et al., 2015). Intriguingly, ε-SG isoform 1 only has been detected in astrocytes and brain vascular tissue, suggesting that ε-SG isoform 2 may be restricted to neurons (Nishiyama et al., 2004). IAPs using either the esg3788 antibody against all ε-SG isoforms or the esg2-1358 antibody specific to ε-SG isoform 2 consistently co-purified β-SG, δ-SG and ζ-SG, indicating that both main ε-SG isoforms present in the brain were incorporated into prototypical εβδζ sarcoglycan heterotetramers (Tables 3.2, 3.4, and 3.5). γ-SG and α-SG were not detected in any ε-SG IAP, which was consistent with reported low expression of their respective genes in the brain (Durbeej and Campbell, 1999, Noguchi et al., 1995, Noguchi et al., 2001, Waite et al., 2016). However, a previous IAP performed in our research group using the esg4990 antibody to purify all ε-SG isoforms from

RIPA-solubilised rat brain did isolate γ-SG (Waite et al., 2016). This esg4990 IAP from rat brain was performed and analysed using an identical protocol to that used for the initial esg2-1358 IAP enriching ε-SG isoform 2 from RIPA-solubilised mouse brain (section 3.2.2.1) (Waite et al., 2016). The co-purification of γ-SG with all ε-SG isoforms but not the possibly neuron-restricted ε-SG isoform 2 (Nishiyama et al., 2004) could indicate that γ-SG interacts only with ε-SG isoform 1 as part of sarcoglycan heterotetramers in non-neuronal tissue. These data were consistent with the presence of multiple DGC-like complexes in different brain cell types (Blake et al., 1999). Absence of γ-SG from subsequent IAPs performed using antibodies against all ε-SG isoforms may reflect a combination of low γ-SG abundance and use of a different protocol for MS analysis of immunoaffinity-purified proteins (section 3.2.5).

The incorporation of both the widely-expressed ε-SG isoform 1 and the brain-specific ε-SG isoform 2 into prototypical sarcoglycan heterotetramers in the brain raises important questions regarding the relevance of these complexes to MD. MD and LGMD are both caused by loss of function mutations leading to deficiency of the mutated sarcoglycan or entire sarcoglycan complex from the cell surface in the affected tissue (Esapa et al., 2007, Sandona and Betto, 2009, Waite et al., 2011). However, no MD-like symptoms have been identified in LGMD patients and *vice versa*. As previously discussed, the early onset and severity of muscle disease in LGMD could mask any MD-like features. Alternatively, differential assembly and trafficking of complexes, or the presence of residual cell surface complexes in the brain could prevent MD-like symptoms in LGMD patients.

When modelled in heterologous cells, alternatively spliced ε-SG isoforms did not affect assembly or trafficking of wild-type brain sarcoglycan heterotetramers (Figures 3.4 and 3.5). Previous studies have highlighted the importance of β-SG for sarcoglycan complex assembly and trafficking, so the LGMD-associated T182A mutant β-SG was expected to result in

intracellular retention of the sarcoglycan complex in heterologous cells (Chen et al., 2006, Draviam et al., 2006a, Shi et al., 2004). Indeed, T182A β-SG abolished trafficking of the βδ-SG core to the cell surface, but reduced levels of an εζ-SG complex could still be detected at the cell surface (Figures 3.4, 3.5 and 3.6). Cell surface ε-SG and ζ-SG has also been identified in the brain of the δ-SG-deficient BIO14.6 hamster (Waite et al., 2016). Therefore, in the brain a pool of ε- and ζ-SG may traffic to the cell surface despite intracellular retention of mutant βδ-SG complexes (Bianchini et al., 2014, Esapa et al., 2007, Gastaldello et al., 2008, Soheili et al., 2012). Low abundance sarcolemmal sarcoglycan complexes have been shown to contribute to membrane stability in *DMD*-deficient mice, so it is possible that a low abundance cell surface sarcoglycan complex in the brain could be sufficient for normal brain function in LGMD patients (Li et al., 2009).

Cai and colleagues showed that the Schwann cell sarcoglycan complex, in common with the brain complex described herein, was composed of an εβδζ sarcoglycan heterotetramer (Cai et al., 2007). In both brain and peripheral nerve, α- and γ-SG are known to be expressed at very low levels, if at all (Cai et al., 2007, Waite et al., 2016). However, Cai and colleagues found a severe reduction of Schwann cell surface ε-SG and ζ-SG in the δ-SG-deficient BIO14.6 hamster (Cai et al., 2007). This contrasted markedly with the residual cell surface ε-SG and ζ-SG detected in the brain of the BIO14.6 hamster and in heterologous cell models of εβδζ sarcoglycan heterotetramers containing an LGMD-associated mutant β-SG (section 3.2.2.2, Figures 3.4 and 3.5) (Waite et al., 2016). Although this may seem paradoxical, cytoplasmic components of the DGC are also preserved in the brains of the *Dmd*^{mdx} and *Dmd*^{mdx3Cv} mouse models of Duchenne Muscular Dystrophy when they are severely reduced in muscle (Blake et al., 1999). Thus, similar sarcoglycan complexes in nerve, muscle and brain appear to be differentially sensitive to the absence of specific sarcoglycans or other components of the DGC.

β-dystroglycan and Dp71 also co-purified with both the widespread ε-SG isoform 1 and brain-specific isoform 2 under mild tissue solubilisation conditions (Figures 3.7 and 3.8). For most IAPs, β-dystroglycan could only be detected by Western blot analysis of immunoaffinity-purified protein; however, two dystroglycan peptides were identified by MS in the esg3788 IAP from C57BL/6J mouse brain tissue (Table 3.6). One of these peptides was derived from α-dystroglycan rather than β-dystroglycan, demonstrating that digitonin did not disrupt the interaction between these two proteins (Table 3.6). Identification of dystroglycan by MS as well as Western blot further confirmed its presence in ε-SG IAPs from brain tissue. Additionally, IAPs using antibodies against β-dystroglycan or C-terminal dystrophin resulted in the co-purification of ε-SG (Figure 3.9). Co-purification of dystroglycan and dystrophin isoforms was lower than would be expected based on the known interaction between these proteins, however (Figure 3.9). The lower than expected co-purification of dystroglycan and dystrophin may have been caused by several factors, including incomplete solubilisation of full-length Dp427, use of monoclonal mouse antibodies on mouse tissue, absence of calcium from the lysis buffer, and partial disruption of the interaction between dystroglycan and dystrophin by the MANDAG2 antibody (Pereboev et al., 2001, Yoon et al., 2012). Overall the above experiments supported incorporation of ε-SG into brain DGC-like complexes, which would implicate DGC-like complexes in MD pathogenesis. DGC-like complexes have crucial roles during brain development and function, including in neuronal migration, synapse organisation and calcium homeostasis (Chapter 1 section 1.4.3) (Anderson et al., 2002, Waite et al., 2012). In particular, perturbations of calcium homeostasis have been directly implicated in both muscular dystrophies and dystonias: paralogues *ANO3* and *ANO5* encode calcium-activated proteins (potentially anion channels) that are mutated in autosomal dominant craniocervical dystonia (*DYT23*) and LGMD type 2L respectively (Bolduc et al., 2010, Charlesworth et al., 2012, Hicks et al., 2011). Loss of cell surface ε-SG in MD may

result in disruption of DGC-like complexes in the brain, leading to neuronal dysfunction through one or more pathways.

A hypothesis-free analysis of immunoaffinity purified ε-SG-containing complexes from mouse brain also identified a non-DGC protein that co-purified with ε-SG isoform 2: tenascin-R (section 3.2.6). Novel proteins interacting with ε-SG in the brain would provide additional insight into its function(s), but identifying these proteins would require careful analysis of IAP-MS datasets to eliminate proteins enriched through contamination, cross-reaction with the antibody used, and/or non-specific interaction with immunoglobulin or the Protein A-agarose support matrix. Ideally, this would be achieved by carrying out a parallel IAP-MS experiment on ε-SG-deficient tissue; only contaminants, cross-reacting and non-specifically interacting proteins would be enriched in an IAP performed on ε-SG deficient tissue using an anti-ε-SG antibody. As tissue from ε-SG-deficient mice could not be obtained for this study, cross-reacting proteins were identified through use of multiple antibodies and comparison of IAPs performed on digitonin versus RIPA-solubilised tissue, while non-specifically interacting proteins were identified in a pre-immune immunoglobulin control. This approach resulted in the identification of tenascin-R as co-purifying with ε-SG using both esg3788 and esg2-1358 antibodies (sections 3.2.5 and 3.2.6).

IAP of tenascin-R-containing complexes from digitonin-solubilised mouse brain resulted in co-purification of ε-SG isoform 2, confirming that tenascin-R did not cross-react with the anti-ε-SG antibodies (Figure 3.16). Although further experiments are required to confirm an interaction between tenascin-R and ε-SG isoform 2 in the brain, this could provide further clues to the function of ε-SG in the brain. Tenascin-R is an extracellular matrix proteoglycan principally involved in organisation of the perineuronal net (PNN) around a subset of CNS neurons (Anlar and Gunel-Ozcan, 2012, Chiovaro et al., 2015, Probstmeier et al., 2001). As part of PNNs, tenascin-R contributes to organisation of synapses including channel and

receptor clustering, synapse stability, extracellular ion concentrations, and presynaptic vesicle content (Anlar and Gunel-Ozcan, 2012, Aspberg et al., 1997, Bruckner et al., 2000, Geissler et al., 2013, Hagihara et al., 1999, Haunso et al., 2000, Morawski et al., 2014, Nikonenko et al., 2003, Oohashi et al., 2015, Weber et al., 1999). Therefore, if confirmed an interaction between ε-SG isoform 2 and tenascin-R might further implicate ε-SG in organisation of the synapses. Perturbation of synapses through disruption of both tenascin-R-containing complexes and DGC-like complexes by ε-SG deficiency might contribute to the pathogenesis of dystonia in *SGCE* mutation-positive MD.

In conclusion, I have shown that prototypical sarcoglycan heterotetramers containing different ε-SG isoforms exist in brain. My data from heterologous cells and previously published data in the δ-SG-deficient BIO14.6 hamster suggest a pool of residual ε-SG and ζ-SG may traffic and function independently in the presence of LGMD-associated sarcoglycan mutations (Waite et al., 2016). This could contribute to the absence of neurological signs in LGMD patients. In addition, ε-SG co-purified with β-dystroglycan and Dp71 from digitonin-solubilised mouse brain; co-purification of these proteins suggests that ε-SG may be a component of DGC-like complexes in the brain. Dysfunction of DGC-like complexes could contribute to the molecular pathogenesis of MD. Finally, ε-SG isoform 2 co-purified with tenascin-R from mouse brain, which could provide a first clue to non-DGC interactions of ε-SG in the brain. Further experiments to verify this result and determine whether tenascin-R and ε-SG interact could include immunohistochemistry to determine co-localisation, or *in vitro* knock-down and co-immunoprecipitation to determine whether these proteins interact and if so what effects disrupting this interaction might have. In addition, it is highly likely that additional ε-SG-interacting proteins exist in the brain. Refinements of the IAP protocol described here – use of different buffers to capture proteins differentially soluble in RIPA versus digitonin buffers, reduction of sample complexity for MS analysis, repetition of the

IAPs to obtain better coverage, and use of fresh tissue for all experiments – might help identify additional interacting proteins. Alternatively, distinct methods such as proximity-dependent labelling could be employed to identify proteins in close physical proximity to ε-SG that might interact with it (Carneiro et al., 2016). Regardless, the present study presents the first identification of proteins that may interact with ε-SG in the brain and provides a framework for future studies.

Chapter 4: Alternative splicing of sarcoglycan pre-mRNA

4.1.Introduction

In the previous chapter I showed that ε-SG interacts with β-, δ- and ζ-SG to form prototypical sarcoglycan complexes in the brain (Chapter 3 sections 3.2.2 and 3.2.4), yet only *SGCE* mutations cause neurological dysfunction (Ozawa et al., 2005). Several possible explanations for this have been proposed, including an as-yet unknown unique function for the brain-specific ε-SG isoform 2 produced through pre-mRNA alternative splicing (Ritz et al., 2011). Alternative splicing of *SGCA, SGCB, SGCD, SGCG* and *SGCZ* pre-mRNA could also contribute to differences in sarcoglycan protein levels and functions between tissues, but has not been studied systematically. Up to 95-100% of mammalian multi-exonic genes are thought to produce alternatively spliced transcripts, so all sarcoglycans were predicted to undergo alternative splicing (Berget et al., 1977, Chow et al., 1977, Early et al., 1980, Kornblihtt et al., 2013, Pan et al., 2008, Wang et al., 2008).

Alternative splicing can have a number of different effects on a transcript's product. If an alternative splicing event affects the transcript coding region without inhibiting translation, the encoded protein isoform might differ from the full-length isoform in properties such as localisation, protein interactions, stability, and protein motifs or domains (Kelemen et al., 2013). This could affect the assembly and function of sarcoglycan complexes. By contrast, alternative splicing events that change the sequence of untranslated regions (UTRs) might affect translation by changing mRNA localisation, stability, translation kinetics, or nuclear export (Kelemen et al., 2013, Zheng and Black, 2013). Finally, alternative splicing events that introduce premature termination codons (PTCs) into the transcript coding region more than 50-55 nucleotides (nt) upstream of the final exon:exon junction can trigger nonsense-mediated decay or translational repression (NMD/R) of the transcript that leaves the

transcript unproductive (untranslatable) (Isken and Maquat, 2007, Lewis et al., 2003, Mendell et al., 2004, Skandalis et al., 2010). While some unproductive transcripts may represent splicing errors, unproductive transcripts and those with UTR sequence changes can modulate protein production from a transcriptionally active gene (Lewis et al., 2003). For example, the splicing factor SRSF4 directly binds and induces inclusion of a so-called "poison" exon into its own transcript, which introduces a PTC and results in reduced protein synthesis (Änkö et al., 2012). Therefore, alternative splicing of sarcoglycan pre-mRNA could affect both sarcoglycan protein levels and functions to modulate the composition and/or function of sarcoglycan complexes.

In order to identify alternatively spliced human *SGCA, SGCB, SGCD, SGCG* and *SGCZ* transcripts, I used an RT-PCR-based approach to screen human adult cerebellum and skeletal muscle tissue for transcripts derived from these genes. Human tissue was used so that alternatively spliced transcripts could be related directly to LGMD and MD. In addition, studies on other genes have suggested that overall levels of alternative splicing may be greater in humans than in mice (Barbosa-Morais et al., 2012, Kim et al., 2007). Cerebellum was selected for its implication in the aetiology of MD by both physiological studies and high ε-SG isoform 2 levels, while skeletal muscle contains a prototypical sarcoglycan heterotetramer and is the tissue predominantly affected by LGMD (Marelli et al., 2008, Ozawa et al., 2005, Popa et al., 2014, Ritz et al., 2011, van der Meer et al., 2012, van der Salm et al., 2013, Yokoi et al., 2012a, Yoshida et al., 1994).

4.2.Results

To identify alternatively spliced *SGCA, SGCB, SGCD, SGCG* and *SGCZ* transcripts, poly-d(T)-primed first strand cDNA synthesis followed by gene-specific PCR was used to generate cDNA mini-libraries from human adult cerebellum and skeletal muscle total RNA (Chapter 2

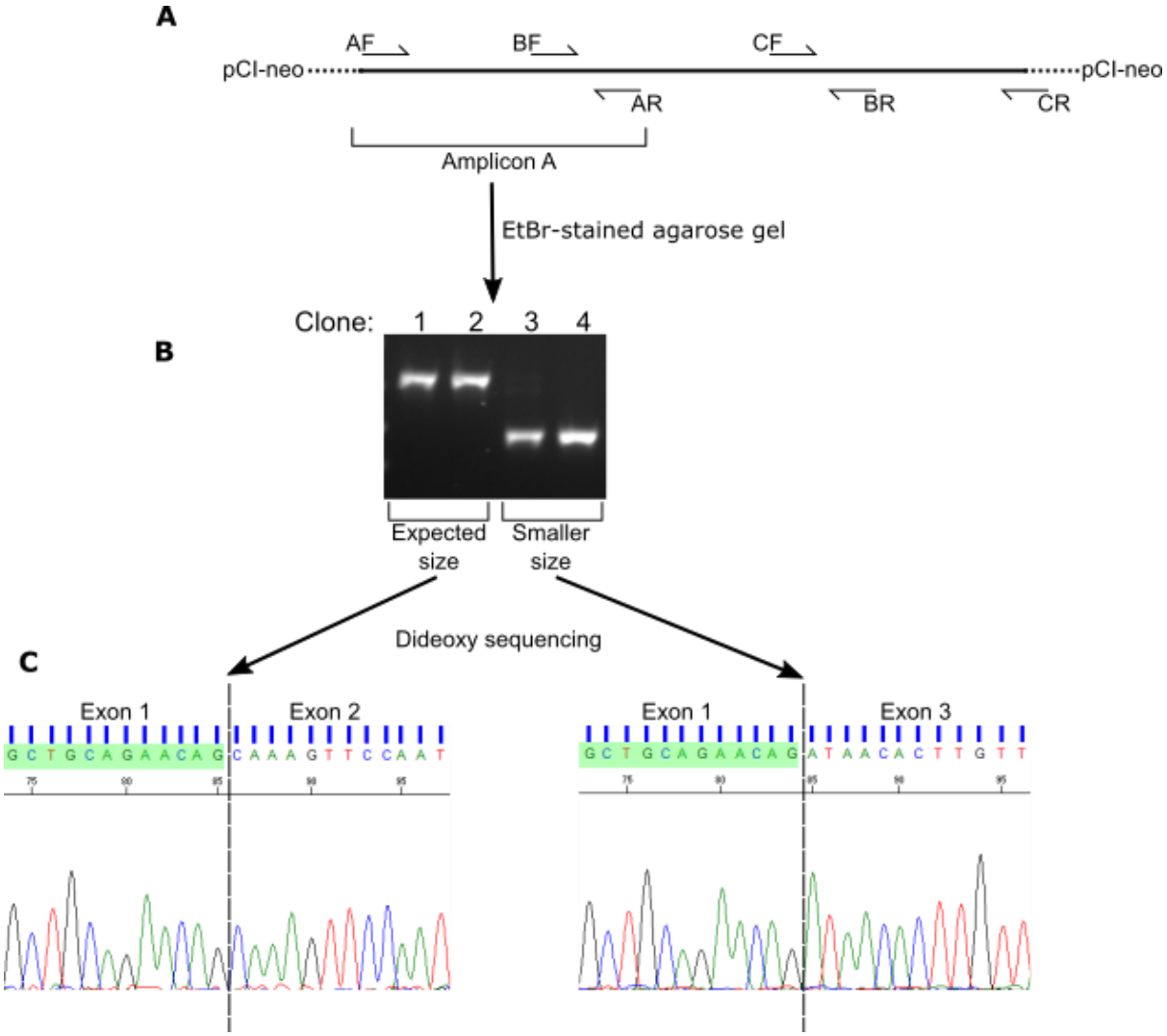


Figure 4.1 Overview of protocol for identification of alternatively spliced sarcoglycan transcripts. Gene-specific RT-PCR of sarcoglycan cDNA was performed to produce mini-libraries of cDNAs corresponding to transcripts from the sarcoglycan genes in skeletal muscle and cerebellum. Mini-library clones were then screened for alternatively spliced transcripts. **A)** Schematic of a cDNA inserted into the plasmid vector pCI-neo; half arrows mark sites of primers for overlapping PCR amplicons covering the entire cDNA insert. "Amplicon A" represents the region amplified using primers AF and AR. **B**) Insert sizes were determined via agarose gel electrophoresis of PCR products. For Amplicon A, two sizes of PCR product were observed: the expected size for clones 1 and 2, and a smaller product for clones 3 and 4. **C**) Clones were sequenced to identify the transcript insert. Clones 1 and 2 represented the full-length major transcript, while clones 3 and 4 represented a transcript lacking exon 2 (gene transcript Δ exon 2).

sections 2.1.1-2.1.3, 2.1.7, 2.2.8-2.2.11). Use of both brain and muscle tissue facilitated screening of the muscle-specific *SGCA* and *SGCG* transcripts in addition to predominantly non-muscle *SGCZ* (Barresi et al., 2000b, Liu et al., 1997, Wheeler et al., 2002). The resultant cDNA mini-libraries were screened by overlapping PCR and then sequencing of cDNA inserts to identify alternatively spliced cDNAs (Chapter 2 sections 2.1.1-2.1.3, 2.1.7, 2.2.8-2.2.11) (Figure 4.1). Promoters and transcription termination sites for each gene were obtained from reference sequences in NCBI RefSeq, ENSEMBL, and Leiden Open Variation

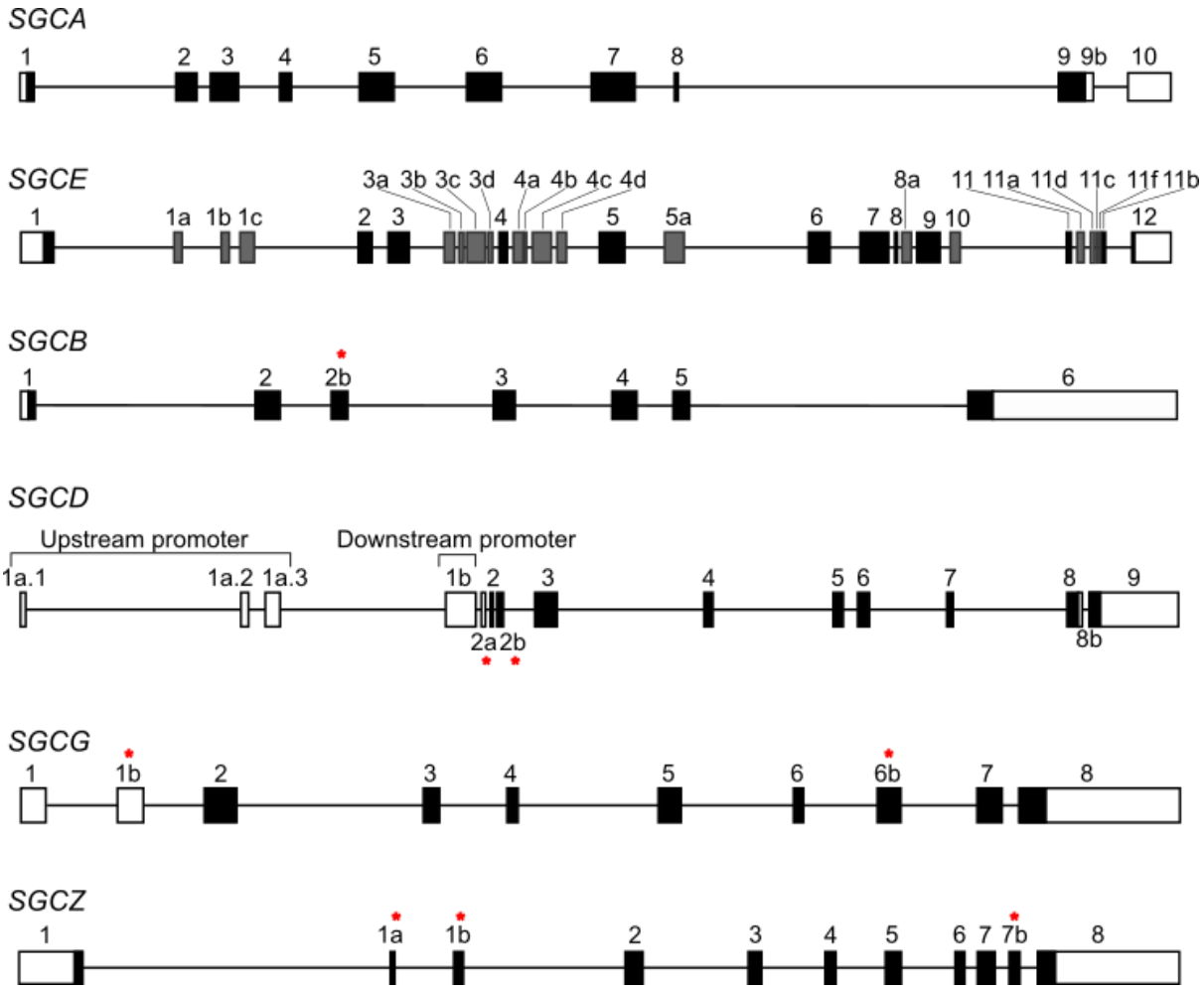


Figure 4.2 Overview of genomic architecture for the human *SGCA, SGCE, SGCB, SGCD, SGCG* and *SGCZ* genes. *SGCE* data from (Ritz et al., 2011). All known exons – both previously annotated and identified in the present study – are depicted. Black or grey fill indicates potentially coding sequence; grey indicates exons typically excluded from the mature transcript. Figure is not to scale, although relative sizes and positions of exons are approximately correct. * denotes novel exons identified in the present study.

Database (LOVD) (Brown et al., 2015, Flicek et al., 2014, Fokkema et al., 2011, Harrow et al., 2012). Figure 4.2 shows the genomic architecture for each sarcoglycan gene. *SGCA, SGCB, SGCG* and *SGCZ* each had one annotated promoter and one transcription termination site, and therefore required a single cDNA mini-library per tissue. By contrast, *SGCD* had two annotated promoters and two transcription termination sites (Figure 4.2). While the termination sites were both well-supported in expressed sequence tag databases, ENCODE annotation for the upstream promoter was weaker than for the downstream (proximal to exon 2) promoter (Figure 4.3) (Rosenbloom et al., 2015, Rosenbloom et al., 2013). However, cDNA from both promoters was detected in the human tissue samples so transcripts from both were screened.

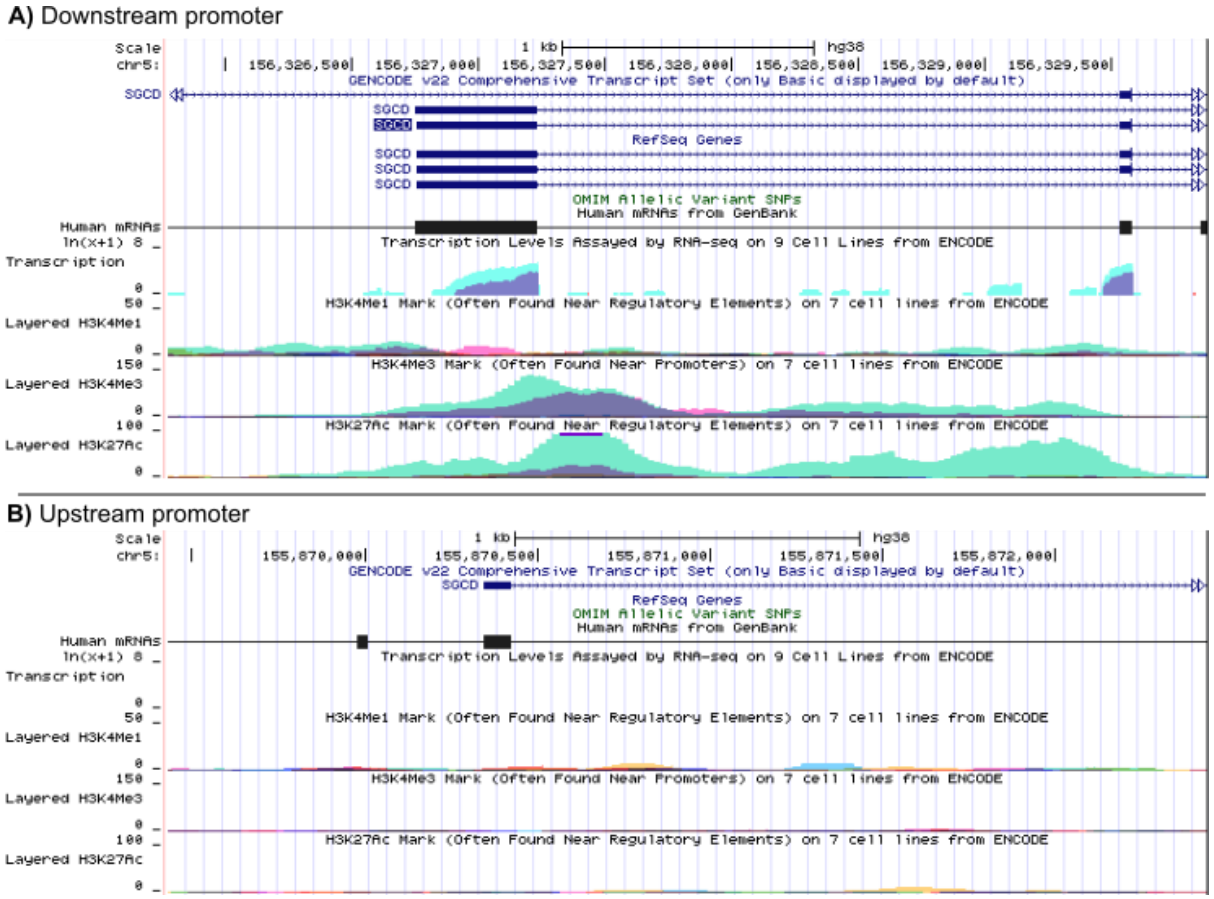


Figure 4.3 Transcription activity markers around the *SGCD* promoters from ENCODE. Screenshot of the UCSC Genome Browser (http://genome.ucsc.edu/) display of annotated histone modifications and transcription levels assessed in cell lines by the ENCODE project (Consortium, 2012, Kent et al., 2002, Rosenbloom et al., 2013). A) Downstream/exon 2 proximal promoter for *SGCD*. B) Upstream/distal promoter.

Each transcript was described relative to the full-length, major transcript from that gene as annotated in LOVD and RefSeq (Brown et al., 2015, Fokkema et al., 2011). A transcript containing the novel exon 1b in addition to the canonical exons would be described as *SGCx* exon 1b+ transcript, whereas a transcript lacking exon 2 would be described as the *SGCx* Δ exon 2 transcript. RT-PCR allowed splicing events to be identified in the context of the entire transcript, rather than in isolation; therefore, complex alternative splicing events involving multiple sequences could be identified. This method also provided an estimate of relative transcript abundance, as transcripts abundant in tissue should be more abundant in a cDNA mini-library derived from that tissue. However, PCR can preferentially amplify shorter targets, which would bias the cDNA mini-library in favour of shorter transcripts. This precluded accurate calculation of transcript abundance.

4.2.1. Few alternatively spliced *SGCB* transcripts were identified in cerebellum and skeletal muscle

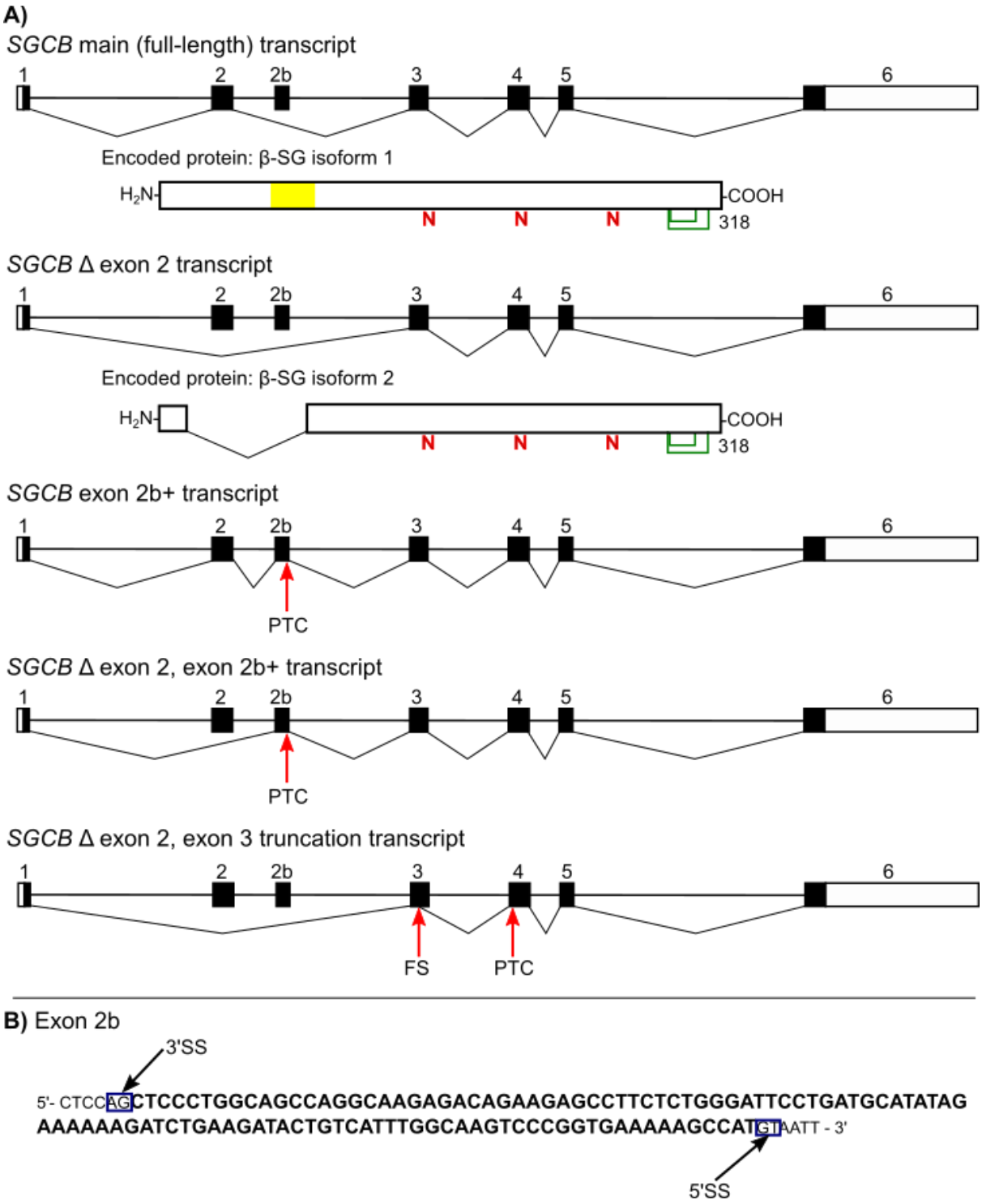


Figure 4.4 *SGCB* transcripts and novel exons identified in human cerebellum and skeletal muscle. A) For each *SGCB* transcript identified in this study, the splicing pattern is displayed against the genomic architecture of the gene. Thin lines connect exons included in the mature transcript. Features such as premature termination codons (PTC) and frameshift sites (FS) are annotated. For translated transcripts, the encoded polypeptide with major features annotated is depicted below the splicing pattern. Yellow indicates a transmembrane domain, N represents N-linked glycosylation sites, and green lines connect cysteine residues involved in disulphide bonds. B) Sequence and splice sites for the novel exon 2b. Abbreviations: 5'SS, 5' splice site; 3'SS, 3' splice site.

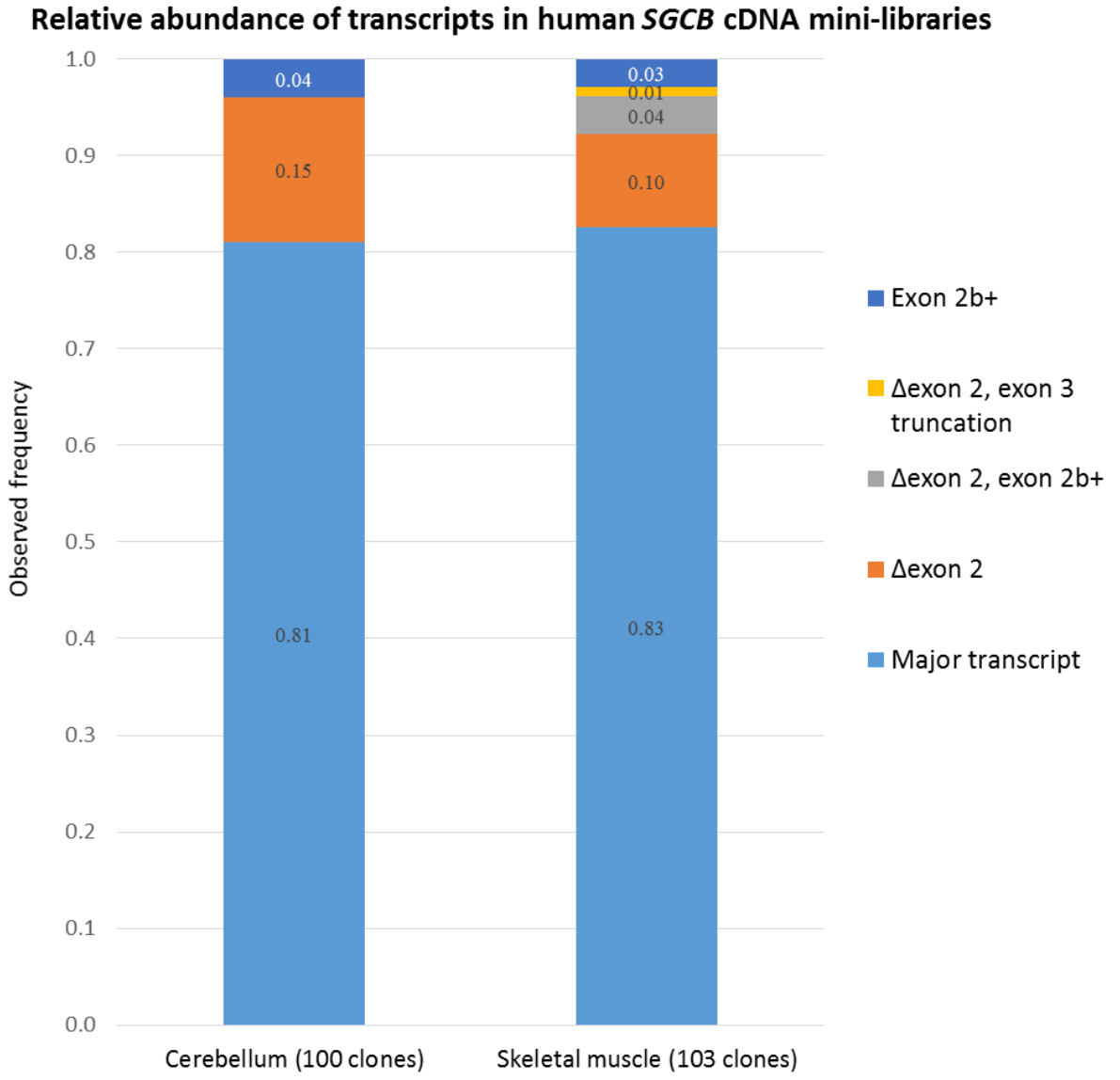


Figure 4.5 *SGCB* transcripts identified in human cerebellum and skeletal muscle. Stacked bar chart representing the abundance of each transcript (described in figure 4.4) identified for the *SGCB* gene in mini-libraries. All transcripts are described relative to the full-length, major transcript.

In total, five distinct alternatively spliced *SGCB* transcripts were identified in 100 cerebellum cDNA mini-library clones and 103 skeletal muscle cDNA mini-library clones (Figure 4.4). The major transcript, RefSeq sequence NM_000232.4, was observed at similar abundance in each mini-library: 81% of cerebellum cDNA mini-library clones, and 83% in skeletal muscle (Figure 4.5). Also present in both mini-libraries was a *SGCB* Δ exon 2 transcript (exon 2 skip) predicted to encode an alternative β-SG isoform 2 without the transmembrane domain found in full-length isoform 1 (Table 4.1, Figure 4.4). A novel exon 2b with sequence similarity to the HAL1 transposable element was detected in 4-8% of clones per mini-library,

but introduced a PTC predicted to stimulate NMD/R of the transcript (Table 4.1, Figures 4.4 and 4.6a) (Kent, 2002, Smit et al., 2013-2015).

SGCB transcript	Effect on mRNA	Type of alternative splicing event	Observed frequency		Predicted effect on encoded protein	Isoform	NCBI RefSeq accession
			Cerebellum (100 clones)	Skeletal muscle (103 clones)			
Major transcript	-	-	0.81	0.83	Full-length protein isoform 1: transmembrane domain at I66-I86; N-linked oligosaccharides at N158, N211 and N258; disulphide bonds between C288:C314 and C290:C307; 318aa	1	**NM_000232.4**
Δ exon 2	Omission of exon 2	Exon skipping	0.15	0.10	Isoform 2 with in-frame loss of Q12-I81 (69aa), removing transmembrane domain but leaving glycosylation sites and disulphide bonds unaffected; 248aa	2	**XM_011534403.1**
Δ exon 2, exon 2b+	Omission of exon 2 and inclusion of novel exon 2b (110nt)	Exon skipping and alt. exon inclusion		0.04	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-
Δ exon 2, exon 3 truncation	Omission of exon 2; omission of 97nt from 3' end exon 3	Exon skipping and use of alternate 5' splice site		0.01	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-
Exon 2b+	Inclusion of novel exon 2b (110nt)	Alt. exon inclusion	0.04	0.03	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-

Table 4.1 Summary of *SGCB* transcripts identified in human cerebellum and skeletal muscle. Over two mini-libraries of human *SGCB* transcripts derived from cerebellum and skeletal muscle total RNA respectively, five distinct transcripts were identified. This table summarises the major characteristics of each transcript: the effect on the mRNA, the type of alternative splicing event involved, the frequency of the transcript in each *SGCB* transcript mini-library, the predicted effect on the protein relative to the major isoform 1, and the predicted protein size. Abbreviations: Alt. exon inclusion, alternative or novel exon inclusion in RNA; PTC, premature termination codon; NMD/R, nonsense-mediated decay/repression; aa, amino acid.

A) *SGCB* exon 2b aligned to HAL1 transposon

```
SGCBexon2b          12 CCAGGCAAGAGAC-AGAAGAGCCTTCTCTGGGATTCCTGAT--GCATATA 58
                              i    i-i      i          ivvi    i--  vi -
HAL1#LINE/L1       636 CCAGGCAGGAGATTGGAAGAGTCTTCTCTGGGGAATCTGACCAGCCCA-A 684

SGCBexon2b          59 GAAAAAAGATCTGAAGATACTGTCATTTGG 88
                         ii     i  i         v    v
HAL1#LINE/L1       685 GAGGAAAGACCTAAAGATACTGACATTAGG 714
```

B) *SGCD* exon 2a aligned to AluSq2 transposon

```
SGCDexon2a            1 ACGGAGTTTCGCTCTTGTTTCCCAGACTGGAGTGCAATGGCATGATCTCT 50
                                        iv     i               ii      -
AluSq2#SINE/A       280 ACGGAGTTTCGCTCTTGTCGCCCAGGCTGGAGTGCAATGGCGCGATCTC- 232

SGCDexon2a           51 GGCTCACATCCACCTCCGCCTCCCGGGTTCAAGCCATTCTTCTGCCTCAG 100
                              vv v                          v     i
AluSq2#SINE/A       231 GGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAG 182

SGCDexon2a          101 CCTCCCGAGTATTAATCCCCTCCTGGGATTACAG 134
                                v----------
AluSq2#SINE/A       181 CCTCCCGAGTAG----------CTGGGATTACAG 158
```

C) *SGCD* exon 2b aligned to AluY transposon

```
SGCDexon2b            1 GCCAGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAG 50
                           i   i
AluY#SINE/Alu         2 GCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAG 51

SGCDexon2b           51 GCAGGTGGATCACGAG 66
                          i  i
AluY#SINE/Alu        52 GCGGGCGGATCACGAG 67
```

D) *SGCZ* exon 1b aligned to ORF2 of the L1M2 transposon

```
SGCZexon1b            1 CACAATTTATTGAAGAGACTGTCTTATCCCCAA----TGTTCTTGGCATC 46
                          v                    i v      ?----           i
L1M2_orf2#LIN      2804 CACCATTTATTGAAGAGACTGTCCTTTCCCCANTGTGTGTTCTTGGCACC 2755

SGCZexon1b           47 TTGGTTGAAAATCAATTGGCT 67
                          v  i        i
L1M2_orf2#LIN      2754 TTTGTCGAAAATCAGTTGGCT 2734
```

E) *SGCG* exon 6b aligned to AluYc transposon

```
SGCGexon6b            1 ACAGAGTCTTGCTGTGTTGCCCAGGCTGGAGTGCAGTGACACGATCTCGG 50
                          i      i   v   i                    i i
AluYc#SINE/Al       266 ACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCGG 217

SGCGexon6b           51 CTCACCGCAAGCTCCGCCTCCTGGGTTCACGCCATTCTCCTGCCTCAGCC 100
                             i              i
AluYc#SINE/Al       216 CTCACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCC 167

SGCGexon6b          101 TCCCGAGTAGCTGGGACTACAGGCGCCCGCCACCACACCCGG 142
                                                    i i
AluYc#SINE/Al       166 TCCCGAGTAGCTGGGACTACAGGCGCCCGCCACCGCGCCCGG 125
```

Figure 4.6 Alignment of novel exons to transposable elements using RepeatMasker. Each novel exon was screened using RepeatMasker to identify any similarity to transposable elements (transposons) (Smit et al., 2013-2015). Five exons showed similarity to transposons, and the alignments are shown here: A) *SGCB* exon 2b to the HAL1 LINE. B) *SGCD* exon 2a to AluSq2. C) *SGCD* exon 2b to AluY. D) *SGCZ* exon 1b to L1M2 retrotransposon ORF2. E) *SGCG* exon 6b to AluYc. Abbreviations: i, transition at that position between transposable element and exon; v, transversion at that position between transposable element and exon.

4.2.2. Numerous alternatively spliced *SGCD* transcripts were identified in cerebellum and skeletal muscle

Since *SGCD* has two promoters and two transcription termination sites as discussed above, four combinations of promoter and transcription termination site were possible for transcripts (Figure 4.2). However, in this study cDNA from only three could be detected: upstream promoter with exon 9 termination site, downstream promoter with exon 9 termination site, and downstream promoter with exon 8b termination site (Figure 4.7). Therefore, three *SGCD* cDNA mini-libraries per tissue were generated and screened for alternative splicing. No transcripts originating in the upstream promoter and terminating in exon 8b could be detected, but may still exist at low levels. Promoter usage only affected the 5' UTR sequence, but *SGCD* transcription termination site altered the C-terminus of the encoded δ-SG protein. Transcripts terminating in exon 8b encoded a δ-SG isoform 2 with a truncated C-terminus compared to δ-SG isoform 1 encoded by transcripts terminating in exon 9 (Figure 4.8).

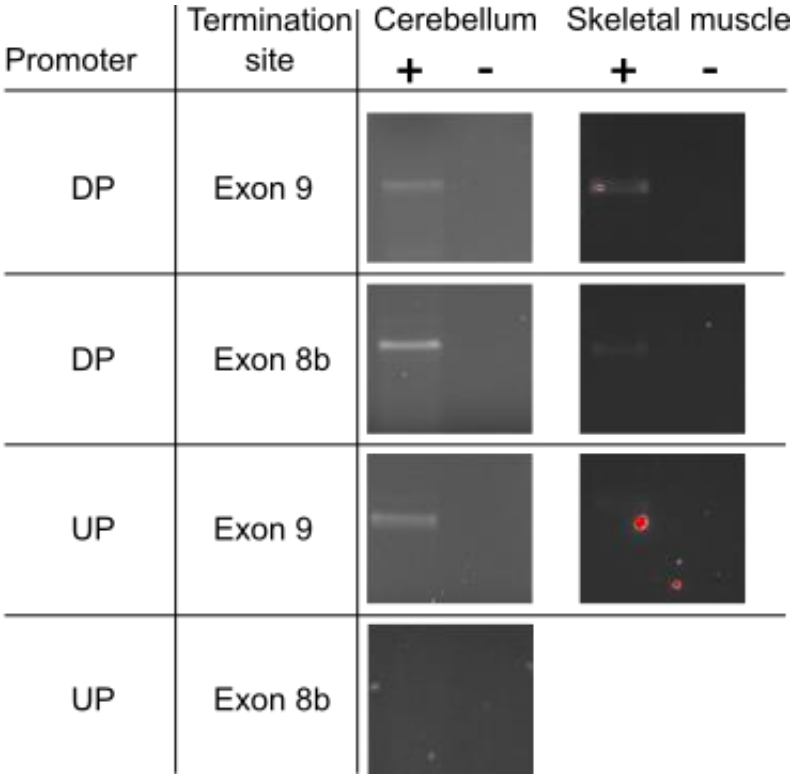
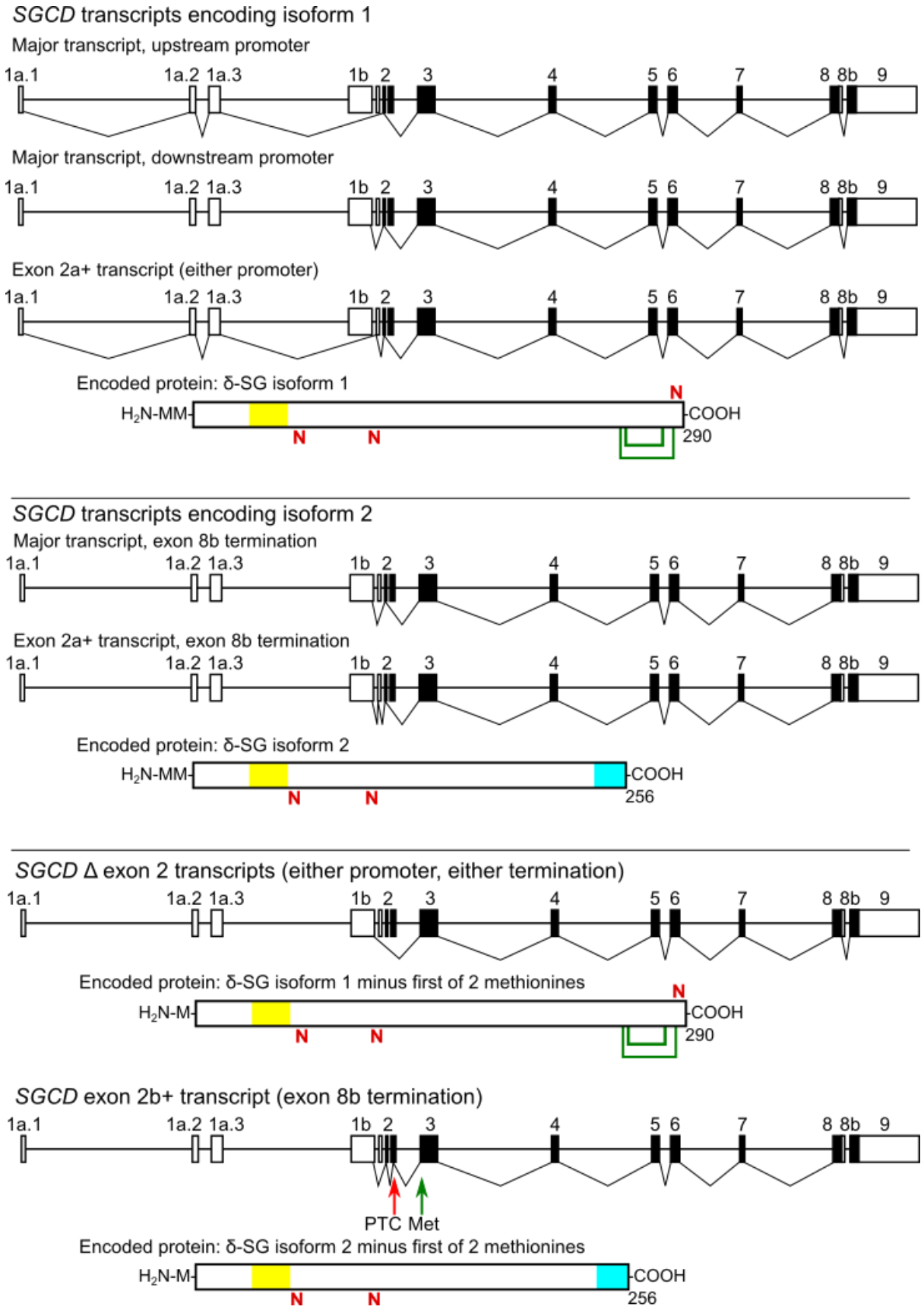


Figure 4.7 RT-PCR of *SGCD* cDNA from human cerebellum and skeletal muscle. *SGCD* cDNA was isolated using promoter- and terminal exon-specific primers from first strand cDNA. Each PCR reaction was resolved on an agarose gel stained using ethidium bromide. Products could be detected using three combinations of primers: downstream promoter and exon 9 termination, downstream promoter and exon 8b termination, upstream promoter and exon 9 termination. The final pair, upstream promoter and exon 8b termination, was not successfully amplified from the sample. Abbreviations: DP, downstream promoter; UP, upstream promoter; +, cDNA in PCR; -, water control PCR.


SGCD transcripts encoding isoform 1
Major transcript, upstream promoter
1a.1 1a.2 1a.3 1b 2 3 4 5 6 7 8 8b 9
Major transcript, downstream promoter
1a.1 1a.2 1a.3 1b 2 3 4 5 6 7 8 8b 9
Exon 2a+ transcript (either promoter)
1a.1 1a.2 1a.3 1b 2 3 4 5 6 7 8 8b 9
Encoded protein: δ-SG isoform 1
H2N-MM
COOH
290
N
N
N
SGCD transcripts encoding isoform 2
Major transcript, exon 8b termination
1a.1 1a.2 1a.3 1b 2 3 4 5 6 7 8 8b 9
Exon 2a+ transcript, exon 8b termination
1a.1 1a.2 1a.3 1b 2 3 4 5 6 7 8 8b 9
Encoded protein: δ-SG isoform 2
H2N-MM
COOH
256
N
N
SGCD Δ exon 2 transcripts (either promoter, either termination)
1a.1 1a.2 1a.3 1b 2 3 4 5 6 7 8 8b 9
Encoded protein: δ-SG isoform 1 minus first of 2 methionines
H2N-M
COOH
290
N
N
N
SGCD exon 2b+ transcript (exon 8b termination)
1a.1 1a.2 1a.3 1b 2 3 4 5 6 7 8 8b 9
PTC Met
Encoded protein: δ-SG isoform 2 minus first of 2 methionines
H2N-M
COOH
256
N
N

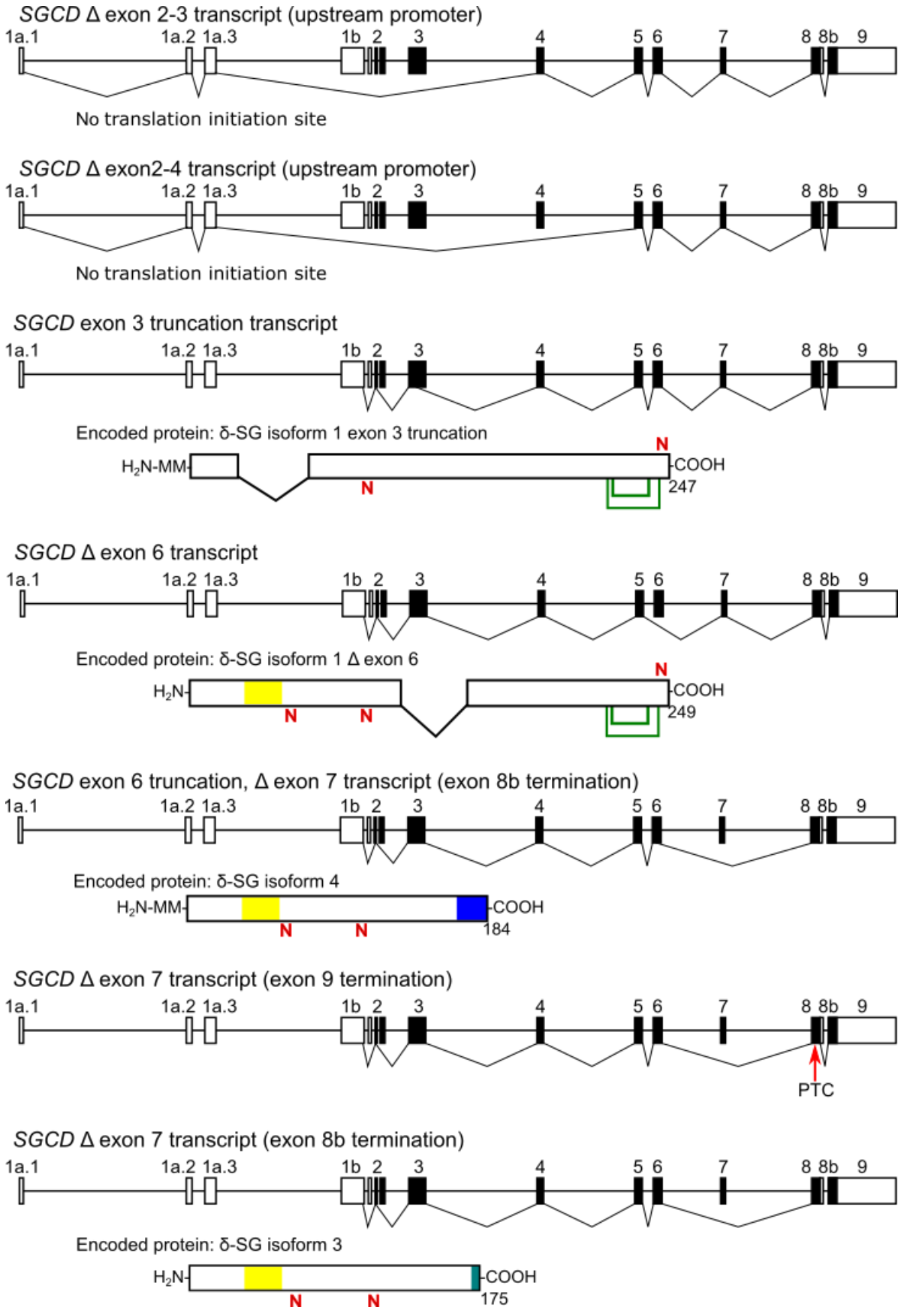


Figure 4.8 *SGCD* transcripts identified in human cerebellum and skeletal muscle cDNA mini-libraries. For each *SGCD* transcript identified in this study, the splicing pattern is displayed against the genomic architecture of the gene. Exons

included in the transcript are connected, while features such as premature termination codons and frameshift sites are annotated. For translated transcripts, the encoded polypeptide with major domains is depicted below the transcript. Yellow indicates a transmembrane domain, N represents N-linked glycosylation sites, blue and turquoise indicate isoform-specific sequences, and green lines connect cysteine residues involved in disulphide bonds. Abbreviations: PTC, premature termination codon; FS, frameshift; Met, methionine translation initiation codon.

All three *SGCD* transcript mini-libraries in both skeletal muscle and cerebellum contained alternatively spliced transcripts (Figure 4.9 and Table 4.2). In both tissues, cDNA mini-libraries of transcripts from the upstream promoter had a greater overall abundance of alternatively spliced transcripts, particularly *SGCD* Δ exon 2 transcripts (Figure 4.9). The *SGCD* Δ exon 2 transcript was in fact the most abundant alternatively spliced transcript overall, and has previously been described (Figure 4.9, Table 4.2) (Jung et al., 1996, Nigro et al., 1996). However, it was not predicted to substantially affect the encoded protein (Figure 4.8). Two additional alternatively spliced transcripts were also detected at greater than 5% frequency in one or more cDNA mini-libraries, suggesting they may be physiologically relevant (Figure 4.9, Table 4.2). The *SGCD* exon 3 truncation transcript lacked the 3' 129nt from exon 3 due to alternative 5' splice site usage, and was predicted to encode a δ-SG isoform missing residues 22-64 including the transmembrane domain (Figure 4.8, Table 4.2). This transcript was identified in 8-12% of clones from upstream promoter cDNA mini-libraries (Figure 4.9, Table 4.2). The *SGCD* Δ exon 7 transcript was less abundant at 8-9% of cerebellum cDNA mini-library clones and only 1% of skeletal muscle cDNA mini-library clones, but was predicted to introduce a frameshift and PTC to the coding region (Figure 4.9, Table 4.2). If the *SGCD* Δ exon 7 transcript terminated in exon 9, the PTC would be more than 50nt upstream of the final exon:exon junction and therefore should induce NMD/R (Figure 4.8). By contrast, if the *SGCD* Δ exon 7 transcript terminated in exon 8b the PTC would be within the final exon and therefore the transcript was predicted to encode a δ-SG isoform 3 with a truncated extracellular region (Figure 4.8).

All other alternatively spliced transcripts identified in *SGCD* cDNA mini-libraries were low abundance, at less than 5% of cDNA mini-library clones each (Table 4.2). In addition, most

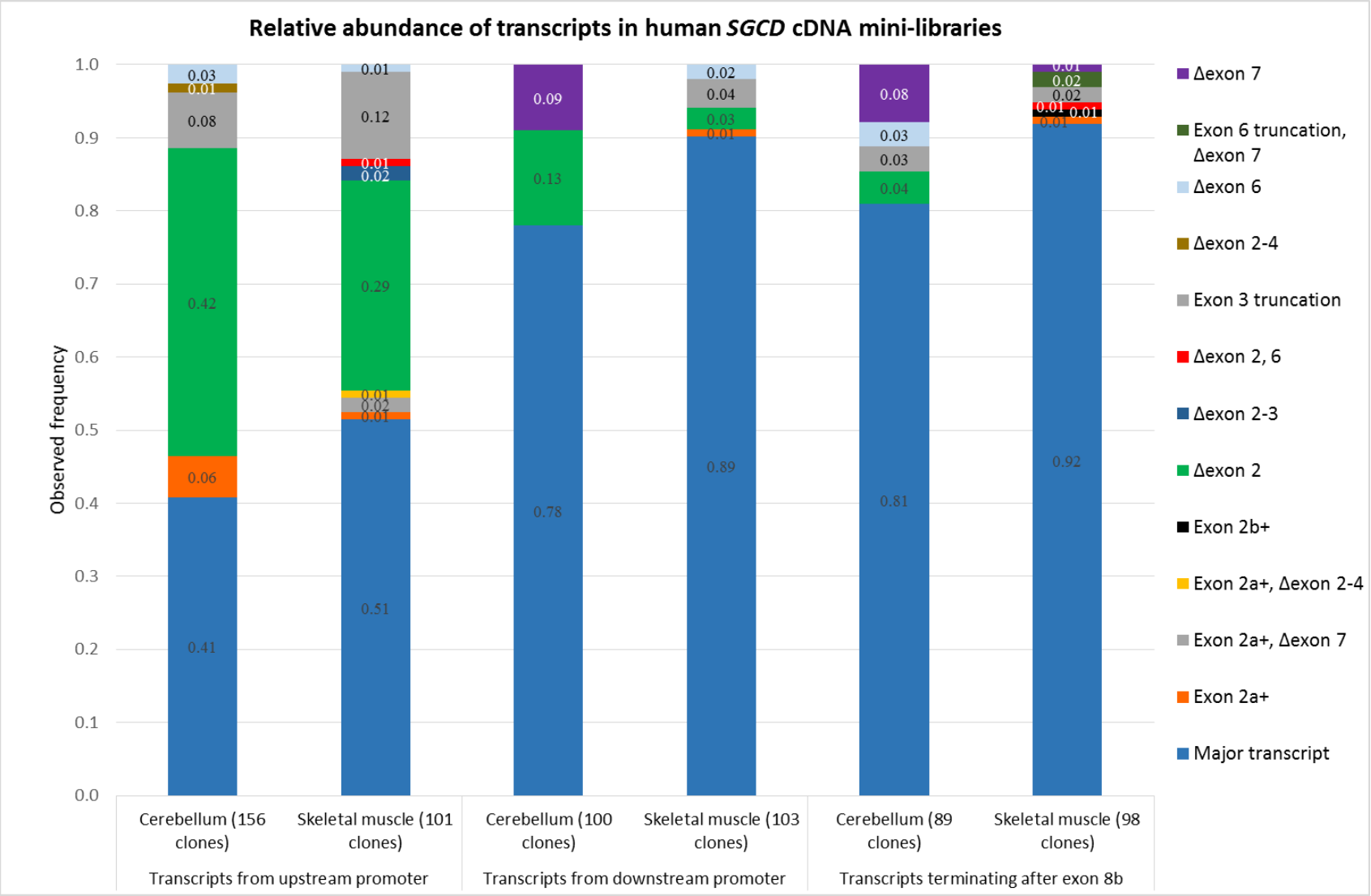


Figure 4.9 Relative abundance of discrete transcripts in *SGCD* cDNA mini-libraries from cerebellum and skeletal muscle. Stacked bar chart representing the abundance of each transcript (described in figure 4.7) identified for the *SGCD* gene in mini-libraries. All transcripts are described relative to the full-length, major transcript.

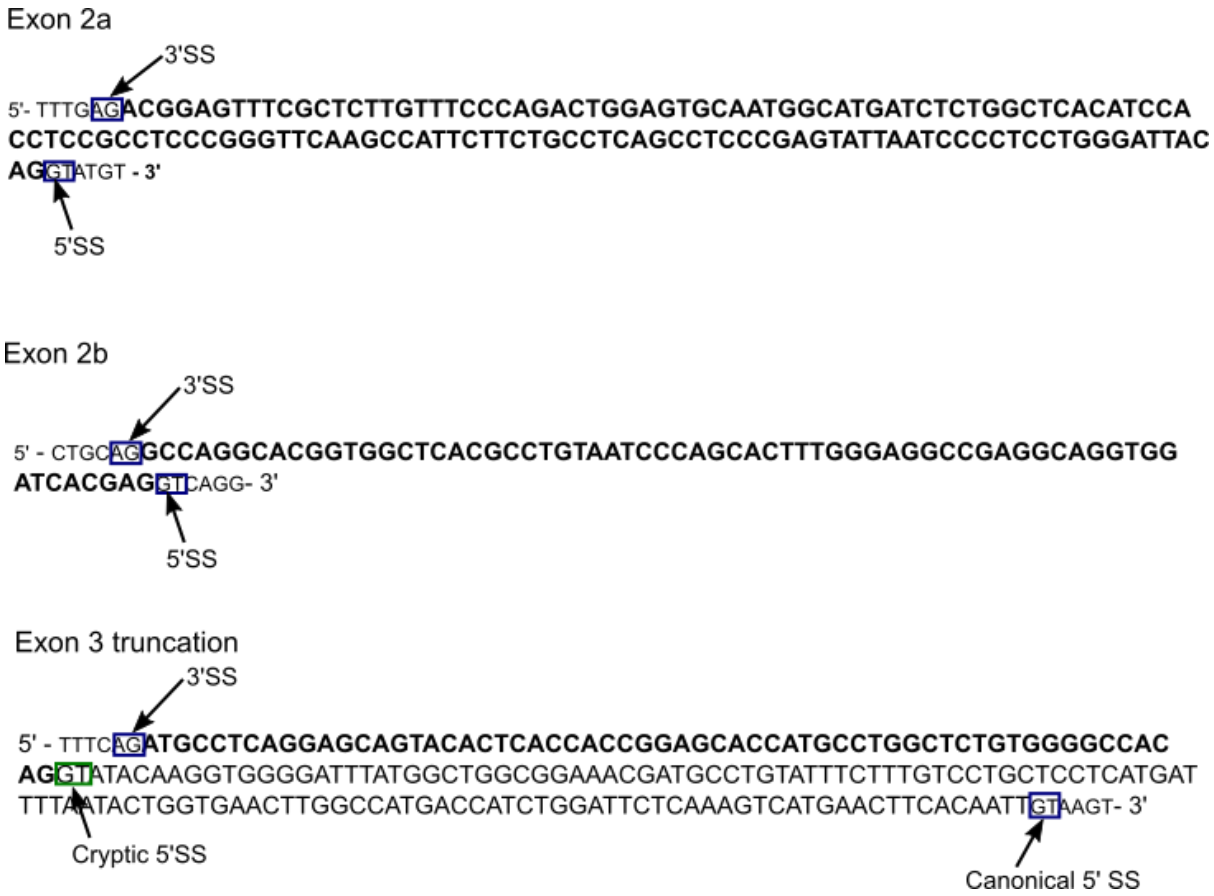


Figure 4.10 Novel exons identified in *SGCD* transcripts from human cerebellum and skeletal muscle. DNA sequences for novel and truncated exons. The 5' splice sites (SS) and 3' splice sites (SS) both cryptic and canonical are indicated in the sequence.

were not predicted to be translated (Figure 4.8). An exception was the *SGCD* Δ exon 6 transcript, which although rare in the cDNA mini-libraries was predicted to encode a δ-SG isoform lacking residues 128-167 from the extracellular region (Figure 4.8, Table 4.2). In addition, six low-abundance (1-2% of cDNA mini-library clones) alternatively spliced transcripts were identified in skeletal muscle but not cerebellum cDNA mini-libraries (Table 4.2). There were also two novel *SGCD* exons among the low abundance alternatively spliced transcripts: exons 2a and 2b (Figure 4.10). Both of these exons showed high similarity to Alu transposable elements, albeit of different families (Figure 4.6b,c) (Kent, 2002). Exon 2b appears to be derived from an Alu element as annotated in the TranspoGene database (Levy et al., 2008). Exon 2a affected only the 5' UTR while exon 2b slightly changed the coding region as summarised in table 4.2. However, these exons were rare in the cDNA mini-libraries.

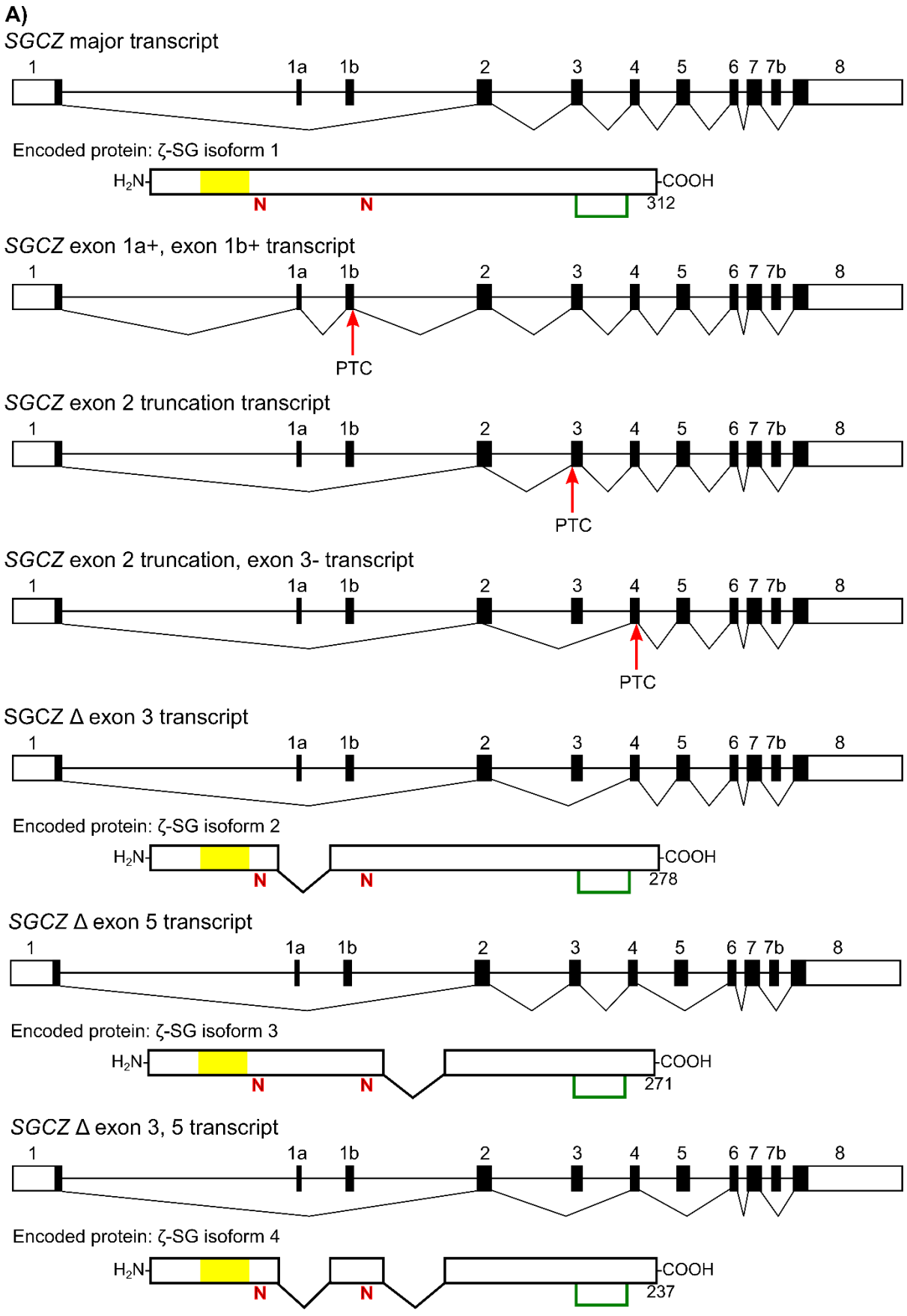
SGCD transcript	Effect on the mRNA	Alternative splicing event	UP		DP		Ex8b ter.		Predicted effect on protein and RefSeq accession	
			CB	SM	CB	SM	CB	SM	Isoform 1 protein (Exon 9 termination)	Isoform 2 protein (Exon 8b termination)
Major transcript	-	-	0.41	0.51	0.78	0.89	0.81	0.92	Full-length **isoform 1**: 290aa; transmembrane domain F37-L57; N-linked glycans at N61, N109 and N285; disulphide bonds at C264:C289 and C266:C282. DP: **NM_000337.5**; UP: **XM_017009724.1**	Full-length **isoform 2**: 256aa; transmembrane domain F37-L57; alternate sequence after E233 compared to isoform 1; N-linked glycans at N61 and N109; no disulphide bonds. **NM_172244.2**
Exon 2a+	Insertion of ~134bp sequence into 5' UTR	Alt. exon inclusion	0.06	0.01		0.01		0.01	Full-length **isoform 1**	Full-length **isoform 2**
Exon 2b+	Insertion of 66bp sequence after exon 2	Alt. exon inclusion						0.01		Introduces PTC immediately following exon 2 initiation site, but translation can initiate from exon 3 for a full-length 255aa **isoform 2** protein lacking the first methionine.
Δ exon 2	Omission of exon 2	Exon skipping	0.42	0.29	0.13	0.03	0.04		Use of exon 3 translation initiation site; protein lacks 1st methionine but is otherwise identical to full-length **isoform 1**; 289aa. **NM_001128209.1**	Use of exon 3 translation initiation site; protein lacks 1st methionine but is otherwise identical to full-length **isoform 2**; 255aa
Δ exon 2-3	Omission of exons 2 and 3	Exon skipping		0.02					No translation initiation site – non-productive transcript	
Δ exon 2-4	Omission of exons 2-4	Exon skipping	0.01						No translation initiation site – non-productive transcript	
Exon 3 truncation	Lacks 3' 129bp from exon 3	Use of alternate 5' splice site	0.08	0.12		0.04	0.03	0.02	Omission of V22-I64 inclusive including 1st N-glycan; lose transmembrane domain; 247aa. **Isoform 1 exon 3 tr**.	Omission of V22-I64 inclusive including 1st N-glycan; lose transmembrane domain; 213aa. **Isoform 2 exon 3 tr**.

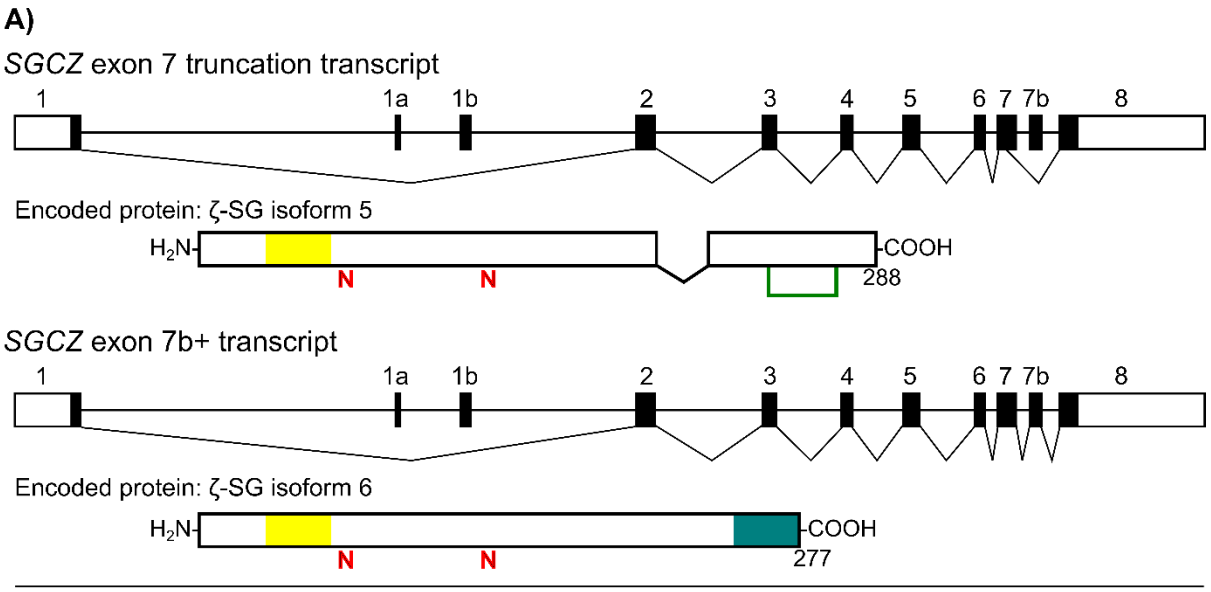
Δ exon 6	Omission of exon 6	Exon skipping	0.03	0.01		0.02	0.03		Omission of G128-L167 inclusive; transmembrane domain, N-linked glycans and disulphide bonds un-affected; 249aa. **Isoform 1 Δ exon 6**	Omission of G128-L167 inclusive; transmembrane domain and N-linked glycans unaffected; 215aa. **Isoform 2 Δ exon 6**
Exon 6 truncation, Δ exon 7	Lacks 3' 82nt from exon 6; omission of exon 7	Use of alternate 5' splice site and exon skipping						0.02		Introduces frameshift in exon 6; unique extracellular domain sequence after V141 with termination after G184; transmembrane domain and first two glycosylation sites maintained; 184aa; **isoform 4**
Δ exon 7	Omission of exon 7	Exon skipping			0.09		0.08	0.01	Introduces frameshift and a PTC >55nt downstream of the final exon:exon junction – NMD/R target	Introduces frameshift; unique extracellular domain sequence after G168 with termination after L175; final N-linked glycosylation site and both disulphide bonds lost; 175aa: **isoform 3**
Exon 2a+, Δ exon 7	Insertion of ~134bp sequence into 5' UTR; omission of exon 7	Alt. exon inclusion; exon skipping		0.02					Frameshift and alternate sequence after G168 with termination after L175; retain N-linked glycans at N61 and N109 but lose final site and both disulphide bonds. **Isoform 1 Δ exon 7**	
Exon 2a+, Δ exon 2-4	Insertion of ~134bp sequence into 5' UTR; omission of exons 2-4	Alt. exon inclusion; exon skipping		0.01					Introduction of frameshift and PTC – NMD/R target, non-productive.	
Δ exon 2, 6	Omission of exons 2 and 6	Exon skipping		0.01				0.01	Omission of 1st methionine and G128-L167 inclusive; transmembrane domain, N-linked glycans and	Omission of 1st methionine and G128-L167 inclusive; transmembrane domain and N-linked glycans unaffected; 215aa. **Isoform 2 Δ exon 2, 6**.

									disulphide bonds remain unaffected; 249aa. **Isoform 1 Δ exon 2, 6.**	

Table 4.2 Summary of *SGCD* transcripts identified in human cerebellum and skeletal muscle. Based on the gene architecture of *SGCD*, three mini-libraries of transcripts were required per tissue. Transcript abundance is provided in each of the six total mini-libraries, alongside the change in the mRNA sequence, the type of alternative splicing event involved, and the predicted effect on the isoform 1 protein and/or the isoform 2 protein depending on which transcript mini-libraries the transcript was identified. Abbreviations: cerebellum, CB; skeletal muscle, SM; transcripts originating from the upstream promoter and terminating in exon 9, UP; transcripts originating from the downstream promoter and terminating in exon 9, DP; transcripts originating from the downstream promoter and terminating in exon 8b, Ex8b ter; Alt. exon inclusion, alternative or novel exon inclusion in RNA; PTC, premature termination codon; NMD/R, nonsense-mediate decay/repression; aa, amino acid.

4.2.3. Alternatively spliced *SGCZ* transcripts in the cerebellum were varied




A)
SGCZ exon 7 truncation transcript
1
1a
1b
2
3
4
5
6
7
7b
8
Encoded protein: ζ-SG isoform 5
H2N-
-COOH
N
N
288
SGCZ exon 7b+ transcript
1
1a
1b
2
3
4
5
6
7
7b
8
Encoded protein: ζ-SG isoform 6
H2N-
-COOH
N
N
277

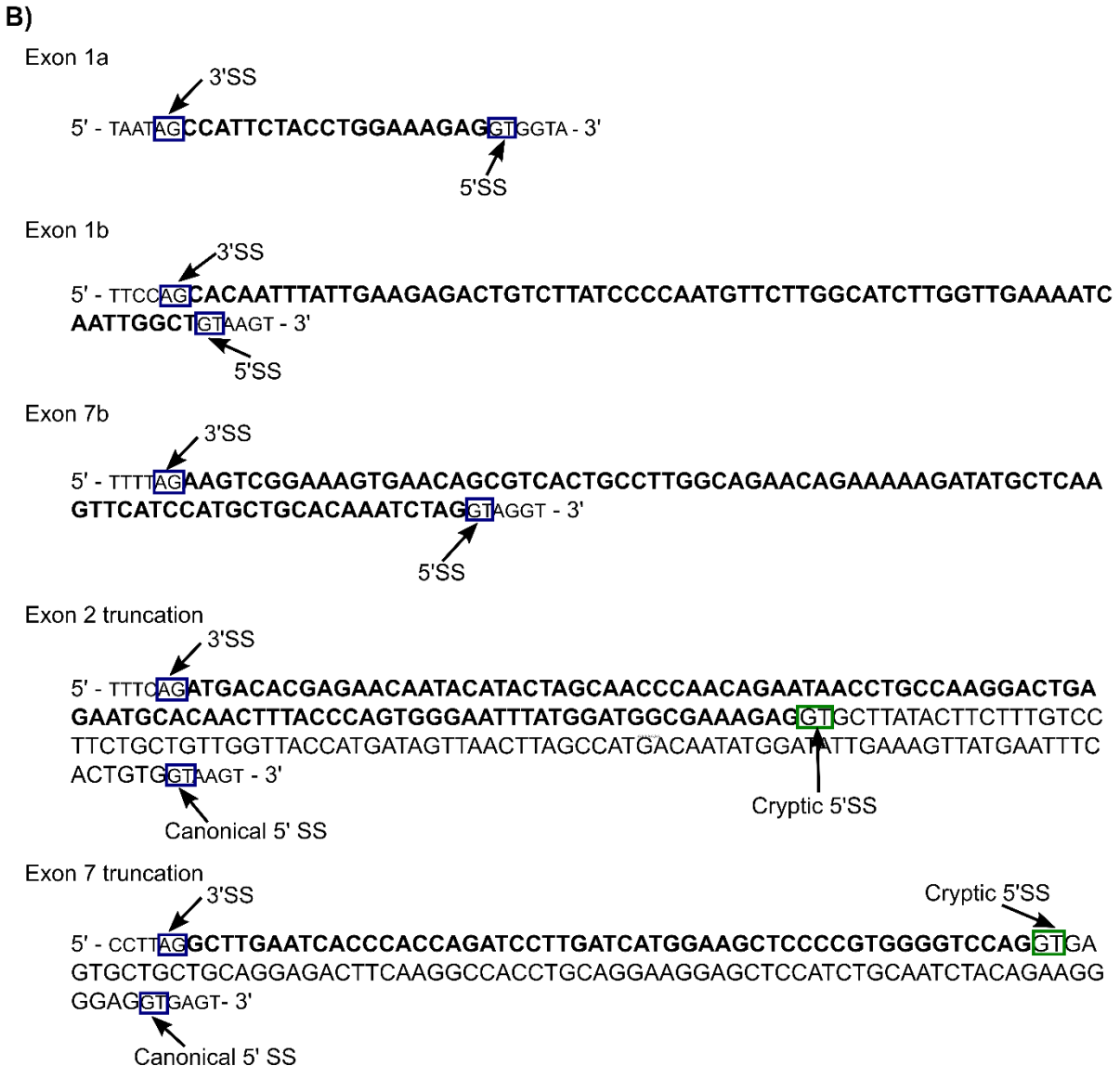

B)
Exon 1a
3'SS
5' - TAATAGCCATTCTACCTGGAAAGAGGTGGTA - 3'
5'SS
Exon 1b
3'SS
5' - TTCCAGCACAATTTATTGAAGAGACTGTCTTATCCCCAATGTTCTTGGCATCTTGGTTGAAAATC
AATTGGCTGTAAGT - 3'
5'SS
Exon 7b
3'SS
5' - TTTTAGAAGTCGGAAAGTGAACAGCGTCACTGCCTTGGCAGAACAGAAAAAGATATGCTCAA
GTTCATCCATGCTGCACAAATCTAGGTAGGT - 3'
5'SS
Exon 2 truncation
3'SS
5' - TTTCAGATGACACGAGAACAATACATACTAGCAACCCAACAGAATAACCTGCCAAGGACTGA
GAATGCACAACTTTACCCAGTGGGAATTTATGGATGGCGAAAGAGGTGCTTATACTTCTTTGTCC
TTCTGCTGTTGGTTACCATGATAGTTAACTTAGCCATGACAATATGGATATTGAAAGTTATGAATTTC
ACTGTGGTAAGT - 3'
Cryptic 5'SS
Canonical 5' SS
Exon 7 truncation
3'SS
Cryptic 5'SS
5' - CCTTAGGCTTGAATCACCCACCAGATCCTTGATCATGGAAGCTCCCCGTGGGGTCCAGGTGA
GTGCTGCTGCAGGAGACTTCAAGGCCACCTGCAGGAAGGAGCTCCATCTGCAATCTACAGAAGG
GGAGGTGAGT- 3'
Canonical 5' SS

Figure 4.11 Transcripts and novel exons identified from *SGCZ* in human cerebellum. A) For each *SGCZ* transcript identified in this study, the splicing pattern is displayed against the genomic architecture of the gene. Exons included in the transcript are connected, while features such as premature termination codons and frameshift sites are annotated. For translated transcripts, the encoded polypeptide with major domains is depicted below the transcript. Yellow indicates a transmembrane domain, N represents N-linked glycosylation sites, and green lines connect cysteine residues involved in disulphide bonds. B) DNA sequences for novel and truncated exons. The 5' splice sites (SS) and 3' splice sites (SS) both cryptic and canonical are indicated in the sequence. Abbreviations: PTC, premature termination codon; FS, frameshift; SS, splice site.

A total of nine different transcripts were identified in 215 clones from the human cerebellum *SGCZ* cDNA mini-library (Figure 4.11). Only 58% of clones screened comprised the canonical, full-length transcript; the remainder were alternatively spliced transcripts (Figure 4.12, Table 4.3). In particular, two alternatively spliced transcripts predicted to encode novel ζ-SG isoforms accounted for nearly a quarter of the cDNA mini-library clones screened: the *SGCZ* Δ exon 3 transcript (15% of clones) and the *SGCZ* Δ exon 5 transcript (9%) of clones (Figure 4.12, Table 4.3). Both transcripts were predicted to encode isoforms lacking in-frame sections from the extracellular region, without affecting post-translational modifications (Figure 4.11). Three other alternatively spliced transcripts were also predicted to encode ζ-SG isoforms, but were each present at or less than 1% of clones and therefore are unlikely to be abundant in tissue (Table 4.3).

The last three alternatively spliced *SGCZ* transcripts were all predicted to be unproductive due to insertion of a PTC more than 50nt upstream of the final exon:exon junction, which would induce NMD/R (Figure 4.11). However, the *SGCZ* exon 2 truncation transcript produced by use of an alternate 5' splice site within exon 2 accounted for 11% of *SGCZ* cDNA mini-library clones (Figure 4.12, Table 4.3). This high abundance may indicate escape from NMD/R, possibly through proximity of the PTC to the translation initiation site or use of an alternate translation initiation site (Figure 4.12 and Table 4.3) (Isken and Maquat, 2007). Another predicted unproductive transcript contained two novel *SGCZ* exons: exons 1a and 1b, the latter of which contained an in-frame termination codon (Figure 4.11). Exon 1b

showed high similarity to the L1M2 transposable element (Figure 4.6d) (Smit et al., 2013-2015).

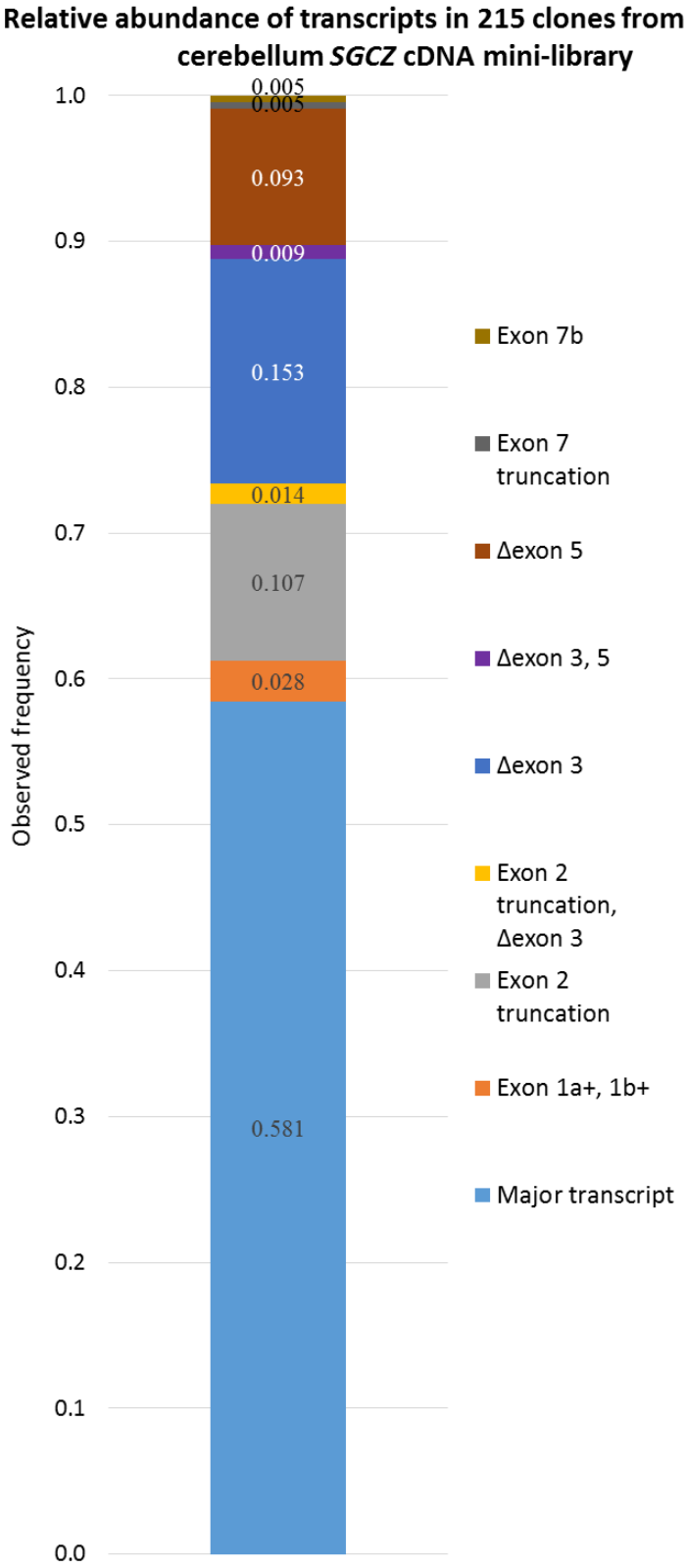
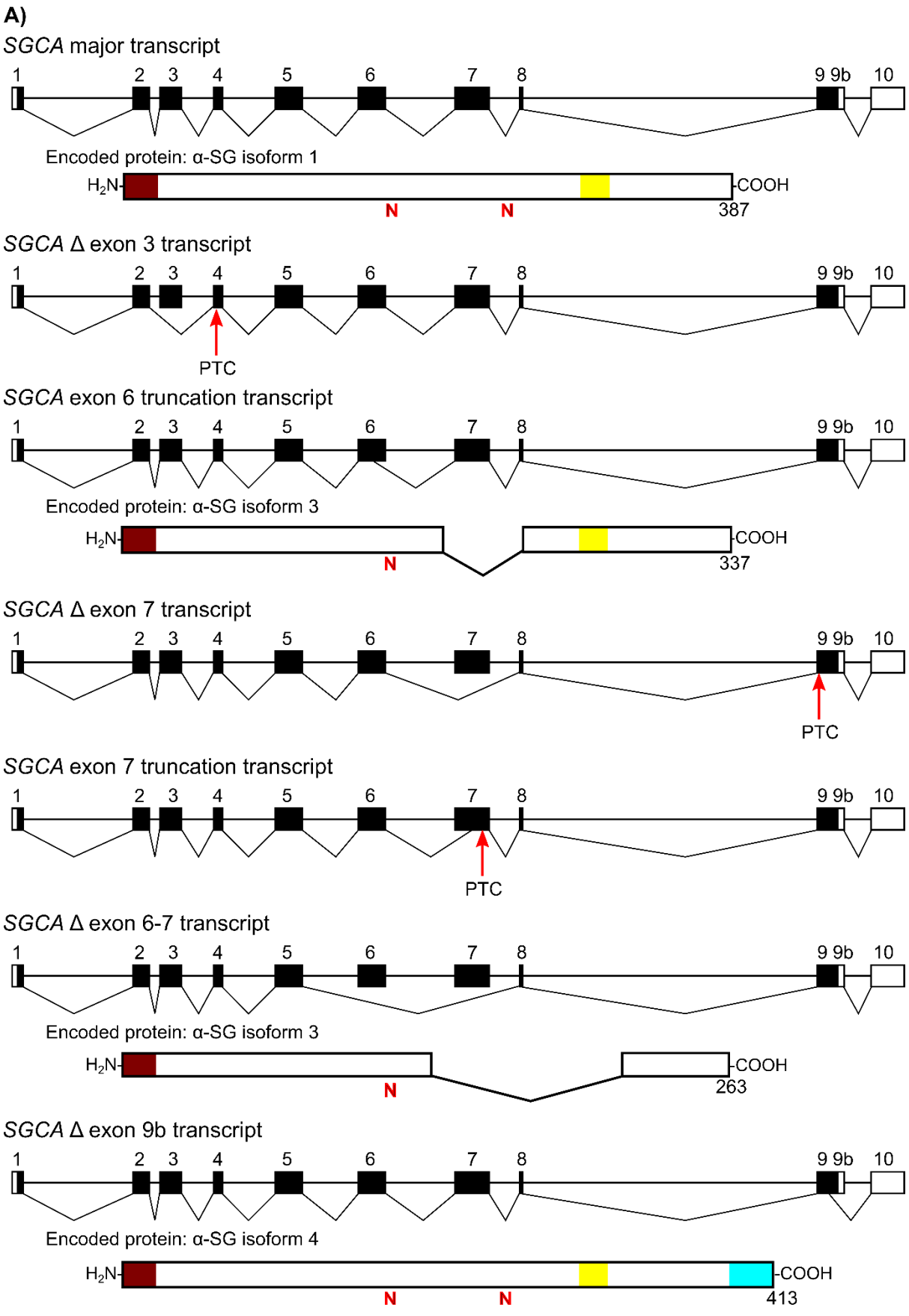


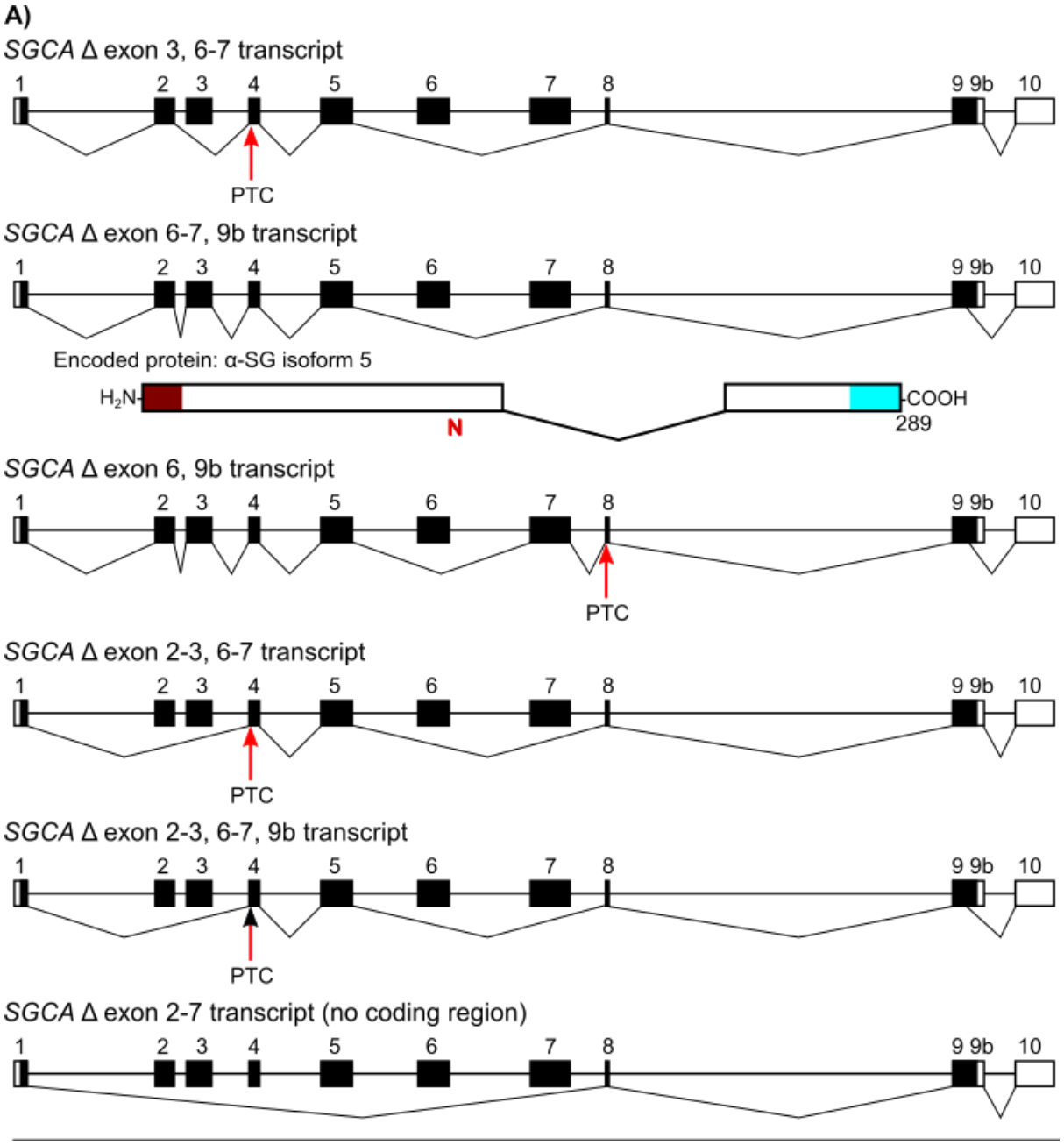
Figure 4.12 Relative abundance of discrete transcripts in *SGCZ* cDNA mini-library from human cerebellum. Stacked bar chart representing the abundance of each transcript (described in figure 4.10) identified for the *SGCZ* gene in mini-libraries. All transcripts are described relative to the full-length, major transcript.

SGCZ transcript	Effect on mRNA	Alternative splicing event	Observed frequency (215 clones)	Predicted effect on protein	Predicted polypeptide size	Protein isoform
Major transcript	-	-	0.58	Full-length protein isoform 1: transmembrane domain F51-L71; N-linked glycans at N75 and N123; disulphide bond at C286:C302.	312aa	1
Exon 1a+, 1b+	Inclusion of 2 novel exons (86nt total) from intron 1	Alt. exon inclusion	0.03	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-
Exon 2 truncation	Loss of 3' 94bp from exon 2	Use of alternate 5' splice site	0.11	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-
Δ exon 3	Omission of exon 3	Exon skipping	0.15	Omission of G80-D113 from the extracellular juxtamembrane domain; N-linked glycosylation sites and disulphide bonds unaffected.	278aa	2
Δ exon 5	Omission of exon 5	Exon skipping	0.09	Omission of A143-G182 from the extracellular domain; N-linked glycosylation sites and disulphide bonds unaffected.	271aa	3
Exon 7 truncation	Loss of 3' 72bp from ex7	Use of alternate 5' splice site	0.005	Omission of V225-E248 from the extracellular domain; N-linked glycosylation sites and disulphide bonds unaffected.	288aa	5
Exon 7b	Insertion of 81bp from intron 7	Alt. exon inclusion	0.005	Frameshift and new termination codon introduced into exon 8; unique sequence after E248 and termination after F277; disulphide bond lost.	277aa	6
Exon 2 truncation, Δ exon 3	Loss of 3' 94bp from exon 2; loss of exon 3	Use of alternate 5' splice site; exon skipping	0.01	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-
Δ exon 3, 5	Omission of exons 3 and 5	Exon skipping	0.01	Omission of G80-D113 and A143-G182; N-linked glycosylation sites and disulphide bonds theoretically unaffected.	237aa	4

Table 4.3 Summary of *SGCZ* transcripts identified in human cerebellum. A total of nine transcripts including the major transcript were identified in clones of human cerebellum *SGCZ* mRNA. The change to the RNA sequence, type of alternative splicing event responsible for the transcript, the observed frequency in the *SGCZ* transcript mini-library, predicted effect on the encoded protein, predicted polypeptide size and isoform designation are provided. Where an alternative translation start size was proposed, the probability that site would actually be used as a translation initiation site based on the NetStart and ATGpr algorithms is provided. Abbreviations: Alt. exon inclusion, alternative or novel exon inclusion; PTC, premature termination codon; NMD/R, nonsense-mediated decay/repression; aa, amino acid.

4.2.4. *SGCA* pre-mRNA undergoes extensive alternative splicing in skeletal muscle





B) Exon 6 truncation

3' SS
Cryptic 5' SS

5' - TCCCAG**GGTATACATTAAG**GTGGGTTCTGCCTCACCTTTTTCTACTTGCCTGAAGATGGTGGCATCC
CCCGATAGCCACGCCCGCTGTGCCCAGGGCCAGCCTCCACTTCTGTCTTGCTACGACACCTTGGCAC
CCCACTTCCGCGTTGACTGGTGCAATGTGACCCTGGTGAGG- 3'

Canonical 5' SS

Exon 7 truncation

Canonical 3' SS
Cryptic 3' SS

5' - GTGCAGGTGGATAAGTCAGTGCCGGAGCCTGCAGATGAGGTGCCCACCCCAGGTGATGGGATCCT
GGAGCATGACCCGTTCTTCTGCCCACCCACTGAG**GCCCCAGACCGTGACTTCTTGGTGGATGCTCTG**
GTCACCCTCCTGGTGCCCCTGCTGGTGGCCCTGCTTCTCACCTTGCTGCTGGCCTATGTCATGTGCT
GCCGGCGGGAGGGAAGGTGAAT- 3'

5' SS

Figure 4.13 *SGCA* transcripts and exons identified in human skeletal muscle. A) For each *SGCA* transcript identified in this study, the splicing pattern is displayed against the genomic architecture of the gene. Exons included in the transcript are connected, while features such as premature termination codons and frameshift sites are annotated. For translated transcripts, the encoded polypeptide with major domains is depicted below the transcript. Yellow indicates a transmembrane domain, N represents N-linked glycosylation sites, and turquoise indicates isoform-specific sequences. B) DNA sequences for truncated exons. The 5' splice sites (SS) and 3' splice sites (SS) both cryptic and canonical are indicated in the sequence. Abbreviations: PTC, premature termination codon; FS, frameshift; SS, splice site.

A total of 12 distinct alternatively spliced *SGCA* transcripts were identified in 97 clones from a skeletal muscle cDNA mini-library, plus the full-length, canonical transcript (Figure 4.13). Less than 50% of the clones screened were the full-length transcript containing all 10 exons; the remainder consisted of alternatively spliced transcripts (Figure 4.14, Table 4.4). There were no novel exons, and all alternative splicing events involved either exon omission of alternate splice site usage (Figure 4.13, Table 4.4). Two of the alternatively spliced transcripts had previously been described: the *SGCA* Δ exon 6-7 transcript was first reported by McNally *et al* (McNally et al., 1994), while the *SGCA* Δ exon 9b transcript is annotated in the LOVD database despite not appearing in published literature (Fokkema et al., 2011). The first of these transcripts, *SGCA* Δ exon 6-7, accounted for 26% of the clones screened in the present study, making it by far the most abundant alternatively spliced *SGCA* transcript (Figure 4.14, Table 4.4). It was predicted to encode a potentially secreted α-SG isoform 2 lacking the transmembrane domain and extracellular juxtamembrane region but retaining the signal peptide (Figure 4.13). The *SGCA* Δ exon 9b transcript was also predicted to be translated, but would encode an α-SG isoform 4 with an extended intracellular region compared to the main α-SG isoform 1; this transcript was rare in the cDNA mini-library at 1% of clones (Figure 4.14, Table 4.4).

Several other alternatively spliced *SGCA* transcripts were newly identified in this study. The majority were predicted to be NMD/R targets due to introduction of a PTC, with one exception (Figure 4.13). This exception was the *SGCA* exon 6 truncation transcript, present in 2% of clones and generated by use of an alternate 5' splice site (Figure 4.14, Table 4.4). It was predicted to encode an α-SG isoform 3 lacking the extracellular juxtamembrane region

(Figure 4.13). In summary, two previously reported and ten novel alternatively spliced *SGCA* transcripts were identified.

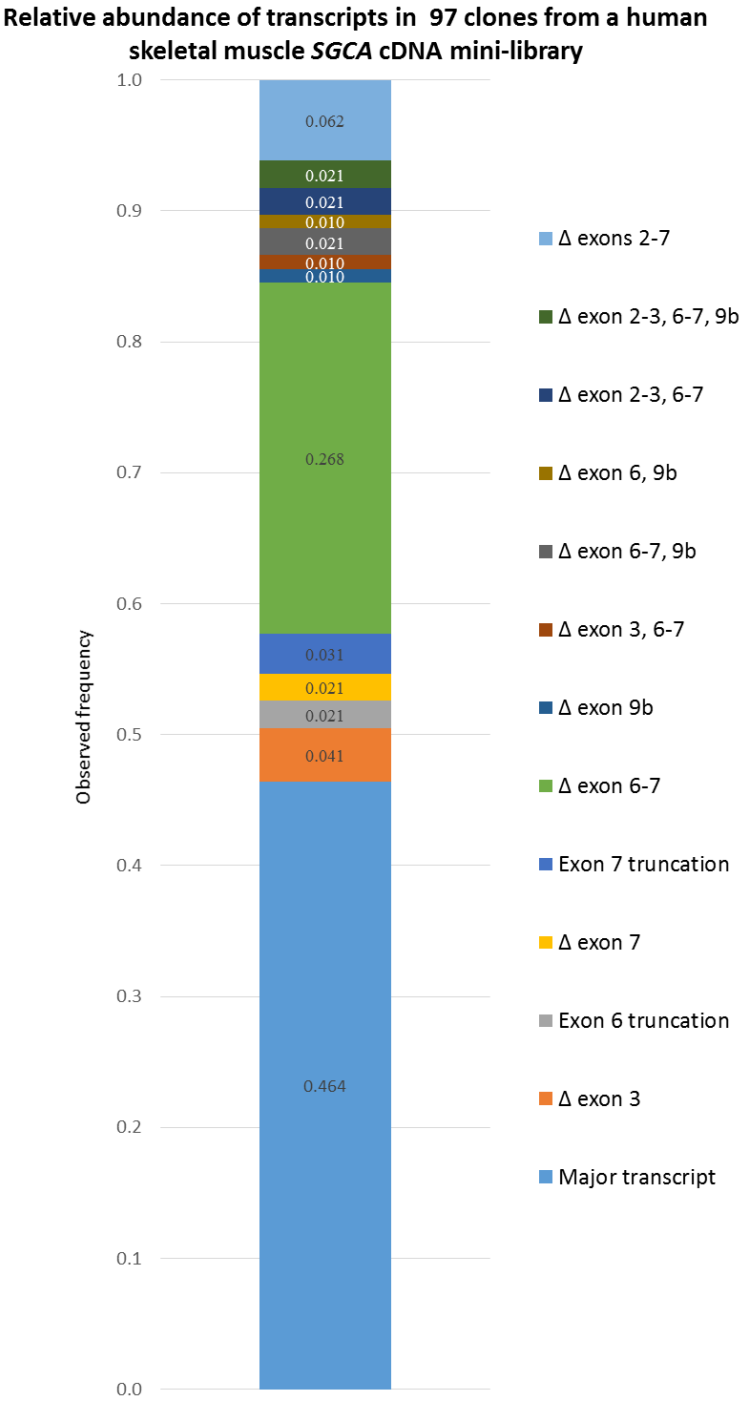


Figure 4.14 Relative abundance of discrete transcripts in a *SGCA* cDNA mini-library from human skeletal muscle. Stacked bar chart representing the abundance of each transcript (described in Figure 4.13) identified for the *SGCA* gene in mini-libraries. All transcripts are described relative to the full-length, major transcript.

Transcript variant	Effect on mRNA	Type of splicing event	Observed frequency (97 clones)	Predicted effect on encoded protein	Predicted polypeptide size	Protein isoform	NCBI accession
Major transcript	-	-	0.46	Full length protein: type I transmembrane; signal peptide M1-A23 (cleaved off); transmembrane domain L291-V311; N-linked oligosaccharides in the extracellular domain at N174 and N246.	387aa	1	**NM_000023.2**
Δ exon 3	Omission of exon 3	Exon skipping	0.04	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-	**XM_011525124.1** (with exon 9b-)
Exon 6 truncation	Loss of 3' 150nt from exon 6	Alternate 5' splice site usage	0.02	Omission of G201-V250 from the extracellular juxtamembrane domain including the second N-linked oligosaccharide; remaining features unaffected.	337aa	3	**XM_011525121.1** (with exon 9b-)
Δ exon 7	Omission of exon 7	Exon skipping	0.02	Introduction of frameshift and PTC in exon 9 >50nt from final exon junction: NMD target – non-productive transcript	-	-	**XM_011525122.1** (with exon 9b-)
Exon 7 truncation	Loss of 5' 92nt from exon 7	Alternate 3' splice site usage	0.03	Introduction of frameshift and PTC – NMD/R target, non-productive.	253aa		-
Δ exon 6-7	Omission of exons 6 and 7	Exon skipping	0.27	Omission of V195-R319 with loss of transmembrane domain, extracellular juxtamembrane domain and second N-linked glycosylation site. Likely secreted.	263aa	2	**NM_001135697.1**
Δ exon 9b	Loss of 3' 74nt from exon 9 (loss of exon 9b)	Alternate 5' splice site usage	0.01	Frameshift after T367 for a unique, extended carboxyl terminus/intracellular domain; signal peptide, transmembrane domain and both N-linked glycosylation sites unaffected.	413aa	4	**XM_011525120.1**
Δ exon 3, 6-7	Omission of exons 3, 6 and 7	Exon skipping	0.01	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-	-

Δ exon 6-7, 9b	Omission of exons 6 and 7; loss of exon 9b	Exon skipping, alternate 5' splice site usage	0.02	Omission of V195-R319 with loss of transmembrane domain, extracellular juxtamembrane domain and second N-linked glycosylation site. Carboxyl terminus extended as in isoform 4. Likely secreted.	289aa	5	-
Δ exon 6, 9b	Omission of exon 6; loss of exon 9b	Exon skipping, alternate 5' splice site usage	0.01	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-	-
Δ exon 2-3, 6-7	Omission of exons 2, 3, 6 and 7	Exon skipping	0.02	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-	-
Δ exon 2-3, 6-7, 9b	Omission of exons 2, 3, 6 and 7; loss exon 9b	Exon skipping, alternate 5' splice site usage	0.02	Introduction of frameshift and PTC – NMD/R target, non-productive.	-	-	-
Δ exon 2-7	Omissions of exons 2 through 7	Exon skipping	0.06	Loss of most of the coding region – non-productive transcript	-	-	-

Table 4.4 Summary of *SGCA* transcripts identified in human skeletal muscle. For each transcript type, the change to the mRNA relative to the major transcript, type of alternative splicing event involved, observed frequency in the *SGCA* skeletal muscle transcript mini-library, predicted effect on the encoded protein, predicted size of the encoded polypeptide, isoform designation, and the corresponding NCBI mRNA sequence if applicable are provided. Abbreviations: PTC, premature termination codon; NMD/R, nonsense-mediated decay/repression; aa, amino acids.

4.2.5. Few alternatively spliced *SGCG* transcripts were identified

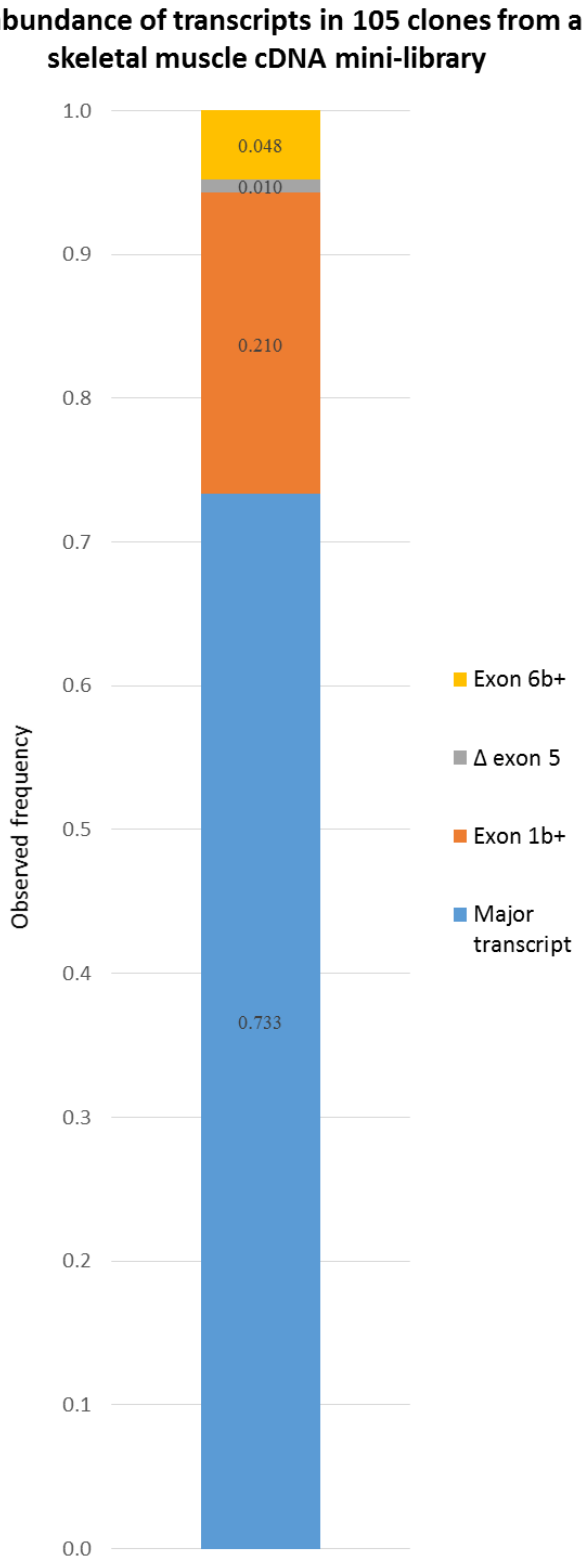
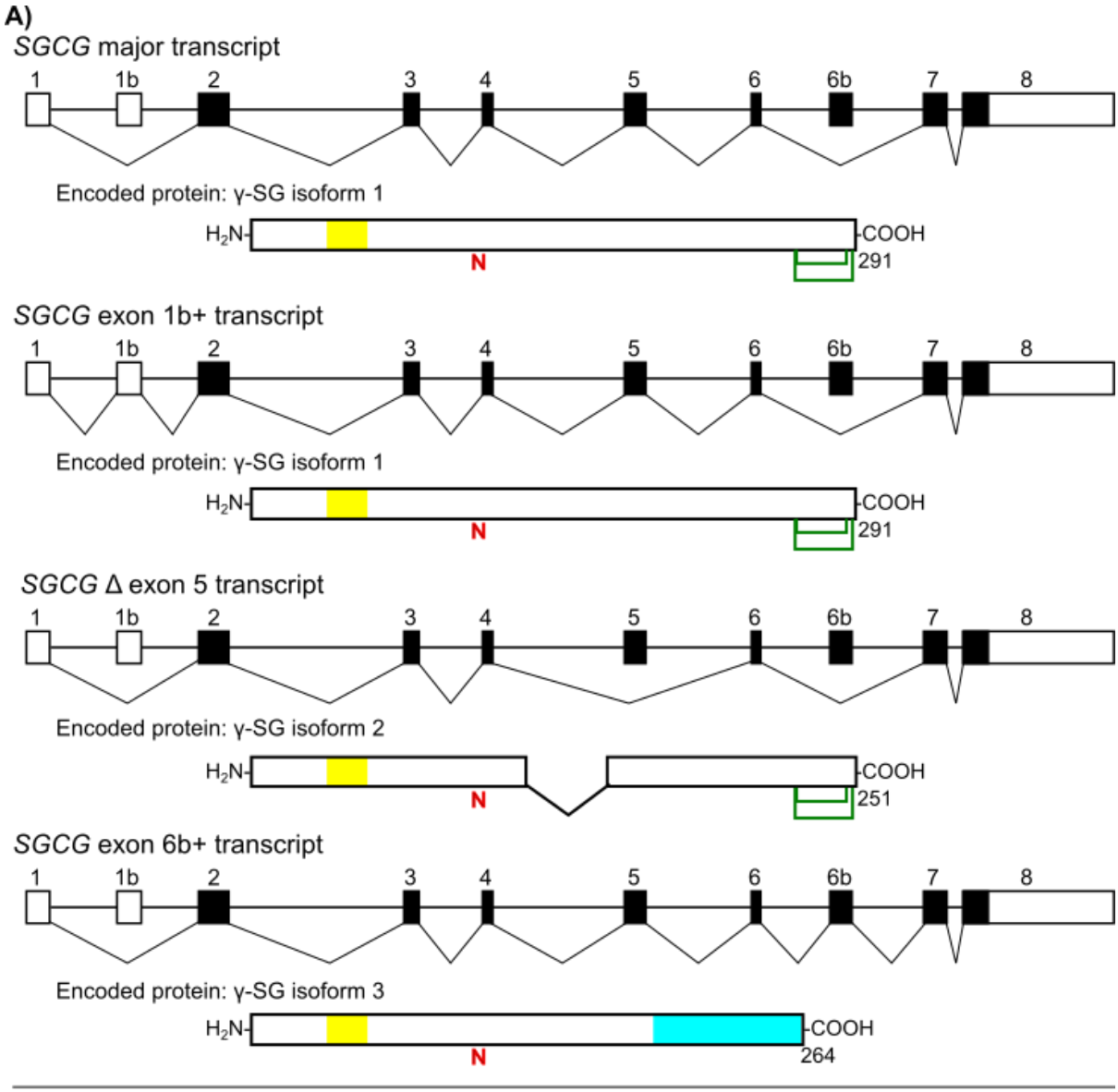


Figure 4.15 Relative abundance of discrete transcripts in *SGCG* cDNA mini-library from human skeletal muscle. Stacked bar chart representing the abundance of each transcript (described in figure 4.16) identified for the *SGCG* gene in mini-libraries. All transcripts are described relative to the full-length, major transcript.

In the skeletal muscle *SGCG* cDNA mini-library, the full-length major transcript accounted for 73% of the 105 clones screened (Figure 4.15). However, an alternatively spliced transcript incorporating a novel exon 1b accounted for another 21% of clones (Figure 4.15). Exon 1b was within the 5' UTR, so this transcript would encode full-length γ-SG isoform 1 (Figure 4.16, Table 4.5). The two remaining alternatively spliced transcripts accounted for only 6% of cDNA mini-library clones together, but both were predicted to encode γ-SG isoforms differing in the

extracellular region. The *SGCG* Δ exon 5 transcript was predicted to encode a γ-SG isoform 2 lacking residues 130-169 from the extracellular region, while inclusion of exon 6b (*SGCG*



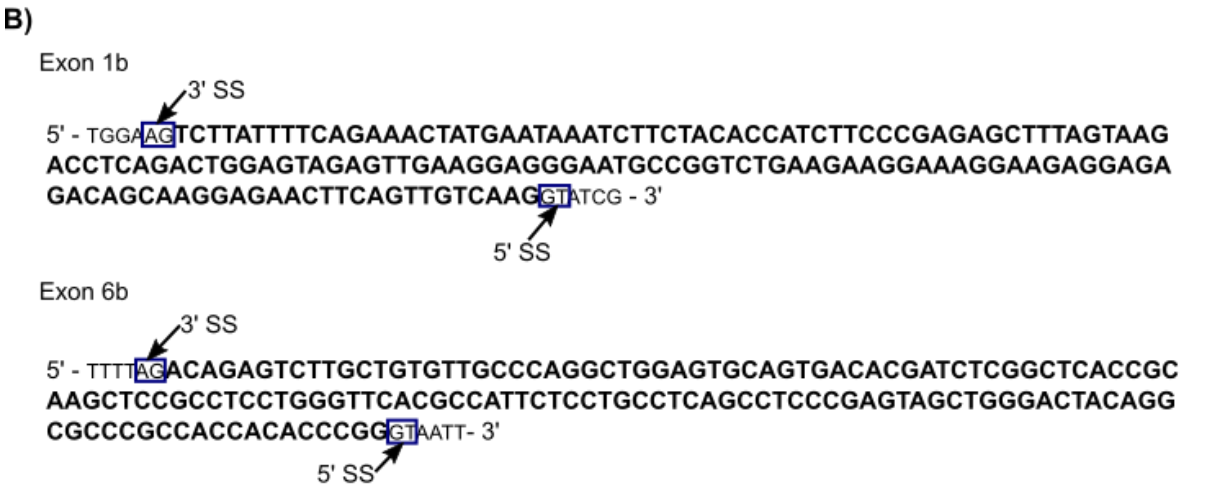


Figure 4.16 Transcripts and novel exons identified for *SGCG* in a cDNA mini-library from human skeletal muscle. A) For each *SGCG* transcript identified in this study, the splicing pattern is displayed against the genomic architecture of the gene. Exons included in the transcript are connected, while features such as premature termination codons and frameshift sites are annotated. For translated transcripts, the encoded polypeptide with major domains is depicted below the transcript. Yellow indicates a transmembrane domain, N represents N-linked glycosylation sites, turquoise marks isoform-specific sequence, and green lines connect cysteine residues involved in disulphide bonds. B) DNA sequences for novel and truncated exons. The 5' splice sites (SS) and 3' splice sites (SS) both cryptic and canonical are indicated in the sequence. Abbreviations: SS, splice site.

exon 6b+ transcript) was predicted to cause a frameshift and encode an isoform 3 with a unique sequence from residue 193 (Figure 4.16, Table 4.5). *SGCG* exon 6b showed strong sequence similarity to the AluYc transposable element (Smit et al., 2013-2015). In summary, although few alternatively spliced *SGCG* transcripts were identified in skeletal muscle all were protein-coding.

Transcript variant	Effect on mRNA	Type of splicing event	Observed frequency (105 clones)	Predicted effect on protein	Predicted polypeptide size	Protein isoform	NCBI mRNA
Major transcript	-	-	0.73	Full-length protein: transmembrane domain L38-L58; N-linked glycan at N110; disulphide bonds at C265:C290 and C267:C283.	291aa	1	**NM_000231.2**
Exon 1b+	Inclusion of 151nt exon 1b after exon 1	Alt. exon inclusion	0.21	Affects only the 5' UTR sequence; full-length isoform 1.	291aa	1	**XM_005266505.2**
Δ exon 5	Omission of exon 5	Exon skipping	0.01	Omission of P130-G169 from the extracellular domain; N-linked glycosylation sites and disulphide bonds maintained.	251aa	2	-
Exon 6b+	Inclusion of 142nt exon 6b after exon 6	Alt. exon inclusion	0.05	Introduces a frameshift and PTC into the conventional CDS, but PTC is within 50nt of the final exon:exon junction so probably escapes NMD; encoded protein has a unique 71aa sequence after R193 and lacks both disulphide bonds.	265aa	3	-

Table 4.5 Overview of *SGCG* transcripts identified in human skeletal muscle. For each transcript type, the change in mRNA sequence relative to the major transcript, type of alternative splicing event responsible, observed frequency in the transcript mini-library, predicted effect on the encoded protein size, protein isoform and relevant NCBI mRNA are provided. Abbreviations: Alt. exon inclusion, alternative or novel exon inclusion into the mRNA; UTR, untranslated region; aa, amino acid.

Gene		Number of alternatively spliced transcripts	Alternatively spliced transcripts affecting UTR only			Alternatively spliced transcripts predicted to be NMD/R targets			Alternatively spliced transcripts encoding protein isoforms			Transcripts only found in skeletal muscle	Transcripts only found in cerebellum
			No.	Freq (CB)	Freq (SM)	No.	Freq (CB)	Freq (SM)	No.	Freq (CB)	Freq (SM)		
SGCB		4	0	0	0	3	4%	8%	1	15%	10%	2	0
SGCD	UP	12	3	48%	30%	4*	1%	3%	6*	11%	16%	5	1
	DP			13%	4%		9%	0%		0%	6%		
	8b			4%	2%		0%	0%		14%	6%		
SGCZ		8	0	0%	NT	3	15%	NT	5	26%	NT	NT	NT
SGCG		3	1	NT	21%	0	NT	0%	2	NT	6%	NT	NT
SGCA		12	0	NT	0%	8	NT	21%	4	NT	32%	NT	NT

Table 4.6 Summary of alternatively spliced transcripts for *SGCB, SGCD, SGCZ, SGCG* and *SGCA*. This table summarises the types of alternatively spliced transcript identified in cDNA mini-libraries of sarcoglycan transcripts from human adult skeletal muscle and cerebellum. For each sarcoglycan gene screened, the number of alternative spliced transcripts identified, the number and frequency in cDNA mini-libraries of alternatively spliced transcripts that only affect an UTR sequence, the number and frequency in cDNA mini-libraries of alternatively spliced transcripts encoding protein isoforms, and the number and frequency in cDNA mini-libraries of alternatively spliced transcripts predicted to be non-productive due to NMD/R are provided. *SGCB* and *SGCD* were screened in both skeletal muscle and cerebellum; the numbers of transcripts identified in only one or other cDNA mini-library are provided. For *SGCD*, data are separated by promoter/transcription termination site: UP, upstream promoter and exon 9 termination; DP, downstream promoter and exon 9 termination; 8b, downstream promoter and exon 8b termination. Abbreviations: UTR, untranslated region; No., number; Freq (CB), frequency of transcripts in cerebellum cDNA mini-library; Freq (SM), frequency of transcripts in skeletal muscle cDNA mini-library; NMD/R, nonsense-mediated decay or translational repression; NT, not tested.*: one alternatively spliced *SGCD* transcript either encoded a novel protein isoform or was predicted to be a NMD/R target depending on the transcription termination site; therefore, it was included in both counts.

4.3. Discussion

While tissue-specific alternative splicing of *SGCE* pre-mRNA was first described in 2004-2005 and has subsequently been studied in depth, there have been only sporadic reports of alternatively spliced transcripts from the other sarcoglycans (McNally et al., 1994, Nishiyama et al., 2004, Ritz et al., 2011, Yokoi et al., 2005). Alternative splicing of *SGCA, SGCB, SGCD, SGCG* and *SGCZ* could contribute to variation in the assembly and function of sarcoglycan complexes, so in this study I systematically identified alternatively spliced transcripts from these genes using RT-PCR. In order to screen all five of these genes, two disease-relevant tissues were examined: cerebellum, which has been implicated in MD, and skeletal muscle, which is affected in sarcoglycan-related LGMD. Use of RT-PCR facilitated identification of complete transcripts rather than isolated alternative splicing events, but could not accurately determine transcript abundance due to PCR amplification bias: shorter transcripts would be preferentially amplified over longer transcripts, resulting in an overestimate of shorter transcripts' abundance. All five of the sarcoglycan genes generated alternatively spliced transcripts, but patterns of splicing differed between sarcoglycan genes as summarised in table 4.6.

Alternatively spliced transcripts encoding novel protein isoforms were uncommon for *SGCB, SGCD* and *SGCG*. Between 81% and 94% of the cDNA mini-library clones screened for each of these genes either encoded the full-length, canonical isoform or encoded an isoform predicted to be indistinguishable from the full-length isoform (i.e. *SGCD* Δ exon 2 transcript) (Tables 4.1, 4.2 and 4.5). Alternatively spliced transcripts predicted to encode novel protein isoforms accounted for 6-15% of cDNA mini-library clones screened, while the remainder consisted of predicted unproductive transcripts (Table 4.6). The large number of unproductive transcripts and those differing in only the 5' UTR sequence from the major transcript suggested that proteome expansion may not be a key result of *SGCB, SGCD* and

SGCG pre-mRNA alternative splicing. Instead, alternative splicing of transcripts from these genes could help regulate protein levels. Alternative splicing to produce unproductive transcripts or transcripts with altered translation kinetics can modulate protein levels without altering overall gene transcription, as has been described for transcription factors in particular (Änkö et al., 2012, Isken and Maquat, 2007, Lewis et al., 2003, Mendell et al., 2004). Both knockout and overexpression of *Sgcg* in mice cause muscular dystrophy, and skeletal muscle is thought to contain equal amounts of each sarcoglycan protein (Hack et al., 1998, Jung et al., 1996, Mizuno et al., 1994, Sandona and Betto, 2009, Zhu et al., 2001). This implies that sarcoglycans must be maintained at the correct level for normal muscle function. Alternative splicing may be one mechanism used to fine-tune sarcoglycan protein levels.

Alternative splicing of *SGCZ* differed strongly from that of its paralogue *SGCG*, however. While 94% of *SGCG* skeletal muscle cDNA mini-library clones encoded the full-length γ-SG isoform 1, only 58% of *SGCZ* cerebellum cDNA mini-library clones encoded ζ-SG isoform 1 (Tables 4.3 and 4.5). Alternatively spliced *SGCZ* transcripts predicted to encode novel ζ-SG isoforms accounted for 26% of *SGCZ* cDNA mini-library clones in total, and just two transcripts encoding protein isoforms comprised 24% of clones (Tables 4.3 and 4.6). Furthermore, no predicted unproductive *SGCG* transcripts were identified in skeletal muscle whereas three unproductive *SGCZ* transcripts accounting for 15% of cerebellum cDNA mini-library clones were identified (Tables 4.3 and 4.5). One alternative splicing event, skipping of exon 5, was conserved between *SGCG* and *SGCZ* but substantially more common in *SGCZ* transcripts compared to *SGCG* (Tables 4.3 and 4.5). These differences in alternative splicing pattern between *SGCG* and *SGCZ* were unexpected, and might reflect differences in function. At present very little is known about ζ-SG, and there is a possibility it may have roles other than those proposed for γ-SG and other sarcoglycans. The ζ-SG isoforms encoded by the alternatively spliced transcripts identified in this study have the capacity to affect functions

and dynamics of sarcoglycan complexes in the brain, which could potentially contribute to the requirement for ε-SG but not other sarcoglycans for normal brain function.

Muscle-specific *SGCA* also exhibited key differences in alternative splicing compared to its more widely expressed paralogue *SGCE*. No alternative splicing events were shared between these two genes, despite highly similar exon/intron boundaries (Figure 4.13) (Ettinger et al., 1997, McNally et al., 1998, Ritz et al., 2011). *SGCE* and *SGCA* appeared to be similar in the abundance of alternative splicing, though: less than half of *SGCA* cDNA mini-library clones comprised the full-length major transcript, while at least a third of *SGCE* transcripts in the brain are reported to be alternatively spliced (Figure 4.14) (Ritz et al., 2011). Both genes also encode alternatively spliced isoforms with extended intracellular regions, though these regions show minimal sequence similarity (data not shown) (Figure 4.13; Chapter 1 Figure 1.2). Intriguingly, the transcript encoding an extended α-SG isoform was detected at low abundance in the skeletal muscle cDNA mini-library, and instead the most abundant transcript was the *SGCA* Δ exons 6-7 transcript, which alone constituted 27% of the mini-library clones screened (Table 4.4). The α-SG isoform 2 encoded by this isoform is predicted to be secreted into the extracellular space, something not previously described for a sarcoglycan. This could suggest as yet unidentified function(s) for α-SG in the skeletal muscle extracellular spaces. Furthermore, eight of the twelve alternatively spliced *SGCA* transcripts were predicted to be unproductive, accounting for 21% of *SGCA* cDNA clones. This suggested alternative splicing could help regulate α-SG protein levels in skeletal muscle as predicted for *SGCB, SGCD* and *SGCG*.

Novel exons in *SGCB, SGCD, SGCZ* and *SGCG* were identified during this study. When searched for similarity to transposable elements, five of these seven novel exons were found to be substantially similar to primate-specific transposable elements (Figure 4.6). *SGCE* exon 10 is also Alu-derived, so five of the 6 sarcoglycan genes have alternatively spliced exons

derived from transposable elements; the exception is *SGCA* (McNally et al., 1998). From one perspective, this was entirely expected: few mutations are required to turn sequences in reverse-orientation Alu elements into alternatively spliced exons, and Alu elements preferentially insert into introns (Rebollo et al., 2012, Sela et al., 2010, Shen et al., 2011a). Consequently, transposable elements are frequently recruited into gene transcripts (Rebollo et al., 2012). Yet even though transposable element-related sequences are abundant in the human genome, only about 5% of known alternatively spliced exons show similarity to transposable elements (de Koning et al., 2011, Lev-Maor et al., 2003, Rebollo et al., 2012, Shen et al., 2011a, Sorek et al., 2002). By this measure, the five transposable element-derived novel exons identified in this study would be unusual. As a result the significance of the novel, transposable element-derived sarcoglycan exons remains unclear.

Several alternatively spliced *SGCB* and *SGCD* transcripts were identified in skeletal muscle but not cerebellum (Tables 4.1 and 4.2). By contrast, the full-length, major transcript was generally more abundant in skeletal muscle cDNA mini-libraries compared to cerebellum cDNA mini-libraries, suggesting that alternative splicing may be less frequent in skeletal muscle. All muscle-specific transcripts were present at low abundance, accounting for less than 5% each and in most cases only 1-2% of the skeletal muscle cDNA mini-library clones screened. Therefore, the apparent tissue specificity of these transcripts could be due to insufficient sampling from each cDNA mini-library. Use of an alternate technique such as RNA sequencing could also help clarify the presence of the simpler muscle-only transcripts, but several of these were complex transcripts involving multiple splicing events that would not be accurately identified with this technique.

In conclusion, this study identified a number of alternatively spliced transcripts from *SGCA, SGCB, SGCD, SGCG* and *SGCZ* in the first systematic screen of these genes' splicing. Therefore alternative splicing is not limited to *SGCE*. This study also demonstrated strong

differences in splicing between paralogous genes, which could reflect functional differences in the encoded proteins. Examination of splicing in additional tissues could further clarify whether any of the alternatively spliced transcripts identified in this study were tissue-specific; the *SGCZ* transcripts would be of particular interest as they could affect sarcoglycan complex composition in the brain. Additionally, use of RNA sequencing or quantitative PCR to determine actual abundance of alternative splicing events in different tissues would reveal any tissue differences in splicing. Additional alternatively spliced transcripts might also be identified, although the present study was expected to identify the most common transcripts. Further investigation of these transcripts and the novel sarcoglycan protein isoforms encoded therein could provide insights into dynamics and functions of sarcoglycan complexes and individual proteins (Chapter 5).

Chapter 5: Characterisation of alternatively spliced sarcoglycan isoforms

5.1. Introduction

While many of the alternatively spliced sarcoglycan gene transcripts identified in Chapter 4 were predicted to be unproductive, several transcripts were predicted to encode novel sarcoglycan isoforms. Most of the alternative splicing events involved omission of an exon from the mature transcript, and accordingly many of the encoded isoforms were predicted to lack regions of the coding sequence present in the full-length protein. In addition, a few alternative splicing events introduced new termination codons into the transcript that should escape nonsense-mediated decay to encode truncated protein isoforms (Amrani et al., 2004, Chang et al., 2007b). Comparing the isoform's amino acid sequence to that of the full-length isoform may provide some initial clues to potential differences in function, but very little is known about the biochemistry of sarcoglycan proteins. Therefore, in order to determine the functional effects of alternative splicing on sarcoglycan proteins, the isoforms needed to be tested *in vitro*. Wild-type, full-length sarcoglycans are known to be glycosylated, and assemble into putatively heterotetrameric complexes that traffic to the cell surface (Holt and Campbell, 1998, Jung et al., 1996, Noguchi et al., 2000). Isoforms generated through alternative splicing might exhibit changes in their ability to interact with some or all other sarcoglycans, as well as their trafficking. In this study I tested a panel of sarcoglycan isoforms for translation, stabilisation by proteasome inhibition, glycosylation, incorporation into sarcoglycan heterotetramers and trafficking to the cell surface in HEK293T cells.

5.2. Results

A series of functional assays was performed for each sarcoglycan isoform. Initially, a plasmid encoding the epitope-tagged isoform was transfected into HEK293T cells in triplicate.

Transfected cells were then either treated with the proteasome inhibitor bortezomib or mock-treated with DMSO, and cell lysates were analysed by Western blot (Chapter 2 section 2.4.4). If protein signal intensity was increased in bortezomib-treated cells compared to DMSO-treated cells, then proteasome inhibition stabilised the isoform. To determine whether isoforms were glycosylated, recombinant protein was treated with the endoglycosidase PNGase F to remove N-linked glycans and then analysed by Western blot (Chapter 2 section 2.6.1). A reduction in mass after PNGase F treatment indicated N-glycosylation of the isoform. The translation, glycosylation and proteasome inhibition stabilisation of each sarcoglycan isoform tested are summarised in Table 5.1.

To determine whether sarcoglycan isoforms were incorporated into sarcoglycan complexes and trafficked to the cell surface, sarcoglycan complex assembly was modelled in HEK293T cells. Isoforms were examined in the context of an εβδζ or αβδγ complex as appropriate. Cells were co-transfected with different combinations of plasmids encoding epitope-tagged sarcoglycans, and plasma membrane sarcoglycans were labelled through cell surface biotinylation (Chapter 2 sections 2.4.9, 2.6.9). Proteins were co-precipitated from cells using M2 anti-FLAG to enrich for FLAG-tagged β-SG, an anti-ε-SG antibody to enrich for ε-SG or NeutrAvidin (Pierce) to enrich for biotinylated protein, and analysed by Western blotting (Chapter 2, sections 2.6.2-3). For each sarcoglycan, the full-length isoform 1 assembled into a heterotetrameric complex that trafficked to the cell surface as previously described (Chapter 3 section 3.2.2.2) (Durbeej and Campbell, 1999, Holt and Campbell, 1998, Jung et al., 1996, Noguchi et al., 2000, Shiga et al., 2006, Straub et al., 1999).

Isoforms were scored on several aspects of incorporation into sarcoglycan complexes and trafficking to the cell surface. First, immunoprecipitation (IP) for β-SG as described above revealed whether the isoform of interest interacted with the βδ-SG core. Co-purification of other sarcoglycans in the β-SG IP alongside the isoform elucidated the composition of

complexes containing that isoform. If all four sarcoglycans co-purified, then the isoform was incorporated into a prototypical sarcoglycan heterotetramer. For β-SG and ζ-SG isoforms, IP using an anti-ε-SG antibody was also carried out to determine whether partial complexes containing ε-SG were formed as previously observed (Chapter 3, section 3.2.2.2). Finally, Western blotting of immunoprecipitated proteins with streptavidin-Alexa Fluor 680 to detect biotinylated proteins demonstrated whether the isoform of interest was trafficked to the cell surface as part of a sarcoglycan complex. Western blot analysis of NeutrAvidin affinity purifications demonstrated whether the isoform was present at the cell surface overall. All experiments were repeated at least twice to confirm results.

5.2.1. Alternatively spliced β-SG isoforms

5.2.1.1. β-SG isoform 2 is degraded by the proteasome.

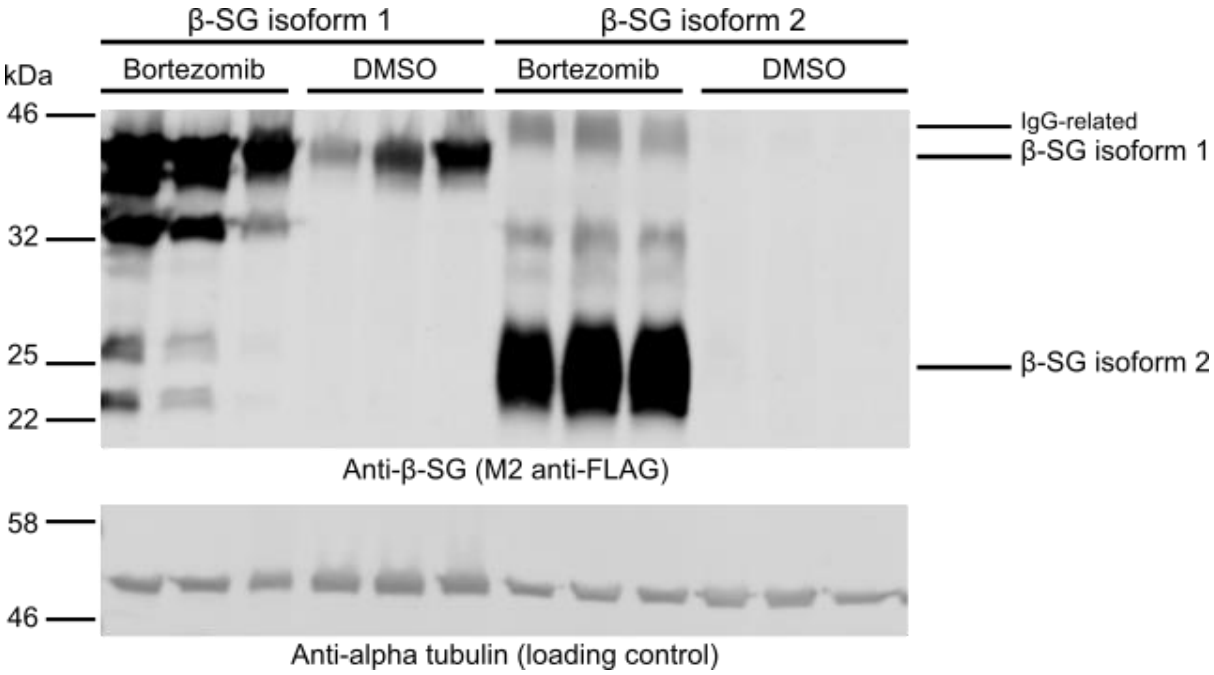


Figure 5.1 β-SG isoform 2 can only be detected in transfected cells with proteasome inhibition. Plasmids encoding isoforms 1 and 2 of β-SG were each transfected in triplicate into HEK293T cells, which were then either treated with the proteasome inhibitor bortezomib or mock-treated with an equal volume of the diluent DMSO. When cell lysates were immunoblotted for β-SG, protein could be detected in all cells transfected with the β-SG isoform 1 plasmid, although the amount of protein present was increased in bortezomib-treated cells compared to DMSO-treated cells as assessed by quantitation of the immunoblot signal. By contrast, β-SG isoform 2 could only be detected in bortezomib-treated cells; no isoform 2 was detected in transfected cells treated with DMSO. Abbreviations: kDa, kilodaltons; IgG, immunoglobulin.

Only one alternatively spliced *SGCB* transcript was predicted to be translated: *SGCB* Δ exon 2 transcript, encoding β-SG isoform 2. This isoform would lack residues Q12-I82 inclusive

relative to isoform 1, including the transmembrane domain. However, it could only be detected in cells treated with proteasome inhibitor (Figure 5.1). By contrast the full-length β-SG isoform 1 could be detected in both bortezomib-treated and DMSO-treated cells, although the signal intensity on Western blot was higher with proteasome inhibition (Figure 5.1, Table 5.1). Additional protein bands were detected in bortezomib-treated cells compared to DMSO-treated cells, reflecting partial β-SG products normally degraded by the proteasome.

5.2.1.2. Lack of β-SG alters sarcoglycan complex assembly

The function of β-SG isoforms was examined in the context of an εβδζ heterotetramer in HEK293T cells. The aim of this experiment was to determine whether co-expression with other sarcoglycans could stabilise β-SG isoform 2, and assess its effect on sarcoglycan complex assembly and trafficking. Since ε-SG/α-SG and ζ-SG/γ-SG are paralogues, results from the εβδζ heterotetramer should apply to other sarcoglycan heterotetramers (McNally et al., 1998, Shiga et al., 2006, Straub et al., 1999). The three β-SG proteins examined were full-length isoform 1, isoform 2, and the LGMD-associated T182A mutant (described in Chapter 3 section 3.2.2.2). β-SG isoform 1 was expected to traffic to the cell surface as part of a heterotetramer, while the T182A mutant would prevent trafficking of the βδ-SG core leaving only a residual εζ-SG complex at the cell surface. While β-SG isoform 1 and T182A mutant β-SG were detected in co-transfected cell lysates, β-SG isoform 2 was not (Figure 5.2, lysate lanes). Therefore, co-expression as part of an εβδζ heterotetramer did not stabilise β-SG isoform 2. As predicted, β-SG isoform 1 was incorporated into prototypical sarcoglycan heterotetramers present at the cell surface (Figure 5.2). By contrast, only residual εζ-SG complexes were present at the surface of cells transfected with T182A mutant β-SG or β-SG isoform 2 (Figure 5.2).

Protein	Isoform	Detected in transfected cells?	Stabilised by proteasome inhibition?	N-linked glycosylation sites	N-linked glycosylated?	Purify with other SGs?	Cell surface?
β-SG	1	Yes	Yes	N158, N211, N258	Yes	Yes	Yes
	2	No	Completely	N88, N141, N188	Not tested	Not tested	Not tested
δ-SG	1	Yes	Yes	N61, N109, N285	Yes	Yes	Yes
	2	Yes	Yes	N61, N109	Yes	Yes	Yes
	Isoform 1 exon 3 truncation	No	Completely	N66, N242	Not tested	Not tested	Not tested
	Isoform 1 Δ exon 6	Yes - weak	Yes	N61, N109, N245	Yes	Not tested	Not tested
	3	Yes	No	N61, N109	Yes	Yes	Yes
ζ-SG	1	Yes	No	N75, N123	Yes	Yes	Yes
	2	Yes	No	N75, N89	Yes	Modified	No
	3	Yes	Yes	N75, N123	Yes	Modified	No
γ-SG	1	Yes	No	N110	Yes	Yes	Yes
	2	Yes	No	N110	Yes	Modified	Yes
	3	Yes	No	N110	Yes	Modified	Yes
α-SG	1	Yes	Yes	N174, N246	Yes	Yes	Yes
	2	No	Not tested	N174	Not tested	Not tested	Not tested
	3	Yes	Yes	N174	Yes	Yes	No
	4	Yes	Yes	N174, N246	Yes	Yes	Yes

Table 5.1 Translation and glycosylation of β-, δ-, ζ-, γ-, and α-SG protein isoforms. The main β-, δ-, ζ-, γ- and α-SG isoforms were each expressed in HEK293T cells, and tested for stability in the absence versus presence of the proteasome inhibitor bortezomib and N-linked glycosylation status via PNGase F treatment of recombinant protein. Experiments were repeated at least twice to ensure results were consistent. The results are tabulated here, alongside the predicted glycosylation sites derived from amino acid sequence. Whether an isoform was stabilised by proteasome inhibition was determined by comparing protein levels in cells treated with bortezomib to cells mock-treated with DMSO through quantitation of immunoblot signal. If an isoform was only detectable in cells treated with bortezomib, it was deemed completely stabilised by proteasome inhibition. If the isoform could be detected in both bortezomib- and DMSO-treated cells, it was deemed stabilised – "Yes" in the table. If protein levels for the isoform were unchanged between DMSO-treated and bortezomib-treated cells, then the isoform was regarded as not stabilised by proteasome inhibition – "No". Abbreviations: N, asparagine.

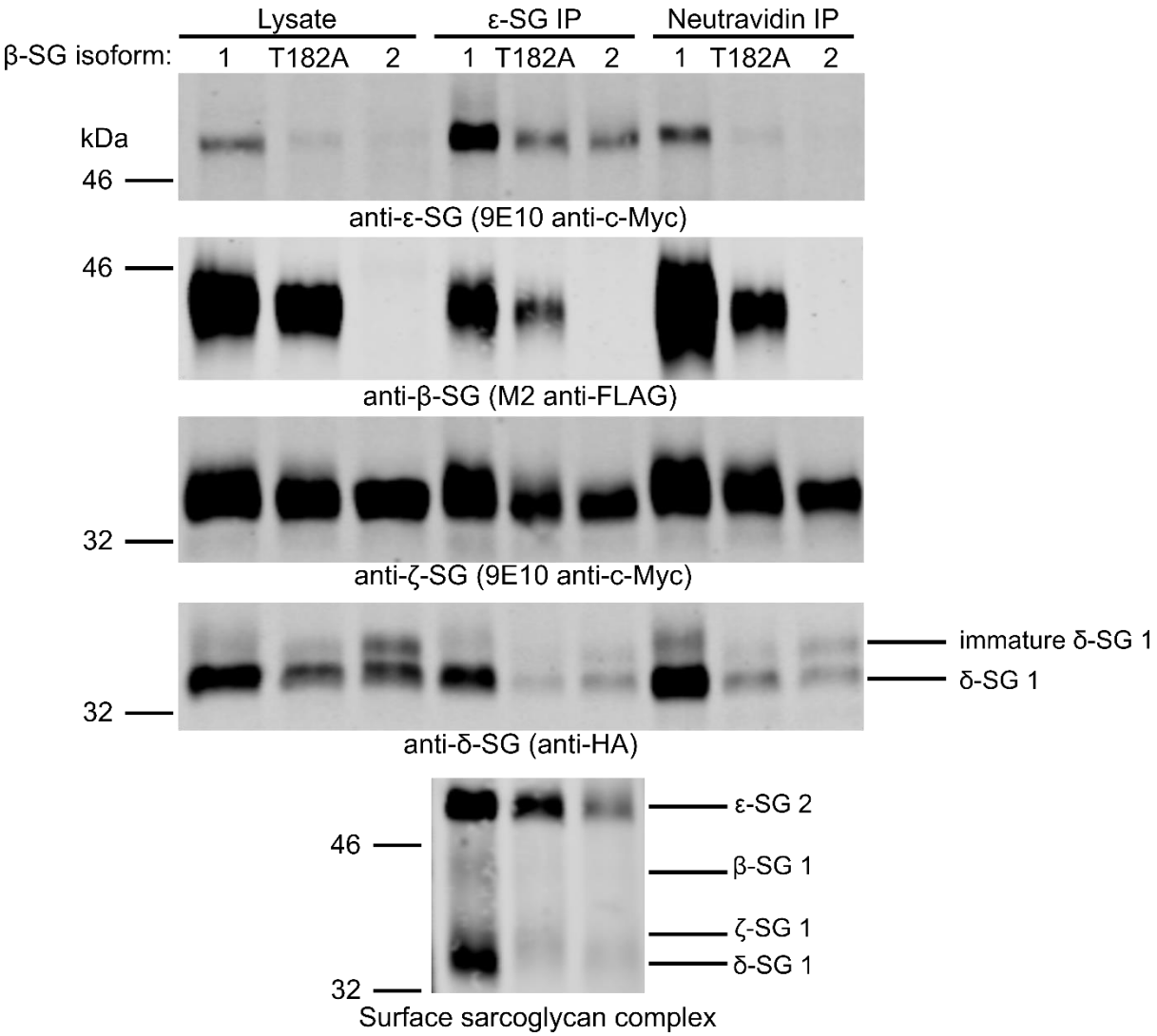


Figure 5.2 Incorporation and trafficking of β-SG isoforms in εβδζ sarcoglycan heterotetramers. HEK293T cells were transfected with plasmids encoding full-length c-Myc-tagged ε-SG isoform 2, HA-tagged δ-SG, c-Myc-tagged ζ-SG, and one of three FLAG-tagged β-SG proteins: wild-type isoform 1, T182A mutant β-SG, or β-SG isoform 2. After surface biotinylation, protein complexes were immunoprecipitated with either anti-ε-SG antibody beads or NeutrAvidin. In cell lysates, β-SG isoform 1 and T182A could be detected but not isoform 2 despite co-transfection with other sarcoglycans. Both isoform 1 and T182A mutant β-SG co-purified with ε-SG, but biotinylation of T182A mutant β-SG was substantially reduced compared to isoform 1 indicating failure to traffic to the cell surface. Biotinylation of δ-SG was also reduced, but some εζ-SG complex could be detected at the cell surface in the presence of T182A mutant β-SG. Co-transfection of β-SG isoform 2 also resulted in presence of only an εζ-SG complex at the cell surface. Abbreviations: kDa, kilodaltons; 1, wild-type β-SG isoform 1; T182A, β-SG with T182A mutation; 2, β-SG isoform 2; IP, immunoprecipitation.

5.2.2. Alternatively spliced δ-SG isoforms

5.2.2.1. Some alternatively spliced *SGCD* transcripts were translated to produce glycosylated δ-SG isoforms

There are two major δ-SG isoforms generated through use of two different transcription termination sites (Chapter 4 section 4.2.2). δ-SG isoform 1 has a full-length, 233aa extracellular region with N-linked glycans at N60, N108 and N284 plus two disulphide bonds (Chapter 4 Figure 4.8). By contrast, δ-SG isoform 2 has a truncated C-terminus with a unique sequence after amino acid residue E233, and therefore lacks the disulphide bonds and final N-linked glycan found in isoform 1 (Chapter 4, Figure 4.8). Both isoforms were detected in transfected cells, and proteasome inhibition increased protein levels compared to DMSO-treated cells as indicated by increased signal intensity on Western blot (Figure 5.3, Table 5.1). Both isoforms were N-glycosylated (Figure 5.4, Table 5.1).

Three alternatively spliced transcripts also encoded δ-SG isoforms. The *SGCD* exon 3 truncation transcript encoded the isoform 1 exon 3 truncation protein lacked amino acid residues V22-I64 including the transmembrane domain (Chapter 4 Figure 4.8). This isoform could only be detected in transfected cells treated with bortezomib, indicating that it would normally be degraded by the proteasome (Figure 5.3, Table 5.1). The *SGCD* Δ exon 6 transcript encoded an isoform lacking residues G128-L167 from the middle of the extracellular region, which was detected in all transfected cells although proteasome inhibition increased protein levels (Figure 5.3, Table 5.1). Finally, *SGCD* Δ exon 7 transcripts terminating in exon 8b encoded δ-SG isoform 3 with a unique, truncated C-terminus from residue G168 (Chapter 4 Figure 4.8). This isoform was detected in all transfected cells, with no effect from bortezomib treatment compared to DMSO treatment (Figure 5.3, Table 5.1).

These isoforms were also N-glycosylated, although de-glycosylated isoform 3 merged with the SDS-PAGE dye front and therefore was incompletely resolved (Figure 5.4, Table 5.1).

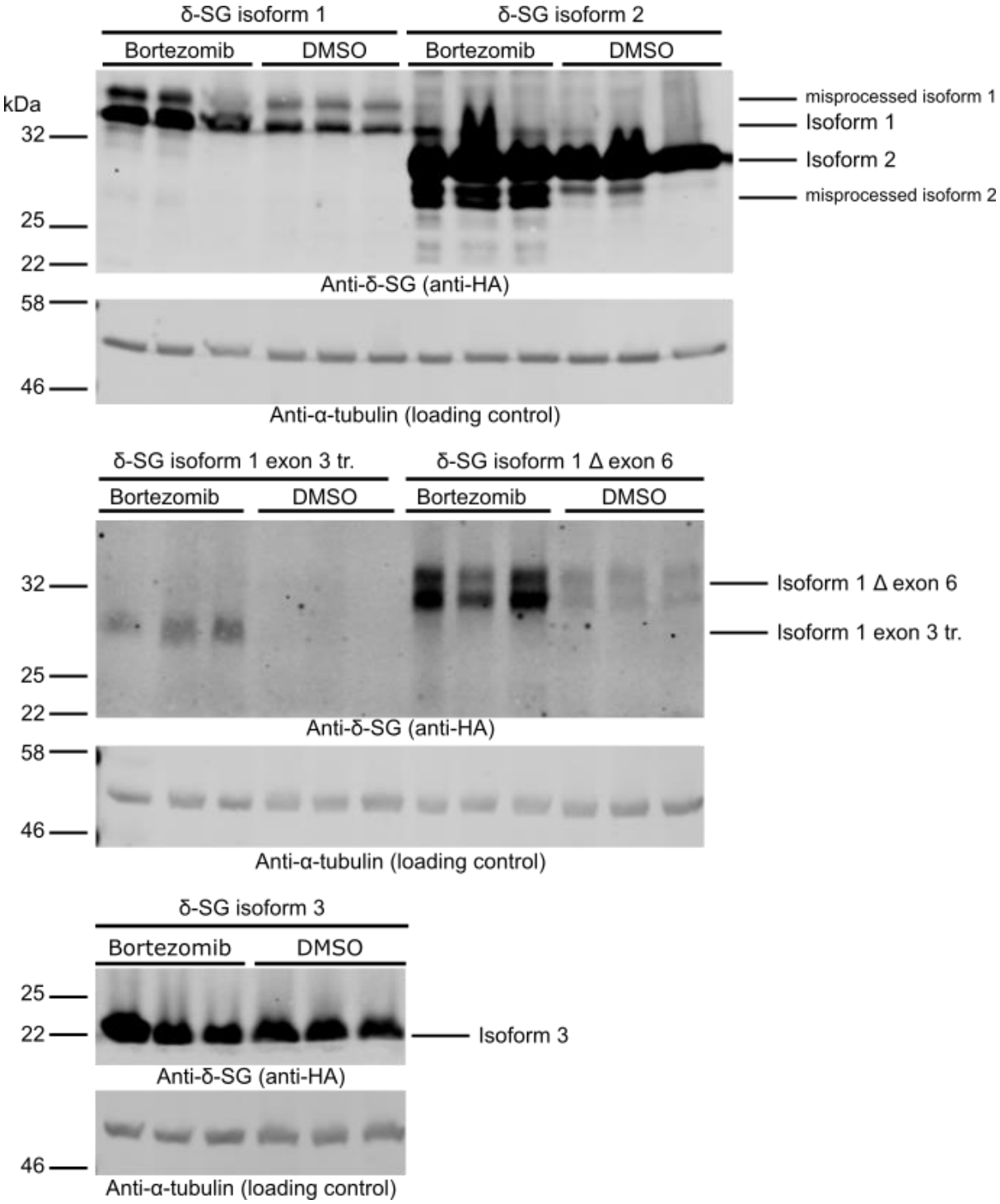


Figure 5.3 Most alternative δ-SG isoforms were stabilised by proteasome inhibition. Plasmids encoding HA-tagged δ-SG isoform 1, isoform 2, isoform 1 exon 3 truncation, isoform 1 Δ exon 6 and isoform 3 were transfected in triplicate into HEK293T cells; these cells were then either treated with the proteasome inhibitor bortezomib or mock-treated with the diluent DMSO. When cell lysates were immunoblotted for δ-SG, all isoforms except for isoform 3 showed an increase in protein levels with bortezomib treatment compared to DMSO treatment, indicating that they were stabilised by protein inhibition. The isoform 1 exon 3 truncation protein could only be detected in cells treated with bortezomib and not at all in DMSO-treated cells; this suggested full degradation of the isoform by the proteasome. Additionally, levels of isoform 1 Δ exon 6 were very low unless cells were treated with bortezomib. Abbreviations: kDa, kilodaltons; IgG, immunoglobulin.

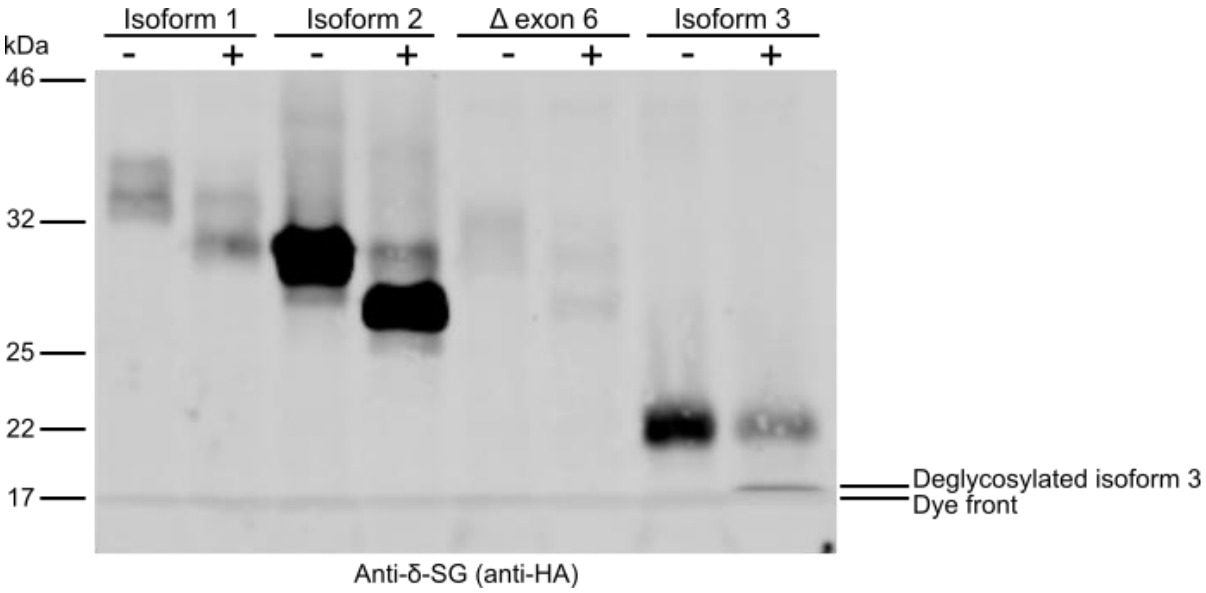


Figure 5.4 Four δ-SG isoforms were N-glycosylated as predicted. Transcripts encoding each HA-tagged δ-SG isoform – isoform 1, isoform 2, isoform 1 Δ exon 6, and isoform 3 – were transfected into HEK293T cells, and half of the cell lysate was treated with PNGase F to remove all N-linked glycans. Immunoblot of the untreated and treated lysates for δ-SG revealed a size shift for all four isoforms after PNGase F treatment, indicating glycosylation of the proteins. Abbreviations: kDa, kilodaltons; -, untreated lysate; +, PNGase F-treated lysate; Δ exon 6, δ-SG isoform 1 Δ exon 6.

5.2.2.2. Alternatively spliced δ-sarcoglycan isoforms do not affect assembly or trafficking of sarcoglycan εβδζ heterotetramers

The effects of δ-SG isoforms on sarcoglycan complex assembly and trafficking were tested in the context of an εβδζ heterotetramer. Two isoforms were not tested: δ-SG isoform 1 Δ exon 6 and δ-SG isoform 1 exon 3 truncation. The former was encoded by a rare transcript (1-3% of *SGCD* cDNA mini-libraries, Chapter 4 section 4.2.2) and poorly translated in HEK293T cells, while the latter required proteasome inhibition for detection in transfected cells (Table 5.1). δ-SG isoform 1 exon 3 truncation was similar to β-SG isoform 2, and therefore co-expression with other sarcoglycans was not predicted to stabilise δ-SG isoform 1 exon 3 truncation either. Therefore, only isoforms 1, 2 and 3 were tested for incorporation into prototypical cell surface sarcoglycan heterotetramers. Complexes lacking δ-SG were also tested for assembly and trafficking.

As predicted, δ-SG isoform 1 was incorporated into a prototypical sarcoglycan heterotetramer that trafficked to the cell surface (Figure 5.5, isoform 1 lanes). δ-SG isoforms

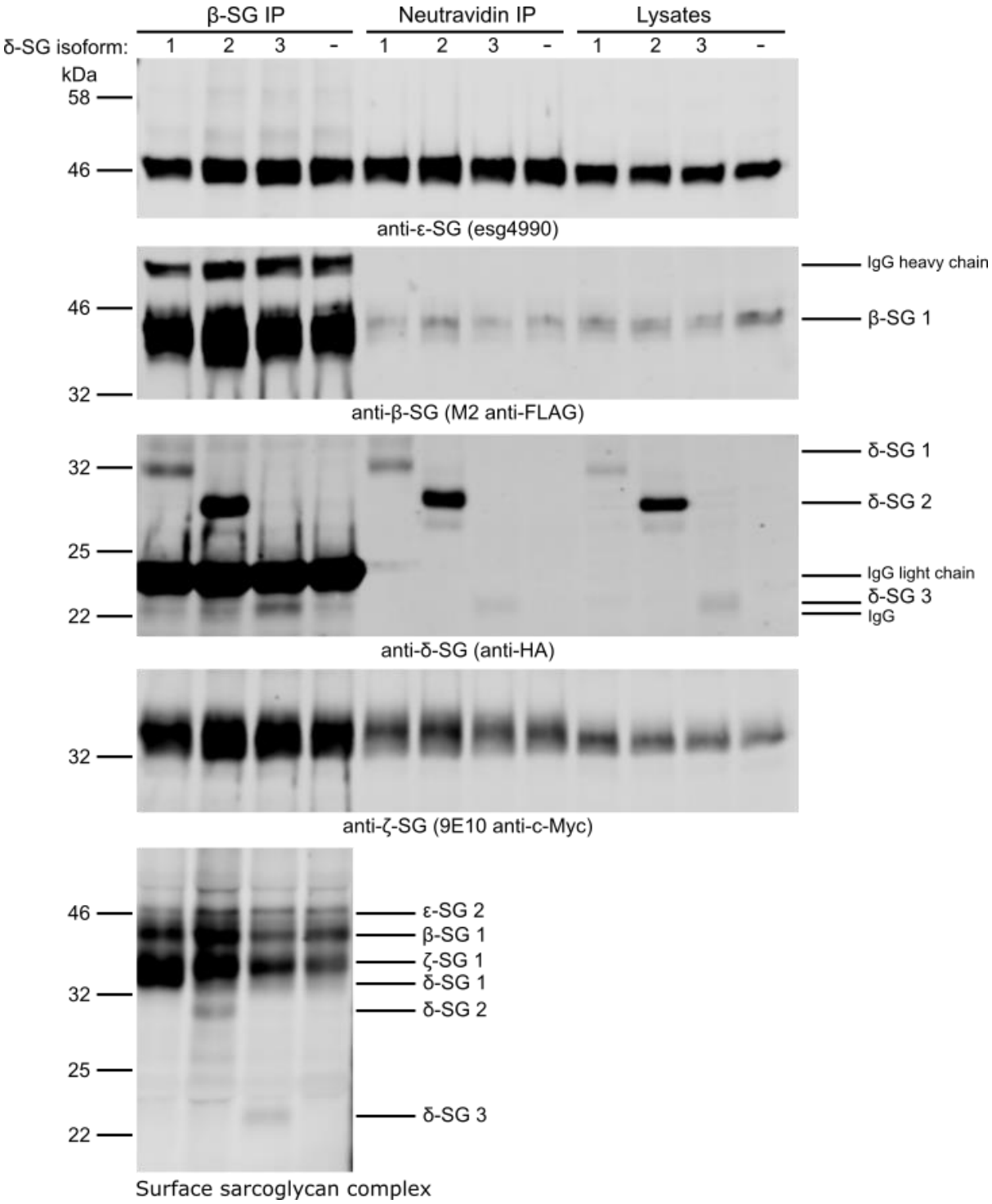


Figure 5.5 Incorporation and trafficking of δ-SG isoforms in εβδζ sarcoglycan heterotetramers. HEK293T cells were transfected with plasmids encoding full-length c-Myc-tagged ε-SG isoform 2, FLAG-tagged β-SG, c-Myc-tagged ζ-SG, and a HA-tagged δ-SG isoform: isoform 1, isoform 2, isoform 3, or no δ-SG at all (-). After surface biotinylation protein complexes were immunoprecipitated with either M2 anti-FLAG antibody resin or NeutrAvidin resin. All three δ-SG isoforms co-purified with the other sarcoglycans, and were biotinylated i.e. at the cell surface. In the absence of δ-SG, the three remaining sarcoglycans co-purified and were biotinylated. Therefore δ-SG isoform did not affect sarcoglycan complex assembly or trafficking. Abbreviations: kDa, kilodaltons; -, no δ-SG transfected.

2 and 3 also co-purified with β-SG, δ-SG and ζ-SG, indicating incorporation into prototypical sarcoglycan heterotetramers (Figure 5.5, isoforms 2 and 3 lanes). These complexes were

biotinylated, indicating that εβδζ heterotetramers incorporating δ-SG isoforms 2 and 3 were present on the cell surface (Figure 5.5, isoforms 2 and 3 lanes). Additionally, biotinylated ε-, β- and ζ-SG co-purified from cells *not* expressing δ-SG (Figure 5.5, - lanes). Therefore in HEK293T cells δ-SG was not required for assembly of an εβζ complex that trafficked to the cell surface, and δ-SG isoforms did not affect sarcoglycan complex assembly or trafficking.

5.2.3. Alternatively spliced ζ-SG isoforms

5.2.3.1. ζ-SG isoforms are translated and N-glycosylated

The alternatively spliced *SGCZ* transcripts *SGCZ* Δ exon 3 and *SGCZ* Δ exon 5 were predicted to encode ζ-SG isoforms 2 and 3 respectively (Chapter 4 section 4.2.3). Isoform 2 lacked residues G80-D113 from the N-terminal extracellular region while isoform 3 lacked residues A143-G182 from the middle of the extracellular region relative to ζ-SG isoform 1 (Chapter 4 Figure 4.11). All three ζ-SG isoforms could be detected in the lysate of transfected cells and were glycosylated (Table 5.1). However, only isoform 3 was stabilised by proteasome inhibition as demonstrated by increased signal intensity on Western blot compared to DMSO-treated cells (Table 5.1).

5.2.3.2. ζ-SG isoforms 2 and 3 interfere with the assembly and trafficking of sarcoglycan heterotetramers

The ζ-SG isoforms were tested in the context of an εβδζ sarcoglycan heterotetramer to identify differences in sarcoglycan heterotetramer assembly and trafficking. All three ζ-SG isoforms co-purified with β-SG and δ-SG using M2 anti-FLAG to enrich for complexes containing FLAG-tagged β-SG (Figure 5.6, β-SG IP lanes). However, co-purification of ε-SG with β-SG was reduced in the presence of ζ-SG isoforms 2 and 3, compared to ζ-SG isoform 1 (Figure 5.6, β-SG IP lanes). Co-purification of ζ-SG isoforms 2 and 3 with the anti-ε-SG antibody was also reduced compared to ζ-SG isoform 1 (Figure 5.6). Neither ζ-SG isoform 2

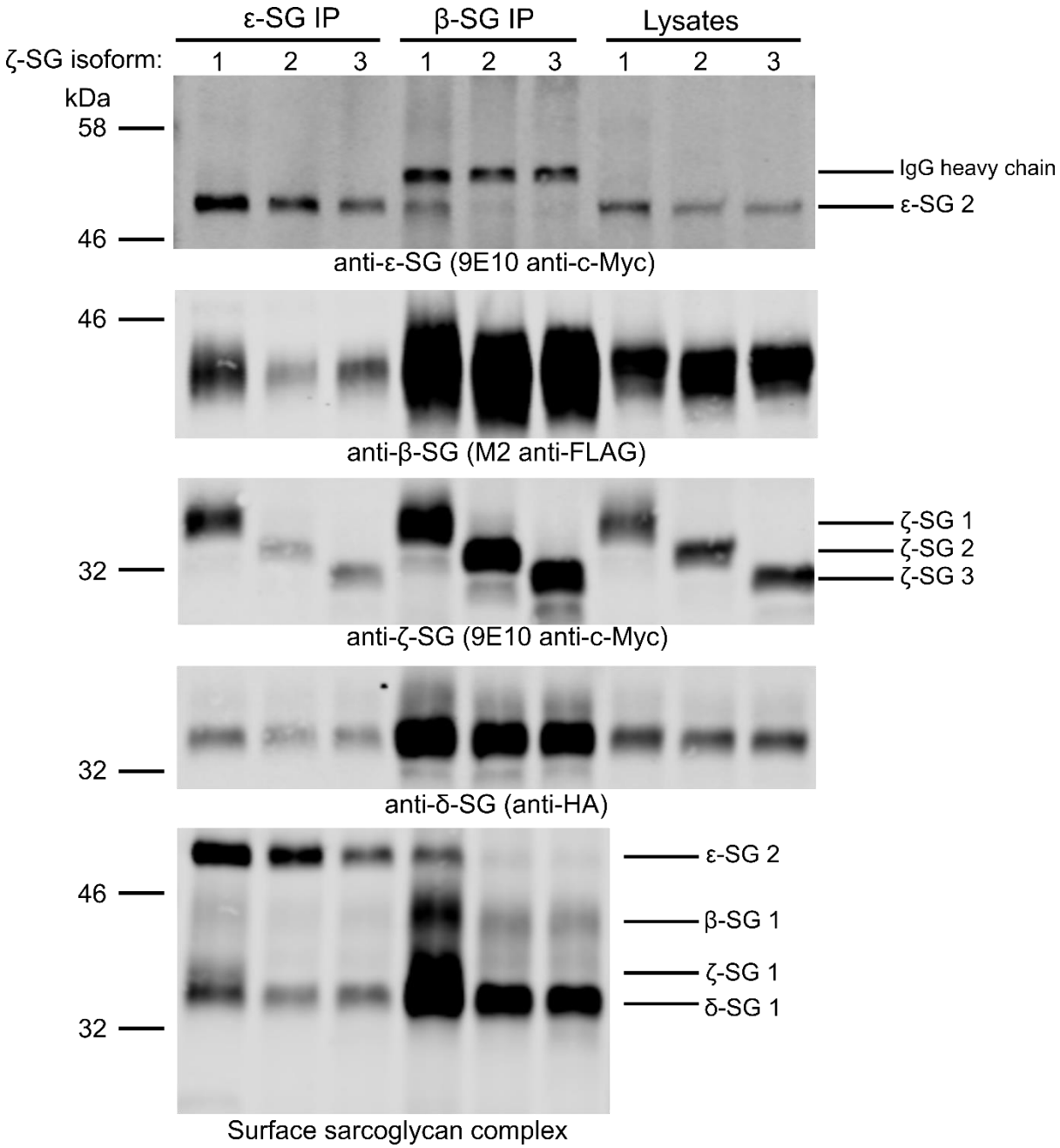


Figure 5.6 Incorporation and trafficking of ζ-SG isoforms in εβδζ sarcoglycan heterotetramers. HEK293T cells were co-transfected with plasmids encoding full-length c-Myc-tagged ε-SG isoform 2, FLAG-tagged β-SG 1, HA-tagged δ-SG 1 and a c-Myc-tagged ζ-SG isoform: isoform 1, 2 or 3. After surface biotinylation protein complexes were immunoprecipitated with either anti-ε-SG antibody or M2 anti-FLAG antibody resin. ζ-SG isoform 1 co-purified with all three other sarcoglycans, and was biotinylated indicating trafficking to the cell surface as part of a heterotetramer. By contrast, co-purification of ζ-SG isoforms 2 and 3 with ε-SG was reduced on both IPs, and led to reduced co-purification of β-SG and δ-SG on ε-SG IP. These isoforms were not biotinylated, but residual ε-SG, β-SG and δ-SG were biotinylated. Therefore, ζ-SG isoforms 2 and 3 inhibited incorporation of ε-SG into the sarcoglycan complex and did not traffic to the cell surface. Abbreviations: kDa, kilodaltons.

nor isoform 3 could be detected at the cell surface, whereas ζ-SG isoform 1 was detected at the cell surface in association with the three other sarcoglycans (Figure 5.6). Therefore, ζ-SG isoform 1 was trafficked to the cell surface as part of a prototypical sarcoglycan heterotetramer, but ζ-SG isoforms 2 and 3 were not. In addition, cell surface β-SG and δ-SG

was reduced in the presence of ζ-SG isoforms 2 and 3 (Figure 5.6). Therefore ζ-SG isoforms 2 and 3 interfered with the assembly and trafficking of sarcoglycan heterotetramers.

5.2.4. Alternatively spliced γ-SG isoforms

5.2.4.1. Alternatively spliced *SGCG* transcripts encode glycosylated γ-SG isoforms.

Two alternatively spliced transcripts were predicted to encode γ-SG isoforms in skeletal muscle. The *SGCG* Δ exon 5 transcript was predicted to encode a γ-SG isoform 2 lacking amino acid residues P130-G169 from the middle of the extracellular region, while the *SGCG* exon 6b+ transcript was predicted to encode an isoform 3 with a truncated, unique C-terminus from residue R193 (Chapter 4 Figure 4.16). All three γ-SG isoforms were expressed and glycosylated when transfected into HEK293T cells, but were not stabilised by proteasome inhibition (Table 5.1).

5.2.4.2. γ-SG isoforms alter sarcoglycan complex assembly and trafficking

The three γ-SG isoforms were examined for assembly and trafficking in the context of αβδγ heterotetramers only. While *SGCE* and *SGCG* expression overlaps in cardiac and smooth muscle, εβδγ heterotetramers are thought to be a minor species compared to αβδγ heterotetramers (Durbeej and Campbell, 1999, Lancioni et al., 2011, Noguchi et al., 1995, Noguchi et al., 2001). As expected, full-length γ-SG isoform 1 interacted with β-SG, δ-SG and α-SG to form heterotetrameric complexes that were present at the cell surface (Figure 5.7 lane 1). Isoforms 2 and 3 also co-purified with β-SG and δ-SG using M2 anti-FLAG to enrich for β-SG (Figure 5.7). However, co-purification of α-SG in β-SG IPs was slightly reduced in the presence of γ-SG isoforms 2 and 3 (Figure 5.7). Therefore, γ-SG isoforms 2 and 3 reduced incorporation of α-SG into the sarcoglycan complex. In the presence of γ-SG isoforms 2 and 3, a reduced amount of biotinylated β-SG, δ-SG and γ-SG was detected in β-

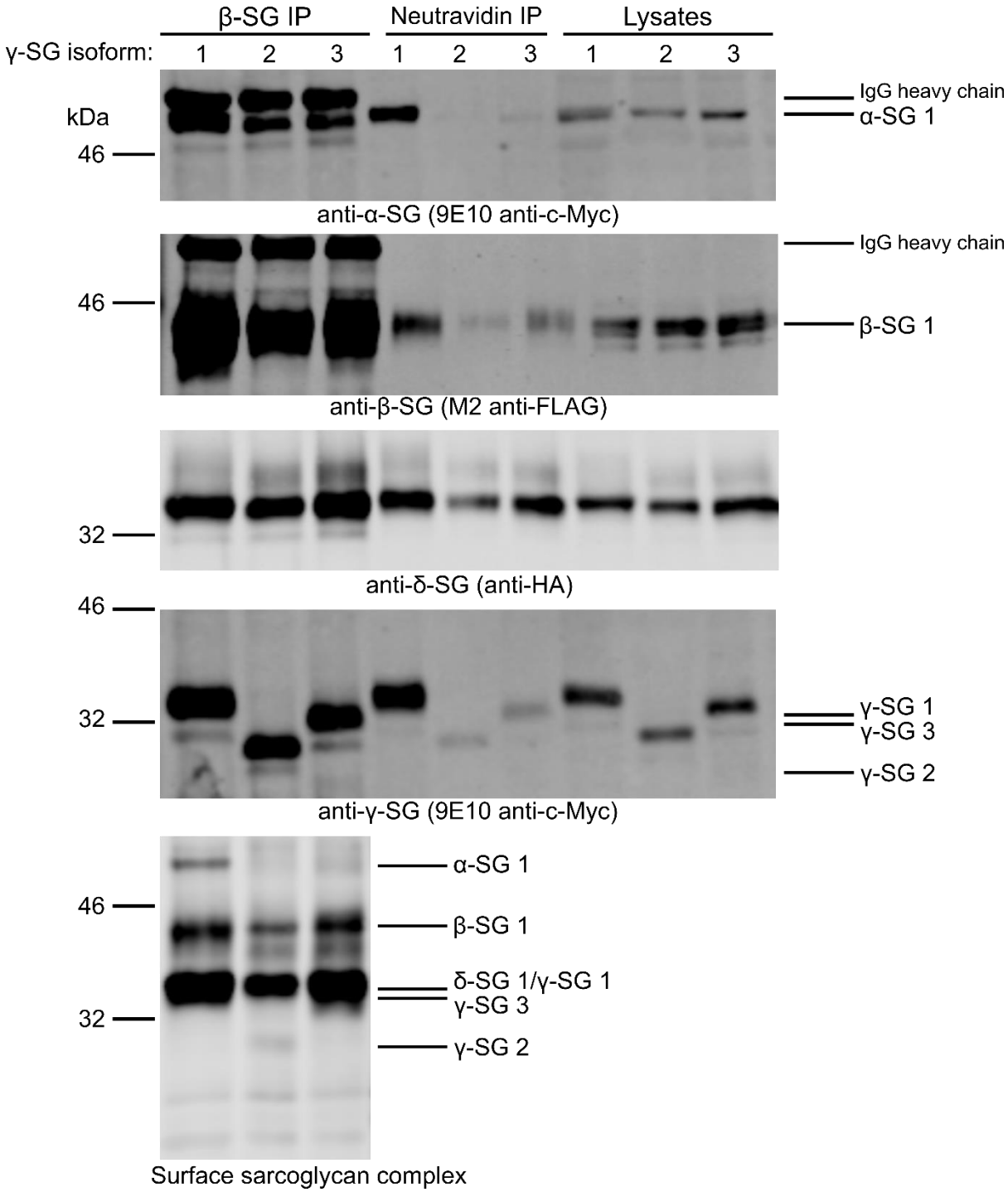


Figure 5.7 Incorporation and trafficking of γ-SG isoforms in αβδγ SGCs. HEK293T cells were co-transfected with full-length c-Myc-tagged α-SG, FLAG-tagged β-SG, HA-tagged δ-SG and a c-Myc-tagged γ-SG isoform: isoform 1, 2 or 3. After cell surface biotinylation proteins were immunoprecipitated with either M2 anti-FLAG antibody resin or NeutrAvidin resin. All three γ-SG isoforms co-purified with α-SG, β-SG and δ-SG; however, co-purification of α-SG was reduced in the presence of γ-SG isoform 2 or 3. The full heterotetramer formed with γ-SG isoform 1 was biotinylated, indicating trafficking to the cell surface. In the presence of γ-SG isoforms 2 and 3, α-SG was not detected at the cell surface but residual β-SG, δ-SG and γ-SG was present at the cell surface. Therefore, γ-SG isoforms 2 and 3 reduce incorporation of α-SG into sarcoglycan complexes, and prevent trafficking of tetramers. However residual complexes containing γ-SG isoforms 2 or 3 remain trafficking competent. Abbreviations: kDa, kilodaltons.

SG IPs (Figure 5.7 isoforms 2 and 3). However, no biotinylated α-SG was present.

NeutrAvidin affinity purification of biotinylated protein also demonstrated that no

biotinylated α-SG was present in cells expressing γ-SG isoforms 2 and 3 (Figure 5.7). Therefore, γ-SG isoforms 2 and 3 reduced incorporation of α-SG into the sarcoglycan complex, and interfered with the trafficking of αβδγ heterotetramers. However, a residual βδγ complex still trafficked to the cell surface.

5.2.5. Alternatively spliced α-SG isoforms

5.2.5.1. Alternative α-SG isoforms were N-glycosylated, and stabilised by the proteasome

Three alternatively spliced *SGCA* transcripts were predicted to be translated (Chapter 4 section 4.2.4). The most abundant in the skeletal cDNA mini-library at 27% of clones was *SGCA* Δ exon 6-7. This transcript was predicted to encode an α-SG isoform 2 lacking residues V195-R319 including the transmembrane domain and extracellular juxtamembrane region compared to isoform 1 (Chapter 4 Figure 4.13). α-SG isoform 2 was predicted to be secreted, but was not examined in the present study because its predicted mass was equal to that of immunoglobulin light chain and therefore would be difficult to detect using the available reagents. The *SGCA* exon 6 truncation transcript was predicted to encode an α-SG isoform 3 lacking amino acid residues G201-V250 from the extracellular juxtamembrane region, while the *SGCA* Δ exon 9b transcript was predicted to encode an isoform 4 with a unique, extended intracellular region from residue T367 (Chapter 4 Figure 4.13). When expressed singly in HEK293T cells, α-SG isoforms 1, 3 and 4 were all substantially stabilised by proteasome inhibition with bortezomib, although they could be detected at very low levels in DMSO-treated cells (Table 5.1). All three isoforms were also N-glycosylated as predicted

from their amino acid sequences (table 5.1). Co-transfection with other sarcoglycans stabilised all α-SG isoforms, as demonstrated in the lysates lanes of Figure 5.8.

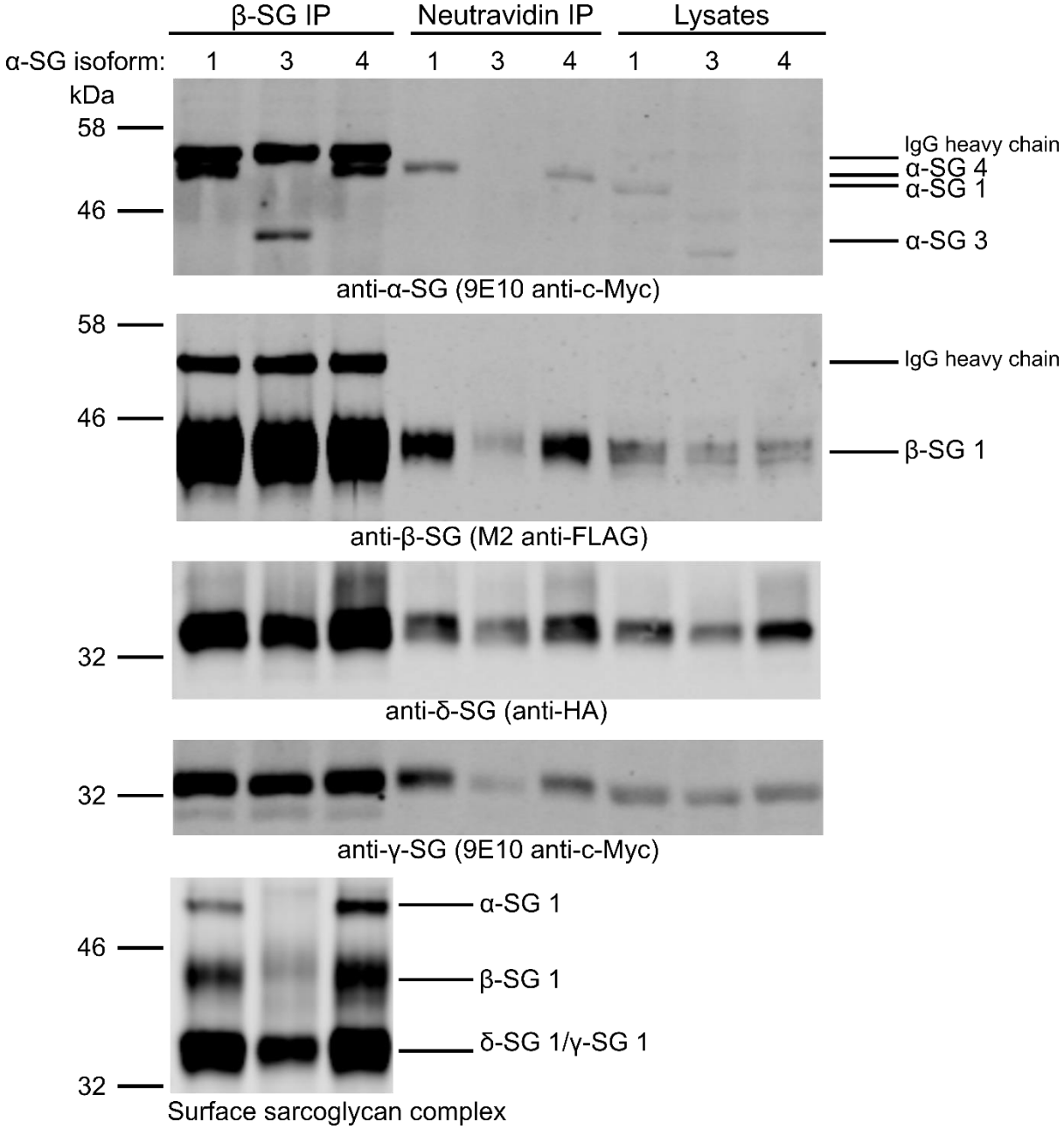


Figure 5.8 Incorporation and trafficking of α-SG isoforms in αβδγ sarcoglycan heterotetramers. HEK293T cells were co-transfected with full-length c-Myc-tagged γ-SG, FLAG-tagged β-SG, HA-tagged δ-SG and a c-Myc-tagged α-SG isoform: isoform 1, 3 or 4. After cell surface biotinylation, protein complexes were immunoprecipitated with either M2 anti-FLAG antibody resin or NeutrAvidin resin. All three isoforms co-purified with the other sarcoglycans. Isoforms 1 and 4 were biotinylated (i.e. trafficked to the cell surface) as part of sarcoglycan complexes. However, isoform 3 was not biotinylated and reduced biotinylation of the other sarcoglycans. Therefore, α-SG isoform 3 impaired trafficking of the complex to the cell surface. Some residual β-SG, δ-SG and γ-SG was still biotinylated, however. Abbreviations: kDa, kilodaltons.

5.2.5.2. Alternatively spliced α-SG isoforms had different effects on the assembly and trafficking of the SGC.

The α-SG isoforms were tested in the context of αβδγ heterotetramers. As expected, full-length α-SG isoform 1 was incorporated into prototypical sarcoglycan heterotetramers that trafficked to the cell surface (Figure 5.8, isoform 1 lanes). Isoform 4 also co-purified with β-SG, δ-SG and γ-SG in the β-SG IP, and all four immunoprecipitated sarcoglycans were biotinylated (Figure 5.8, isoform 4 lanes). Therefore, α-SG isoform 4 was incorporated into a prototypical sarcoglycan heterotetramer that trafficked to the cell surface. By contrast, while isoform 3 co-purified with the other sarcoglycans it was not biotinylated (Figure 5.8, isoform 3 lanes). Biotinylation of β-SG and γ-SG was also reduced in the presence of α-SG isoform 3 (Figure 5.8, isoform 3 lanes). Therefore, α-SG isoform 3 could assemble into a prototypical sarcoglycan heterotetramer but impaired trafficking of that heterotetramer to the cell surface.

5.3. Discussion

While several protein-coding alternatively spliced transcripts were identified for *SGCB, SGCD, SGCZ, SGCG* and *SGCZ* in Chapter 4, the functionality of the encoded isoforms was unknown. Therefore, the aim of this study was to examine several key attributes of those sarcoglycan isoforms to assess their functionality. Initially the stability and glycosylation of isoforms was examined and compared to the full-length isoform 1 of each sarcoglycan. Although β-SG isoform 2 and δ-SG isoform 1 exon 3 truncation could only be detected in transfected cells when the proteasome was inhibited, all other isoforms were expressed and N-glycosylated as predicted (Table 5.1). Subsequently, a panel of sarcoglycan isoforms were tested for incorporation into prototypical sarcoglycan heterotetramers and trafficking to the cell surface. Isoforms were tested in physiologically relevant heterotetramers: *SGCZ* expression does not overlap with *SGCA* so ζ-SG isoforms were tested in the context of εβδζ

heterotetramers, while α-SG and γ-SG isoforms were tested in the context of αβδγ heterotetramers (Ettinger et al., 1997, Noguchi et al., 2001, Roberds et al., 1993a, Shiga et al., 2006, Wheeler et al., 2002, Yamamoto et al., 1994). β-SG and δ-SG are present in all prototypical sarcoglycan heterotetramers, so isoforms of these proteins were tested in the context of εβδζ heterotetramers following the premise that complexes containing paralogues were functionally equivalent (McNally et al., 1998, Shiga et al., 2006, Straub et al., 1999).

The failure of wild-type and alternatively spliced δ-SG isoforms to affect trafficking of sarcoglycan heterotetramers was unexpected (Figure 5.5). A previous study found that residues 57-92 of δ-SG were required for the interaction with β-SG, while the extracellular C-terminal disulphide bonds and N-linked glycosylation were required for trafficking to the plasma membrane in heterologous cells (Chen et al., 2006). δ-SG isoforms 2 and 3 lack the disulphide bonds and final N-linked glycan found in δ-SG isoform 1 (Chapter 4 Figure 4.7). Therefore, these isoforms were expected to interact normally with other sarcoglycans but not traffic to the plasma membrane. However, when expressed alongside ε-SG, β-SG and ζ-SG in HEK293T cells, δ-SG isoforms 2 and 3 were detected at the cell surface as part of εβδζ heterotetramers (Figure 5.5). This contradictory result may be due to the system used to model sarcoglycan interactions and trafficking. Importantly, the previous study examined the interaction and trafficking of β-SG and δ-SG in isolation, without other sarcoglycans (Chen et al., 2006). By contrast, in the present study I examined δ-SG isoforms in a physiologically relevant, heterotetrameric context. Inclusion of ε-SG and ζ-SG in the sarcoglycan complex alongside β-SG and δ-SG might be compensating for the loss of disulphide bonds in δ-SG and allow the complex to traffic to the cell surface. In particular, the high similarity between the paralogues δ-SG and ζ-SG could permit ζ-SG to compensate for the putative trafficking defect in δ-SG isoforms 2 and 3 (Shiga et al., 2006, Wheeler et al., 2002). Further research

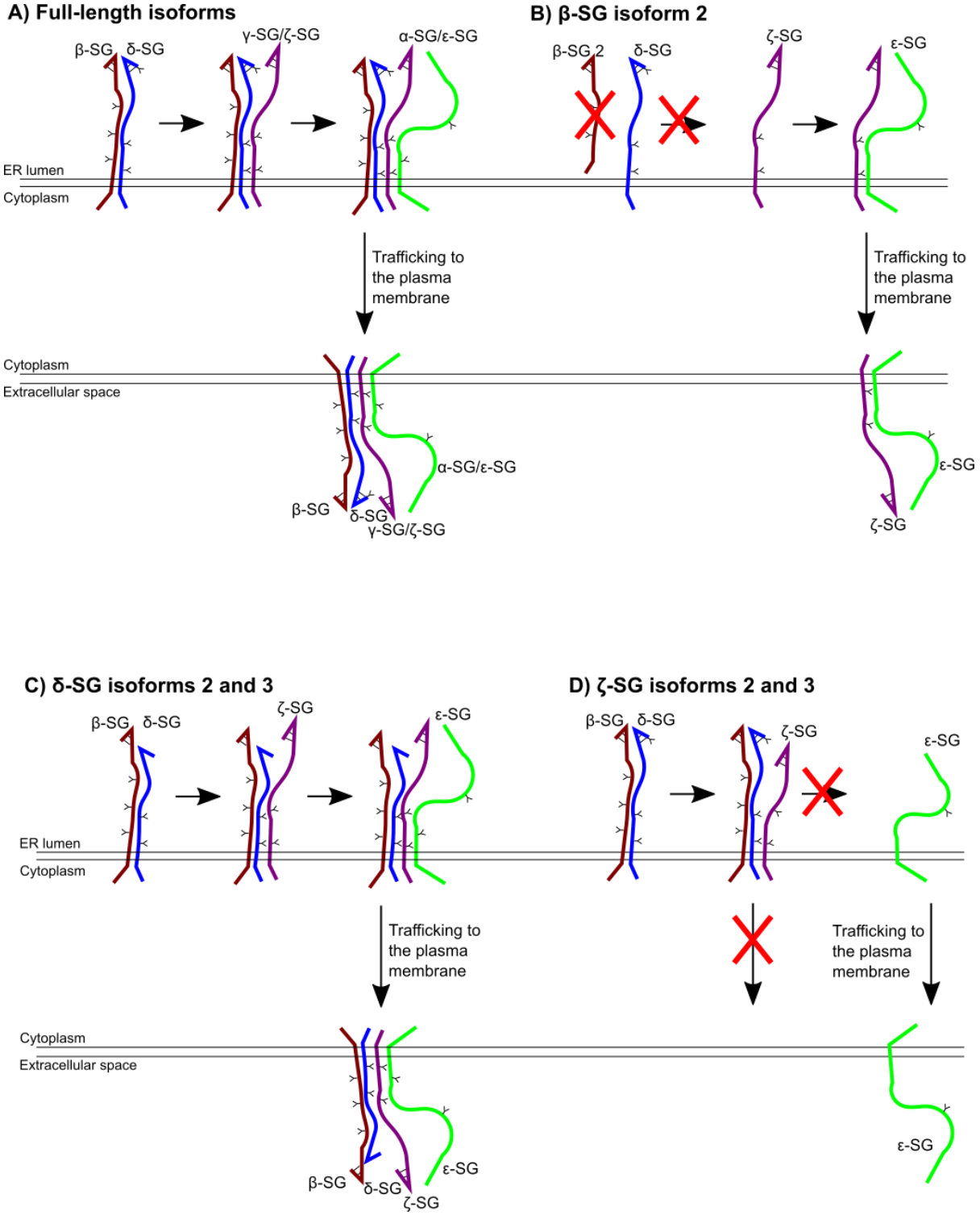
will be required to determine whether this is the case, and also to examine the behaviour of muscle-type αβδγ heterotetramers containing δ-SG isoforms 2 and 3.

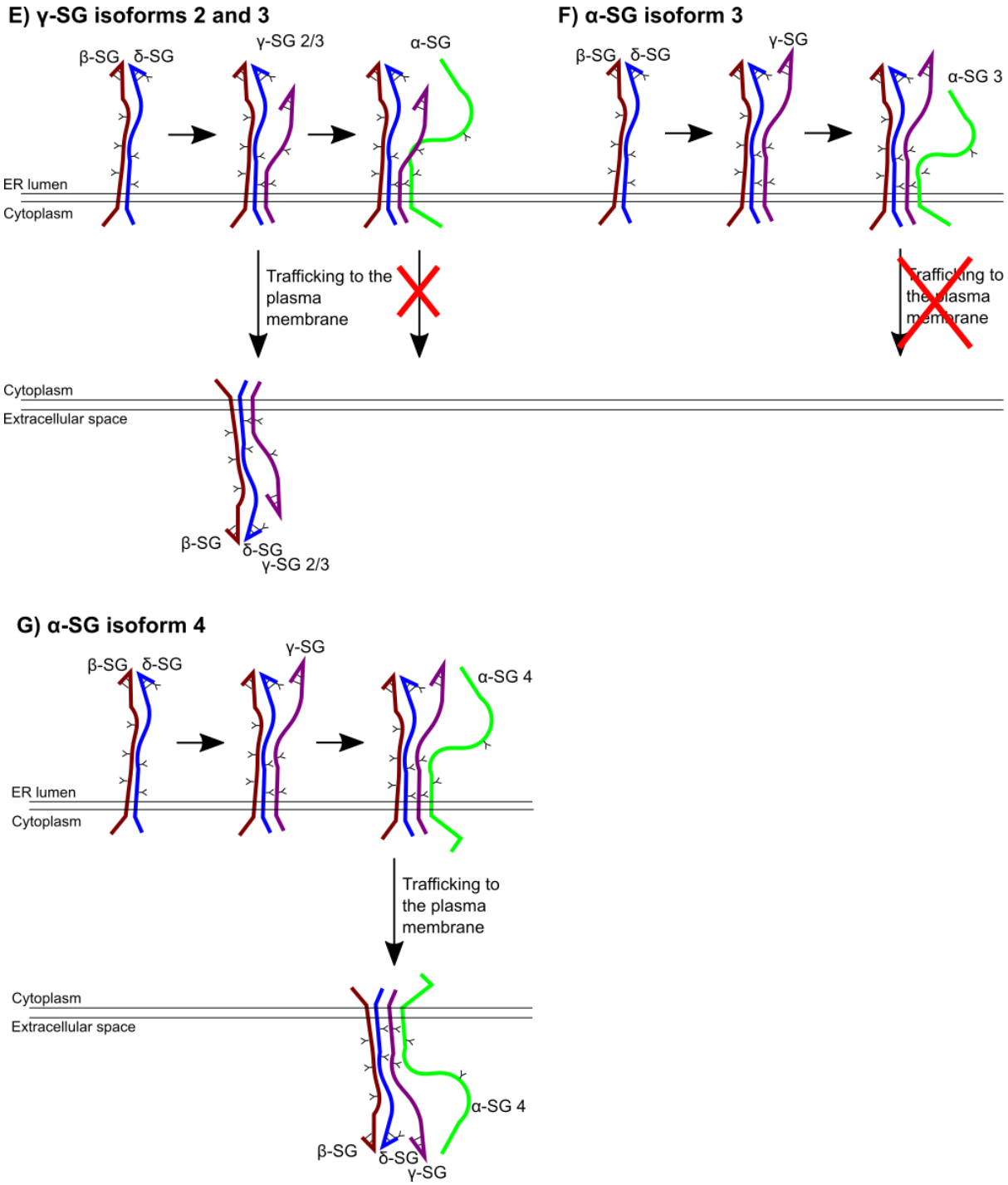
Also unexpected were the differences in cell surface trafficking between ζ-SG and γ-SG isoforms. As mentioned above, ζ-SG and γ-SG are paralogues, and generally thought to be functionally equivalent (Ozawa et al., 2005, Shiga et al., 2006, Wheeler et al., 2002). In addition, both *SGCZ* and *SGCG* produce Δ exon 5 transcripts. *SGCZ* Δ exon 5 transcripts encode ζ-SG isoform 3 lacking residues A143-G182 from the middle of the extracellular region, while *SGCG* Δ exon 5 transcripts encode γ-SG isoform 2 lacking the equivalent residues P130-G169 (Chapter 4 Tables 4.3 and 4.5). Both ζ-SG isoforms and γ-SG isoforms produced through alternative splicing reduced incorporation of α-SG/ε-SG into the sarcoglycan complex (Figures 5.6 and 5.7). However, residual γ-SG isoforms 2 and 3 could still be detected at the cell surface as part of βδγ complexes, whereas ζ-SG isoforms 2 and 3 were completely absent from the cell surface. This difference in trafficking capacity of directly equivalent ζ-SG and γ-SG isoforms suggested that these proteins may not be completely interchangeable. Inclusion of ζ-SG versus γ-SG in a sarcoglycan heterotetramer could affect trafficking and perhaps protein interactions in as yet unidentified ways.

The α-SG isoforms studied had distinct effects on trafficking of sarcoglycan heterotetramers. As described above, the extracellular juxtamembrane region of δ-SG is thought to mediate interaction with other sarcoglycans (Chen et al., 2006). However, α-SG isoform 3 lacked the extracellular juxtamembrane domain and yet was able to assemble into a sarcoglycan heterotetramer (Figure 5.8). Instead, loss of amino acid residues 201-250 impaired its ability to traffic to the plasma membrane. By contrast the elongation of the intracellular domain in α-SG isoform 4 did not affect sarcoglycan heterotetramer assembly or trafficking compared to α-SG isoform 1. This parallels the identical sarcoglycan heterotetramer assembly and trafficking of ε-SG isoforms with differences in the intracellular region (Chapter 3 section

3.2.2.2). For these paralogues, the intracellular domain does not affect assembly or trafficking.

The observed effects of β-, δ-, ζ-, γ- and α-SG isoforms on sarcoglycan heterotetramer assembly and trafficking in HEK293T cells generally support the current model of





sarcoglycan complex assembly and trafficking as illustrated in Figure 5.9 (Chan et al., 1998, Hack et al., 2000, Noguchi et al., 2000, Shi et al., 2004, Yoshida et al., 1997, Yoshida et al., 1994). In this model, sarcoglycan heterotetramer assembly is thought to occur sequentially: β-SG and δ-SG form a core dimer with which γ-SG or ζ-SG interacts and then α-SG or ε-SG interacts more loosely with the resultant trimer (Chan et al., 1998, Noguchi et al., 2000, Yoshida et al., 1997, Yoshida et al., 1994). Co-IP experiments have indicated that γ-SG/ζ-SG

Figure 5.9 Summary of the proposed model for sarcoglycan complex assembly and trafficking, with the effects of different alternative sarcoglycan isoforms. A) Model of canonical sarcoglycan complex assembly. On the ER membrane, β-SG and δ-SG interact first to form the βδ-SG core. Either γ-SG or ζ-SG then interacts with the βδ-SG core. Finally, either α-SG or ε-SG interact with the complex – apparently directly through γ-SG/ζ-SG – to complete the heterotetramer. This complex then traffics to the plasma membrane via the Golgi apparatus (not shown on this model). B) β-SG isoform 2, which lacks a transmembrane domain, is degraded by the proteasome and so the βδ-SG core cannot form. However, ε-SG and ζ-SG can still interact to form a residual complex that traffics to the plasma membrane. C) The alternative δ-SG isoforms 2 and 3 still interact with the other sarcoglycans to form a heterotetramer that traffics to the cell surface despite lacking the disulphide bonds of isoform 1. D) ζ-SG isoforms 2 and 3 lack sections of the extracellular domain, and interact with the βδ-SG core to form a βδζ complex. However, ε-SG is not incorporated into this complex and the complex does not traffic to the cell surface. ε-SG traffics independently to the cell surface, though. E) Alternative γ-SG isoforms 2 and 3 have alterations to the extracellular domain compared to isoform 1, and interact with the other sarcoglycans. However, the full tetramer with α-SG does not traffic to the cell surface; instead, partial complexes containing β-SG, δ-SG and γ-SG isoform 2/3 can be detected at the cell surface. F) α-SG isoform 3 lacks the extracellular juxtamembrane domain compared to isoform 1, and interacts with the βδγ sub-complex. However, the resultant heterotetramer does not traffic to the cell surface. G) α-SG isoform 4 has an extended intracellular domain but incorporates into a cell surface complex as in A.

may mediate the interaction between α-SG/ε-SG and the remainder of the sarcoglycan complex (Chan et al., 1998, Hack et al., 2000, Noguchi et al., 2000, Shi et al., 2004). This would suggest a direct interaction between γ-SG/ζ-SG and α-SG/ε-SG. The specific disruption of α-SG/ε-SG incorporation into the sarcoglycan complex by alternative γ-SG and ζ-SG isoforms supports that hypothesis (Figure 5.9d, e). This implicates ζ-SG residues 80-113, ζ-SG residues 143-182 or γ-SG residues 130-169, and the C-terminus of γ-SG in the association with α-SG/ε-SG, as described below.

Based on the data presented in this study, δ-SG seems to be non-essential in assembly of the sarcoglycan complex (Figure 5.5, 5.9c). Even in the absence of δ-SG, ε-SG, β-SG and ζ-SG interacted to form a complex that trafficked to the cell surface (Figure 5.5). However humans and mice carrying *SGCD* mutations develop muscular dystrophy with loss of the entire SGC from the sarcolemma, indicating a requirement for δ-SG *in vivo* (Coral-Vazquez et al., 1999, Duggan et al., 1997b). While the reasons for this apparent contradiction are unclear, one possibility is that *in vivo* additional mechanisms ensure partial complexes lacking δ-SG cannot traffic to the plasma membrane. These mechanisms may not be captured in experiments carried out using heterologous cells and might involve δ-SG-specific protein interactions, or conformation of the sarcoglycan complex as a whole.

The sarcoglycan isoforms also provide additional insight into the protein regions required for interactions between sarcoglycans. Chen et al. (2006) identified residues 57-92 in mouse δ-SG as required for the interaction with β-SG. None of the δ-SG isoforms studied affected this region, and accordingly they all interacted normally with β-SG (Figure 5.5). However, Chen et al. (2006) also found that γ-SG residues 94-194 were required for the interaction with the βδ-SG core. γ-SG isoform 2 lacked amino acid residues P130-P169, whereas ζ-SG isoform 2 removed amino acid residues G80-D113 that correspond to γ-SG residues G69-D100 (Chapter 4 Figure 4.10). Both of these isoforms still interacted with the βδ-SG core, and therefore the interaction between γ-SG and the βδ-SG core most likely requires γ-SG residues 101-129 or a more distal region (Figures 5.6-5.7). This assumes that γ-SG and ζ-SG interact with the βδ-SG core in the same way. However, all three regions of γ-SG and ζ-SG affected in alternatively spliced isoforms were found to modulate the interaction with α-SG/ε-SG, however. This suggested the overall extracellular conformation may be more important for the interaction with α-SG/ε-SG than any specific region, since most of the γ-SG/ζ-SG extracellular region was altered in at least one isoform. Finally, Chen et al. (2006) proposed that the extracellular juxtamembrane region of α-SG might be important for its interaction with γ-SG based on clustering of LGMD 2D-associated mutations in that region. This region is lost in α-SG isoform 3, but the isoform can be incorporated into a prototypical sarcoglycan heterotetramer (Figure 5.8). Instead, trafficking of the complex to the cell surface was impaired. This suggested that the α-SG extracellular juxtamembrane region may be required for trafficking rather than sarcoglycan complex interactions.

The prevalence of sarcoglycan isoforms described in this study also have implications for the molecular pathogenesis and treatment of sarcoglycan-associated LGMD and MD. A recent proof-of-concept study examined the use of induced exon skipping as a means to treat LGMD 2C, which is caused by autosomal recessive *SGCG* mutations as covered in Chapter 1 section

1.3.2 (Gao et al., 2015). Although the "Mini-Gamma" protein generated through skipping exons 4-7 improved the phenotype of *Sgcg* knockout mice, mice still had signs of muscular dystrophy (Gao et al., 2015). Based on the observed effects of γ-SG isoforms 2 and 3 described in this study, which like "Mini-Gamma" lacked sections of the extracellular domain, "Mini-Gamma" is likely to interfere with inclusion of α-SG into muscle sarcoglycan complexes and this could contribute to the incomplete phenotype rescue observed in mice (Gao et al., 2015). The observed differences between similar γ-SG and ζ-SG also demonstrate that sarcoglycan heterotetramer composition may be more important for function than previously thought. Paralogous sarcoglycans may be able to replace each other in some functions but not others, and incorporation of ζ-SG versus γ-SG in brain complexes could impact trafficking of those complexes. As discussed in Chapter 4 (section 4.2.3), transcripts encoding ζ-SG isoforms 2 and 3 were relatively abundant in the cerebellum *SGCZ* cDNA mini-library. ζ-SG isoforms 2 and 3 impair sarcoglycan complex assembly and trafficking to the cell surface, so their presence in the brain could lead to cell surface ε-SG that is not associated with a sarcoglycan complex. Distinct functions of this pool of sarcoglycan complex-independent ε-SG could contribute to the specifically neurological phenotype of *SGCE* mutations. Since this independent cell surface ε-SG would not be affected by LGMD-associated sarcoglycan mutations, it could also contribute to the absence of neurological signs in LGMD.

In conclusion, the sarcoglycan isoforms produced from alternatively spliced transcripts described in Chapter 4 had a range of effects on the assembly and trafficking of the sarcoglycan complex. One potentially abundant isoform could not be studied, however: α-SG isoform 2. This isoform is approximately the same size as mouse immunoglobulin light chain, and so could not be resolved from immunoglobulin on Western blots using the available reagents. Production of additional antibodies raised in non-mouse species would

facilitate the study of this isoform. Another outstanding question is whether the sarcoglycan isoforms studied here are actually present *in vivo*. Due to the sarcoglycan complex-disrupting effects of several isoforms, novel antibodies against the sarcoglycans would also resolve this question. However, tissue source may be a difficulty: the transcripts encoding these isoforms were identified in human tissue, and it is not clear whether they are also present in mice which would be the most accessible animal model. Indeed, previous research suggested that overall alternative splicing is less abundant in mice compared to humans (Kim et al., 2007). This could complicate efforts to detect these isoforms *in vivo*. However, the present study does provide a basic characterisation of several novel sarcoglycan isoforms, and highlights the potential importance of ζ-SG isoforms in the brain that may be relevant to the molecular pathogenesis of MD.

Chapter 6: General discussion

6.1.Introduction

The overall aim of this study was to identify proteins interacting with ε-SG in the brain, and from that gain insight into the molecular pathogenesis of *SGCE* mutation-positive MD. To achieve this aim, three major objectives were pursued:

- To determine whether the ubiquitous and brain-specific ε-SG isoforms are part of DGC-like complexes in the brain.
- To identify additional, non-DGC proteins with which ε-SG interacts in the brain.
- To systematically examine alternative splicing of *SGCA, SGCB, SGCD, SGCG* and *SGCZ* to identify transcripts encoding novel sarcoglycan isoforms that could affect sarcoglycan complex function.

Mutations in *SGCE* were first identified as a cause of MD in 2001, but little is known about the function of the encoded ε-SG protein even now. Most pathogenic *SGCE* mutations result in cell surface ε-SG deficiency, suggesting cell surface ε-SG is essential for normal brain function (Carecchio et al., 2013, Esapa et al., 2007, Grunewald et al., 2008, Waite et al., 2011). In addition, the contrast between the neurological phenotype of ε-SG deficiency and the muscle-only LGMD phenotype induced by deficiency of the related α-, β-, δ- and γ-SG proteins remains unexplained. A unique function for the brain-specific ε-SG isoform 2, generated through alternative splicing, has been proposed to underlie the specifically neurological phenotype of *SGCE* mutations (Nishiyama et al., 2004, Ritz et al., 2011, Yokoi et al., 2005). In smooth muscle and peripheral nerve, ε-SG interacts with β-SG, δ-SG and γ-SG or ζ-SG to form heterotetrameric sarcoglycan complexes that are part of the DGC or DGC-like complexes (Cai et al., 2007, Durbeej and Campbell, 1999, Imamura et al., 2000,

Straub et al., 1999). Immunoaffinity purification of ε-SG, including of isoform 2 specifically, from mouse brain tissue resulted in co-purification of β-SG, δ-SG and ζ-SG (Chapter 3). This co-purification indicated that in the brain both ε-SG isoforms 1 and 2 are part of prototypical sarcoglycan complexes comprising εβδζ (Chapter 3). Brain ε-SG also interacts with dystroglycan and Dp71, suggesting that it contributes to DGC-like complexes. Finally, the extracellular matrix proteoglycan tenascin-R co-purified specifically with ε-SG isoform 2 from mouse brain.

Alternatively spliced *SGCA, SGCB, SGCD, SGCG* and *SGCZ* transcripts encoding novel sarcoglycan protein isoforms may also contribute to differences in sarcoglycan complexes between tissues. A systematic screen for alternatively spliced transcripts from these genes in human cerebellum and skeletal muscle was performed using RT-PCR (Chapter 4). All five genes produced alternatively spliced transcripts, some of which encoded novel sarcoglycan isoforms which differentially affected assembly and/or trafficking of sarcoglycan complexes in heterologous cells (Chapter 5).

6.2.ε-SG protein interactions in the brain

Chapter 3 describes experiments designed to capture protein complexes containing ε-SG from mouse brain. IAPs using antibodies that bound to either all ε-SG isoforms or only ε-SG isoform 2 were performed and then analysed by Western blotting and mass spectrometry. ε-SG co-purified with β-SG, δ-SG and ζ-SG using both antibodies. Based on previous studies of sarcoglycan complexes, this pattern of co-purification would support the presence of εβδζ complexes including either ε-SG isoform 1 or isoform 2 in the brain (Chan et al., 1998, Holt and Campbell, 1998, Jung et al., 1996, Noguchi et al., 2000, Shiga et al., 2006). However, ε-SG isoforms 1 and 2 would be part of different complexes based on the absence of ε-SG isoform 1 from esg2-1358 IAPs that specifically enriched ε-SG isoform 2. Intriguingly, γ-SG

did not co-purify with ε-SG from mouse brain using either antibody despite its co-purification with an antibody against all ε-SG isoforms from rat brain in a previous study (Waite et al., 2016). This most likely reflects the low abundance of γ-SG in the brain (Chapter 3).

The identification of ε-SG-containing prototypical sarcoglycan complexes in the brain was unexpected given the absence of LGMD-like features in MD and vice versa. Both disorders are caused by loss-of-function mutations resulting in cell surface deficiency of the individual sarcoglycan or entire sarcoglycan complex, yet β-SG or δ-SG deficiency affects only muscle function and not the CNS (Esapa et al., 2007, Sandona and Betto, 2009, Waite et al., 2011). Modelling of sarcoglycan complex assembly in heterologous cells demonstrated that residual εζ-SG complexes could traffic to the cell surface despite the presence of LGMD-associated mutant β-SG that impeded trafficking of the βδ-SG core; residual cell surface ε-SG and ζ-SG have also been identified in the brain of δ-SG-deficient BIO14.6 hamsters (Waite et al., 2016). Residual sarcoglycans are functional in muscle, so residual cell surface ε-SG and ζ-SG in the brain could be sufficient for normal brain function in sarcoglycan-deficient LGMD patients (Bianchini et al., 2014, Esapa et al., 2007, Gastaldello et al., 2008, Li et al., 2009, Soheili et al., 2012, Waite et al., 2011, Waite et al., 2016). Differential trafficking of partial sarcoglycan complexes in different tissues may contribute to the distinct phenotypes of sarcoglycan-associated LGMD and MD.

Under mild tissue lysis conditions, Dp71 and β-dystroglycan were also found to co-purify with ε-SG from brain (Chapter 3 section 3.2.3). If tissue was lysed using harsher RIPA buffer, this co-purification was lost; a similar result was observed using lung tissue in which ε-SG is known to be a component of DGCs (Chapter 3). Therefore, ε-SG was part of DGC-like complexes in the brain but the protein interactions required for these complexes were disrupted by RIPA buffer. By extension, RIPA buffer may also disrupt other protein

interactions of ε-SG and other sarcoglycan proteins, which could have contributed to previous failed attempts at identifying ε-SG interactors.

Incorporation of ε-SG into DGC-like complexes in the brain suggests that perturbation of DGC-like complexes may contribute to the *SGCE* mutation-positive MD phenotype. DGC-like complexes in the brain are thought to have a number of functions, including organisation of neurotransmitter receptors at the synapse and calcium homeostasis (Brünig et al., 2002, Knuesel et al., 1999, Krasowska et al., 2014, Waite et al., 2012). In particular, DGC-like complexes help organise GABA receptors on the postsynaptic membrane; as previously discussed, GABAergic signalling may contribute to surround inhibition which is impaired in dystonia (Chapter 1 section 1.2.3.1) (Garibotto et al., 2011, Herath et al., 2010, Levy and Hallett, 2002, Waite et al., 2012). As part of DGC-like complexes, ε-SG might contribute to neurotransmitter receptor clustering at the synapse. Deficiency of ε-SG might perturb DGC-like complexes, resulting in aberrant synapse organisation. Specific effects of disrupting DGC-like complexes on GABAergic signalling could contribute to the molecular pathogenesis of dystonia.

As part of DGC-like complexes, ε-SG may also contribute to calcium homeostasis. Numerous lines of evidence support a role for the DGC generally and sarcoglycans specifically in maintaining normal calcium levels in muscle, and abnormal brain calcium levels have also been described in patients with *DMD* mutations (Chapter 1 section 1.4.3) (Allen et al., 2016, Anderson et al., 2002, Andersson et al., 2012, Waite et al., 2012). Therefore, as part of DGC-like complexes ε-SG could contribute to maintenance of normal intracellular calcium levels. Three other monogenic dystonia genes (*ANO3, PRRT2* and *HPCA*) are thought to directly perturb calcium signalling/homeostasis: the ANO3 protein is calcium activated, PRRT2 protein is regulated through binding to calcium sensor proteins, and hippocalcin is itself a calcium sensor (Charlesworth et al., 2015, Huang et al., 2013,

Valente et al., 2016). More recently, mutations in *CACNA1B* which encodes a calcium channel pore-forming subunit have been linked to dystonia as well (Ament et al., 2015, Groen et al., 2015a). Mutations in these genes are thought to contribute to dystonia by altering the cellular response to calcium fluctuations or by directly changing intracellular calcium concentration. If ε-SG deficiency led to disruption of DGC-like complexes, *SGCE* mutations could also affect calcium homeostasis.

In addition, the extracellular matrix protein tenascin-R co-purified with ε-SG from brain (Chapter 3). This was not a consequence of non-specific interactions or direct binding of tenascin-R to an anti-ε-SG antibody, as a reciprocal IAP using an anti-tenascin-R antibody resulted in co-purification of ε-SG. Surprisingly, only ε-SG isoform 2 and not isoform 1 was detected in the tenascin-R IAP. A functional difference between the widespread (isoform 1) and brain-specific (isoform 2) ε-SG isoforms has previously been proposed to contribute to the brain-specific effects of ε-SG deficiency, but to date no such differences between isoforms have been identified (Ritz et al., 2011). As described above, both isoforms are incorporated into prototypical sarcoglycan heterotetramers and interact with other DGC proteins. Furthermore, when examined in heterologous cells all ε-SG isoforms exhibited identical sarcoglycan complex assembly and trafficking (Chapter 3 section 3.2.2.2). Co-purification of only ε-SG isoform 2 with tenascin-R is the first sign there might be a genuine functional difference between ε-SG isoforms, and provides addition evidence in favour of independent localisation of the two major ε-SG isoforms in the brain (Chan et al., 2005, Nishiyama et al., 2004).

Although the co-purification of ε-SG isoform 2 and tenascin-R does not provide conclusive evidence of an interaction between these proteins, the specificity of the co-purification is suggestive. In addition, there are similarities in the phenotypes of tenascin-R deficient mice and ε-SG-deficient mice. Mice deficient in tenascin-R have no gross abnormalities, but

display several neurological phenotypes including increased anxiety, impaired motor learning and impaired coordination (Freitag et al., 2003, Montag-Sallaz and Montag, 2003). ε-SG-deficient mice also have an anxiety-like behavioural phenotype, and impaired motor coordination (Chapter 1 section 1.3.2.4) (Yokoi et al., 2006, Zhang et al., 2012). Recently, an exome sequencing study identified several rare *TNR* variants in familial Parkinson disease; this provides additional support for dysfunction of tenascin-R in movement disorders (Farlow et al., 2016). As a central organiser of perineuronal nets (PNNs), tenascin-R contributes to normal clustering of neurotransmitter receptors and ion channels at the synapse (Anlar and Gunel-Ozcan, 2012, Geissler et al., 2013, Oohashi et al., 2015). Therefore, an interaction between ε-SG and tenascin-R could further implicate the former in synapse organisation.

6.3. Alternative splicing of *SGCA, SGCB, SGCD, SGCG* and *SGCZ* pre-mRNA

The experiments described in chapter 4 identified alternatively spliced transcripts from human *SGCA, SGCB, SGCD, SGCG* and *SGCZ* using PCR-based screening of cDNA mini-libraries derived from adult human skeletal muscle and cerebellum. Chapter 5 described initial characterisation of novel sarcoglycan isoforms encoded by some of these alternatively spliced transcripts. While *SGCE* alternative splicing has already been described in depth, this was the first systematic study of alternative splicing in transcripts from the other five sarcoglycans (Ritz et al., 2011). All five genes were found to produce alternatively spliced transcripts with varying effects on protein synthesis, but there were substantial differences in alternative splicing between paralogous sarcoglycans. No alternative splicing events were shared between *SGCE* and *SGCA* at all, despite highly similar genes and encoded proteins (Chapter 4 section 4.2.4) (Dickens et al., 2002, Ettinger et al., 1997, McNally et al., 1998, Ritz et al., 2011). By contrast, paralogues *SGCG* and *SGCZ* shared one alternative splicing event, omission of exon 5; this was substantially more common in *SGCZ* cDNA than *SGCG*, however (Chapter 4 sections 4.2.3 and 4.2.5). The γ-SG and ζ-SG isoforms 2 and 3 encoded

by Δ exon 5 transcripts from *SGCG* and *SGCZ* respectively both reduced incorporation of α-SG/ε-SG into the sarcoglycan complex in heterologous cells, but showed a strong difference in cell surface trafficking (Chapters 4 and 5). While γ-SG isoform 3 was detectable at the cell surface, ζ-SG isoform 2 was not (Chapter 5 sections 5.2.3.2 and 5.2.4.2). Therefore, these similar alternatively spliced isoforms of paralogous sarcoglycans had very different trafficking despite similar effects on sarcoglycan complex assembly.

The distinct alternative splicing of paralogous sarcoglycans and the different effects of identical alternative splicing events on different sarcoglycans suggest hitherto unappreciated divergence between sarcoglycans. In particular, a *SGCA* transcript encoding a potentially secreted α-SG isoform was abundant in the skeletal muscle *SGCA* cDNA mini-library at 27% of clones, but no equivalent for *SGCE* has been described (Chapter 4 Figure 4.12). This difference could indicate a unique function for α-SG in the extracellular matrix not shared by its paralogue ε-SG. Additionally, the different effects on sarcoglycan complex trafficking of highly similar alternatively spliced γ-SG and ζ-SG isoforms hint that composition could strongly affect sarcoglycan complex functionality. Since *SGCZ* expression is enriched in the brain whereas *SGCG* is enriched in muscle, this might contribute to the different sarcoglycan requirements in brain versus muscle (Chapter 1 section 1.4.1) (Barresi et al., 2000b, Shiga et al., 2006, Yamamoto et al., 1994). Moreover, transcripts encoding ζ-SG isoforms that impair sarcoglycan complex assembly accounted for a quarter of *SGCZ* cerebellum cDNA mini-library clones (Chapter 4 section 4.2.3). If accurate, these data suggest a substantial proportion of ζ-SG in the brain may be assembled into ε-SG-deficient partial complexes that do not traffic to the cell surface. This would leave a pool of cell surface ε-SG not associated with other sarcoglycans, which might have a distinct function(s) important for MD pathogenesis that is not affected in sarcoglycan-deficient LGMD.

6.4. Future directions

This thesis provided evidence for incorporation of brain ε-SG into DGC-like complexes. However, these experiments did not reveal whether ε-SG is a component of all brain DGC-like complexes or only certain subtypes, and what effect ε-SG deficiency might have on these complexes. Of particular interest might be whether ε-SG alone is part of brain DGC-like complexes, or whether the entire prototypical sarcoglycan complex participates in these. This would in turn provide additional insight into the similarities and differences between brain and muscle prototypical sarcoglycan complexes. Clarification of ε-SG deficiency's effects on DGC-like complex function would further extend our understanding of ε-SG's function in the brain. IAPs performed as part of this study also identified co-purification between ε-SG and the extracellular matrix proteoglycan tenascin-R; as described in Chapter 3 (section 3.3) further experiments are required to determine whether ε-SG isoform 2 genuinely interacts with tenascin-R. There are also likely to be additional proteins that interact with ε-SG in the brain. Finally, the novel alternatively spliced sarcoglycan isoforms described in Chapters 4 and 5 were only studied in heterologous cells and total RNA from human tissue. They will need to be identified *in vivo* before their effect on sarcoglycan complex function can be understood. Of particular interest for MD will be the ζ-SG alternatively spliced isoforms, which based on the experiments performed in heterologous cells were unlikely to co-purify with ε-SG. Verification of these isoforms and confirmation that they do not interact with cell surface ε-SG as observed in heterologous cells could reveal a sarcoglycan complex-independent role for ε-SG in the brain. Overall, these data could further contribute to our understanding of the molecular pathogenesis of MD, and the different phenotypes of sarcoglycan-associated LGMD and MD.

Bibliography

Adams, M. E., Mueller, H. A. & Froehner, S. C. 2001. In vivo requirement of the α-syntrophin PDZ domain for the sarcolemmal localization of nNOS and aquaporin-4. *The Journal of Cell Biology,* 155**,** 113-122.

Ainscough, J. F., Rahman, F. A., Sercombe, H., Sedo, A., Gerlach, B. & Coverley, D. 2007. C-terminal domains deliver the DNA replication factor Ciz1 to the nuclear matrix. *Journal of Cell Science,* 120**,** 115-24.

Albanese, A., Bhatia, K., Bressman, S. B., Delong, M. R., Fahn, S., Fung, V. S., Hallett, M., Jankovic, J., Jinnah, H. A., Klein, C., Lang, A. E., Mink, J. W. & Teller, J. K. 2013. Phenomenology and classification of dystonia: a consensus update. *Movement Disorders,* 28**,** 863-73.

Alexander, D. C. 1987. An efficient vector-primer cDNA cloning system. *Methods in Enzymology,* 154**,** 41-64.

Allen, D. G., Whitehead, N. P. & Froehner, S. C. 2016. Absence of Dystrophin Disrupts Skeletal Muscle Signaling: Roles of Ca2+, Reactive Oxygen Species, and Nitric Oxide in the Development of Muscular Dystrophy. *Physiological Reviews,* 96**,** 253-305.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. 1990. Basic local alignment search tool. *Journal of molecular biology,* 215**,** 403-410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research,* 25**,** 3389-3402.

Ament, S. A., Szelinger, S., Glusman, G., Ashworth, J., Hou, L., Akula, N., Shekhtman, T., Badner, J. A., Brunkow, M. E., Mauldin, D. E., Stittrich, A.-B., Rouleau, K., Detera-Wadleigh, S. D., Nurnberger, J. I., Edenberg, H. J., Gershon, E. S., Schork, N., The Bipolar Genome, S., Price, N. D., Gelinas, R., Hood, L., Craig, D., Mcmahon, F. J., Kelsoe, J. R. & Roach, J. C. 2015. Rare variants in neuronal excitability genes influence risk for bipolar disorder. *Proceedings of the National Academy of Sciences,* 112**,** 3576-3581.

Amenta, A. R., Creely, H. E., Mercado, M. L., Hagiwara, H., Mckechnie, B. A., Lechner, B. E., Rossi, S. G., Wang, Q., Owens, R. T., Marrero, E., Mei, L., Hoch, W., Young, M. F., Mcquillan, D. J., Rotundo, R. L. & Fallon, J. R. 2012. Biglycan is an extracellular MuSK binding protein important for synapse stability. *Journal of Neuroscience,* 32**,** 2324-34.

Amici, M., Doherty, A., Jo, J., Jane, D., Cho, K., Collingridge, G. & Dargan, S. 2009. Neuronal calcium sensors and synaptic plasticity. *Biochemical Society Transactions,* 37**,** 1359-1363.

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D., Ghosh, S. & Jacobson, A. 2004. A faux 3[prime]-UTR promotes aberrant termination and triggers nonsense- mediated mRNA decay. *Nature,* 432**,** 112-118.

Anderson, J. L., Head, S. I., Rae, C. & Morley, J. W. 2002. Brain function in Duchenne muscular dystrophy. *Brain,* 125**,** 4-13.

Andersson, D. C., Meli, A. C., Reiken, S., Betzenhauser, M. J., Umanskaya, A., Shiomi, T., D'armiento, J. & Marks, A. R. 2012. Leaky ryanodine receptors in beta-sarcoglycan deficient mice: a potential common defect in muscular dystrophy. *Skeletal Muscle,* 2**,** 9.

Änkö, M.-L., Müller-Mcnicoll, M., Brandl, H., Curk, T., Gorup, C., Henry, I., Ule, J. & Neugebauer, K. M. 2012. The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. *Genome Biology,* 13**,** 1-17.

Anlar, B. & Gunel-Ozcan, A. 2012. Tenascin-R: Role in the central nervous system. *The International Journal of Biochemistry & Cell Biology,* 44**,** 1385-1389.

Araishi, K., Sasaoka, T., Imamura, M., Noguchi, S., Hama, H., Wakabayashi, E., Yoshida, M., Hori, T. & Ozawa, E. 1999. Loss of the sarcoglycan complex and sarcospan leads to muscular dystrophy in beta-sarcoglycan-deficient mice. *Human molecular genetics,* 8**,** 1589-98.

Argyelan, M., Carbon, M., Niethammer, M., Ulug, A. M., Voss, H. U., Bressman, S. B., Dhawan, V. & Eidelberg, D. 2009. Cerebellothalamocortical connectivity regulates penetrance in dystonia. *Journal of Neuroscience,* 29**,** 9740-7.

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., De Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G., Xenarios, I. & Stockinger, H. 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic acids research,* 40**,** W597-W603.

Asmus, F., Hjermind, L. E., Dupont, E., Wagenstaller, J., Haberlandt, E., Munz, M., Strom, T. M. & Gasser, T. 2007. Genomic deletion size at the epsilon-sarcoglycan locus determines the clinical phenotype. *Brain,* 130**,** 2736-45.

Asmus, F., Langseth, A., Doherty, E., Nestor, T., Munz, M., Gasser, T., Lynch, T. & King, M. D. 2009. "Jerky" dystonia in children: spectrum of phenotypes and genetic testing. *Movement Disorders,* 24**,** 702-9.

Asmus, F., Zimprich, A., Naumann, M., Berg, D., Bertram, M., Ceballos-Baumann, A., Pruszak-Seel, R., Kabus, C., Dichgans, M., Fuchs, S., Muller-Myhsok, B. & Gasser, T. 2001. Inherited Myoclonus-dystonia syndrome: narrowing the 7q21-q31 locus in German families. *Annals of Neurology,* 49**,** 121-4.

Asmus, F., Zimprich, A., Tezenas Du Montcel, S., Kabus, C., Deuschl, G., Kupsch, A., Ziemann, U., Castro, M., Kuhn, A. A., Strom, T. M., Vidailhet, M., Bhatia, K. P., Durr, A., Wood, N. W., Brice, A. & Gasser, T. 2002. Myoclonus-dystonia syndrome: epsilon-sarcoglycan mutations and phenotype. *Annals of Neurology,* 52**,** 489-92.

Aspberg, A., Miura, R., Bourdoulous, S., Shimonaka, M., Heinegard, D., Schachner, M., Ruoslahti, E. & Yamaguchi, Y. 1997. The C-type lectin domains of lecticans, a family of aggregating chondroitin sulfate proteoglycans, bind tenascin-R by protein-protein interactions independent of carbohydrate moiety. *Proceedings of the National Academy of Sciences,* 94**,** 10116-21.

Augood, S. J., Martin, D. M., Ozelius, L. J., Breakefield, X. O., Penney, J. B., Jr. & Standaert, D. G. 1999. Distribution of the mRNAs encoding torsinA and torsinB in the normal adult human brain. *Annals of Neurology,* 46**,** 761-9.

Augood, S. J., Penney, J. B., Jr., Friberg, I. K., Breakefield, X. O., Young, A. B., Ozelius, L. J. & Standaert, D. G. 1998. Expression of the early-onset torsion dystonia gene (DYT1) in human brain. *Annals of Neurology,* 43**,** 669-73.

Balint, B. & Bhatia, K. P. 2014. Dystonia: an update on phenomenology, classification, pathogenesis and treatment. *Current Opinion in Neurology,* 27**,** 468-76.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. 2007. Quantitative mass spectrometry in proteomics: a critical review. *Analytical & Bioanalytical Chemistry,* 389**,** 1017-31.

Bara-Jimenez, W., Shelton, P. & Hallett, M. 2000a. Spatial discrimination is abnormal in focal hand dystonia. *Neurology,* 55**,** 1869-73.

Bara-Jimenez, W., Shelton, P., Sanger, T. D. & Hallett, M. 2000b. Sensory discrimination capabilities in patients with focal hand dystonia. *Annals of Neurology,* 47**,** 377-380.

Barbosa-Morais, N. L., Irimia, M., Pan, Q., Xiong, H. Y., Gueroussov, S., Lee, L. J., Slobodeniuc, V., Kutter, C., Watt, S., Çolak, R., Kim, T., Misquitta-Ali, C. M., Wilson, M. D., Kim, P. M., Odom, D. T., Frey, B. J. & Blencowe, B. J. 2012. The Evolutionary Landscape of Alternative Splicing in Vertebrate Species. *Science,* 338**,** 1587-1593.

Barresi, R., Di Blasi, C., Negri, T., Brugnoni, R., Vitali, A., Felisari, G., Salandi, A., Daniel, S., Cornelio, F., Morandi, L. & Mora, M. 2000a. Disruption of heart sarcoglycan complex and severe cardiomyopathy caused by beta sarcoglycan mutations. *Journal of medical genetics,* 37**,** 102-7.

Barresi, R., Moore, S. A., Stolle, C. A., Mendell, J. R. & Campbell, K. P. 2000b. Expression of gamma -sarcoglycan in smooth muscle and its interaction with the smooth muscle sarcoglycan-sarcospan complex. *Journal of Biological Chemistry,* 275**,** 38554-60.

Bartoli, M., Gicquel, E., Barrault, L., Soheili, T., Malissen, M., Malissen, B., Vincent-Lacaze, N., Perez, N., Udd, B., Danos, O. & Richard, I. 2008. Mannosidase I inhibition rescues the human alpha-sarcoglycan R77C recurrent mutation. *Human molecular genetics,* 17**,** 1214-21.

Barton, E. R. 2006. Impact of sarcoglycan complex on mechanical signal transduction in murine skeletal muscle. *American journal of physiology. Cell physiology,* 290**,** C411-9.

Barton, E. R. 2010. Restoration of gamma-sarcoglycan localization and mechanical signal transduction are independent in murine skeletal muscle. *The Journal of biological chemistry,* 285**,** 17263-70.

Berardelli, A., Rothwell, J. C., Day, B. L. & Marsden, C. D. 1985. Pathophysiology of blepharospasm and oromandibular dystonia. *Brain,* 108 (Pt 3)**,** 593-608.

Berget, S. M., Moore, C. & Sharp, P. A. 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences,* 74**,** 3171-5.

Bessiere, D., Lacroix, C., Campagne, S., Ecochard, V., Guillet, V., Mourey, L., Lopez, F., Czaplicki, J., Demange, P., Milon, A., Girard, J. P. & Gervais, V. 2008. Structure-function analysis of the THAP zinc finger of THAP1, a large C2CH DNA-binding module linked to Rb/E2F pathways. *The Journal of biological chemistry,* 283**,** 4352-63.

Betto, R., Senter, L., Ceoldo, S., Tarricone, E., Biral, D. & Salviati, G. 1999. Ecto-ATPase activity of alpha-sarcoglycan (adhalin). *The Journal of biological chemistry,* 274**,** 7907-12.

Beukers, R. J., Foncke, E. M., Van Der Meer, J. N., Nederveen, A. J., De Ruiter, M. B., Bour, L. J., Veltman, D. J. & Tijssen, M. A. 2010. Disorganized sensorimotor integration in mutation-positive myoclonus-dystonia: a functional magnetic resonance imaging study. *Archives of Neurology,* 67**,** 469-74.

Beukers, R. J., Foncke, E. M., Van Der Meer, J. N., Veltman, D. J. & Tijssen, M. A. 2011. Functional magnetic resonance imaging evidence of incomplete maternal imprinting in myoclonus-dystonia. *Archives of Neurology,* 68**,** 802-5.

Bhatia, K. P. & Marsden, C. D. 1994. The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain,* 117 (Pt 4)**,** 859-76.

Bhatia, K. P. & Schneider, S. A. 2012. Identification of PRRT2 as the causative gene of paroxysmal kinesigenic dyskinesia. *Movement Disorders,* 27**,** 707.

Bianchini, E., Fanin, M., Mamchaoui, K., Betto, R. & Sandona, D. 2014. Unveiling the degradative route of the V247M alpha-sarcoglycan mutant responsible for LGMD-2D. *Human Molecular Genetics,* 23**,** 3746-58.

Blake, D. J., Hawkes, R., Benson, M. A. & Beesley, P. W. 1999. Different Dystrophin-like Complexes Are Expressed in Neurons and Glia. *The Journal of Cell Biology,* 147**,** 645-658.

Blake, D. J., Love, D. R., Tinsley, J., Morris, G. E., Turley, H., Gatter, K., Dickson, G., Edwards, Y. H. & Davies, K. E. 1992. Characterization of a 4.8kb transcript from the Duchenne muscular dystrophy locus expressed in Schwannoma cells. *Human Molecular Genetics,* 1**,** 103-9.

Blake, D. J., Nawrotzki, R., Loh, N. Y., Gorecki, D. C. & Davies, K. E. 1998. beta-dystrobrevin, a member of the dystrophin-related protein family. *Proceedings of the National Academy of Sciences,* 95**,** 241-6.

Blanchard, A., Ea, V., Roubertie, A., Martin, M., Coquart, C., Claustres, M., Beroud, C. & Collod-Beroud, G. 2011. DYT6 dystonia: review of the literature and creation of the UMD Locus-Specific Database (LSDB) for mutations in the THAP1 gene. *Human mutation,* 32**,** 1213-24.

Blanco-Arias, P., Einholm, A. P., Mamsa, H., Concheiro, C., Gutierrez-De-Teran, H., Romero, J., Toustrup-Jensen, M. S., Carracedo, A., Jen, J. C., Vilsen, B. & Sobrido, M. J. 2009. A C-terminal mutation of ATP1A3 underscores the crucial role of sodium affinity in the pathophysiology of rapid-onset dystonia-parkinsonism. *Human Molecular Genetics,* 18**,** 2370-7.

Bolduc, V., Marlow, G., Boycott, K. M., Saleki, K., Inoue, H., Kroon, J., Itakura, M., Robitaille, Y., Parent, L., Baas, F., Mizuta, K., Kamata, N., Richard, I., Linssen, W. H., Mahjneh, I., De Visser, M., Bashir, R. & Brais, B. 2010. Recessive mutations in the putative calcium-activated chloride channel Anoctamin 5 cause proximal LGMD2L and distal MMD3 muscular dystrophies. *American journal of human genetics,* 86**,** 213-21.

Bottger, P., Tracz, Z., Heuck, A., Nissen, P., Romero-Ramos, M. & Lykke-Hartmann, K. 2011. Distribution of Na/K-ATPase alpha 3 isoform, a sodium-potassium P-type pump associated with rapid-onset of dystonia parkinsonism (RDP) in the adult mouse brain. *Journal of Comparative Neurology,* 519**,** 376-404.

Boulay, A. C., Saubamea, B., Cisternino, S., Mignon, V., Mazeraud, A., Jourdren, L., Blugeon, C. & Cohen-Salmon, M. 2015. The Sarcoglycan complex is expressed in the cerebrovascular system and is specifically regulated by astroglial Cx30 channels. *Frontiers in Cellular Neuroscience,* 9**,** 9.

Bowe, M. A., Mendis, D. B. & Fallon, J. R. 2000. The small leucine-rich repeat proteoglycan biglycan binds to alpha-dystroglycan and is upregulated in dystrophic muscle. *Journal of Cell Biology,* 148**,** 801-10.

Bozic, D., Engel, J. & Brancaccio, A. 1998. Sequence analysis suggests the presence of an IG-like domain in the N-terminal region of alpha-dystroglycan which was crystallized after mutation of a protease susceptible site (Arg168-->His). *Matrix Biology,* 17**,** 495-500.

Brashear, A., Dobyns, W. B., De Carvalho Aguiar, P., Borg, M., Frijns, C. J., Gollamudi, S., Green, A., Guimaraes, J., Haake, B. C., Klein, C., Linazasoro, G., Munchau, A., Raymond, D., Riley, D., Saunders-Pullman, R., Tijssen, M. A., Webb, D., Zaremba, J., Bressman, S. B. & Ozelius, L. J. 2007. The phenotypic spectrum of rapid-onset dystonia-parkinsonism (RDP) and mutations in the ATP1A3 gene. *Brain,* 130**,** 828-35.

Breakefield, X. O., Blood, A. J., Li, Y., Hallett, M., Hanson, P. I. & Standaert, D. G. 2008. The pathophysiological basis of dystonias. *Nature Reviews Neuroscience,* 9**,** 222-34.

Brenman, J. E., Chao, D. S., Gee, S. H., Mcgee, A. W., Craven, S. E., Santillano, D. R., Wu, Z., Huang, F., Xia, H., Peters, M. F., Froehner, S. C. & Bredt, D. S. 1996. Interaction of Nitric Oxide Synthase with the Postsynaptic Density Protein PSD-95 and α1-Syntrophin Mediated by PDZ Domains. *Cell,* 84**,** 757-767.

Bressman, S. B., De Leon, D., Brin, M. F., Risch, N., Burke, R. E., Greene, P. E., Shale, H. & Fahn, S. 1989. Idiopathic dystonia among ashkenazi jews: Evidence for autosomal dominant inheritance. *Annals of Neurology,* 26**,** 612-620.

Bressman, S. B., Sabatti, C., Raymond, D., De Leon, D., Klein, C., Kramer, P. L., Brin, M. F., Fahn, S., Breakefield, X., Ozelius, L. J. & Risch, N. J. 2000. The DYT1 phenotype and guidelines for diagnostic testing. *Neurology,* 54**,** 1746-52.

Brown, G. R., Hem, V., Katz, K. S., Ovetsky, M., Wallin, C., Ermolaeva, O., Tolstoy, I., Tatusova, T., Pruitt, K. D., Maglott, D. R. & Murphy, T. D. 2015. Gene: a gene-centered information resource at NCBI. *Nucleic Acids Research,* 43**,** D36-D42.

Brown, R. S., Zhao, C., Chase, A. R., Wang, J. & Schlieker, C. 2014. The mechanism of Torsin ATPase activation. *Proceedings of the National Academy of Sciences,* 111**,** E4822-31.

Bruckner, G., Grosche, J., Schmidt, S., Hartig, W., Margolis, R. U., Delpech, B., Seidenbecher, C. I., Czaniera, R. & Schachner, M. 2000. Postnatal development of perineuronal nets in wild-type mice and in a mutant deficient in tenascin-R. *Journal of Comparative Neurology,* 428**,** 616-29.

Brünig, I., Suter, A., Knuesel, I., Lüscher, B. & Fritschy, J.-M. 2002. GABAergic Terminals Are Required for Postsynaptic Clustering of Dystrophin But Not of GABAA Receptors and Gephyrin. *The Journal of Neuroscience,* 22**,** 4805-4813.

Bruno, M. K., Lee, H. Y., Auburger, G. W., Friedman, A., Nielsen, J. E., Lang, A. E., Bertini, E., Van Bogaert, P., Averyanov, Y., Hallett, M., Gwinn-Hardy, K., Sorenson, B., Pandolfo, M., Kwiecinski, H., Servidei, S., Fu, Y. H. & Ptacek, L. 2007. Genotype-phenotype correlation of paroxysmal nonkinesigenic dyskinesia. *Neurology,* 68**,** 1782-9.

Cai, H., Erdman, R. A., Zweier, L., Chen, J., Shaw, J. H. T., Baylor, K. A., Stecker, M. M., Carey, D. J. & Chan, Y. M. 2007. The sarcoglycan complex in Schwann cells and its role in myelin stability. *Experimental Neurology,* 205**,** 257-69.

Caldwell, G. A., Cao, S., Sexton, E. G., Gelwix, C. C., Bevel, J. P. & Caldwell, K. A. 2003. Suppression of polyglutamine-induced protein aggregation in Caenorhabditis elegans by torsin proteins. *Human Molecular Genetics,* 12**,** 307-19.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics,* 10**,** 421.

Camargos, S., Scholz, S., Simon-Sanchez, J., Paisan-Ruiz, C., Lewis, P., Hernandez, D., Ding, J., Gibbs, J. R., Cookson, M. R., Bras, J., Guerreiro, R., Oliveira, C. R., Lees, A., Hardy, J., Cardoso, F. & Singleton, A. B. 2008. DYT16, a novel young-onset dystonia-parkinsonism disorder: identification of a segregating mutation in the stress-response protein PRKRA. *Lancet Neurol,* 7**,** 207-15.

Campagne, S., Muller, I., Milon, A. & Gervais, V. 2012. Towards the classification of DYT6 dystonia mutants in the DNA-binding domain of THAP1. *Nucleic Acids Res,* 40**,** 9927-40.

Campbell, D. B., North, J. B. & Hess, E. J. 1999. Tottering Mouse Motor Dysfunction Is Abolished on the Purkinje Cell Degeneration (pcd) Mutant Background. *Experimental Neurology,* 160**,** 268-278.

Carbon, M., Kingsley, P. B., Tang, C., Bressman, S. & Eidelberg, D. 2008. Microstructural white matter changes in primary torsion dystonia. *Movement Disorders,* 23**,** 234-9.

Carbon, M., Niethammer, M., Peng, S., Raymond, D., Dhawan, V., Chaly, T., Ma, Y., Bressman, S. & Eidelberg, D. 2009. Abnormal striatal and thalamic dopamine neurotransmission: Genotype-related features of dystonia. *Neurology,* 72**,** 2097-103.

Carbon, M., Raymond, D., Ozelius, L., Saunders-Pullman, R., Frucht, S., Dhawan, V., Bressman, S. & Eidelberg, D. 2013. Metabolic changes in DYT11 myoclonus-dystonia. *Neurology,* 80**,** 385-391.

Carbon, M., Su, S., Dhawan, V., Raymond, D., Bressman, S. & Eidelberg, D. 2004. Regional metabolism in primary torsion dystonia: effects of penetrance and genotype. *Neurology,* 62**,** 1384-90.

Carecchio, M., Magliozzi, M., Copetti, M., Ferraris, A., Bernardini, L., Bonetti, M., Defazio, G., Edwards, M. J., Torrente, I., Pellegrini, F., Comi, C., Bhatia, K. P. & Valente, E. M. 2013. Defining the epsilon-sarcoglycan (SGCE) gene phenotypic signature in myoclonus-dystonia: a reappraisal of genetic testing criteria. *Movement Disorders,* 28**,** 787-94.

Carneiro, D. G., Clarke, T., Davies, C. C. & Bailey, D. 2016. Identifying novel protein interactions: Proteomic methods, optimisation approaches and data analysis pipelines. *Methods,* 95**,** 46-54.

Cayrol, C., Lacroix, C., Mathe, C., Ecochard, V., Ceribelli, M., Loreau, E., Lazar, V., Dessen, P., Mantovani, R., Aguilar, L. & Girard, J. P. 2007. The THAP-zinc finger protein THAP1 regulates endothelial cell proliferation through modulation of pRB/E2F cell-cycle target genes. *Blood,* 109**,** 584-94.

Chan, P., Gonzalez-Maeso, J., Ruf, F., Bishop, D. F., Hof, P. R. & Sealfon, S. C. 2005. Epsilon-sarcoglycan immunoreactivity and mRNA expression in mouse brain. *Journal of Comparative Neurology,* 482**,** 50-73.

Chan, Y. M., Bonnemann, C. G., Lidov, H. G. & Kunkel, L. M. 1998. Molecular organization of sarcoglycan complex in mouse myotubes in culture. *The Journal of cell biology,* 143**,** 2033-44.

Chang, B. S., Duzcan, F., Kim, S., Cinbis, M., Aggarwal, A., Apse, K. A., Ozdel, O., Atmaca, M., Zencir, S., Bagci, H. & Walsh, C. A. 2007a. The role of RELN in lissencephaly and neuropsychiatric disease. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* 144B**,** 58-63.

Chang, Y., Imam, S. & Wilkinson, M. 2007b. The Nonsense-Mediated Decay RNA Surveillance Pathway. *Annual Review of Biochemistry,* 76**,** 51-74.

Charlesworth, G., Angelova, Plamena r., Bartolomé-Robledo, F., Ryten, M., Trabzuni, D., Stamelou, M., Abramov, Andrey y., Bhatia, Kailash p. & Wood, Nicholas w. 2015. Mutations in HPCA Cause Autosomal-Recessive Primary Isolated Dystonia. *The American Journal of Human Genetics,* 96**,** 657-665.

Charlesworth, G., Plagnol, V., Holmstrom, K. M., Bras, J., Sheerin, U. M., Preza, E., Rubio-Agusti, I., Ryten, M., Schneider, S. A., Stamelou, M., Trabzuni, D., Abramov, A. Y., Bhatia, K. P. & Wood, N. W. 2012. Mutations in ANO3 cause dominant craniocervical dystonia: ion channel implicated in pathogenesis. *American journal of human genetics,* 91**,** 1041-50.

Chawla, G., Lin, C. H., Han, A., Shiue, L., Ares, M., Jr. & Black, D. L. 2009. Sam68 regulates a set of alternatively spliced exons during neurogenesis. *Molecular & Cellular Biology,* 29**,** 201-13.

Chen, J., Shi, W., Zhang, Y., Sokol, R., Cai, H., Lun, M., Moore, B. F., Farber, M. J., Stepanchick, J. S., Bonnemann, C. G. & Chan, Y. M. 2006. Identification of

functional domains in sarcoglycans essential for their interaction and plasma membrane targeting. *Experimental Cell Research,* 312**,** 1610-25.

Chen, J., Skinner, M. A., Shi, W., Yu, Q.-C., Wildeman, A. G. & Chan, Y.-M. M. 2007. The 16 kDa subunit of vacuolar H+-ATPase is a novel sarcoglycan-interacting protein. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease,* 1772**,** 570-579.

Chen, P., Burdette, A. J., Porter, J. C., Ricketts, J. C., Fox, S. A., Nery, F. C., Hewett, J. W., Berkowitz, L. A., Breakefield, X. O., Caldwell, K. A. & Caldwell, G. A. 2010. The early-onset torsion dystonia-associated protein, torsinA, is a homeostatic regulator of endoplasmic reticulum stress response. *Human Molecular Genetics,* 19**,** 3502-15.

Cheng, F. B., Wan, X. H., Feng, J. C., Ma, L. Y., Hou, B., Feng, F., Wang, L. & Yang, Y. M. 2012. Subcellular distribution of THAP1 and alterations in the microstructure of brain white matter in DYT6 dystonia. *Parkinsonism and Related Disorders,* 18**,** 978-82.

Cheng, L., Guo, X.-F., Yang, X.-Y., Chong, M., Cheng, J., Li, G., Gui, Y.-H. & Lu, D.-R. 2006. Delta-sarcoglycan is necessary for early heart and muscle development in zebrafish. *Biochemical and biophysical research communications,* 344**,** 1290-9.

Chia, R., Achilli, F., Festing, M. F. W. & Fisher, E. M. C. 2005. The origins and uses of mouse outbred stocks. *Nature Genetics,* 37**,** 1181-1186.

Chiovaro, F., Chiquet-Ehrismann, R. & Chiquet, M. 2015. Transcriptional regulation of tenascin genes. *Cell adhesion & migration,* 9**,** 34-47.

Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. 1977. An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. *Cell,* 12**,** 1-8.

Chung, W. & Campanelli, J. T. 1999. WW and EF hand domains of dystrophin-family proteins mediate dystroglycan binding. *Molecular Cell Biology Research Communications,* 2**,** 162-71.

Clouaire, T., Roussigne, M., Ecochard, V., Mathe, C., Amalric, F. & Girard, J. P. 2005. The THAP domain of THAP1 is a large C2CH module with zinc-dependent sequence-specific DNA-binding activity. *Proceedings of the National Academy of Sciences,* 102**,** 6907-12.

Connors, N. C., Adams, M. E., Froehner, S. C. & Kofuji, P. 2004. The potassium channel Kir4.1 associates with the dystrophin-glycoprotein complex via alpha-syntrophin in glia. *Journal of Biological Chemistry,* 279**,** 28387-92.

Consortium, E. P. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489**,** 57-74.

Copeland, N. A., Sercombe, H. E., Ainscough, J. F. & Coverley, D. 2010. Ciz1 cooperates with cyclin-A-CDK2 to activate mammalian DNA replication in vitro. *Journal of Cell Science,* 123**,** 1108-15.

Copeland, N. A., Sercombe, H. E., Wilson, R. H. & Coverley, D. 2015. Cyclin-A-CDK2-mediated phosphorylation of CIZ1 blocks replisome formation and initiation of mammalian DNA replication. *Journal of Cell Science,* 128**,** 1518-27.

Coral-Vazquez, R., Cohn, R. D., Moore, S. A., Hill, J. A., Weiss, R. M., Davisson, R. L., Straub, V., Barresi, R., Bansal, D., Hrstka, R. F., Williamson, R. & Campbell, K. P. 1999. Disruption of the sarcoglycan-sarcospan complex in vascular smooth muscle: a novel mechanism for cardiomyopathy and muscular dystrophy. *Cell,* 98**,** 465-74.

Corvol, J. C., Studler, J. M., Schonn, J. S., Girault, J. A. & Herve, D. 2001. Galpha(olf) is necessary for coupling D1 and A2a receptors to adenylyl cyclase in the striatum. *Journal of Neurochemistry,* 76**,** 1585-8.

Coverley, D., Marr, J. & Ainscough, J. 2005. Ciz1 promotes mammalian DNA replication. *Journal of Cell Science,* 118**,** 101-12.

Crosbie, R. H., Barresi, R. & Campbell, K. P. 2002. Loss of sarcolemma nNOS in sarcoglycan-deficient muscle. *The FASEB Journal,* 16**,** 1786-1791.

Crosbie, R. H., Heighway, J., Venzke, D. P., Lee, J. C. & Campbell, K. P. 1997. Sarcospan, the 25-kDa transmembrane component of the dystrophin-glycoprotein complex. *Journal of Biological Chemistry,* 272**,** 31221-4.

Crosbie, R. H., Lebakken, C. S., Holt, K. H., Venzke, D. P., Straub, V., Lee, J. C., Grady, R. M., Chamberlain, J. S., Sanes, J. R. & Campbell, K. P. 1999. Membrane targeting and stabilization of sarcospan is mediated by the sarcoglycan subcomplex. *Journal of Cell Biology,* 145**,** 153-65.

Crosbie, R. H., Lim, L. E., Moore, S. A., Hirano, M., Hays, A. P., Maybaum, S. W., Collin, H., Dovico, S. A., Stolle, C. A., Fardeau, M., Tome, F. M. & Campbell, K. P. 2000. Molecular and genetic characterization of sarcospan: insights into sarcoglycan-sarcospan interactions. *Human molecular genetics,* 9**,** 2019-27.

Dale, R. C., Nasti, J. J. & Peters, G. B. 2011. Familial 7q21.3 microdeletion involving epsilon-sarcoglycan causing myoclonus dystonia, cognitive impairment, and psychosis. *Movement Disorders,* 26**,** 1774-5.

Dauer, W. 2014. Inherited isolated dystonia: clinical genetics and gene function. *Neurotherapeutics,* 11**,** 807-16.

De Carvalho Aguiar, P., Sweadner, K. J., Penniston, J. T., Zaremba, J., Liu, L., Caton, M., Linazasoro, G., Borg, M., Tijssen, M. A., Bressman, S. B., Dobyns, W. B., Brashear, A. & Ozelius, L. J. 2004. Mutations in the Na+/K+ -ATPase alpha3 gene ATP1A3 are associated with rapid-onset dystonia parkinsonism. *Neuron,* 43**,** 169-75.

De Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A. & Pollock, D. D. 2011. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLoS Genetics,* 7**,** e1002384.

De Rosa, M. C., Pirolli, D., Bozzi, M., Sciandra, F., Giardina, B. & Brancaccio, A. 2011. A second Ig-like domain identified in dystroglycan by molecular modelling and dynamics. *Journal of Molecular Graphics and Modelling,* 29**,** 1015-1024.

Deandrade, M. P., Yokoi, F., Van Groen, T., Lingrel, J. B. & Li, Y. 2011. Characterization of Atp1a3 mutant mice as a model of rapid-onset dystonia with parkinsonism. *Behavioural Brain Research,* 216**,** 659-65.

Deberardinis, R. J., Conforto, D., Russell, K., Kaplan, J., Kollros, P. R., Zackai, E. H. & Emanuel, B. S. 2003. Myoclonus in a patient with a deletion of the epsilon-sarcoglycan locus on chromosome 7q21. *American Journal of Medical Genetics. Part A. 121A,* 1**,** 31-6.

Deng, H., Le, W.-D. & Jankovic, J. 2008. Genetic study of an American family with DYT3 dystonia (lubag). *Neuroscience letters,* 448**,** 180-3.

Dickens, N. J., Beatson, S. & Ponting, C. P. 2002. Cadherin-like domains in alpha-dystroglycan, alpha/epsilon-sarcoglycan and yeast and bacterial proteins. *Current Biology,* 12**,** 19.

Dobricic, V., Kresojevic, N., Marjanovic, A., Tomic, A., Svetel, M., Novakovic, I. & Kostic, V. S. 2016. HPCA-related dystonia: Too rare to be found? *Movement Disorders*.

Doheny, D., Danisi, F., Smith, C., Morrison, C., Velickovic, M., De Leon, D., Bressman, S. B., Leung, J., Ozelius, L., Klein, C., Breakefield, X. O., Brin, M. F. & Silverman, J. M. 2002. Clinical findings of a myoclonus-dystonia family with two distinct mutations. *Neurology,* 59**,** 1244-6.

Domingo, A., Westenberger, A., Lee, L. V., Braenne, I., Liu, T., Vater, I., Rosales, R., Jamora, R. D., Pasco, P. M., Cutiongco-Dela Paz, E. M., Freimann, K., Schmidt, T. G., Dressler, D., Kaiser, F. J., Bertram, L., Erdmann, J., Lohmann, K. & Klein, C. 2015. New insights into the genetics of X-linked dystonia-parkinsonism (XDP, DYT3). *European Journal of Human Genetics,* 23**,** 1334-40.

Draganski, B., Schneider, S. A., Fiorio, M., Klöppel, S., Gambarin, M., Tinazzi, M., Ashburner, J., Bhatia, K. P. & Frackowiak, R. S. J. 2009. Genotype–phenotype interactions in primary dystonias revealed by differential changes in brain structure. *NeuroImage,* 47**,** 1141-1147.

Draviam, R., Billington, L., Senchak, A., Hoffman, E. P. & Watkins, S. C. 2001. Confocal analysis of the dystrophin protein complex in muscular dystrophy. *Muscle & nerve,* 24**,** 262-72.

Draviam, R. A., Shand, S. H. & Watkins, S. C. 2006a. The β-δ-core of sarcoglycan is essential for deposition at the plasma membrane. *Muscle & Nerve,* 34**,** 691-701.

Draviam, R. A., Wang, B., Shand, S. H., Xiao, X. & Watkins, S. C. 2006b. α-Sarcoglycan is Recycled from the Plasma Membrane in the Absence of Sarcoglycan Complex Assembly. *Traffic,* 7**,** 793-810.

Drinnan, S. L., Hope, B. T., Snutch, T. P. & Vincent, S. R. 1991. Golf in the basal ganglia. *Molecular and Cellular Neuroscience,* 2**,** 66-70.

Dubridge, R. B., Tang, P., Hsia, H. C., Leong, P. M., Miller, J. H. & Calos, M. P. 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Molecular & Cellular Biology,* 7**,** 379-87.

Duclos, F., Straub, V., Moore, S. A., Venzke, D. P., Hrstka, R. F., Crosbie, R. H., Durbeej, M., Lebakken, C. S., Ettinger, A. J., Van Der Meulen, J., Holt, K. H., Lim, L. E., Sanes, J. R., Davidson, B. L., Faulkner, J. A., Williamson, R. & Campbell, K. P. 1998. Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. *Journal of Cell Biology,* 142**,** 1461-71.

Dufke, C., Hauser, A. K., Sturm, M., Fluhr, S., Wachter, T., Leube, B., Auburger, G., Ott, T., Bauer, P., Gasser, T. & Grundmann, K. 2015. Mutations in CIZ1 are not a major cause for dystonia in Germany. *Movement Disorders,* 30**,** 740-3.

Duggan, D. J., Gorospe, J. R., Fanin, M., Hoffman, E. P. & Angelini, C. 1997a. Mutations in the sarcoglycan genes in patients with myopathy. *The New England journal of medicine,* 336**,** 618-24.

Duggan, D. J. & Hoffman, E. P. 1996. Autosomal recessive muscular dystrophy and mutations of the sarcoglycan complex. *Neuromuscular disorders : NMD,* 6**,** 475-82.

Duggan, D. J., Manchester, D., Stears, K. P., Mathews, D. J., Hart, C. & Hoffman, E. P. 1997b. Mutations in the delta-sarcoglycan gene are a rare cause of autosomal recessive limb-girdle muscular dystrophy (LGMD2). *Neurogenetics,* 1**,** 49-58.

Durbeej, M. & Campbell, K. 1999. Biochemical Characterization of the Epithelial Dystroglycan Complex. *Journal of Biological Chemistry,* 274**,** 26609-26616.

Durbeej, M., Cohn, R. D., Hrstka, R. F., Moore, S. A., Allamand, V., Davidson, B. L., Williamson, R. A. & Campbell, K. P. 2000. Disruption of the beta-sarcoglycan gene reveals pathogenetic complexity of limb-girdle muscular dystrophy type 2E. *Molecular Cell,* 5**,** 141-51.

Durr, A., Tassin, J., Vidailhet, M., Durif, F., Jedynak, P., Agid, Y. & Brice, A. 2000. D2 dopamine receptor gene in myoclonic dystonia and essential myoclonus. *Annals of Neurology,* 48**,** 127-8.

Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. 1980. Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. *Cell,* 20**,** 313-9.

Edwards, M. J., Huang, Y.-Z., Mir, P., Rothwell, J. C. & Bhatia, K. P. 2006. Abnormalities in motor cortical plasticity differentiate manifesting and nonmanifesting DYT1 carriers. *Movement Disorders,* 21**,** 2181-2186.

Egger, K., Mueller, J., Schocke, M., Brenneis, C., Rinnerthaler, M., Seppi, K., Trieb, T., Wenning, G. K., Hallett, M. & Poewe, W. 2007. Voxel based morphometry reveals

specific gray matter changes in primary dystonia. *Movement Disorders,* 22**,** 1538-1542.

Einholm, A. P., Toustrup-Jensen, M. S., Holm, R., Andersen, J. P. & Vilsen, B. 2010. The Rapid-onset Dystonia Parkinsonism Mutation D923N of the Na+,K+-ATPase α3 Isoform Disrupts Na+ Interaction at the Third Na+ Site. *Journal of Biological Chemistry,* 285**,** 26245-26254.

Eliuk, S. & Makarov, A. 2015. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annual Review of Analytical Chemistry,* 8**,** 61-80.

Eng, J. K., Mccormack, A. L. & Yates, J. R. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry,* 5**,** 976-89.

Esapa, C. T., Bentham, G. R., Schroder, J. E., Kroger, S. & Blake, D. J. 2003. The effects of post-translational processing on dystroglycan synthesis and trafficking. *FEBS Letters,* 555**,** 209-16.

Esapa, C. T., Waite, A., Locke, M., Benson, M. A., Kraus, M., Mcilhinney, R. A., Sillitoe, R. V., Beesley, P. W. & Blake, D. J. 2007. SGCE missense mutations that cause myoclonus-dystonia syndrome impair epsilon-sarcoglycan trafficking to the plasma membrane: modulation by ubiquitination and torsinA. *Human Molecular Genetics,* 16**,** 327-42.

Ettinger, A. J., Feng, G. & Sanes, J. R. 1997. epsilon-Sarcoglycan, a broadly expressed homologue of the gene mutated in limb-girdle muscular dystrophy 2D.[Erratum appears in J Biol Chem 1998 Jul 31;273(31):19922]. *Journal of Biological Chemistry,* 272**,** 32534-8.

Fanin, M., Duggan, D. J., Mostacciuolo, M. L., Martinello, F., Freda, M. P., Soraru, G., Trevisan, C. P., Hoffman, E. P. & Angelini, C. 1997. Genetic epidemiology of muscular dystrophies resulting from sarcoglycan gene mutations. *Journal of medical genetics,* 34**,** 973-7.

Farlow, J. L., Robak, L. A., Hetrick, K., Bowling, K., Boerwinkle, E., Coban-Akdemir, Z. H., Gambin, T., Gibbs, R. A., Gu, S., Jain, P., Jankovic, J., Jhangiani, S., Kaw, K., Lai, D., Lin, H., Ling, H., Liu, Y., Lupski, J. R., Muzny, D., Porter, P., Pugh, E., White, J., Doheny, K., Myers, R. M., Shulman, J. M. & Foroud, T. 2016. Whole-Exome Sequencing in Familial Parkinson Disease. *JAMA Neurology,* 73**,** 68-75.

Fayssoil, A. 2010. Cardiac diseases in sarcoglycanopathies. *International Journal of Cardiology,* 144**,** 67-68.

Ferrari-Toninelli, G., Paccioretti, S., Francisconi, S., Uberti, D. & Memo, M. 2004. TorsinA negatively controls neurite outgrowth of SH-SY5Y human neuronal cell line. *Brain Research,* 1012**,** 75-81.

Flicek, P., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Girón, C. G., Gordon, L., Hourlier, T., Hunt, S., Johnson, N., Juettemann, T., Kähäri, A. K., Keenan, S., Kulesha, E., Martin, F. J., Maurel, T., Mclaren, W. M., Murphy, D. N., Nag, R., Overduin, B., Pignatelli, M., Pritchard, B., Pritchard, E., Riat, H. S., Ruffier, M., Sheppard, D., Taylor, K., Thormann, A., Trevanion, S. J., Vullo, A., Wilder, S. P., Wilson, M., Zadissa, A., Aken, B. L., Birney, E., Cunningham, F., Harrow, J., Herrero, J., Hubbard, T. J. P., Kinsella, R., Muffato, M., Parker, A., Spudich, G., Yates, A., Zerbino, D. R. & Searle, S. M. J. 2014. Ensembl 2014. *Nucleic Acids Research,* 42**,** D749-D755.

Fokkema, I. F. a. C., Taschner, P. E. M., Schaafsma, G. C. P., Celli, J., Laros, J. F. J. & Den Dunnen, J. T. 2011. LOVD v.2.0: the next generation in gene variant databases. *Human Mutation,* 32**,** 557-563.

Folsom, T. D. & Fatemi, S. H. 2013. The involvement of Reelin in neurodevelopmental disorders. *Neuropharmacology,* 68**,** 122-135.

Foncke, E. M., Gerrits, M. C., Van Ruissen, F., Baas, F., Hedrich, K., Tijssen, C. C., Klein, C. & Tijssen, M. A. 2006. Distal myoclonus and late onset in a large Dutch family with myoclonus-dystonia. *Neurology,* 67**,** 1677-80.

Fraysse, B., Nagi, S. M., Boher, B., Ragot, H., Laine, J., Salmon, A., Fiszman, M. Y., Toussaint, M. & Fromes, Y. 2010. Ca2+ overload and mitochondrial permeability transition pore activation in living delta-sarcoglycan-deficient cardiomyocytes. *American journal of physiology. Cell physiology,* 299**,** C706-13.

Freitag, S., Schachner, M. & Morellini, F. 2003. Behavioral alterations in mice deficient for the extracellular matrix glycoprotein tenascin-R. *Behavioural brain research,* 145**,** 189-207.

Fremont, R., Calderon, D. P., Maleki, S. & Khodakhah, K. 2014. Abnormal high-frequency burst firing of cerebellar neurons in rapid-onset dystonia-parkinsonism. *Journal of Neuroscience,* 34**,** 11723-32.

Fremont, R., Tewari, A. & Khodakhah, K. 2015. Aberrant Purkinje cell activity is the cause of dystonia in a shRNA-based mouse model of Rapid Onset Dystonia-Parkinsonism. *Neurobiology of Disease,* 82**,** 200-12.

Frucht, S. J. 2013. The definition of dystonia: Current concepts and controversies. *Movement Disorders,* 28**,** 884-888.

Fuchs, T., Gavarini, S., Saunders-Pullman, R., Raymond, D., Ehrlich, M. E., Bressman, S. B. & Ozelius, L. J. 2009. Mutations in the THAP1 gene are responsible for DYT6 primary torsion dystonia. *Nature Genetics,* 41**,** 286-8.

Fuchs, T., Saunders-Pullman, R., Masuho, I., Luciano, M. S., Raymond, D., Factor, S., Lang, A. E., Liang, T. W., Trosch, R. M., White, S., Ainehsazan, E., Herve, D., Sharma, N., Ehrlich, M. E., Martemyanov, K. A., Bressman, S. B. & Ozelius, L. J. 2013. Mutations in GNAL cause primary torsion dystonia. *Nature Genetics,* 45**,** 88-92.

Fung, V. S., Jinnah, H. A., Bhatia, K. & Vidailhet, M. 2013. Assessment of patients with isolated or combined dystonia: an update on dystonia syndromes. *Movement Disorders,* 28**,** 889-98.

Furukawa, Y., Kapatos, G., Haycock, J. W., Worsley, J., Wong, H., Kish, S. J. & Nygaard, T. G. 2002. Brain biopterin and tyrosine hydroxylase in asymptomatic dopa-responsive dystonia. *Annals of Neurology,* 51**,** 637-641.

Furukawa, Y., Nygaard, T. G., Gutlich, M., Rajput, A. H., Pifl, C., Distefano, L., Chang, L. J., Price, K., Shimadzu, M., Hornykiewicz, O., Haycock, J. W. & Kish, S. J. 1999. Striatal biopterin and tyrosine hydroxylase protein reduction in dopa-responsive dystonia. *Neurology,* 53**,** 1032-41.

Gao, Q. Q., Wyatt, E., Goldstein, J. A., Lopresti, P., Castillo, L. M., Gazda, A., Petrossian, N., Earley, J. U., Hadhazy, M., Barefield, D. Y., Demonbreun, A. R., Bönnemann, C., Wolf, M. & Mcnally, E. M. 2015. Reengineering a transmembrane protein to treat muscular dystrophy using exon skipping. *The Journal of Clinical Investigation,* 125**,** 4186-4195.

García, M. A., Meurs, E. F. & Esteban, M. 2007. The dsRNA protein kinase PKR: Virus and cell control. *Biochimie,* 89**,** 799-811.

Garibotto, V., Romito, L. M., Elia, A. E., Soliveri, P., Panzacchi, A., Carpinelli, A., Tinazzi, M., Albanese, A. & Perani, D. 2011. In vivo evidence for GABA(A) receptor changes in the sensorimotor system in primary dystonia. *Movement Disorders,* 26**,** 852-7.

Gastaldello, S., D'angelo, S., Franzoso, S., Fanin, M., Angelini, C., Betto, R. & Sandona, D. 2008. Inhibition of proteasome activity promotes the correct localization of disease-

causing alpha-sarcoglycan mutants in HEK-293 cells constitutively expressing beta-, gamma-, and delta-sarcoglycan. *The American journal of pathology,* 173**,** 170-81.

Gavarini, S., Cayrol, C., Fuchs, T., Lyons, N., Ehrlich, M. E., Girard, J. P. & Ozelius, L. J. 2010. Direct interaction between causative genes of DYT1 and DYT6 primary dystonia. *Annals of Neurology,* 68**,** 549-53.

Geissler, M., Gottschling, C., Aguado, A., Rauch, U., Wetzel, C. H., Hatt, H. & Faissner, A. 2013. Primary hippocampal neurons, which lack four crucial extracellular matrix molecules, display abnormalities of synaptic structure and function and severe deficits in perineuronal net formation. *Journal of neuroscience,* 33**,** 7742-55.

Gerrits, M. C., Foncke, E. M., De Haan, R., Hedrich, K., Van De Leemput, Y. L., Baas, F., Ozelius, L. J., Speelman, J. D., Klein, C. & Tijssen, M. A. 2006. Phenotype-genotype correlation in Dutch patients with myoclonus-dystonia. *Neurology,* 66**,** 759-61.

Geyer, H. L. & Bressman, S. B. 2006. The diagnosis of dystonia. *Lancet Neurology,* 5**,** 780-90.

Ghezzi, D., Canavese, C., Kovacevic, G., Zamurovic, D., Barzaghi, C., Giorgi, C., Zorzi, G., Zeviani, M., Pinton, P., Garavaglia, B. & Nardocci, N. 2015. A family with paroxysmal nonkinesigenic dyskinesias (PNKD): evidence of mitochondrial dysfunction. *European Journal of Paediatric Neurology,* 19**,** 64-8.

Ghezzi, D., Viscomi, C., Ferlini, A., Gualandi, F., Mereghetti, P., Degrandis, D. & Zeviani, M. 2009. Paroxysmal non-kinesigenic dyskinesia is caused by mutations of the MR-1 mitochondrial targeting sequence. *Human Molecular Genetics,* 18**,** 1058-64.

Gonzalez-Alegre, P. & Paulson, H. L. 2004. Aberrant cellular behavior of mutant torsinA implicates nuclear envelope dysfunction in DYT1 dystonia. *Journal of neuroscience,* 24**,** 2593-601.

Goodchild, R. E. & Dauer, W. T. 2004. Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation. *Proceedings of the National Academy of Sciences,* 101**,** 847-52.

Goto, S., Lee, L. V., Munoz, E. L., Tooyama, I., Tamiya, G., Makino, S., Ando, S., Dantes, M. B., Yamada, K., Matsumoto, S., Shimazu, H., Kuratsu, J.-I., Hirano, A. & Kaji, R. 2005. Functional anatomy of the basal ganglia in X-linked recessive dystonia-parkinsonism. *Annals of neurology,* 58**,** 7-17.

Grabowski, M., Zimprich, A., Lorenz-Depiereux, B., Kalscheuer, V., Asmus, F., Gasser, T., Meitinger, T. & Strom, T. M. 2003. The epsilon-sarcoglycan gene (SGCE), mutated in myoclonus-dystonia syndrome, is maternally imprinted. *European Journal of Human Genetics,* 11**,** 138-44.

Granata, A., Watson, R., Collinson, L. M., Schiavo, G. & Warner, T. T. 2008. The dystonia-associated protein torsinA modulates synaptic vesicle recycling. *Journal of Biological Chemistry,* 283**,** 7568-79.

Green, M. R. & Sambrook, J. 2012. *Molecular Cloning: A Laboratory Manual,* New York, Cold Spring Harbor Laboratory Press.

Groen, J., Van Rootselaar, A. F., Van Der Salm, S. M., Bloem, B. R. & Tijssen, M. 2011. A new familial syndrome with dystonia and lower limb action myoclonus. *Movement Disorders,* 26**,** 896-900. doi: 10.1002/mds.23557. Epub 2011 Mar 2.

Groen, J. L., Andrade, A., Ritz, K., Jalalzadeh, H., Haagmans, M., Bradley, T. E. J., Jongejan, A., Verbeek, D. S., Nürnberg, P., Denome, S., Hennekam, R. C. M., Lipscombe, D., Baas, F. & Tijssen, M. a. J. 2015a. CACNA1B mutation is linked to unique myoclonus-dystonia syndrome. *Human Molecular Genetics,* 24**,** 987-993.

Groen, J. L., Ritz, K., Jalalzadeh, H., Van Der Salm, S. M. A., Jongejan, A., Mook, O. R., Haagmans, M. A., Zwinderman, A. H., Motazacker, M. M., Hennekam, R. C., Baas,

F. & Tijssen, M. a. J. 2015b. RELN rare variants in myoclonus-dystonia. *Movement Disorders*, n/a-n/a.

Groh, S., Zong, H., Goddeeris, M. M., Lebakken, C. S., Venzke, D., Pessin, J. E. & Campbell, K. P. 2009. Sarcoglycan complex: implications for metabolic defects in muscular dystrophies. *Journal of Biological Chemistry,* 284**,** 19178-82.

Grunewald, A., Djarmati, A., Lohmann-Hedrich, K., Farrell, K., Zeller, J. A., Allert, N., Papengut, F., Petersen, B., Fung, V., Sue, C. M., O'sullivan, D., Mahant, N., Kupsch, A., Chuang, R. S., Wiegers, K., Pawlack, H., Hagenah, J., Ozelius, L. J., Stephani, U., Schuit, R., Lang, A. E., Volkmann, J., Munchau, A. & Klein, C. 2008. Myoclonus-dystonia: significance of large SGCE deletions. *Human Mutation,* 29**,** 331-2.

Guyon, J. R., Kudryashova, E., Potts, A., Dalkilic, I., Brosius, M. A., Thompson, T. G., Beckmann, J. S., Kunkel, L. M. & Spencer, M. J. 2003. Calpain 3 cleaves filamin C and regulates its ability to interact with gamma- and delta-sarcoglycans. *Muscle & nerve,* 28**,** 472-83.

Hack, A. A., Lam, M. Y., Cordier, L., Shoturma, D. I., Ly, C. T., Hadhazy, M. A., Hadhazy, M. R., Sweeney, H. L. & Mcnally, E. M. 2000. Differential requirement for individual sarcoglycans and dystrophin in the assembly and function of the dystrophin-glycoprotein complex. *Journal of Cell Science,* 113**,** 2535-44.

Hack, A. A., Ly, C. T., Jiang, F., Clendenin, C. J., Sigrist, K. S., Wollmann, R. L. & Mcnally, E. M. 1998. Gamma-sarcoglycan deficiency leads to muscle membrane defects and apoptosis independent of dystrophin. *Journal of Cell Biology,* 142**,** 1279-87.

Hagihara, K., Miura, R., Kosaki, R., Berglund, E., Ranscht, B. & Yamaguchi, Y. 1999. Immunohistochemical evidence for the brevican-tenascin-R interaction: colocalization in perineuronal nets suggests a physiological role for the interaction in the adult rat brain. *Journal of Comparative Neurology,* 410**,** 256-64.

Hallett, M. 2011. Neurophysiology of dystonia: The role of inhibition. *Neurobiology of Disease,* 42**,** 177-84.

Han, F., Lang, A. E., Racacho, L., Bulman, D. E. & Grimes, D. A. 2003. Mutations in the epsilon-sarcoglycan gene found to be uncommon in seven myoclonus-dystonia families. *Neurology,* 61**,** 244-6.

Harrow, J., Frankish, A., Gonzalez, J., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-Reyes, G., Saunders, G., Steward, C., Harte, R., Lin, M., Howald, C., Tanzer, A., Derrien, T., Chrast, J., Walters, N., Balasubramanian, S., Pei, B., Tress, M., Rodriguez, J. M., Ezkurdia, I., Van Baren, J., Brent, M., Haussler, D., Kellis, M., Valencia, A., Reymond, A., Gerstein, M., Guigó, R. & Hubbard, T. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome research,* 22**,** 1760-1774.

Hassoni, A. A. & Cullen, M. J. 1999. Calcium homeostasis and ultrastructural studies in a patient with limb girdle muscular dystrophy type 2C. *Neuropathology and applied neurobiology,* 25**,** 244-53.

Haunso, A., Ibrahim, M., Bartsch, U., Letiembre, M., Celio, M. R. & Menoud, P. 2000. Morphology of perineuronal nets in tenascin-R and parvalbumin single and double knockout mice. *Brain Research,* 864**,** 142-5.

Hayashi, K., Wakayama, Y., Inoue, M., Kojima, H., Shibuya, S., Jimi, T., Hara, H. & Oniki, H. 2006. Sarcospan: ultrastructural localization and its relation to the sarcoglycan subcomplex. *Micron (Oxford, England : 1993),* 37**,** 591-6.

Helliwell, T. R., Nguyen, T. M. & Morris, G. E. 1994. Expression of the 43 kDa dystrophin-associated glycoprotein in human neuromuscular disease. *Neuromuscular Disorders,* 4**,** 101-13.

Herath, P., Gallea, C., Van Der Veen, J. W., Horovitz, S. G. & Hallett, M. 2010. In vivo neurochemistry of primary focal hand dystonia: a magnetic resonance spectroscopic neurometabolite profiling study at 3T. *Movement Disorders,* 25**,** 2800-8.

Heron, S. E., Grinton, B. E., Kivity, S., Afawi, Z., Zuberi, S. M., Hughes, J. N., Pridmore, C., Hodgson, B. L., Iona, X., Sadleir, L. G., Pelekanos, J., Herlenius, E., Goldberg-Stern, H., Bassan, H., Haan, E., Korczyn, A. D., Gardner, A. E., Corbett, M. A., Gecz, J., Thomas, P. Q., Mulley, J. C., Berkovic, S. F., Scheffer, I. E. & Dibbens, L. M. 2012. PRRT2 mutations cause benign familial infantile epilepsy and infantile convulsions with choreoathetosis syndrome. *American journal of human genetics,* 90**,** 152-60.

Hersheson, J., Mencacci, N. E., Davis, M., Macdonald, N., Trabzuni, D., Ryten, M., Pittman, A., Paudel, R., Kara, E., Fawcett, K., Plagnol, V., Bhatia, K. P., Medlar, A. J., Stanescu, H. C., Hardy, J., Kleta, R., Wood, N. W. & Houlden, H. 2013. Mutations in the autoregulatory domain of beta-tubulin 4a cause hereditary dystonia. *Annals of Neurology,* 73**,** 546-53.

Herve, D., Le Moine, C., Corvol, J. C., Belluscio, L., Ledent, C., Fienberg, A. A., Jaber, M., Studler, J. M. & Girault, J. A. 2001. Galpha(olf) levels are regulated by receptor usage and control dopamine and adenosine action in the striatum. *Journal of Neuroscience,* 21**,** 4390-9.

Herve, D., Levi-Strauss, M., Marey-Semper, I., Verney, C., Tassin, J. P., Glowinski, J. & Girault, J. A. 1993. G(olf) and Gs in rat basal ganglia: possible involvement of G(olf) in the coupling of dopamine D1 receptor with adenylyl cyclase. *Journal of Neuroscience,* 13**,** 2237-48.

Herzfeld, T., Nolte, D. & Muller, U. 2007. Structural and functional analysis of the human TAF1/DYT3 multiple transcript system. *Mammalian genome : official journal of the International Mammalian Genome Society,* 18**,** 787-95.

Hess, C. W., Raymond, D., Aguiar Pde, C., Frucht, S., Shriberg, J., Heiman, G. A., Kurlan, R., Klein, C., Bressman, S. B., Ozelius, L. J. & Saunders-Pullman, R. 2007. Myoclonus-dystonia, obsessive-compulsive disorder, and alcohol dependence in SGCE mutation carriers. *Neurology,* 68**,** 522-4.

Hewett, J. W., Zeng, J., Niland, B. P., Bragg, D. C. & Breakefield, X. O. 2006. Dystonia-causing mutant torsinA inhibits cell adhesion and neurite extension through interference with cytoskeletal dynamics. *Neurobiology of Disease,* 22**,** 98-111.

Heydemann, A., Demonbreun, A., Hadhazy, M., Earley, J. U. & Mcnally, E. M. 2007. Nuclear sequestration of delta-sarcoglycan disrupts the nuclear localization of lamin A/C and emerin in cardiomyocytes. *Human molecular genetics,* 16**,** 355-63.

Heydemann, A., Huber, J. M., Kakkar, R., Wheeler, M. T. & Mcnally, E. M. 2004. Functional nitric oxide synthase mislocalization in cardiomyopathy. *Journal of molecular and cellular cardiology,* 36**,** 213-23.

Hicks, D., Sarkozy, A., Muelas, N., Koehler, K., Huebner, A., Hudson, G., Chinnery, P. F., Barresi, R., Eagle, M., Polvikoski, T., Bailey, G., Miller, J., Radunovic, A., Hughes, P. J., Roberts, R., Krause, S., Walter, M. C., Laval, S. H., Straub, V., Lochmuller, H. & Bushby, K. 2011. A founder mutation in Anoctamin 5 is a major cause of limb-girdle muscular dystrophy. *Brain,* 134**,** 171-82.

Hisatsune, C., Miyamoto, H., Hirono, M., Yamaguchi, N., Sugawara, T., Ogawa, N., Ebisui, E., Ohshima, T., Yamada, M., Hensch, T. K., Hattori, M. & Mikoshiba, K. 2013. IP3R1 deficiency in the cerebellum/brainstem causes basal ganglia-independent dystonia by triggering tonic Purkinje cell firings in mice. *Frontiers in Neural Circuits,* 7**,** 156.

Hjermind, L. E., Vissing, J., Asmus, F., Krag, T., Lochmuller, H., Walter, M. C., Erdal, J., Blake, D. J. & Nielsen, J. E. 2008. No muscle involvement in myoclonus-dystonia

caused by epsilon-sarcoglycan gene mutations. *European Journal of Neurology,* 15**,** 525-9.

Holt, K. H. & Campbell, D. L. 1998. Assembly of the Sarcoglycan Complex. INSIGHTS FOR MUSCULAR DYSTROPHY. *Journal of Biological Chemistry,* 273**,** 34667-34670.

Holt, K. H., Crosbie, R. H., Venzke, D. P. & Campbell, K. P. 2000. Biosynthesis of dystroglycan: processing of a precursor propeptide. *FEBS Letters,* 468**,** 79-83.

Hong, S. E., Shugart, Y. Y., Huang, D. T., Shahwan, S. A., Grant, P. E., Hourihane, J. O., Martin, N. D. & Walsh, C. A. 2000. Autosomal recessive lissencephaly with cerebellar hypoplasia is associated with human RELN mutations. *Nature Genetics,* 26**,** 93-6.

Houlden, H., Schneider, S. A., Paudel, R., Melchers, A., Schwingenschuh, P., Edwards, M., Hardy, J. & Bhatia, K. P. 2010. THAP1 mutations (DYT6) are an additional cause of early-onset dystonia. *Neurology,* 74**,** 846-50.

Howell, B. W., Herrick, T. M. & Cooper, J. A. 1999. Reelin-induced tyrosine [corrected] phosphorylation of disabled 1 during neuronal positioning. *Genes and Development,* 13**,** 643-8.

Howell, B. W., Herrick, T. M., Hildebrand, J. D., Zhang, Y. & Cooper, J. A. 2000. Dab1 tyrosine phosphorylation sites relay positional signals during mouse brain development. *Current Biology,* 10**,** 877-885.

Huang, F., Wang, X., Ostertag, E. M., Nuwal, T., Huang, B., Jan, Y. N., Basbaum, A. I. & Jan, L. Y. 2013. TMEM16C facilitates Na(+)-activated K+ currents in rat sensory neurons and regulates pain processing. *Nature neuroscience,* 16**,** 1284-90.

Hubsch, C., Vidailhet, M., Rivaud-Pechoux, S., Pouget, P., Brochard, V., Degos, B., Pelisson, D., Golmard, J. L., Gaymard, B. & Roze, E. 2011. Impaired saccadic adaptation in DYT11 dystonia. *Journal of Neurology, Neurosurgery & Psychiatry,* 82**,** 1103-6.

Iafrati, J., Orejarena, M. J., Lassalle, O., Bouamrane, L. & Chavis, P. 2014. Reelin, an extracellular matrix protein linked to early onset psychiatric diseases, drives postnatal development of the prefrontal cortex via GluN2B-NMDARs and the mTOR pathway. *Molecular Psychiatry,* 19**,** 417-426.

Ibraghimov-Beskrovnaya, O., Ervasti, J. M., Leveille, C. J., Slaughter, C. A., Sernett, S. W. & Campbell, K. P. 1992. Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature,* 355**,** 696-702.

Ikeda, K., Satake, S., Onaka, T., Sugimoto, H., Takeda, N., Imoto, K. & Kawakami, K. 2013. Enhanced inhibitory neurotransmission in the cerebellar cortex of Atp1a3-deficient heterozygous mice. *Journal of physiology,* 591**,** 3433-49.

Ikeda, Y., Gu, Y., Iwanaga, Y., Hoshijima, M., Oh, S. S., Giordano, F. J., Chen, J., Nigro, V., Peterson, K. L., Chien, K. R. & Ross, J., Jr. 2002. Restoration of deficient membrane proteins in the cardiomyopathic hamster by in vivo cardiac gene transfer. *Circulation,* 105**,** 502-8.

Ilsley, J. L., Sudol, M. & Winder, S. J. 2001. The interaction of dystrophin with β-dystroglycan is regulated by tyrosine phosphorylation. *Cellular Signalling,* 13**,** 625-632.

Imamura, M., Araishi, K., Noguchi, S. & Ozawa, E. 2000. A sarcoglycan-dystroglycan complex anchors Dp116 and utrophin in the peripheral nervous system. *Human Molecular Genetics,* 9**,** 3091-100.

Isken, O. & Maquat, L. 2007. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes & Development,* 21**,** 1833-3856.

Iwata, Y., Katanosaka, Y., Arai, Y., Komamura, K., Miyatake, K. & Shigekawa, M. 2003. A novel mechanism of myocyte degeneration involving the Ca2+-permeable growth factor-regulated channel. *The Journal of cell biology,* 161**,** 957-67.

Iwata, Y., Katanosaka, Y., Shijun, Z., Kobayashi, Y., Hanada, H., Shigekawa, M. & Wakabayashi, S. 2005. Protective effects of Ca2+ handling drugs against abnormal Ca2+ homeostasis and cell damage in myopathic skeletal muscle cells. *Biochemical pharmacology,* 70**,** 740-51.

Jo, J., Son, G. H., Winters, B. L., Kim, M. J., Whitcomb, D. J., Dickinson, B. A., Lee, Y.-B., Futai, K., Amici, M., Sheng, M., Collingridge, G. L. & Cho, K. 2010. Muscarinic receptors induce LTD of NMDAR EPSCs via a mechanism involving hippocalcin, AP2 and PSD-95. *Nature neuroscience,* 13**,** 1216-1224.

Jung, D., Duclos, F., Apostol, B., Straub, V., Lee, J. C., Allamand, V., Venzke, D. P., Sunada, Y., Moomaw, C. R., Leveille, C. J., Slaughter, C. A., Crawford, T. O., Mcpherson, J. D. & Campbell, K. P. 1996. Characterization of delta-sarcoglycan, a novel component of the oligomeric sarcoglycan complex involved in limb-girdle muscular dystrophy. *The Journal of biological chemistry,* 271**,** 32321-9.

Kagi, G., Katschnig, P., Fiorio, M., Tinazzi, M., Ruge, D., Rothwell, J. & Bhatia, K. P. 2013. Sensory tricks in primary cervical dystonia depend on visuotactile temporal discrimination. *Movement Disorders,* 28**,** 356-61.

Kaiser, F. J., Osmanoric, A., Rakovic, A., Erogullari, A., Uflacker, N., Braunholz, D., Lohnau, T., Orolicki, S., Albrecht, M., Gillessen-Kaesbach, G., Klein, C. & Lohmann, K. 2010. The dystonia gene DYT1 is repressed by the transcription factor THAP1 (DYT6). *Annals of Neurology,* 68**,** 554-9.

Kaji, R., Rothwell, J. C., Katayama, M., Ikeda, T., Kubori, T., Kohara, N., Mezaki, T., Shibasaki, H. & Kimura, J. 1995. Tonic vibration reflex and muscle afferent block in writer's cramp. *Annals of Neurology,* 38**,** 155-162.

Kamm, C., Boston, H., Hewett, J., Wilbur, J., Corey, D. P., Hanson, P. I., Ramesh, V. & Breakefield, X. O. 2004. The early onset dystonia protein torsinA interacts with kinesin light chain 1. *Journal of Biological Chemistry,* 279**,** 19882-92.

Kasahara, K., Kawakami, Y., Kiyono, T., Yonemura, S., Kawamura, Y., Era, S., Matsuzaki, F., Goshima, N. & Inagaki, M. 2014. Ubiquitin-proteasome system controls ciliogenesis at the initial step of axoneme extension. *Nature Communications,* 5.

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M. & Stamm, S. 2013. Function of alternative splicing. *Gene,* 514**,** 1-30.

Kent, W. J. 2002. BLAT—The BLAST-Like Alignment Tool. *Genome Research,* 12**,** 656-664.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. & Haussler, D. 2002. The human genome browser at UCSC. *Genome Research,* 12**,** 996-1006.

Khan, N. L., Wood, N. W. & Bhatia, K. P. 2003. Autosomal recessive, DYT2-like primary torsion dystonia: a new family. *Neurology,* 61**,** 1801-3.

Kim, E., Magen, A. & Ast, G. 2007. Different levels of alternative splicing among eukaryotes. *Nucleic Acids Research,* 35**,** 125-31.

Kim, Kwang s., Kobayashi, M., Takamatsu, K. & Tzingounis, Anastasios v. 2012. Hippocalcin and KCNQ Channels Contribute to the Kinetics of the Slow Afterhyperpolarization. *Biophysical Journal,* 103**,** 2446-2454.

Kinugawa, K., Vidailhet, M., Clot, F., Apartis, E., Grabli, D. & Roze, E. 2009. Myoclonus-dystonia: an update. *Movement Disorders,* 24**,** 479-89.

Klein, C., Brin, M. F., Kramer, P., Sena-Esteves, M., De Leon, D., Doheny, D., Bressman, S., Fahn, S., Breakefield, X. O. & Ozelius, L. J. 1999. Association of a missense change

in the D2 dopamine receptor with myoclonus dystonia. *Proceedings of the National Academy of Sciences,* 96**,** 5173-5176.

Klein, C., Gurvich, N., Sena-Esteves, M., Bressman, S., Brin, M. F., Ebersole, B. J., Fink, S., Forsgren, L., Friedman, J., Grimes, D., Holmgren, G., Kyllerman, M., Lang, A. E., De Leon, D., Leung, J., Prioleau, C., Raymond, D., Sanner, G., Saunders-Pullman, R., Vieregge, P., Wahlstrom, J., Breakefield, X. O., Kramer, P. L., Ozelius, L. J. & Sealfon, S. C. 2000a. Evaluation of the role of the D2 dopamine receptor in myoclonus dystonia. *Annals of Neurology,* 47**,** 369-73.

Klein, C., Liu, L., Doheny, D., Kock, N., Muller, B., De Carvalho Aguiar, P., Leung, J., De Leon, D., Bressman, S. B., Silverman, J., Smith, C., Danisi, F., Morrison, C., Walker, R. H., Velickovic, M., Schwinger, E., Kramer, P. L., Breakefield, X. O., Brin, M. F. & Ozelius, L. J. 2002. Epsilon-sarcoglycan mutations found in combination with other dystonia gene mutations. *Annals of Neurology,* 52**,** 675-9.

Klein, C., Schilling, K., Saunders-Pullman, R. J., Garrels, J., Breakefield, X. O., Brin, M. F., Deleon, D., Doheny, D., Fahn, S., Fink, J. S., Forsgren, L., Friedman, J., Frucht, S., Harris, J., Holmgren, G., Kis, B., Kurlan, R., Kyllerman, M., Lang, A. E., Leung, J., Raymond, D., Robishaw, J. D., Sanner, G., Schwinger, E., Tabamo, R. E. & Tagliati, M. 2000b. A major locus for myoclonus-dystonia maps to chromosome 7q in eight families. *American journal of human genetics,* 67**,** 1314-9.

Klinge, L., Dekomien, G., Aboumousa, A., Charlton, R., Epplen, J. T., Barresi, R., Bushby, K. & Straub, V. 2008. Sarcoglycanopathies: Can muscle immunoanalysis predict the genotype? *Neuromuscular Disorders,* 18**,** 934-941.

Knuesel, I., Mastrocola, M., Zuellig, R. A., Bornhauser, B., Schaub, M. C. & Fritschy, J.-M. 1999. Altered synaptic clustering of GABAA receptors in mice lacking dystrophin (mdx mice). *European Journal of Neuroscience,* 11**,** 4457-4462.

Kobayashi, M., Masaki, T., Hori, K., Masuo, Y., Miyamoto, M., Tsubokawa, H., Noguchi, H., Nomura, M. & Takamatsu, K. 2005. Hippocalcin-deficient mice display a defect in cAMP response element-binding protein activation associated with impaired spatial and associative memory. *Neuroscience,* 133**,** 471-484.

Kojovic, M., Cordivari, C. & Bhatia, K. 2011. Myoclonic disorders: a practical approach for diagnosis and treatment. *Therapeutic Advances in Neurological Disorders,* 4**,** 47-62.

Konakova, M., Huynh, D. P., Yong, W. & Pulst, S. M. 2001. Cellular distribution of torsin A and torsin B in normal human brain. *Archives of Neurology,* 58**,** 921-7.

Konakova, M. & Pulst, S. M. 2001. Immunocytochemical characterization of torsin proteins in mouse brain. *Brain research,* 922**,** 1-8.

Kornblihtt, A., Schor, I., Alló, M., Dujardin, G., Petrillo, E. & Muñoz, M. 2013. Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nature reviews. Molecular cell biology,* 14**,** 153-165.

Krasowska, E., Zabłocki, K., Górecki, D. C. & Swinny, J. D. 2014. Aberrant Location of Inhibitory Synaptic Marker Proteins in the Hippocampus of Dystrophin-Deficient Mice: Implications for Cognitive Impairment in Duchenne Muscular Dystrophy. *PLoS ONE,* 9**,** e108364.

Kull, B., Svenningsson, P. & Fredholm, B. B. 2000. Adenosine A(2A) receptors are colocalized with and activate g(olf) in rat striatum. *Molecular Pharmacology,* 58**,** 771-7.

Kunzelmann, K., Cabrita, I., Wanitchakool, P., Ousingsawat, J., Sirianant, L., Benedetto, R. & Schreiber, R. 2016. Modulating Ca(2+) signals: a common theme for TMEM16, Ist2, and TMC. *Pflugers Arch,* 468**,** 475-90.

Lai, Y., Zhao, J., Yue, Y. & Duan, D. 2013. α2 and α3 helices of dystrophin R16 and R17 frame a microdomain in the α1 helix of dystrophin R17 for neuronal NOS binding. *Proceedings of the National Academy of Sciences,* 110**,** 525-530.

Lakatosova, S. & Ostatnikova, D. 2012. Reelin and its complex involvement in brain development and function. *The International Journal of Biochemistry & Cell Biology,* 44**,** 1501-1504.

Lancioni, A., Rotundo, I. L., Kobayashi, Y. M., D'orsi, L., Aurino, S., Nigro, G., Piluso, G., Acampora, D., Cacciottolo, M., Campbell, K. P. & Nigro, V. 2011. Combined deficiency of alpha and epsilon sarcoglycan disrupts the cardiac dystrophin complex. *Human Molecular Genetics,* 20**,** 4644-54.

Ledoux, M. S. 2011. Animal models of dystonia: Lessons from a mutant rat. *Neurobiology of Disease,* 42**,** 152-61.

Ledoux, M. S. & Brady, K. A. 2003. Secondary cervical dystonia associated with structural lesions of the central nervous system. *Movement Disorders,* 18**,** 60-9.

Ledoux, M. S., Lorden, J. F. & Ervin, J. M. 1993. Cerebellectomy Eliminates the Motor Syndrome of the Genetically Dystonic Rat. *Experimental Neurology,* 120**,** 302-310.

Ledoux, M. S., Lorden, J. F. & Meinzen-Derr, J. 1995. Selective elimination of cerebellar output in the genetically dystonic rat. *Brain Research,* 697**,** 91-103.

Lee, H. Y., Huang, Y., Bruneau, N., Roll, P., Roberson, E. D., Hermann, M., Quinn, E., Maas, J., Edwards, R., Ashizawa, T., Baykan, B., Bhatia, K., Bressman, S., Bruno, M. K., Brunt, E. R., Caraballo, R., Echenne, B., Fejerman, N., Frucht, S., Gurnett, C. A., Hirsch, E., Houlden, H., Jankovic, J., Lee, W. L., Lynch, D. R., Mohammed, S., Muller, U., Nespeca, M. P., Renner, D., Rochette, J., Rudolf, G., Saiki, S., Soong, B. W., Swoboda, K. J., Tucker, S., Wood, N., Hanna, M., Bowcock, A. M., Szepetowski, P., Fu, Y. H. & Ptacek, L. J. 2012a. Mutations in the gene PRRT2 cause paroxysmal kinesigenic dyskinesia with infantile convulsions. *Cell Reports,* 1**,** 2-12.

Lee, H. Y., Nakayama, J., Xu, Y., Fan, X., Karouani, M., Shen, Y., Pothos, E. N., Hess, E. J., Fu, Y. H., Edwards, R. H. & Ptacek, L. J. 2012b. Dopamine dysregulation in a mouse model of paroxysmal nonkinesigenic dyskinesia. *Journal of Clinical Investigation,* 122**,** 507-18.

Lee, H. Y., Xu, Y., Huang, Y., Ahn, A. H., Auburger, G. W., Pandolfo, M., Kwiecinski, H., Grimes, D. A., Lang, A. E., Nielsen, J. E., Averyanov, Y., Servidei, S., Friedman, A., Van Bogaert, P., Abramowicz, M. J., Bruno, M. K., Sorensen, B. F., Tang, L., Fu, Y. H. & Ptacek, L. J. 2004. The gene for paroxysmal non-kinesigenic dyskinesia encodes an enzyme in a stress response pathway. *Human Molecular Genetics,* 13**,** 3161-70.

Lee, L. V., Maranon, E., Demaisip, C., Peralta, O., Borres-Icasiano, R., Arancillo, J., Rivera, C., Munoz, E., Tan, K. & Reyes, M. T. 2002. The natural history of sex-linked recessive dystonia parkinsonism of Panay, Philippines (XDP). *Parkinsonism & related disorders,* 9**,** 29-38.

Lee, L. V., Rivera, C., Teleg, R. A., Dantes, M. B., Pasco, P. M., Jamora, R. D., Arancillo, J., Villareal-Jordan, R. F., Rosales, R. L., Demaisip, C., Maranon, E., Peralta, O., Borres, R., Tolentino, C., Monding, M. J. & Sarcia, S. 2011. The unique phenomenology of sex-linked dystonia parkinsonism (XDP, DYT3, "Lubag"). *International Journal of Neuroscience,* 121 Suppl 1**,** 3-11.

Lee, Y., Hur, I., Park, S.-Y., Kim, Y.-K., Suh, M. & Kim, N. 2006. The role of PACT in the RNA silencing pathway. *The EMBO Journal,* 25**,** 522-532.

Lehéricy, S., Tijssen, M. a. J., Vidailhet, M., Kaji, R. & Meunier, S. 2013. The anatomical basis of dystonia: Current view using neuroimaging. *Movement Disorders,* 28**,** 944-957.

Leung, J., Klein, C., Friedman, J., Vieregge, P., Jacobs, H., Doheny, D., Kamm, C., Deleon, D., Pramstaller, P., Penney, J., Eisengart, M., Jankovic, J., Gasser, T., Bressman, S., Corey, D., Kramer, P., Brin, M., Ozelius, L. & Breakefield, X. 2001. Novel mutation in the TOR1A (DYT1) gene in atypical, early onset dystonia and polymorphisms in dystonia and early onset parkinsonism. *Neurogenetics,* 3**,** 133-143.

Lev-Maor, G., Sorek, R., Shomron, N. & Ast, G. 2003. The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. *Science,* 300.

Levy, A., Sela, N. & Ast, G. 2008. TranspoGene and microTranspoGene: transposed elements influence on the transcriptome of seven vertebrates and invertebrates. *Nucleic Acids Research,* 36**,** D47-52.

Levy, L. M. & Hallett, M. 2002. Impaired brain GABA in focal dystonia. *Annals of Neurology,* 51**,** 93-101.

Lewis, B. P., Green, R. E. & Brenner, S. E. 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proceedings of the National Academy of Sciences,* 100**,** 189-92.

Li, D., Long, C., Yue, Y. & Duan, D. 2009. Sub-physiological sarcoglycan expression contributes to compensatory muscle protection in mdx mice. *Human molecular genetics,* 18**,** 1209-20.

Li, J. Y., Cunic, D. I., Paradiso, G., Gunraj, C., Pal, P. K., Lang, A. E. & Chen, R. 2008. Electrophysiological features of myoclonus-dystonia. *Movement Disorders,* 23**,** 2055-61.

Li, M., Niu, F., Zhu, X., Wu, X., Shen, N., Peng, X. & Liu, Y. 2015. PRRT2 Mutant Leads to Dysfunction of Glutamate Signaling. *International journal of molecular sciences,* 16**,** 9134-51.

Lim, L. E., Duclos, F., Broux, O., Bourg, N., Sunada, Y., Allamand, V., Meyer, J., Richard, I., Moomaw, C., Slaughter, C. & Et Al. 1995. Beta-sarcoglycan: characterization and role in limb-girdle muscular dystrophy linked to 4q12. *Nature genetics,* 11**,** 257-65.

Lin, Y. C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van De Peer, Y., Tavernier, J. & Callewaert, N. 2014. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature Communications,* 5**,** 4767.

Liu, L., Vachon, P. H., Kuang, W., Xu, H., Wewer, U. M., Kylsten, P. & Engvall, E. 1997. Mouse adhalin: primary structure and expression during late stages of muscle differentiation in vitro. *Biochemical and biophysical research communications,* 235**,** 227-35.

Liu, L. A. & Engvall, E. 1999. Sarcoglycan isoforms in skeletal muscle. *The Journal of biological chemistry,* 274**,** 38171-38176.

Liu, Q., Niu, N., Wada, Y. & Liu, J. 2016. The Role of Cdkn1A-Interacting Zinc Finger Protein 1 (CIZ1) in DNA Replication and Pathophysiology. *International journal of molecular sciences,* 17.

Lohmann, K. & Klein, C. 2013. Genetics of dystonia: what's known? What's new? What's next? *Movement Disorders,* 28**,** 899-905.

Lohmann, K., Wilcox, R. A., Winkler, S., Ramirez, A., Rakovic, A., Park, J. S., Arns, B., Lohnau, T., Groen, J., Kasten, M., Bruggemann, N., Hagenah, J., Schmidt, A., Kaiser, F. J., Kumar, K. R., Zschiedrich, K., Alvarez-Fischer, D., Altenmuller, E., Ferbert, A., Lang, A. E., Munchau, A., Kostic, V., Simonyan, K., Agzarian, M., Ozelius, L. J., Langeveld, A. P., Sue, C. M., Tijssen, M. A. & Klein, C. 2013. Whispering dysphonia (DYT4 dystonia) is caused by a mutation in the TUBB4 gene. *Annals of Neurology,* 73**,** 537-45.

Loos, M., Li, K. W., Van Der Schors, R., Gouwenberg, Y., Van Der Loo, R., Williams, R. W., Smit, A. B. & Spijker, S. 2016. Impact of genetic variation on synaptic protein levels in genetically diverse mice. *PROTEOMICS,* 16**,** 1123-1130.

Ma, K., Babij, R., Cortes, E., Vonsattel, J. P. & Louis, E. D. 2012. Cerebellar pathology of a dual clinical diagnosis: patients with essential tremor and dystonia. *Tremor Other Hyperkinet Mov (N Y),* 2.

Ma, L. Y., Wang, L., Yang, Y. M., Feng, T. & Wan, X. H. 2015. Mutations in ANO3 and GNAL gene in thirty-three isolated dystonia families. *Movement Disorders,* 30**,** 743-4.

Makino, S., Kaji, R., Ando, S., Tomizawa, M., Yasuno, K., Goto, S., Matsumoto, S., Tabuena, M. D., Maranon, E., Dantes, M., Lee, L. V., Ogasawara, K., Tooyama, I., Akatsu, H., Nishimura, M. & Tamiya, G. 2007. Reduced neuron-specific expression of the TAF1 gene is associated with X-linked dystonia-parkinsonism. *American journal of human genetics,* 80**,** 393-406.

Marelli, C., Canafoglia, L., Zibordi, F., Ciano, C., Visani, E., Zorzi, G., Garavaglia, B., Barzaghi, C., Albanese, A., Soliveri, P., Leone, M., Panzica, F., Scaioli, V., Pincherle, A., Nardocci, N. & Franceschetti, S. 2008. A neurophysiological study of myoclonus in patients with DYT11 myoclonus-dystonia syndrome. *Movement Disorders,* 23**,** 2041-8.

Marshall, J. L., Chou, E., Oh, J., Kwok, A., Burkin, D. J. & Crosbie-Watson, R. H. 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**,** 4378-4393.

Matsumura, M., Sawaguchi, T. & Kubota, K. 1992. GABAergic inhibition of neuronal activity in the primate motor and premotor cortex during voluntary movement. *Journal of Neurophysiology,* 68**,** 692-702.

Matsumura, M., Sawaguchi, T., Oishi, T., Ueki, K. & Kubota, K. 1991. Behavioral deficits induced by local injection of bicuculline and muscimol into the primate motor and premotor cortex. *Journal of Neurophysiology,* 65**,** 1542-53.

Mcgrail, K. M., Phillips, J. M. & Sweadner, K. J. 1991. Immunofluorescent localization of three Na,K-ATPase isozymes in the rat central nervous system: both neurons and glia can express more than one Na,K-ATPase. *Journal of Neuroscience,* 11**,** 381-91.

Mclean, P. J., Kawamata, H., Shariff, S., Hewett, J., Sharma, N., Ueda, K., Breakefield, X. O. & Hyman, B. T. 2002. TorsinA and heat shock proteins act as molecular chaperones: suppression of alpha-synuclein aggregation. *Journal of Neurochemistry,* 83**,** 846-54.

Mcnally, E. M., Ly, C. T. & Kunkel, L. M. 1998. Human epsilon-sarcoglycan is highly related to alpha-sarcoglycan (adhalin), the limb girdle muscular dystrophy 2D gene. *FEBS letters,* 422**,** 27-32.

Mcnally, E. M., Passos-Bueno, M. R., Bonnemann, C. G., Vainzof, M., De Sa Moreira, E., Lidov, H. G., Othmane, K. B., Denton, P. H., Vance, J. M., Zatz, M. & Kunkel, L. M. 1996. Mild and severe muscular dystrophy caused by a single gamma-sarcoglycan mutation. *American journal of human genetics,* 59**,** 1040-7.

Mcnally, E. M., Yoshida, M., Mizuno, Y., Ozawa, E. & Kunkel, L. M. 1994. Human adhalin is alternatively spliced and the gene is located on chromosome 17q21. *Proceedings of the National Academy of Sciences,* 91**,** 9690-4.

Mencacci, N. E., R'bibo, L., Bandres-Ciga, S., Carecchio, M., Zorzi, G., Nardocci, N., Garavaglia, B., Batla, A., Bhatia, K. P., Pittman, A. M., Hardy, J., Weissbach, A., Klein, C., Gasser, T., Lohmann, E. & Wood, N. W. 2015a. The CACNA1B R1389H variant is not associated with myoclonus-dystonia in a large European multicentric cohort. *Human Molecular Genetics,* 24**,** 5326-9.

Mencacci, N. E., Rubio-Agusti, I., Zdebik, A., Asmus, F., Ludtmann, M. H., Ryten, M., Plagnol, V., Hauser, A. K., Bandres-Ciga, S., Bettencourt, C., Forabosco, P., Hughes, D., Soutar, M. M., Peall, K., Morris, H. R., Trabzuni, D., Tekman, M., Stanescu, H. C., Kleta, R., Carecchio, M., Zorzi, G., Nardocci, N., Garavaglia, B., Lohmann, E., Weissbach, A., Klein, C., Hardy, J., Pittman, A. M., Foltynie, T., Abramov, A. Y., Gasser, T., Bhatia, K. P. & Wood, N. W. 2015b. A missense mutation in KCTD17 causes autosomal dominant myoclonus-dystonia. *American journal of human genetics,* 96**,** 938-47.

Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F. & Dietz, H. C. 2004. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nature Genetics,* 36**,** 1073-8.

Mercado, M. L., Amenta, A. R., Hagiwara, H., Rafii, M. S., Lechner, B. E., Owens, R. T., Mcquillan, D. J., Froehner, S. C. & Fallon, J. R. 2006. Biglycan regulates the expression and sarcolemmal localization of dystrobrevin, syntrophin, and nNOS. *FASEB J,* 20**,** 1724-6.

Meunier, S., Lourenco, G., Roze, E., Apartis, E., Trocello, J. M. & Vidailhet, M. 2008. Cortical excitability in DYT-11 positive myoclonus dystonia. *Movement Disorders,* 23**,** 761-4.

Meunier, S., Russmann, H., Shamim, E., Lamy, J.-C. & Hallett, M. 2012. Plasticity of cortical inhibition in dystonia is impaired after motor learning and paired-associative stimulation. *European Journal of Neuroscience,* 35**,** 975-986.

Millay, D. P., Goonasekera, S. A., Sargent, M. A., Maillet, M., Aronow, B. J. & Molkentin, J. D. 2009. Calcium influx is sufficient to induce muscular dystrophy through a TRPC-dependent mechanism. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 19023-8.

Miltgen, M., Blanchard, A., Mathieu, H., Kreisler, A., Jean Pierre, D., Salgado, D., Roubertie, A., Barre, L., Rai, G., Blanck, V., Frederic, M., Douay, X., Mazzolenni, R., Charpentier, P., Gonzalez, V., Destee, A., Beroud, C. & Collod-Beroud, G. 2016. Novel heterozygous mutation in ANO3 responsible for craniocervical dystonia. *Movement Disorders*.

Mitsui, K., Matsumoto, A., Ohtsuka, S., Ohtsubo, M. & Yoshimura, A. 1999. Cloning and characterization of a novel p21(Cip1/Waf1)-interacting zinc finger protein, ciz1. *Biochemical and Biophysical Research Communications,* 264**,** 457-64.

Mizuno, Y., Noguchi, S., Yamamoto, H., Yoshida, M., Suzuki, A., Hagiwara, Y., Hayashi, Y. K., Arahata, K., Nonaka, I., Hirai, S. & Et Al. 1994. Selective defect of sarcoglycan complex in severe childhood autosomal recessive muscular dystrophy muscle. *Biochemical & Biophysical Research Communications,* 203**,** 979-83.

Molloy, F. M., Carr, T. D., Zeuner, K. E., Dambrosia, J. M. & Hallett, M. 2003. Abnormalities of spatial discrimination in focal and generalized dystonia. *Brain,* 126**,** 2175-2182.

Montag-Sallaz, M. & Montag, D. 2003. Severe cognitive and motor coordination deficits in tenascin-R-deficient mice. *Genes, Brain and Behavior,* 2**,** 20-31.

Morawski, M., Dityatev, A., Hartlage-Rubsamen, M., Blosa, M., Holzer, M., Flach, K., Pavlica, S., Dityateva, G., Grosche, J., Bruckner, G. & Schachner, M. 2014. Tenascin-R promotes assembly of the extracellular matrix of perineuronal nets via clustering of aggrecan. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 369**,** 20140046.

Morel, N. & Poea-Guyon, S. 2015. The membrane domain of vacuolar H(+)ATPase: a crucial player in neurotransmitter exocytotic release. *Cellular and Molecular Life Sciences,* 72**,** 2561-73.

Morganti, M. C., Taylor, J., Pesheva, P. & Schachner, M. 1990. Oligodendrocyte-derived J1-160/180 extracellular matrix glycoproteins are adhesive or repulsive depending on the partner cell type and time of interaction. *Experimental Neurology,* 109**,** 98-110.

Muller, B., Hedrich, K., Kock, N., Dragasevic, N., Svetel, M., Garrels, J., Landt, O., Nitschke, M., Pramstaller, P. P., Reik, W., Schwinger, E., Sperner, J., Ozelius, L., Kostic, V. & Klein, C. 2002. Evidence that paternal expression of the epsilon-sarcoglycan gene accounts for reduced penetrance in myoclonus-dystonia. *American journal of human genetics,* 71**,** 1303-11.

Muller, U., Herzfeld, T. & Nolte, D. 2007. The TAF1/DYT3 multiple transcript system in X-linked dystonia-parkinsonism. *American journal of human genetics,* 81**,** 415-8.

Nadler, J. J., Zou, F., Huang, H., Moy, S. S., Lauder, J., Crawley, J. N., Threadgill, D. W., Wright, F. A. & Magnuson, T. R. 2006. Large-Scale Gene Expression Differences Across Brain Regions and Inbred Strains Correlate With a Behavioral Phenotype. *Genetics,* 174**,** 1229-1236.

Naismith, T. V., Dalal, S. & Hanson, P. I. 2009. Interaction of torsinA with its major binding partners is impaired by the dystonia-associated DeltaGAG deletion. *Journal of Biological Chemistry,* 284**,** 27866-74.

Naismith, T. V., Heuser, J. E., Breakefield, X. O. & Hanson, P. I. 2004. TorsinA in the nuclear envelope. *Proceedings of the National Academy of Sciences,* 101**,** 7612-7.

Nakamura, T. Y., Iwata, Y., Sampaolesi, M., Hanada, H., Saito, N., Artman, M., Coetzee, W. A. & Shigekawa, M. 2001. Stretch-activated cation channels in skeletal muscle myotubes from sarcoglycan-deficient hamsters. *American journal of physiology. Cell physiology,* 281**,** C690-9.

Nakashima, K., Rothwell, J. C., Day, B. L., Thompson, P. D., Shannon, K. & Marsden, C. D. 1989. Reciprocal inhibition between forearm muscles in patients with writer's cramp and other occupational cramps, symptomatic hemidystonia and hemiparesis due to stroke. *Brain,* 112 (Pt 3)**,** 681-97.

Nardocci, N., Zorzi, G., Barzaghi, C., Zibordi, F., Ciano, C., Ghezzi, D. & Garavaglia, B. 2008. Myoclonus-dystonia syndrome: clinical presentation, disease course, and genetic features in 11 families. *Movement Disorders,* 23**,** 28-34.

Nastase, M. V., Young, M. F. & Schaefer, L. 2012. Biglycan: A Multivalent Proteoglycan Providing Structure and Signals. *Journal of Histochemistry & Cytochemistry,* 60**,** 963-975.

Nery, F. C., Armata, I. A., Farley, J. E., Cho, J. A., Yaqub, U., Chen, P., Da Hora, C. C., Wang, Q., Tagaya, M., Klein, C., Tannous, B., Caldwell, K. A., Caldwell, G. A., Lencer, W. I., Ye, Y. & Breakefield, X. O. 2011. TorsinA participates in endoplasmic reticulum-associated degradation. *Nature Communications,* 2**,** 393.

Nery, F. C., Zeng, J., Niland, B. P., Hewett, J., Farley, J., Irimia, D., Li, Y., Wiche, G., Sonnenberg, A. & Breakefield, X. O. 2008. TorsinA binds the KASH domain of nesprins and participates in linkage between nuclear envelope and cytoskeleton. *Journal of Cell Science,* 121**,** 3476-86.

Neychev, V. K., Fan, X., Mitev, V. I., Hess, E. J. & Jinnah, H. A. 2008. The basal ganglia and cerebellum interact in the expression of dystonic movement. *Brain,* 131**,** 2499-2509.

Neychev, V. K., Gross, R. E., Lehericy, S., Hess, E. J. & Jinnah, H. A. 2011. The functional neuroanatomy of dystonia. *Neurobiology of Disease,* 42**,** 185-201.

Nguyen, T. M., Ginjaar, I. B., Van Ommen, G. J. & Morris, G. E. 1992. Monoclonal antibodies for dystrophin analysis. Epitope mapping and improved binding to SDS-treated muscle sections. *Biochemical Journal,* 288 (Pt 2)**,** 663-8.

Nigro, V., Okazaki, Y., Belsito, A., Piluso, G., Matsuda, Y., Politano, L., Nigro, G., Ventura, C., Abbondanza, C., Molinari, A. M., Acampora, D., Nishimura, M., Hayashizaki, Y. & Puca, G. A. 1997. Identification of the Syrian hamster cardiomyopathy gene. *Human molecular genetics,* 6**,** 601-7.

Nigro, V., Piluso, G., Belsito, A., Politano, L., Puca, A. A., Papparella, S., Rossi, E., Viglietto, G., Esposito, M. G., Abbondanza, C., Medici, N., Molinari, A. M., Nigro, G. & Puca, G. A. 1996. Identification of a novel sarcoglycan gene at 5q33 encoding a sarcolemmal 35 kDa glycoprotein. *Human molecular genetics,* 5**,** 1179-86.

Nikonenko, A., Schmidt, S., Skibo, G., Bruckner, G. & Schachner, M. 2003. Tenascin-R-deficient mice show structural alterations of symmetric perisomatic synapses in the CA1 region of the hippocampus. *Journal of Comparative Neurology,* 456**,** 338-49.

Nishikawa, T., Ota, T. & Isogai, T. 2000. Prediction whether a human cDNA sequence contains initiation codon by combining statistical information and similarity with protein sequences. *Bioinformatics,* 16**,** 960-7.

Nishiyama, A., Endo, T., Takeda, S. & Imamura, M. 2004. Identification and characterization of epsilon-sarcoglycans in the central nervous system. *Brain Research. Molecular Brain Research,* 125**,** 1-12.

Noguchi, S., Mcnally, E. M., Ben Othmane, K., Hagiwara, Y., Mizuno, Y., Yoshida, M., Yamamoto, H., Bonnemann, C. G., Gussoni, E., Denton, P. H., Kyriakides, T., Middleton, L., Hentati, F., Ben Hamida, M., Nonaka, I., Vance, J. M., Kunkel, L. M. & Ozawa, E. 1995. Mutations in the dystrophin-associated protein gamma-sarcoglycan in chromosome 13 muscular dystrophy. *Science (New York, N.Y.),* 270**,** 819-22.

Noguchi, S., Wakabayashi-Takai, E., Sasaoka, T. & Ozawa, E. 2001. Analysis of the spatial, temporal and tissue-specific transcription of gamma-sarcoglycan gene using a transgenic mouse. *FEBS letters,* 495**,** 77-81.

Noguchi, S., Wakabayashi, E., Imamura, M., Yoshida, M. & Ozawa, E. 2000. Formation of sarcoglycan complex with differentiation in cultured myocytes. *European Journal of Biochemistry,* 267**,** 640-648.

Nygaard, T. G., Raymond, D., Chen, C., Nishino, I., Greene, P. E., Jennings, D., Heiman, G. A., Klein, C., Saunders-Pullman, R. J., Kramer, P., Ozelius, L. J. & Bressman, S. B. 1999. Localization of a gene for myoclonus-dystonia to chromosome 7q21-q31. *Annals of Neurology,* 46**,** 794-8.

Oh, D.-Y., Yon, C., Oh, K.-J., Lee, K. S. & Han, J.-S. 2006. Hippocalcin increases phospholipase D2 expression through extracellular signal-regulated kinase activation and lysophosphatidic acid potentiates the hippocalcin-induced phospholipase D2 expression. *Journal of Cellular Biochemistry,* 97**,** 1052-1065.

Oohashi, T., Edamatsu, M., Bekku, Y. & Carulli, D. 2015. The hyaluronan and proteoglycan link proteins: Organizers of the brain extracellular matrix and key molecules for neuronal function and plasticity. *Experimental Neurology,* 274, Part B**,** 134-144.

Ozawa, E., Mizuno, Y., Hagiwara, Y., Sasaoka, T. & Yoshida, M. 2005. Molecular and cell biology of the sarcoglycan complex. *Muscle Nerve,* 32**,** 563-576.

Ozelius, L. J., Hewett, J. W., Page, C. E., Bressman, S. B., Kramer, P. L., Shalish, C., De Leon, D., Brin, M. F., Raymond, D., Corey, D. P., Fahn, S., Risch, N. J., Buckler, A. J., Gusella, J. F. & Breakefield, X. O. 1997. The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. *Nature Genetics,* 17**,** 40-8.

Palmer, C. L., Lim, W., Hastie, P. G. R., Toward, M., Korolchuk, V. I., Burbidge, S. A., Banting, G., Collingridge, G. L., Isaac, J. T. R. & Henley, J. M. 2005. Hippocalcin Functions as a Calcium Sensor in Hippocampal LTD. *Neuron,* 47**,** 487-494.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics,* 40**,** 1413-5.

Paronetto, M., Messina, V., Barchi, M., Geremia, R., Richard, S. & Sette, C. 2011. Sam68 marks the transcriptionally active stages of spermatogenesis and modulates alternative splicing in male germ cells. *Nucleic Acids Research,* 39**,** 4961-4974.

Parsons, S. A., Millay, D. P., Sargent, M. A., Naya, F. J., Mcnally, E. M., Sweeney, H. L. & Molkentin, J. D. 2007. Genetic disruption of calcineurin improves skeletal muscle pathology and cardiac disease in a mouse model of limb-girdle muscular dystrophy. *The Journal of biological chemistry,* 282**,** 10068-78.

Patel, C. V., Handy, I., Goldsmith, T. & Patel, R. C. 2000. PACT, a stress-modulated cellular activator of interferon-induced double-stranded RNA-activated protein kinase, PKR. *Journal of Biological Chemistry,* 275**,** 37993-8.

Patel, N., Jankovic, J. & Hallett, M. 2014. Sensory aspects of movement disorders. *The Lancet Neurology,* 13**,** 100-112.

Paterlini, M., Revilla, V., Grant, A. L. & Wisden, W. 2000. Expression of the neuronal calcium sensor protein family in the rat brain. *Neuroscience,* 99**,** 205-216.

Peall, K. J., Dijk, J. M., Saunders-Pullman, R., Dreissen, Y. E., Van Loon, I., Cath, D., Kurian, M. A., Owen, M. J., Foncke, E. M., Morris, H. R., Gasser, T., Bressman, S., Asmus, F. & Tijssen, M. A. 2016. Psychiatric disorders, myoclonus dystonia and SGCE: an international study. *Annals of clinical and translational neurology,* 3**,** 4-11.

Peall, K. J., Kurian, M. A., Wardle, M., Waite, A. J., Hedderly, T., Lin, J. P., Smith, M., Whone, A., Pall, H., White, C., Lux, A., Jardine, P. E., Lynch, B., Kirov, G., O'riordan, S., Samuel, M., Lynch, T., King, M. D., Chinnery, P. F., Warner, T. T., Blake, D. J., Owen, M. J. & Morris, H. R. 2014. SGCE and myoclonus dystonia: motor characteristics, diagnostic criteria and clinical predictors of genotype. *Journal of Neurology,* 261**,** 2296-304.

Peall, K. J., Smith, D. J., Kurian, M. A., Wardle, M., Waite, A. J., Hedderly, T., Lin, J. P., Smith, M., Whone, A., Pall, H., White, C., Lux, A., Jardine, P., Bajaj, N., Lynch, B., Kirov, G., O'riordan, S., Samuel, M., Lynch, T., King, M. D., Chinnery, P. F., Warner, T. T., Blake, D. J., Owen, M. J. & Morris, H. R. 2013. SGCE mutations cause psychiatric disorders: clinical and genetic characterization. *Brain,* 136**,** 294-303.

Peall, K. J., Waite, A. J., Blake, D. J., Owen, M. J. & Morris, H. R. 2011. Psychiatric disorders, myoclonus dystonia, and the epsilon-sarcoglycan gene: a systematic review. *Movement Disorders,* 26**,** 1939-42.

Pearson, T. S., Akman, C., Hinton, V. J., Engelstad, K. & De Vivo, D. C. 2013. Phenotypic spectrum of glucose transporter type 1 deficiency syndrome (Glut1 DS). *Current neurology and neuroscience reports,* 13**,** 342.

Pedersen, A. G. & Nielsen, H. 1997. Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. *Proc Int Conf Intell Syst Mol Biol,* 5**,** 226-33.

Pereboev, A. V., Ahmed, N., Thi Man, N. & Morris, G. E. 2001. Epitopes in the interacting regions of beta-dystroglycan (PPxY motif) and dystrophin (WW domain). *Biochimica et Biophysica Acta,* 1527**,** 54-60.

Peters, G. A., Li, S. & Sen, G. C. 2006. Phosphorylation of specific serine residues in the PKR activation domain of PACT is essential for its ability to mediate apoptosis. *Journal of Biological Chemistry,* 281**,** 35129-36.

Petersen, T. N., Brunak, S., Von Heijne, G. & Nielsen, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods,* 8**,** 785-786.

Petrof, B. J., Shrager, J. B., Stedman, H. H., Kelly, A. M. & Sweeney, H. L. 1993. Dystrophin protects the sarcolemma from stresses developed during muscle contraction. *Proceedings of the National Academy of Sciences,* 90**,** 3710-3714.

Phukan, J., Albanese, A., Gasser, T. & Warner, T. 2011. Primary dystonia and dystonia-plus syndromes: clinical characteristics, diagnosis, and pathogenesis. *The Lancet Neurology,* 10**,** 1074-1085.

Picollo, A., Malvezzi, M. & Accardi, A. 2015. TMEM16 Proteins: Unknown Structure and Confusing Functions. *Journal of Molecular Biology,* 427**,** 94-105.

Piras, G., El Kharroubi, A., Kozlov, S., Escalante-Alcalde, D., Hernandez, L., Copeland, N. G., Gilbert, D. J., Jenkins, N. A. & Stewart, C. L. 2000. Zac1 (Lot1), a potential tumor suppressor gene, and the gene for epsilon-sarcoglycan are maternally imprinted genes: identification by a subtractive screen of novel uniparental fibroblast lines. *Molecular & Cellular Biology,* 20**,** 3308-15.

Pizoli, C. E., Jinnah, H. A., Billingsley, M. L. & Hess, E. J. 2002. Abnormal cerebellar signaling induces dystonia in mice. *Journal of Neuroscience,* 22**,** 7825-33.

Poisson, A., Krack, P., Thobois, S., Loiraud, C., Serra, G., Vial, C. & Broussolle, E. 2012. History of the 'geste antagoniste' sign in cervical dystonia. *Journal of neurology,* 259**,** 1580-4.

Politano, L., Nigro, V., Passamano, L., Petretta, V., Comi, L. I., Papparella, S., Nigro, G., Rambaldi, P. F., Raia, P., Pini, A., Mora, M., Giugliano, M. A., Esposito, M. G. & Nigro, G. 2001. Evaluation of cardiac and respiratory involvement in sarcoglycanopathies. *Neuromuscular Disorders,* 11**,** 178-85.

Popa, T., Milani, P., Richard, A., Hubsch, C., Brochard, V., Tranchant, C., Sadnicka, A., Rothwell, J., Vidailhet, M., Meunier, S. & Roze, E. 2014. The neurophysiological features of myoclonus-dystonia and differentiation from other dystonias. *JAMA Neurology,* 71**,** 612-9.

Prins, K. W., Humston, J. L., Mehta, A., Tate, V., Ralston, E. & Ervasti, J. M. 2009. Dystrophin is a microtubule-associated protein. *The Journal of Cell Biology,* 186**,** 363-369.

Probstmeier, R., Nellen, J., Gloor, S., Wernig, A. & Pesheva, P. 2001. Tenascin-R is expressed by Schwann cells in the peripheral nervous system. *Journal of Neuroscience Research,* 64**,** 70-8.

Prudente, C. N., Pardo, C. A., Xiao, J., Hanfelt, J., Hess, E. J., Ledoux, M. S. & Jinnah, H. A. 2013. Neuropathology of cervical dystonia. *Experimental Neurology,* 241**,** 95-104.

Quadri, M., Olgiati, S., Sensi, M., Gualandi, F., Groppo, E., Rispoli, V., Graafland, J., Breedveld, G. J., Fabbrini, G., Berardelli, A. & Bonifati, V. 2016. PRKRA Mutation Causing Early-Onset Generalized Dystonia-Parkinsonism (DYT16) in an Italian Family. *Movement Disorders,* 31**,** 765-7.

Quartarone, A., Bagnato, S., Rizzo, V., Morgante, F., Sant'angelo, A., Crupi, D., Romano, M., Messina, C., Berardelli, A. & Girlanda, P. 2005. Corticospinal excitability during motor imagery of a simple tonic finger movement in patients with writer's cramp. *Movement Disorders,* 20**,** 1488-95.

Quartarone, A., Bagnato, S., Rizzo, V., Siebner, H. R., Dattola, V., Scalfari, A., Morgante, F., Battaglia, F., Romano, M. & Girlanda, P. 2003. Abnormal associative plasticity of the human motor cortex in writer's cramp. *Brain,* 126**,** 2586-2596.

Quartarone, A. & Hallett, M. 2013. Emerging concepts in the physiological basis of dystonia. *Movement Disorders,* 28**,** 958-967.

Quartarone, A., Morgante, F., Sant'angelo, A., Rizzo, V., Bagnato, S., Terranova, C., Siebner, H. R., Berardelli, A. & Girlanda, P. 2008. Abnormal plasticity of sensorimotor

circuits extends beyond the affected body part in focal dystonia. *journal of Neurology, Neurosurgery & Psychiatry,* 79**,** 985-90.

Quinn, N. P. 1996. Essential myoclonus and myoclonus dystonia. *Movement Disorders,* 11**,** 119-124.

Rafii, M. S., Hagiwara, H., Mercado, M. L., Seo, N. S., Xu, T., Dugan, T., Owens, R. T., Hook, M., Mcquillan, D. J., Young, M. F. & Fallon, J. R. 2006. Biglycan binds to alpha- and gamma-sarcoglycan and regulates their expression during development. *Journal of Cellular Physiology,* 209**,** 439-47.

Raike, R. S., Pizoli, C. E., Weisz, C., Van Den Maagdenberg, A. M. J. M., Jinnah, H. A. & Hess, E. J. 2013. Limited regional cerebellar dysfunction induces focal dystonia in mice. *Neurobiology of Disease,* 49**,** 200-210.

Rainier, S., Thomas, D., Tokarz, D., Ming, L., Bui, M., Plein, E., Zhao, X., Lemons, R., Albin, R., Delaney, C., Alvarado, D. & Fink, J. K. 2004. Myofibrillogenesis regulator 1 gene mutations cause paroxysmal dystonic choreoathetosis. *Archives of Neurology,* 61**,** 1025-9.

Ramos, V. F., Karp, B. I. & Hallett, M. 2014. Tricks in dystonia: ordering the complexity. *Journal of Neurology, Neurosurgery & Psychiatry,* 85**,** 987-93.

Raymond, D., Saunders-Pullman, R., De Carvalho Aguiar, P., Schule, B., Kock, N., Friedman, J., Harris, J., Ford, B., Frucht, S., Heiman, G. A., Jennings, D., Doheny, D., Brin, M. F., De Leon Brin, D., Multhaupt-Buell, T., Lang, A. E., Kurlan, R., Klein, C., Ozelius, L. & Bressman, S. 2008. Phenotypic spectrum and sex effects in eleven myoclonus-dystonia families with epsilon-sarcoglycan mutations. *Movement Disorders,* 23**,** 588-92.

Rebollo, R., Romanish, M. T. & Mager, D. L. 2012. Transposable Elements: An Abundant and Natural Source of Regulatory Sequences for Host Genes. *Annual Review of Genetics,* 46**,** 21-42.

Ridding, M. C., Sheean, G., Rothwell, J. C., Inzelberg, R. & Kujirai, T. 1995. Changes in the balance between motor cortical excitation and inhibition in focal, task specific dystonia. *J Neurol Neurosurg Psychiatry,* 59**,** 493-8.

Risch, N. J., Bressman, S. B., Deleon, D., Brin, M. F., Burke, R. E., Greene, P. E., Shale, H., Claus, E. B., Cupples, L. A. & Fahn, S. 1990. Segregation analysis of idiopathic torsion dystonia in Ashkenazi Jews suggests autosomal dominant inheritance. *American journal of human genetics,* 46**,** 533-8.

Risch, N. J., Bressman, S. B., Senthil, G. & Ozelius, L. J. 2007. Intragenic Cis and Trans modification of genetic susceptibility in DYT1 torsion dystonia. *American journal of human genetics,* 80**,** 1188-93.

Ritz, K., Gerrits, M. C., Foncke, E. M., Van Ruissen, F., Van Der Linden, C., Vergouwen, M. D., Bloem, B. R., Vandenberghe, W., Crols, R., Speelman, J. D., Baas, F. & Tijssen, M. A. 2009. Myoclonus-dystonia: clinical and genetic evaluation of a large cohort. *Journal of Neurology, Neurosurgery & Psychiatry,* 80**,** 653-8.

Ritz, K., Van Schaik, B. D., Jakobs, M. E., Van Kampen, A. H., Aronica, E., Tijssen, M. A. & Baas, F. 2011. SGCE isoform characterization and expression in human brain: implications for myoclonus-dystonia pathogenesis? *European Journal of Human Genetics,* 19**,** 438-44.

Roberds, S. L., Anderson, R. D., Ibraghimov-Beskrovnaya, O. & Campbell, K. P. 1993a. Primary structure and muscle-specific expression of the 50-kDa dystrophin-associated glycoprotein (adhalin). *Journal of Biological Chemistry,* 268**,** 23739-42.

Roberds, S. L., Ervasti, J. M., Anderson, R. D., Ohlendieck, K., Kahl, S. D., Zoloto, D. & Campbell, K. P. 1993b. Disruption of the dystrophin-glycoprotein complex in the cardiomyopathic hamster. *Journal of Biological Chemistry,* 268**,** 11496-9.

Rodacker, V., Toustrup-Jensen, M. & Vilsen, B. 2006. Mutations Phe785Leu and Thr618Met in Na+,K+-ATPase, Associated with Familial Rapid-onset Dystonia Parkinsonism, Interfere with Na+ Interaction by Distinct Mechanisms. *Journal of Biological Chemistry,* 281**,** 18539-18548.

Romaniello, R., Arrigoni, F., Bassi, M. T. & Borgatti, R. 2015. Mutations in alpha- and beta-tubulin encoding genes: implications in brain malformations. *Brain and Development,* 37**,** 273-80.

Rona, S., Berardelli, A., Vacca, L., Inghilleri, M. & Manfredi, M. 1998. Alterations of motor cortical inhibition in patients with dystonia. *Movement Disorders,* 13**,** 118-124.

Rosales, R. L. 2010. X-linked dystonia parkinsonism: clinical phenotype, genetics and therapeutics. *Journal of Movement Disorders,* 3**,** 32-8.

Rosenbloom, K. R., Armstrong, J., Barber, G. P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T. R., Fujita, P. A., Guruvadoo, L., Haeussler, M., Harte, R. A., Heitner, S., Hickey, G., Hinrichs, A. S., Hubley, R., Karolchik, D., Learned, K., Lee, B. T., Li, C. H., Miga, K. H., Nguyen, N., Paten, B., Raney, B. J., Smit, A. F., Speir, M. L., Zweig, A. S., Haussler, D., Kuhn, R. M. & Kent, W. J. 2015. The UCSC Genome Browser database: 2015 update. *Nucleic Acids Research,* 43**,** D670-81.

Rosenbloom, K. R., Sloan, C. A., Malladi, V. S., Dreszer, T. R., Learned, K., Kirkup, V. M., Wong, M. C., Maddren, M., Fang, R., Heitner, S. G., Lee, B. T., Barber, G. P., Harte, R. A., Diekhans, M., Long, J. C., Wilder, S. P., Zweig, A. S., Karolchik, D., Kuhn, R. M., Haussler, D. & Kent, W. J. 2013. ENCODE Data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Research,* 41**,** D56-D63.

Roussigne, M., Kossida, S., Lavigne, A. C., Clouaire, T., Ecochard, V., Glories, A., Amalric, F. & Girard, J. P. 2003. The THAP domain: a novel protein motif with similarity to the DNA-binding domain of P element transposase. *Trends in Biochemical Sciences,* 28**,** 66-9.

Roze, E., Apartis, E., Clot, F., Dorison, N., Thobois, S., Guyant-Marechal, L., Tranchant, C., Damier, P., Doummar, D., Bahi-Buisson, N., Andre-Obadia, N., Maltete, D., Echaniz-Laguna, A., Pereon, Y., Beaugendre, Y., Dupont, S., De Greslan, T., Jedynak, C. P., Ponsot, G., Dussaule, J. C., Brice, A., Durr, A. & Vidailhet, M. 2008. Myoclonus-dystonia: clinical and electrophysiologic pattern related to SGCE mutations. *Neurology,* 70**,** 1010-6.

Rozen, S. & Skaletsky, H. 1999. Primer3 on the WWW for General Users and for Biologist Programmers. *In:* MISENER, S. & KRAWETZ, S. (eds.) *Bioinformatics Methods and Protocols.* Humana Press.

Sabogal, A., Lyubimov, A. Y., Corn, J. E., Berger, J. M. & Rio, D. C. 2010. THAP proteins target specific DNA sites through bipartite recognition of adjacent major and minor grooves. *Nature Structural & Molecular Biology,* 17**,** 117-23.

Sadnicka, A., Hamada, M., Bhatia, K. P., Rothwell, J. C. & Edwards, M. J. 2014. A reflection on plasticity research in writing dystonia. *Movement Disorders,* 29**,** 980-7.

Saitoh, S., Takamatsu, K., Kobayashi, M. & Noguchi, T. 1993. Distribution of hippocalcin mRNA and immunoreactivity in rat brain. *Neuroscience Letters,* 157**,** 107-110.

Sako, W., Morigaki, R., Kaji, R., Tooyama, I., Okita, S., Kitazato, K., Nagahiro, S., Graybiel, A. M. & Goto, S. 2011. Identification and localization of a neuron-specific isoform of TAF1 in rat brain: implications for neuropathology of DYT3 dystonia. *Neuroscience,* 189**,** 100-7.

Salamov, A. A., Nishikawa, T. & Swindells, M. B. 1998. Assessing protein coding region integrity in cDNA sequencing projects. *Bioinformatics,* 14**,** 384-90.

Sandberg, R., Yasuda, R., Pankratz, D. G., Carter, T. A., Del Rio, J. A., Wodicka, L., Mayford, M., Lockhart, D. J. & Barlow, C. 2000. Regional and strain-specific gene

expression mapping in the adult mouse brain. *Proceedings of the National Academy of Sciences,* 97**,** 11038-11043.

Sandona, D. & Betto, R. 2009. Sarcoglycanopathies: molecular pathogenesis and therapeutic prospects. *Expert reviews in molecular medicine,* 11**,** e28.

Sandona, D., Gastaldello, S., Martinello, T. & Betto, R. 2004. Characterization of the ATP-hydrolysing activity of alpha-sarcoglycan. *The Biochemical journal,* 381**,** 105-12.

Saunders-Pullman, R., Md, M. P. H., Shriberg, J., Ms, M. P. H., Heiman, G., Raymond, D., Wendt, K., Kramer, P., Schilling, K., Kurlan, R., Klein, C., Ozelius, L., Risch, N. & Bressman, S. 2002. Myoclonus dystonia: Possible association with obsessive-compulsive disorder and alcohol dependence. *Neurology,* 58**,** 242-245.

Schneider, S. A., Paisan-Ruiz, C., Garcia-Gorostiaga, I., Quinn, N. P., Weber, Y. G., Lerche, H., Hardy, J. & Bhatia, K. P. 2009. GLUT1 gene mutations cause sporadic paroxysmal exercise-induced dyskinesias. *Movement Disorders,* 24**,** 1684-8.

Schule, B., Kock, N., Svetel, M., Dragasevic, N., Hedrich, K., De Carvalho Aguiar, P., Liu, L., Kabakci, K., Garrels, J., Meyer, E. M., Berisavac, I., Schwinger, E., Kramer, P. L., Ozelius, L. J., Klein, C. & Kostic, V. 2004. Genetic heterogeneity in ten families with myoclonus-dystonia. *Journal of Neurology, Neurosurgery & Psychiatry,* 75**,** 1181-5.

Scontrini, A., Conte, A., Defazio, G., Fiorio, M., Fabbrini, G., Suppa, A., Tinazzi, M. & Berardelli, A. 2009. Somatosensory temporal discrimination in patients with primary focal dystonia. *Journal of Neurology, Neurosurgery & Psychiatry,* 80**,** 1315-1319.

Sela, N., Mersch, B., Hotz-Wagenblatt, A. & Ast, G. 2010. Characteristics of transposable element exonization within human and mouse. *PLoS One,* 5**,** e10907.

Sengel, C., Gavarini, S., Sharma, N., Ozelius, L. J. & Bragg, D. C. 2011. Dimerization of the DYT6 dystonia protein, THAP1, requires residues within the coiled-coil domain. *Journal of Neurochemistry,* 118**,** 1087-100.

Shashidharan, P., Kramer, B. C., Walker, R. H., Olanow, C. W. & Brin, M. F. 2000. Immunohistochemical localization and distribution of torsinA in normal human and rat brain. *Brain Research,* 853**,** 197-206.

Shen, S., Lin, L., Cai, J. J., Jiang, P., Kenkel, E. J., Stroik, M. R., Sato, S., Davidson, B. L. & Xing, Y. 2011a. Widespread establishment and regulatory impact of Alu exons in human genes. *Proceedings of the National Academy of Sciences,* 108**,** 2837-2842.

Shen, Y., Ge, W. P., Li, Y., Hirano, A., Lee, H. Y., Rohlmann, A., Missler, M., Tsien, R. W., Jan, L. Y., Fu, Y. H. & Ptacek, L. J. 2015. Protein mutated in paroxysmal dyskinesia interacts with the active zone protein RIM and suppresses synaptic vesicle exocytosis. *Proceedings of the National Academy of Sciences,* 112**,** 2935-41.

Shen, Y., Lee, H. Y., Rawson, J., Ojha, S., Babbitt, P., Fu, Y. H. & Ptacek, L. J. 2011b. Mutations in PNKD causing paroxysmal dyskinesia alters protein cleavage and stability. *Human Molecular Genetics,* 20**,** 2322-32.

Shi, W., Chen, Z., Schottenfeld, J., Stahl, R. C., Kunkel, L. M. & Chan, Y.-M. 2004. Specific assembly pathway of sarcoglycans is dependent on beta- and delta-sarcoglycan. *Muscle & Nerve,* 29**,** 409-419.

Shiga, K., Yoshioka, H., Matsumiya, T., Kimura, I., Takeda, S. & Imamura, M. 2006. Zeta-sarcoglycan is a functional homologue of gamma-sarcoglycan in the formation of the sarcoglycan complex. *Experimental Cell Research,* 312**,** 2083-92.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., Mcwilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology,* 7.

Silveira-Moriyama, L., Gardiner, A. R., Meyer, E., King, M. D., Smith, M., Rakshi, K., Parker, A., Mallick, A. A., Brown, R., Vassallo, G., Jardine, P. E., Guerreiro, M. M.,

Lees, A. J., Houlden, H. & Kurian, M. A. 2013. Clinical features of childhood-onset paroxysmal kinesigenic dyskinesia with PRRT2 gene mutations. *Developmental Medicine & Child Neurology,* 55**,** 327-34.

Singh, M., Castillo, D., Patel, C. V. & Patel, R. C. 2011. Stress-induced phosphorylation of PACT reduces its interaction with TRBP and leads to PKR activation. *Biochemistry,* 50**,** 4550-60.

Sitburana, O., Wu, L. J., Sheffield, J. K., Davidson, A. & Jankovic, J. 2009. Motor overflow and mirror dystonia. *Parkinsonism Relat Disord,* 15**,** 758-61.

Skandalis, A., Frampton, M., Seger, J. & Richards, M. H. 2010. The adaptive significance of unproductive alternative splicing in primates. *RNA,* 16**,** 2014-2022.

Smit, A. F. A., Hubley, R. & Green, P. 2013-2015. *RepeatMasker Open-4.0* [Online]. Available: http://www.repeatmasker.org [Accessed 2016].

Snow, W. M., Anderson, J. E. & Jakobson, L. S. 2013. Neuropsychological and neurobehavioral functioning in Duchenne muscular dystrophy: A review. *Neuroscience & Biobehavioral Reviews,* 37**,** 743-752.

Soheili, T., Gicquel, E., Poupiot, J., N'guyen, L., Le Roy, F., Bartoli, M. & Richard, I. 2012. Rescue of sarcoglycan mutations by inhibition of endoplasmic reticulum quality control is associated with minimal structural modifications. *Human mutation,* 33**,** 429-39.

Solares-Perez, A., Alvarez, R., Crosbie, R. H., Vega-Moreno, J., Medina-Monares, J., Estrada, F. J., Ortega, A. & Coral-Vazquez, R. 2010a. Altered calcium pump and secondary deficiency of gamma-sarcoglycan and microspan in sarcoplasmic reticulum membranes isolated from delta-sarcoglycan knockout mice. *Cell calcium,* 48**,** 28-36.

Solares-Perez, A., Sanchez, J. A., Zentella-Dehesa, A., Garcia, M. C. & Coral-Vazquez, R. M. 2010b. Intracellular Ca2+ transients in delta-sarcoglycan knockout mouse skeletal muscle. *Biochimica et biophysica acta,* 1800**,** 373-9.

Sorek, R., Ast, G. & Graur, D. 2002. Alu-Containing Exons are Alternatively Spliced. *Genome Research,* 12**,** 1060-1067.

Sosa, B. A., Demircioglu, F. E., Chen, J. Z., Ingram, J., Ploegh, H. L. & Schwartz, T. U. 2014. How lamina-associated polypeptide 1 (LAP1) activates Torsin. *Elife,* 3**,** e03239.

Sotgia, F., Lee, H., Bedford, M. T., Petrucci, T., Sudol, M. & Lisanti, M. P. 2001. Tyrosine Phosphorylation of β-Dystroglycan at Its WW Domain Binding Motif, PPxY, Recruits SH2 Domain Containing Proteins. *Biochemistry,* 40**,** 14585-14592.

Spinazzola, J. M., Smith, T. C., Liu, M., Luna, E. J. & Barton, E. R. 2015. Gamma-sarcoglycan is required for the response of archvillin to mechanical stimulation in skeletal muscle. *Human Molecular Genetics,* 24**,** 2470-81.

Stamelou, M., Charlesworth, G., Cordivari, C., Schneider, S. A., Kagi, G., Sheerin, U. M., Rubio-Agusti, I., Batla, A., Houlden, H., Wood, N. W. & Bhatia, K. P. 2014. The phenotypic spectrum of DYT24 due to ANO3 mutations. *Movement Disorders,* 29**,** 928-34.

Straub, V., Duclos, F., Venzke, D. P., Lee, J. C., Cutshall, S., Leveille, C. J. & Campbell, K. P. 1998. Molecular pathogenesis of muscle degeneration in the delta-sarcoglycan-deficient hamster. *The American journal of pathology,* 153**,** 1623-30.

Straub, V., Ettinger, A. J., Durbeej, M., Venzke, D. P., Cutshall, S., Sanes, J. R. & Campbell, K. P. 1999. epsilon-sarcoglycan replaces alpha-sarcoglycan in smooth muscle to form a unique dystrophin-glycoprotein complex. *The Journal of biological chemistry,* 274**,** 27989-27996.

Straub, V., Rafael, J. A., Chamberlain, J. S. & Campbell, K. P. 1997. Animal Models for Muscular Dystrophy Show Different Patterns of Sarcolemmal Disruption. *The Journal of Cell Biology,* 139**,** 375-385.

Sugimoto, H., Ikeda, K. & Kawakami, K. 2014. Heterozygous mice deficient in Atp1a3 exhibit motor deficits by chronic restraint stress. *Behavioural Brain Research,* 272**,** 100-10.

Suls, A., Dedeken, P., Goffin, K., Van Esch, H., Dupont, P., Cassiman, D., Kempfle, J., Wuttke, T. V., Weber, Y., Lerche, H., Afawi, Z., Vandenberghe, W., Korczyn, A. D., Berkovic, S. F., Ekstein, D., Kivity, S., Ryvlin, P., Claes, L. R., Deprez, L., Maljevic, S., Vargas, A., Van Dyck, T., Goossens, D., Del-Favero, J., Van Laere, K., De Jonghe, P. & Van Paesschen, W. 2008. Paroxysmal exercise-induced dyskinesia and epilepsy is due to mutations in SLC2A1, encoding the glucose transporter GLUT1. *Brain,* 131**,** 1831-44.

Suzuki, J., Fujii, T., Imao, T., Ishihara, K., Kuba, H. & Nagata, S. 2013. Calcium-dependent phospholipid scramblase activity of TMEM16 protein family members. *Journal of Biological Chemistry,* 288**,** 13305-16.

Suzuki, Y., Mizoguchi, S., Kiyosawa, M., Mochizuki, M., Ishiwata, K., Wakakura, M. & Ishii, K. 2007. Glucose hypermetabolism in the thalamus of patients with essential blepharospasm. *Journal of neurology,* 254**,** 890-6.

Sweney, M. T., Newcomb, T. M. & Swoboda, K. J. 2015. The expanding spectrum of neurological phenotypes in children with ATP1A3 mutations, Alternating Hemiplegia of Childhood, Rapid-onset Dystonia-Parkinsonism, CAPOS and beyond. *Pediatric neurology,* 52**,** 56-64.

Swiderski, K., Shaffer, S. A., Gallis, B., Odom, G. L., Arnett, A. L., Scott Edgar, J., Baum, D. M., Chee, A., Naim, T., Gregorevic, P., Murphy, K. T., Moody, J., Goodlett, D. R., Lynch, G. S. & Chamberlain, J. S. 2014. Phosphorylation within the cysteine-rich region of dystrophin enhances its association with β-dystroglycan and identifies a potential novel therapeutic target for skeletal muscle wasting. *Human Molecular Genetics,* 23**,** 6697-6711.

Tadayoni, R., Rendon, A., Soria-Jasso, L. E. & Cisneros, B. 2012. Dystrophin Dp71: the smallest but multifunctional product of the Duchenne muscular dystrophy gene. *Molecular Neurobiology,* 45**,** 43-60.

Tan, L. C., Methawasin, K., Teng, E. W., Ng, A. R., Seah, S. H., Au, W. L., Liu, J. J., Foo, J. N., Zhao, Y. & Tan, E. K. 2014. Clinico-genetic comparisons of paroxysmal kinesigenic dyskinesia patients with and without PRRT2 mutations. *European Journal of Neurology,* 21**,** 674-8.

Tezenas Du Montcel, S., Clot, F., Vidailhet, M., Roze, E., Damier, P., Jedynak, C. P., Camuzat, A., Lagueny, A., Vercueil, L., Doummar, D., Guyant-Marechal, L., Houeto, J. L., Ponsot, G., Thobois, S., Cournelle, M. A., Durr, A., Durif, F., Echenne, B., Hannequin, D., Tranchant, C., Brice, A. & French Dystonia, N. 2006. Epsilon sarcoglycan mutations and phenotype in French patients with myoclonic syndromes. *Journal of Medical Genetics,* 43**,** 394-400.

Thompson, T. G., Chan, Y.-M., Hack, A. A., Brosius, M., Rajala, M., Lidov, H. G. W., Mcnally, E. M., Watkins, S. & Kunkel, L. M. 2000. Filamin 2 (FLN2): A Muscle-specific Sarcoglycan Interacting Protein. *The Journal of Cell Biology,* 148**,** 115-126.

Tisch, S., Limousin, P., Rothwell, J. C., Asselman, P., Quinn, N., Jahanshahi, M., Bhatia, K. P. & Hariz, M. 2006a. Changes in blink reflex excitability after globus pallidus internus stimulation for dystonia. *Movement Disorders,* 21**,** 1650-1655.

Tisch, S., Limousin, P., Rothwell, J. C., Asselman, P., Zrinzo, L., Jahanshahi, M., Bhatia, K. P. & Hariz, M. I. 2006b. Changes in forearm reciprocal inhibition following pallidal stimulation for dystonia. *Neurology,* 66**,** 1091-3.

Tissir, F. & Goffinet, A. M. 2003. Reelin and brain development. *Nature Reviews Neuroscience,* 4**,** 496-505.

Torres, G. E., Sweeney, A. L., Beaulieu, J. M., Shashidharan, P. & Caron, M. G. 2004. Effect of torsinA on membrane proteins reveals a loss of function and a dominant-negative phenotype of the dystonia-associated DeltaE-torsinA mutant. *Proceedings of the National Academy of Sciences,* 101**,** 15650-5.

Toustrup-Jensen, M. S., Einholm, A. P., Schack, V. R., Nielsen, H. N., Holm, R., Sobrido, M. J., Andersen, J. P., Clausen, T. & Vilsen, B. 2014. Relationship between intracellular Na+ concentration and reduced Na+ affinity in Na+,K+-ATPase mutants causing neurological disease. *Journal of Biological Chemistry,* 289**,** 3186-97.

Trinkle-Mulcahy, L., Boulon, S., Lam, Y., Urcia, R., Boisvert, F.-M., Vandermoere, F., Morrice, N., Swift, S., Rothbauer, U., Leonhardt, H. & Lamond, A. 2008. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *The Journal of Cell Biology,* 183**,** 223-239.

Turk, R., T Hoen, P. A., Sterrenburg, E., De Menezes, R. X., De Meijer, E. J., Boer, J. M., Van Ommen, G. J. & Den Dunnen, J. T. 2004. Gene expression variation between mouse inbred strains. *BMC Genomics,* 5**,** 57.

Tzingounis, A. V., Kobayashi, M., Takamatsu, K. & Nicoll, R. A. 2007. Hippocalcin Gates the Calcium Activation of the Slow Afterhyperpolarization in Hippocampal Pyramidal Cells. *Neuron,* 53**,** 487-493.

Uitti, R. J. & Maraganore, D. M. 1993. Adult onset familial cervical dystonia: report of a family including monozygotic twins. *Movement Disorders,* 8**,** 489-94.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B., Remm, M. & Rozen, S. 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Research,* 40**,** e115-e115.

Vacher, H., Mohapatra, D. P. & Trimmer, J. S. 2008. Localization and targeting of voltage-dependent ion channels in mammalian central neurons. *Physiological Reviews,* 88**,** 1407-47.

Vainzof, M., Passos-Bueno, M. R., Canovas, M., Moreira, E. S., Pavanello, R. C., Marie, S. K., Anderson, L. V., Bonnemann, C. G., Mcnally, E. M., Nigro, V., Kunkel, L. M. & Zatz, M. 1996. The sarcoglycan complex in the six autosomal recessive limb-girdle muscular dystrophies. *Human Molecular Genetics,* 5**,** 1963-9.

Valente, P., Castroflorio, E., Rossi, P., Fadda, M., Sterlini, B., Cervigni, R. I., Prestigio, C., Giovedi, S., Onofri, F., Mura, E., Guarnieri, F. C., Marte, A., Orlando, M., Zara, F., Fassio, A., Valtorta, F., Baldelli, P., Corradi, A. & Benfenati, F. 2016. PRRT2 Is a Key Component of the Ca(2+)-Dependent Neurotransmitter Release Machinery. *Cell reports,* 15**,** 117-31.

Van Der Meer, J. N., Beukers, R. J., Van Der Salm, S. M., Caan, M. W., Tijssen, M. A. & Nederveen, A. J. 2012. White matter abnormalities in gene-positive myoclonus-dystonia. *Movement Disorders,* 27**,** 1666-72.

Van Der Salm, S. M., Van Der Meer, J. N., Nederveen, A. J., Veltman, D. J., Van Rootselaar, A. F. & Tijssen, M. A. 2013. Functional MRI study of response inhibition in myoclonus dystonia. *Experimental Neurology,* 247**,** 623-9.

Van Der Salm, S. M., Van Rootselaar, A. F., Foncke, E. M., Koelman, J. H., Bour, L. J., Bhatia, K. P., Rothwell, J. C. & Tijssen, M. A. 2009. Normal cortical excitability in Myoclonus-Dystonia--a TMS study. *Experimental Neurology,* 216**,** 300-5.

Van Tricht, M. J., Dreissen, Y. E., Cath, D., Dijk, J. M., Contarino, M. F., Van Der Salm, S. M., Foncke, E. M., Groen, J. L., Schmand, B. & Tijssen, M. A. 2012. Cognition and psychopathology in myoclonus-dystonia. *Journal of Neurology, Neurosurgery & Psychiatry,* 83**,** 814-20.

Vander Heyden, A. B., Naismith, T. V., Snapp, E. L., Hodzic, D. & Hanson, P. I. 2009. LULL1 retargets TorsinA to the nuclear envelope revealing an activity that is

impaired by the DYT1 dystonia mutation. *Molecular & Cellular Biology,* 20**,** 2661-72.

Vaughn, L. S., Bragg, D. C., Sharma, N., Camargos, S., Cardoso, F. & Patel, R. C. 2015. Altered activation of protein kinase PKR and enhanced apoptosis in dystonia cells carrying a mutation in PKR activator protein PACT. *Journal of Biological Chemistry,* 290**,** 22543-57.

Vemula, S. R., Puschmann, A., Xiao, J., Zhao, Y., Rudzinska, M., Frei, K. P., Truong, D. D., Wszolek, Z. K. & Ledoux, M. S. 2013. Role of Galpha(olf) in familial and sporadic adult-onset primary dystonia. *Human Molecular Genetics,* 22**,** 2510-9.

Vidailhet, M., Tassin, J., Durif, F., Nivelon-Chevallier, A., Agid, Y., Brice, A. & Durr, A. 2001. A major locus for several phenotypes of myoclonus-dystonia on chromosome 7q. *Neurology,* 56**,** 1213-1216.

Vulinovic, F., Lohmann, K., Rakovic, A., Capetian, P., Alvarez-Fischer, D., Schmidt, A., Weissbach, A., Erogullari, A., Kaiser, F. J., Wiegers, K., Ferbert, A., Rolfs, A., Klein, C. & Seibler, P. 2014. Unraveling cellular phenotypes of novel TorsinA/TOR1A mutations. *Human Mutation,* 35**,** 1114-22.

Waite, A. 2009. The molecular genetics of myoclonus-dystonia syndrome. *Oxford University,* PhD Thesis.

Waite, A., Brown, S. C. & Blake, D. J. 2012. The dystrophin-glycoprotein complex in brain development and disease. *Trends in Neurosciences,* 35**,** 487-96.

Waite, A., De Rosa, M. C., Brancaccio, A. & Blake, D. J. 2011. A gain-of-glycosylation mutation associated with myoclonus-dystonia syndrome affects trafficking and processing of mouse epsilon-sarcoglycan in the late secretory pathway. *Human Mutation,* 32**,** 1246-58.

Waite, A., Tinsley, C. L., Locke, M. & Blake, D. J. 2009. The neurobiology of the dystrophin-associated glycoprotein complex. [Review] [128 refs]. *Annals of medicine,* 41**,** 344-59.

Waite, A. J., Carlisle, F. A., Chan, Y. M. & Blake, D. J. 2016. Myoclonus dystonia and muscular dystrophy: epsilon-sarcoglycan is part of the dystrophin-associated protein complex in brain. *Movement Disorders*.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P. & Burge, C. B. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature,* 456**,** 470-6.

Wang, J. L., Cao, L., Li, X. H., Hu, Z. M., Li, J. D., Zhang, J. G., Liang, Y., San, A., Li, N., Chen, S. Q., Guo, J. F., Jiang, H., Shen, L., Zheng, L., Mao, X., Yan, W. Q., Zhou, Y., Shi, Y. T., Ai, S. X., Dai, M. Z., Zhang, P., Xia, K., Chen, S. D. & Tang, B. S. 2011. Identification of PRRT2 as the causative gene of paroxysmal kinesigenic dyskinesias. *Brain,* 134**,** 3493-3501.

Warder, D. E. & Keherly, M. J. 2003. Ciz1, Cip1 interacting zinc finger protein 1 binds the consensus DNA sequence ARYSR(0-2)YYAC. *Journal of Biomedical Science,* 10**,** 406-17.

Waters, C. H., Faust, P. L., Powers, J., Vinters, H., Moskowitz, C., Nygaard, T., Hunt, A. L. & Fahn, S. 1993. Neuropathology of lubag (x-linked dystonia parkinsonism). *Movement Disorders,* 8**,** 387-390.

Way, M., Pope, B., Cross, R. A., Kendrick-Jones, J. & Weeds, A. G. 1992. Expression of the N-terminal domain of dystrophin in E. coli and demonstration of binding to F-actin. *FEBS Letters,* 301**,** 243-245.

Weber, P., Bartsch, U., Rasband, M. N., Czaniera, R., Lang, Y., Bluethmann, H., Margolis, R. U., Levinson, S. R., Shrager, P., Montag, D. & Schachner, M. 1999. Mice Deficient for Tenascin-R Display Alterations of the Extracellular Matrix and

Decreased Axonal Conduction Velocities in the CNS. *The Journal of Neuroscience,* 19**,** 4245-4262.

Weber, Y. G., Kamm, C., Suls, A., Kempfle, J., Kotschet, K., Schule, R., Wuttke, T. V., Maljevic, S., Liebrich, J., Gasser, T., Ludolph, A. C., Van Paesschen, W., Schols, L., De Jonghe, P., Auburger, G. & Lerche, H. 2011. Paroxysmal choreoathetosis/spasticity (DYT9) is caused by a GLUT1 defect. *Neurology,* 77**,** 959-64.

Weber, Y. G., Storch, A., Wuttke, T. V., Brockmann, K., Kempfle, J., Maljevic, S., Margari, L., Kamm, C., Schneider, S. A., Huber, S. M., Pekrun, A., Roebling, R., Seebohm, G., Koka, S., Lang, C., Kraft, E., Blazevic, D., Salvo-Vargas, A., Fauler, M., Mottaghy, F. M., Munchau, A., Edwards, M. J., Presicci, A., Margari, F., Gasser, T., Lang, F., Bhatia, K. P., Lehmann-Horn, F. & Lerche, H. 2008. GLUT1 mutations are a cause of paroxysmal exertion-induced dyskinesias and induce hemolytic anemia by a cation leak. *Journal of Clinical Investigation,* 118**,** 2157-68.

Weise, D., Schramm, A., Beck, M., Reiners, K. & Classen, J. 2011. Loss of topographic specificity of LTD-like plasticity is a trait marker in focal dystonia. *Neurobiology of Disease,* 42**,** 171-176.

Weise, D., Schramm, A., Stefan, K., Wolters, A., Reiners, K., Naumann, M. & Classen, J. 2006. The two sides of associative plasticity in writer's cramp. *Brain,* 129**,** 2709-2721.

Weissbach, A., Kasten, M., Grunewald, A., Bruggemann, N., Trillenberg, P., Klein, C. & Hagenah, J. 2013. Prominent psychiatric comorbidity in the dominantly inherited movement disorder myoclonus-dystonia. *Parkinsonism Relat Disord,* 19**,** 422-5.

Wheeler, M. T., Zarnegar, S. & Mcnally, E. M. 2002. zeta-Sarcoglycan, a novel component of the sarcoglycan complex, is reduced in muscular dystrophy. *Human Molecular Genetics,* 11**,** 2147-2154.

Wijemanne, S. & Jankovic, J. 2015. Dopa-responsive dystonia--clinical and genetic heterogeneity. *Nature reviews neurology,* 11**,** 414-24.

Wilcox, R. A., Winkler, S., Lohmann, K. & Klein, C. 2011. Whispering dysphonia in an Australian family (DYT4): a clinical and genetic reappraisal. *Movement Disorders,* 26**,** 2404-8.

Wilson, B. K. & Hess, E. J. 2013. Animal models for dystonia. *Movement Disorders,* 28**,** 982-989.

Wu, L., Tang, H. D., Huang, X. J., Zheng, L., Liu, X. L., Wang, T., Wang, J. Y., Cao, L. & Chen, S. D. 2014. PRRT2 truncated mutations lead to nonsense-mediated mRNA decay in Paroxysmal Kinesigenic Dyskinesia. *Parkinsonism and Related Disorders,* 20**,** 1399-404.

Xiao, J. & Ledoux, M. S. 2003. Cloning, developmental regulation and neural localization of rat epsilon-sarcoglycan. *Brain research. Molecular brain research,* 119**,** 132-43.

Xiao, J., Uitti, R. J., Zhao, Y., Vemula, S. R., Perlmutter, J. S., Wszolek, Z. K., Maraganore, D. M., Auburger, G., Leube, B., Lehnhoff, K. & Ledoux, M. S. 2012. Mutations in CIZ1 cause adult onset primary cervical dystonia. *Annals of Neurology,* 71**,** 458-469.

Xiao, J., Vemula, S. R., Xue, Y., Khan, M. M., Carlisle, F. A., Waite, A. J., Blake, D. J., Dragatsis, I., Zhao, Y. & Ledoux, M. S. 2017. Role of major and brain-specific Sgce isoforms in the pathogenesis of myoclonus-dystonia syndrome. *Neurobiology of Disease,* 98**,** 52-65.

Xiao, Z. C., Taylor, J., Montag, D., Rougon, G. & Schachner, M. 1996. Distinct effects of recombinant tenascin-R domains in neuronal cell functions and identification of the domain interacting with the neuronal recognition molecule F3/11. *European Journal of Neuroscience,* 8**,** 766-82.

Xiromerisiou, G., Houlden, H., Scarmeas, N., Stamelou, M., Kara, E., Hardy, J., Lees, A. J., Korlipara, P., Limousin, P., Paudel, R., Hadjigeorgiou, G. M. & Bhatia, K. P. 2012. THAP1 mutations and dystonia phenotypes: genotype phenotype correlations. *Movement Disorders,* 27**,** 1290-4.

Yamamoto, H., Mizuno, Y., Hayashi, K., Nonaka, I., Yoshida, M. & Ozawa, E. 1994. Expression of dystrophin-associated protein 35DAG (A4) and 50DAG (A2) is confined to striated muscles. *Journal of Biochemistry,* 115**,** 162-7.

Yates, J. R., 3rd, Eng, J. K., Mccormack, A. L. & Schieltz, D. 1995. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Analytical Chemistry,* 67**,** 1426-36.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. & Madden, T. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics,* 13**,** 134.

Yokoi, F., Dang, M. T., Li, J. & Li, Y. 2006. Myoclonus, motor deficits, alterations in emotional responses and monoamine metabolism in epsilon-sarcoglycan deficient mice. *Journal of biochemistry,* 140**,** 141-146.

Yokoi, F., Dang, M. T., Mitsui, S. & Li, Y. 2005. Exclusive paternal expression and novel alternatively spliced variants of epsilon-sarcoglycan mRNA in mouse brain. *FEBS Letters,* 579**,** 4822-8.

Yokoi, F., Dang, M. T., Yang, G., Li, J., Doroodchi, A., Zhou, T. & Li, Y. 2012a. Abnormal nuclear envelope in the cerebellar Purkinje cells and impaired motor learning in DYT11 myoclonus-dystonia mouse models. *Behavioural Brain Research,* 227**,** 12-20.

Yokoi, F., Dang, M. T., Zhou, T. & Li, Y. 2012b. Abnormal nuclear envelopes in the striatum and motor deficits in DYT11 myoclonus-dystonia mouse models. *Human Molecular Genetics,* 21**,** 916-925.

Yoon, J. H., Johnson, E., Xu, R., Martin, L. T., Martin, P. T. & Montanaro, F. 2012. Comparative Proteomic Profiling of Dystroglycan-Associated Proteins in Wild Type, mdx, and Galgt2 Transgenic Mouse Skeletal Muscle. *Journal of Proteome Research,* 11**,** 4413-4424.

Yoshida, M., Hama, H., Ishikawa-Sakurai, M., Imamura, M., Mizuno, Y., Araishi, K., Wakabayashi-Takai, E., Noguchi, S., Sasaoka, T. & Ozawa, E. 2000. Biochemical evidence for association of dystrobrevin with the sarcoglycan-sarcospan complex as a basis for understanding sarcoglycanopathy. *Human molecular genetics,* 9**,** 1033-1040.

Yoshida, M., Noguchi, S., Wakabayashi, E., Piluso, G., Belsito, A., Nigro, V. & Ozawa, E. 1997. The fourth component of the sarcoglycan complex. *FEBS Letters,* 403**,** 143-148.

Yoshida, M., Suzuki, A., Yamamoto, H., Noguchi, S., Mizuno, Y. & Ozawa, E. 1994. Dissociation of the complex of dystrophin and its associated proteins into several unique groups by n-octyl β-d-glucoside. *European Journal of Biochemistry,* 222**,** 1055-1061.

Yoshida, T., Pan, Y., Hanada, H., Iwata, Y. & Shigekawa, M. 1998. Bidirectional Signaling between Sarcoglycans and the Integrin Adhesion System in Cultured L6 Myocytes. *Journal of Biological Chemistry,* 273**,** 1583-1590.

Zaki, M., Shehab, M., El-Aleem, A. A., Abdel-Salam, G., Koeller, H. B., Ilkin, Y., Ross, M. E., Dobyns, W. B. & Gleeson, J. G. 2007. Identification of a novel recessive RELN mutation using a homozygous balanced reciprocal translocation. *American Journal of Medical Genetics Part A,* 143a**,** 939-44.

Zech, M., Castrop, F., Schormair, B., Jochim, A., Wieland, T., Gross, N., Lichtner, P., Peters, A., Gieger, C., Meitinger, T., Strom, T. M., Oexle, K., Haslinger, B. & Winkelmann,

J. 2014a. DYT16 revisited: exome sequencing identifies PRKRA mutations in a European dystonia family. *Movement Disorders,* 29**,** 1504-10.

Zech, M., Gross, N., Jochim, A., Castrop, F., Kaffe, M., Dresel, C., Lichtner, P., Peters, A., Gieger, C., Meitinger, T., Haslinger, B. & Winkelmann, J. 2014b. Rare sequence variants in ANO3 and GNAL in a primary torsion dystonia series and controls. *Movement Disorders,* 29**,** 143-7.

Zhang, L., Yokoi, F., Parsons, D. S., Standaert, D. G. & Li, Y. 2012. Alteration of striatal dopaminergic neurotransmission in a mouse model of DYT11 myoclonus-dystonia. *PLoS One,* 7**,** e33669.

Zhao, C., Brown, R. S., Chase, A. R., Eisele, M. R. & Schlieker, C. 2013a. Regulation of Torsin ATPases by LAP1 and LULL1. *Proceedings of the National Academy of Sciences,* 110**,** E1545-54.

Zhao, Y., Xiao, J., Gong, S., Clara, J. A. & Ledoux, M. S. 2013b. Neural expression of the transcription factor THAP1 during development in rat. *Neuroscience,* 231**,** 282-95.

Zheng, S. & Black, D. L. 2013. Alternative pre-mRNA splicing in neurons: growing up and extending its reach. *Trends in Genetics,* 29**,** 442-448.

Zhu, X., Hadhazy, M., Groh, M. E., Wheeler, M. T., Wollmann, R. & Mcnally, E. M. 2001. Overexpression of gamma-sarcoglycan induces severe muscular dystrophy. Implications for the regulation of Sarcoglycan assembly. *The Journal of biological chemistry,* 276**,** 21785-90.

Zhuang, X., Belluscio, L. & Hen, R. 2000. G(olf)alpha mediates dopamine D1 receptor signaling. *Journal of Neuroscience,* 20**,** Rc91.

Zimprich, A., Grabowski, M., Asmus, F., Naumann, M., Berg, D., Bertram, M., Scheidtmann, K., Kern, P., Winkelmann, J., Muller-Myhsok, B., Riedel, L., Bauer, M., Muller, T., Castro, M., Meitinger, T., Strom, T. M. & Gasser, T. 2001. Mutations in the gene encoding epsilon-sarcoglycan cause myoclonus-dystonia syndrome. *Nature Genetics,* 29**,** 66-9.

Appendix I: Oligonucleotide primers

All primer sequences are written 5' to 3', and were obtained from Sigma-Aldrich.

Species	Gene	Exon	Name	Purpose	Restriction site	Epitope tag	Sequence
Human	*SGCE*	2	HESGMYCPFUSE-F	RT-PCR	*Nco*I	c-Myc	CGGCCATGGTTGAGCAAAAGCTCATTTCTGAAGAGGACTTGG ATCGGAGTGTATACCCATC
Human	*SGCE*	12	HESGMYCPFUSE-R	RT-PCR	*Nco*I	-	TAACCATGGTGTGATGTAACTGCTATATTGTC
Mouse	*Kcna3*	1	cmKcna3F	RT-PCR	*Sal*I	-	AGTTGTCGACGCCAGACATGACCGTGGTGC
Mouse	*Kcna3*	1	cmKcna3R	RT-PCR	*Not*I	-	ATGAGCGGCCGCTCAGCACAGAATTGGCAACCGTA
Mouse	*Tnr*	1	cmTnR_F	RT-PCR	*Sal*I	-	AGTTGTCGACATGCTGGCTACCACTGAGAG
Mouse	*Tnr*	21	cmTnR_R	RT-PCR	*Not*I	FLAG	ATGAGCGGCCGCTCACTTGTCGTCATCGTCTTTGTAGTCGAAT TTCAAGGCTCGCCGT
Human	*SGCB*	1	hSGCB_cF	RT-PCR	*Sal*I	-	AGTTGTCGACGCGGGAAGATGGCGGCA
Human	*SGCB*	6	hSGCB_cR	RT-PCR	*Not*I	FLAG	ATGAGCGGCCGCTCACTTGTCGTCATCGTCTTTGTAGTCATGA GTGTTTCCACAGGGGTTG
Human	*SGCB*	1	hSGCB_Ex1F	Colony PCR	-	-	GCGGGAAGATGGCGGCA
Human	*SGCB*	4	hSGCB_Ex3R	Colony PCR	-	-	GCTTTGTTGTCCCTTGCTGA
Human	*SGCB*	2	hSGCB_Ex2aF	Colony PCR	-	-	GAAGTCCATGCGTGAGAAGG
Human	*SGCB*	5	hSGCB_Ex4R	Colony PCR	-	-	ACACCTTCATTTCCACGCAC
Human	*SGCB*	4	hSGCB_Ex3F	Colony PCR	-	-	GCAAGGGACAACAAAGCTCA
Human	*SGCB*	6	hSGCB_Ex5R	Colony PCR	-	-	CCAGTCACCACTACCCAACT
Human	*SGCD*	1a.3	hSGCD_164F	RT-PCR	*Sal*I	-	AGTTGTCGACGTGTGAAGGCTGAGACAACC
Human	*SGCD*	1b	hSGCD_476F	RT-PCR	*Sal*I	-	AGTTGTCGACTTTCATCGGCCGGTTTGTGA
Human	*SGCD*	8b	hSGCD_475R	RT-PCR	*Not*I	HA	ATGAGCGGCCGCTCAAGCGTAATCTGGAACATCGTATGGGTA ACCAGTTATGGGAAGGGTTGG
Human	*SGCD*	9	hSGCD_388R	RT-PCR	*Not*I	HA	ATGAGCGGCCGCTCAAGCGTAATCTGGAACATCGTATGGGTA GAGGCAGACACTTGTGTTTATCTG

Human	*SGCD*	8	hSGCD_7skR	Recloning for HA tag	*Not*I	HA	ATGAGCGGCCGCTCAAGCGTAATCTGGAACATCGTATGGGTA GAGACCGGGTTGGGGACTCCAAC
Human	*SGCD*	3	hSGCD_Ex3R	Colony PCR	-	-	TCCTCAGGTTTCCCATTCCA
Human	*SGCD*	2	hSGCD_Ex2aF	Colony PCR	-	-	GTTGAGTGAAGGGACCAGGT
Human	*SGCD*	5	hSGCD_Ex4R	Colony PCR	-	-	GGCAGACTTGAAGTACAGGG
Human	*SGCD*	4	hSGCD_Ex3F	Colony PCR	-	-	TCTTACAACCTCTCTACGCCA
Human	*SGCD*	7	hSGCD_Ex6R	Colony PCR	-	-	AGATTTAGGGAACACTGTGCC
Human	*SGCD*	6	hSGCD_Ex5F	Colony PCR	-	-	TGAAGTGGTAGTAGGAGCTGA
Human	*SGCD*	8b	hSGCD_Ex7aR	Colony PCR	-	-	TCTCATCCCTCACCTCTCCA
Human	*SGCD*	9	hSGCD_Ex7bR	Colony PCR	-	-	AGGCGTGTAGGATCCATGAG
Human	*SGCZ*	1	hSGCZ_cLNF	RT-PCR	*Sal*I	-	AGTTGTCGACAGTTGCGCTCCATGGACAG
Human	*SGCZ*	8	hSGCZ_cR	RT-PCR	*Not*I	c-Myc	ATGAGCGGCCGCTCACAAGTCCTCTTCAGAAATGAGCTTTTG CTCGCTCCACAGGCAGATGTTG
Human	*SGCZ*	1	hSGCZ_ex1aF	Colony PCR	-	-	AGTTGCGCTCCATGGACAG
Human	*SGCZ*	5	hSGCZ_ex4R	Colony PCR	-	-	GCCATCTTCACTGGCTCTCA
Human	*SGCZ*	4	hSGCZ_ex3F	Colony PCR	-	-	ATCACATGGGGCAGTTAACC
Human	*SGCZ*	7	hSGCZ_Ex6R	Colony PCR	-	-	CTTGAAGTCTCCTGCAGCAG
Human	*SGCZ*	5	hSGCZ_Ex5F	Colony PCR	-	-	ACATCAGAGCAGAGCCATCC
Human	*SGCZ*	8	hSGCZ_Ex7R	Colony PCR	-	-	TGAGGAGCTGGGTGAAGAAG
Human	*SGCZ*	2	hSGCZ_2nR	Colony PCR	-	-	GCATTCTCAGTCCTTGGCAG
Human	*SGCZ*	2	hSGCZ_2nF	Colony PCR	-	-	CCAAGGACTGAGAATGCACA
Human	*SGCZ*	4	hSGCZ_4nR	Colony PCR	-	-	GCCCCATGTGATTTCTTGCA
Human	*SGCZ*	6	hSGCZ_6nR	Colony PCR	-	-	CACAGAGTGCCCAAATACGG
Human	*SGCA*	1	hSGCA_cF	RT-PCR	*Sal*I	-	AGTTGTCGACCTCTGTCACTCACCGGGC
Human	*SGCA*	10	hSGCA_cR	RT-PCR	*Not*I	-	ATGAGCGGCCGCCCCCTCTCCCTGCTTGTTTA
Human	*SGCA*	4	hSGCA_4R	Colony PCR	-	-	TGGTATCAAAGCTGTCCCGA
Human	*SGCA*	3	hSGCA_3F	Colony PCR	-	-	TGGGCTCCAGGTCATTGAG
Human	*SGCA*	7	hSGCA_7R	Colony PCR	-	-	CTCCGGCACTGACTTATCCA
Human	*SGCA*	6	hSGCA_6F	Colony PCR	-	-	CGTTGACTGGTGCAATGTGA

Human	*SGCA*	8	hSGCA_8R	Colony PCR	-	-	TAGCCAGGTCTCTCTTCAGC
Human	*SGCA*	8	hSGCA_8F	Colony PCR	-	-	GCTGAAGAGAGACCTGGCTA
Human	*SGCA*	5	hSGCA_5F	Colony PCR	-	-	CTTCAGCTGCTCAACGTCAC
Human	*SGCA*	9	hSGCA_9cR	Recloning for c-Myc tag	*Not*I	c-Myc	ATGAGCGGCCGCTCACAAGTCCTCTTCAGAAATGAGCTTTTG CTCGTGCTGGTCCAGAATGAGGG
Human	*SGCA*	10	hSGCA_9bR	Recloning for c-Myc tag	*Not*I	c-Myc	ATGAGCGGCCGCTCACAAGTCCTCTTCAGAAATGAGCTTTTG CTCCCCCACCCCACCCCCTCTCCCTGCTTGTTTA
Human	*SGCG*	1	hSGCG_cF	RT-PCR	*Sal*I	-	AGTTGTCGACCCAGCTGTAGTTCATTCGCC
Human	*SGCG*	8	hSGCG_cR	RT-PCR	*Not*I	c-Myc	ATGAGCGGCCGCTCACAAGTCCTCTTCAGAAATGAGCTTTTG CTCGAGGCAGATGTGGCTGTG
Human	*SGCG*	5	hSGCG_5R	Colony PCR	-	-	CCGTCGTTGGAGTTGATCTG
Human	*SGCG*	4	hSGCG_4F	Colony PCR	-	-	TCACAGGCAGGTTAAAAGTCG
Human	*SGCG*	7	hSGCG_7R	Colony PCR	-	-	CATCCATGCTTAGACTCCGAG
Human	*SGCG*	5	hSGCG_5F	Colony PCR	-	-	CAGATCAACTCCAACGACGG
Human	*SGCG*	2	hSGCG_2cF	Recloning into pCMV-myc	*Sal*I	-	AGTTGTCGACGTTGCGCATGGTGCGTGAGCAGTAC
Human	*SGCG*	8	hSGCG_2Rcmv	Recloning into pCMV-myc	*Not*I	-	ATGAGCGGCCGCTCAGAGGCAGATGTGGCTGTG
Human	*SGCZ*	-	hZlnP267S_F	Mutagenesis for PCR error	-	-	TTCTTCACCCAGCTCCTCAAGTTCTCGACAGAC
Human	*SGCZ*	-	hzlnP267S_R		-	-	GTCTGTCGAGAACTTGAGGAGCTGGGTGAAGAA
Human	*SGCZ*	-	hZlnN113D_F	Mutagenesis for PCR error	-	-	AAGAAATTCATTCTCGAAAGGATAGTCCGCTGGTCTTACAG
Human	*SGCZ*	-	hZlnN113D_R		-	-	CTGTAAGACCAGCGGACTATCCTTTCGAGAATGAATTTCTT
Human	*SGCZ*	-	hZlnT304S_F	Mutagenesis for PCR error	-	-	TAGGTTCCACTTGTCAGTCCAGTAGCAACATCTGC
Human	*SGCZ*	-	hZlnT304S_R		-	-	GCAGATGTTGCTACTGGACTGACAAGTGGAACCTA
Human	*SGCD*	-	hD2A8T_F	Mutagenesis for PCR error	-	-	CTCAGGAGCAGTACACTCACCACCGGAGC
Human	*SGCD*	-	hD2A8T_R		-	-	GCTCCGGTGGTGAGTGTACTGCTCCTGAG
Human	*SGCD*	-	hD2A220T_F	Mutagenesis for PCR error	-	-	GGCAATATGGAAGCCACCTGCAGGACAGAGC
Human	*SGCD*	-	hD2A220T_R		-	-	GCTCTGTCCTGCAGGTGGCTTCCATATTGCC
Human	*SGCD*	-	hD3R9H_F	Mutagenesis for PCR error	-	-	GAGCAGTACACTCACCACCGGAGCACC
Human	*SGCD*	-	hD3R9H_R		-	-	GGTGCTCCGGTGGTGAGTGTACTGCTC

Human	*SGCD*	-	hD3R176K_F	Mutagenesis for PCR error	-	-	GAGGGCACAGTGTTCCCTAAATCTATAGAAACACCTAA
Human	*SGCD*	-	hD3R176K_R		-	-	TTAGGTGTTTCTATAGATTTAGGGAACACTGTGCCCTC
Human	*SGCA*	-	hA237tc_F	Mutagenesis for PCR error	-	-	TGCCTCCCCGCGTGGACAGCGCCCAG
Human	*SGCA*	-	hA237tc_R		-	-	CTGGGCGCTGTCCACGCGGGGAGGCA
Human	*SGCA*	-	hA34ex5ct_F	Mutagenesis for PCR error	-	-	CGGAGGAGGTGCTGCCCTCAACACCTGCC
Human	*SGCA*	-	hA34ex5ct_R		-	-	GGCAGGTGTTGAGGGCAGCACCTCCTCCG
Human	*SGCA*	-	hA34_ex7_tc_F	Mutagenesis for PCR error	-	-	CCTGGTGCCCCTGCTGGTGGCCCT
Human	*SGCA*	-	hA34_ex7_tc_R		-	-	AGGGCCACCAGCAGGGGCACCAGG
Human	*SGCA*	-	hAT3_ex10_ct_F	Mutagenesis for PCR error	-	-	CTGATTCCAGCTCCTGGCCCTCCTGG
Human	*SGCA*	-	hAT3_ex10_ct_R		-	-	CCAGGAGGGCCAGGAGCTGGAATCAG

Appendix II: Pre-immune immunoglobulin IAP, CD-1 mice

Table of proteins identified in the pre-immune immunoglobulin IAP from digitonin-solubilised wild-type CD-1 mouse brain. Proteins are ranked by protein identity score, an output of the SEQUEST algorithm that provides the likelihood that the protein in question was genuinely present in the sample. Also provided are the percent of the protein covered by peptides identified in the sample (Coverage), the number of distinct peptides identified for that protein in the sample (Number of unique peptides), the number of mass spectra corresponding to that protein (Spectral count), and the molecular weight of the protein in kilodaltons (Molecular weight (kDa)).

Gene symbol	Protein name	Protein identity score	Coverage (%)	Number of unique peptides	Spectral count	Molecular weight (kDa)
Cltc	Clathrin heavy chain 1	261.93	44	50	87	191.4
Tuba1a	Tubulin alpha-1A chain	260.42	64.3	3	76	50.1
Tubb2b	Tubulin beta-2B chain	228.72	75.96	1	84	49.9
Tubb2a	Tubulin beta-2A chain	227.69	75.96	2	85	49.9
Tubb4b	Tubulin beta-4B chain	220.09	77.75	1	82	49.8
tuba	Tubulin alpha chain	219.85	64.37	1	64	49.8
Tubb4a	Tubulin beta-4A chain	215.56	78.83	4	80	49.6
Gapdh	Glyceraldehyde-3-phosphate dehydrogenase	212.71	69.97	15	66	35.8
Tubb5	Tubulin beta-5 chain	203.45	77.93	4	76	49.6
Tuba4a	Tubulin alpha-4A chain	166.71	62.28	4	52	49.9
Atp5b	ATP synthase subunit beta, mitochondrial	121.2	48.39	16	37	56.3
Tubb3	Tubulin beta-3 chain	117.2	50.89	5	42	50.4
Syn1	Synapsin-1	115.32	47.88	17	39	74.1
Dpysl2	Dihydropyrimidinase-related protein 2	109.22	66.26	18	35	62.2
KRT1	Keratin, type II cytoskeletal 1	101.02	45.34	17	27	66
Actb	Actin, cytoplasmic 1	99.78	57.07	1	33	41.7
ACTB	Actin, cytoplasmic 1	93.6	57.6	1	31	41.7

Hspd1	60 kDa heat shock protein, mitochondrial	88.28	34.03	13	24	60.9
Syn2	Synapsin-2	86.42	38.05	11	28	63.3
Eef1a2	Elongation factor 1-alpha 2	85.7	30.24	4	26	50.4
Eef1a1	Elongation factor 1-alpha 1	85.2	30.3	4	25	50.1
Atp1a3	Sodium/potassium-transporting ATPase subunit alpha-3	81.75	25.27	8	25	111.6
Mb21d2	Protein MB21D2 (Mab-21 domain-containing protein 2)	79.22	39.71	12	25	55.6
Slc1a3	Excitatory amino acid transporter 1	73.11	15.1	5	20	59.6
Cnp	2',3'-cyclic-nucleotide 3'-phosphodiesterase	71.52	41.9	12	24	47.1
Atp5a1	ATP synthase subunit alpha, mitochondrial	67.95	38.17	14	24	54.6
Nsf	Vesicle-fusing ATPase	66.67	36.16	19	26	82.6
Dnm1	Dynamin-1	62.77	23.99	11	20	97.7
Hsp90aa1	Heat shock protein HSP 90-alpha	60.95	26.47	9	22	84.7
Ighg2b	Immunoglobulin gamma heavy chain	59.99	47.62	2	18	36.7
Stxbp1	Syntaxin-binding protein 1	58.94	34.68	12	17	67.5
Hspa8	Heat shock cognate 71 kDa protein	54.69	42.05	11	17	50.4
Ighg	Immunoglobulin gamma-3	50.41	33.83	1	18	51.9
Atp1a1	Sodium/potassium-transporting ATPase subunit alpha-1	48.7	16.62	4	15	112.9
HSP90AB1	Heat shock protein HSP 90-beta	48.16	26.24	8	19	83.2
Ighg	Immunoglobulin gamma-3	46.61	38.3	1	19	25.8
KRT1	Keratin, type II cytoskeletal 1	46.37	27.8	11	14	66
Ina	Alpha-internexin	46.07	27.15	9	14	55.3
Igkv8-30	Immunoglobulin kappa light chain	45.46	34.58	1	19	26.6
Igkc	Immunoglobulin kappa light chain	45.26	39.73	1	19	24.1
Ap2b1	AP-2 complex subunit beta	45.17	16.43	8	13	101.3
Atp1a2	Sodium/potassium-transporting ATPase subunit alpha-2	44.09	15.21	2	14	103.5
Nefm	Neurofilament medium polypeptide	43.33	22.17	10	14	95.9
Ighg1	Immunoglobulin heavy chain	43.18	46.6	2	16	35.7
Tcp1	T-complex protein 1 subunit alpha	43	24.82	7	14	60.4
Actbl2	Beta-actin-like protein 2	40.74	26.33	1	12	42

Idh3a	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	40.73	21.04	5	12	39.6
Ighg1	Immunoglobulin heavy chain	39.1	46.6	1	14	35.7
Hnrnpk	Heterogeneous nuclear ribonucleoprotein K	38.14	25.91	8	17	48.5
Pkm	Pyruvate kinase PKM	35.8	12.05	4	10	57.8
Ncdn	Neurochondrin	34.87	24.01	11	12	78.8
Vdac1	Voltage-dependent anion-selective channel protein 1	34.15	56.08	9	12	32.3
Gls	Glutaminase kidney isoform, mitochondrial	34	14.54	5	9	73.9
Pfkp	6-phosphofructokinase	33.77	15.18	5	9	85.5
Prss1	Trypsin-1	33.08	12.2	2	13	26.1
Hspa12a	Heat shock 70 kDa protein 12A	32.72	23.41	11	11	74.8
Sirt2	NAD-dependent protein deacetylase sirtuin-2	31.64	25.07	6	8	39.4
Ywhaz	14-3-3 protein zeta/delta	31.37	28.98	5	9	27.8
Eno1	Alpha-enolase	30.02	22.35	5	7	47.1
Nefl	Neurofilament light polypeptide	29.51	19.02	5	9	57.8
Ap2a2	AP-2 complex subunit alpha-2	28.57	18.66	5	10	104
Crmp1	Dihydropyrimidinase-related protein 1	28.26	20.28	5	8	62.1
Rap1gds	Rap1 GTPase-GDP dissociation stimulator	27.7	17.83	5	7	59.1
Arf3	ADP-ribosylation factor 3	27.62	54.7	1	10	20.6
Slc1a2	Excitatory amino acid transporter 2	27.24	13.82	4	7	60.6
Cfl1	Cofilin-1	26.11	50.6	6	7	18.5
Pygb	Alpha-1,4 glucan phosphorylase	25.39	14.71	7	11	96.6
Sh3glb2	Endophilin-B2	24.94	23.51	6	11	44.9
Hnrnpl	Heterogeneous nuclear ribonucleoprotein L	24.71	7.17	2	6	63.9
Sept6	Septin-6	24.66	19.59	5	7	49.6
Atp6v1b2	V-type proton ATPase subunit B, brain isoform	24.14	20.94	6	7	56.5
Dnm3	Dynamin-3	22.87	5.01	1	6	92.8
Cpsf7	Cleavage and polyadenylation specificity factor subunit 7	22.82	5.31	1	5	52
Uba1	Ubiquitin-like modifier-activating enzyme 1	22.7	9.07	5	10	117.7
Tmod2	Tropomodulin-2	21.42	16.52	2	5	39.5

Ddx3x	ATP-dependent RNA helicase DDX3X	21.36	8.32	3	6	73
Tubb1	Tubulin beta-1 chain	21.25	13.75	1	9	50.4
Rab3a	Ras-related protein Rab-3A	21.22	35	4	7	25
Atp6v1a	V-type proton ATPase catalytic subunit A	21.21	9.4	4	6	68.3
Ppp2r1a	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	21.15	16.47	6	8	65.3
Sept8	Septin-8	21.11	18.88	4	7	49.8
Gfap	Glial fibrillary acidic protein	21.11	21.86	7	8	49.9
Sept7	Septin-7	20.88	19.69	4	7	45
Ywhag	14-3-3 protein gamma	20.56	25.1	3	5	28.3
Ddx1	ATP-dependent RNA helicase DDX1	20.3	5.14	2	6	82.4
Sept5	Septin-5	20.01	26.29	5	9	42.7
Acat1	Acetyl-CoA acetyltransferase, mitochondrial	19.94	15.57	4	6	44.8
Snap91	Clathrin coat assembly protein AP180	19.8	13.25	5	6	86
Ckmt1	Creatine kinase U-type, mitochondrial	19.68	11.24	2	5	47
Acly	ATP-citrate synthase	19.4	8.03	2	5	47.9
Psmd2	26S proteasome non-ATPase regulatory subunit 2	19.24	6.67	2	6	66.8
Matr3	Matrin-3	19.09	7.09	3	5	94.6
Rmdn3	Regulator of microtubule dynamics protein 3	18.96	4.89	1	4	52
Atp2b4	Plasma membrane calcium-transporting ATPase 4	18.28	6.59	1	9	122.2
Ogdh	2-oxoglutarate dehydrogenase, mitochondrial	18.06	7.56	2	7	116
Cct5	T-complex protein 1 subunit epsilon	18.06	14.6	4	6	59.6
Ogdhl	2-oxoglutarate dehydrogenase-like, mitochondrial	17.8	9.6	2	8	114.5
Xpnpep1	Xaa-Pro aminopeptidase 1	17.49	13.01	5	6	66
Mbp	Myelin basic protein	17.38	32.64	6	7	20.9
Slc4a4	Electrogenic sodium bicarbonate cotransporter 1	17.14	2.34	1	4	115.1
Etfa	Electron transfer flavoprotein subunit alpha, mitochondrial	17.06	13.21	3	5	35
Atp2b2	Plasma membrane calcium-transporting ATPase 2	16.92	5.63	3	5	127.2
Stip1	Stress-induced-phosphoprotein 1	16.92	10.5	4	5	62.5

Gnao1	Guanine nucleotide-binding protein G(o) subunit alpha	16.85	22.88	5	6	40.1
Kpnb1	Importin subunit beta-1	16.84	4	2	4	97.1
Pi4ka	Phosphatidylinositol 4-kinase alpha	16.83	2.89	3	5	231.2
Amph	Amphiphysin	16.56	5.45	2	4	69.6
Hadha	Trifunctional enzyme subunit alpha, mitochondrial	16.44	9.57	3	4	82.6
Gpd1	Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic	16.4	9.17	1	3	37.5
Gda	Guanine deaminase	16.39	7.27	2	4	50.9
Arf4	ADP-ribosylation factor 4	16.32	38.89	1	6	20.4
Park7	Protein deglycase DJ-1	16.29	18.52	2	4	20
Ap2a1	AP-2 complex subunit alpha-1	16.21	9.11	2	6	107.6
Cse1l	Exportin-2	15.95	8.2	4	5	103.8
ARF1	ADP-ribosylation factor 1	15.83	43.65	1	6	20.6
Ppp3ca	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform	15.27	14.97	4	7	58.6
Eef2	Elongation factor 2	14.96	5.22	3	4	93.5
Cct6a	T-complex protein 1 subunit zeta	14.91	23.39	5	5	42.9
PLP1	Myelin proteolipid	14.9	20.22	5	6	30.1
Nono	Non-POU domain-containing octamer-binding protein	14.83	4.86	1	8	54.4
Add1	Alpha-adducin	14.82	13.79	5	5	72.9
Actr1a	Alpha-centractin	14.65	16.49	1	4	42.6
Pfkl	ATP-dependent 6-phosphofructokinase, liver type	14.53	6.41	2	4	85.3
Vps35	Vacuolar protein sorting-associated protein 35	14.4	6.16	3	4	83.9
Dpysl3	Dihydropyrimidinase-related protein 3	14.4	10.74	1	4	61.7
Ctbp1	C-terminal-binding protein 1	14.25	14.29	3	4	47.7
Ahsg	Alpha-2-HS-glycoprotein	14.21	15.36	2	5	37.3
Ak1	Adenylate kinase isoenzyme 1	14.17	19.07	3	4	21.5
Ywhab	14-3-3 protein beta/alpha	14.07	19.92	2	4	28.1
Gad2	Glutamate decarboxylase 2	14.01	7.35	2	3	65.2
Osbpl10	Oxysterol-binding protein-related protein 10	13.95	7.11	2	4	65.8

Ywhae	14-3-3 protein epsilon	13.68	16.08	3	4	29.2
Map6	Microtubule-associated protein 6	13.56	15.32	5	5	54.7
Hspa9	Stress-70 protein, mitochondrial	13.55	10.6	5	5	73.4
Vdac2	Voltage-dependent anion-selective channel protein 2	13.3	28.98	6	6	30.4
Try10	MCG140784	13.21	14.63	1	8	26.2
Acaa1a	3-ketoacyl-CoA thiolase A, peroxisomal	13.12	20.41	1	4	10.2
Actr1b	Beta-centractin	12.86	45.4	2	4	19.6
Ap1b1	AP-1 complex subunit beta-1	12.83	8.52	2	4	101.1
Camk2a	Calcium/calmodulin-dependent protein kinase type II subunit alpha	12.77	9.21	3	4	54.1
Ahcyl1	Putative adenosylhomocysteinase 2	12.72	15.09	4	7	58.9
C1qb	Complement C1q subcomponent subunit B	12.67	15.02	2	4	26.7
Hspa1L	Heat shock protein 1-like protein	12.62	6.86	1	4	70.6
Dnm1l	Dynamin-1-like protein	12.62	6.88	2	3	79.5
Cadps	Calcium-dependent secretion activator 1	12.59	5.39	5	5	153
Sars	Serine--tRNA ligase, cytoplasmic	12.51	8.61	2	3	58.2
Pdhb	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	12.4	16.43	3	4	38.9
Dnajc6	Putative tyrosine-protein phosphatase auxilin	12.09	6.93	2	5	102.2
Sept11	Septin-11	12.06	12.06	2	4	49.7
Igh	Immunoglobulin H	12.03	8.81	1	4	52.2
Gpd1l	Glycerol-3-phosphate dehydrogenase 1-like protein	12	9.12	1	3	38.2
Rasal1	RasGAP-activating-like protein 1	11.76	3	1	3	89.3
Idh3b	Isocitrate dehydrogenase [NAD] subunit, mitochondrial	11.7	8.07	2	3	42.2
Tomm70a	Mitochondrial import receptor subunit TOM70	11.66	2.95	1	3	67.5
Cand1	Cullin-associated NEDD8-dissociated protein 1	11.57	4.63	4	4	136.2
C1qa	Complement C1q subcomponent subunit A	11.51	14.29	2	3	26
Cpsf6	Cleavage and polyadenylation-specificity factor subunit 6	11.49	13.74	1	3	21.2
Tufm	Elongation factor Tu	11.48	13.05	4	4	49.5
Slc25a5	ADP/ATP translocase 2	11.35	25.84	2	5	32.9
ACT1	Actin 1	11.3	15.92	1	6	41.7

Acadsb	Short/branched chain specific acyl-CoA dehydrogenase, mitochondrial	11.26	6.18	1	3	37.7
Cct2	T-complex protein 1 subunit beta	11.11	9.35	3	4	57.4
Atp2b1	Plasma membrane calcium-transporting ATPase 1	11.05	5.33	1	4	134.7
Tubal3	Tubulin alpha chain-like 3	11.03	7.4	1	5	50
Calb1	Calbindin	10.73	13.03	2	3	30
Git1	ARF GTPase-activating protein GIT1	10.71	5.52	2	3	84.1
Nefh	Neurofilament heavy polypeptide	10.65	3.92	2	4	112.5
Sfpq	Splicing factor, proline- and glutamine-rich	10.63	30.33	2	3	13.9
Tppp	Tubulin polymerization-promoting protein	10.43	13.76	2	3	23.6
Sccpdh	Saccharopine dehydrogenase-like oxidoreductase	10.4	17.06	1	3	18.6
Hspa5	78 kDa glucose-regulated protein	10.39	5.77	1	3	56.9
Atp5c1	ATP synthase subunit gamma, mitochondrial	10.22	26.32	2	3	18.7
Eif3f	Translation initiation factor-3 subunit 5	10.14	12.84	1	2	16.1
Krt5	Keratin, type II cytoskeletal 5	10.06	6.03	2	3	61.7
Hnrnph1	Heterogeneous nuclear ribonucleoprotein H	10.05	10.47	3	3	49.2
Prss1	Trypsin-1	10.04	8.13	1	4	26.1
Vdac3	Voltage-dependent anion-selective channel protein 3	9.87	12.01	2	4	30.7
Snrpd1	Small nuclear ribonucleoprotein Sm D1	9.75	16.81	1	2	13.3
Slc25a12	Calcium-binding mitochondrial carrier protein Aralar1	9.75	2.51	1	2	74.5
Atp2b3	Plasma membrane calcium-transporting ATPase 3	9.73	4.81	1	5	125.6
Pde1b	Calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase 1B	9.67	9.17	3	3	52.5
Dlat	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	9.64	3.74	2	3	67.9
Cct7	T-complex protein 1 subunit eta	9.58	10.66	3	3	59.6
Prss2	Anionic trypsin-2	9.51	15.45	1	3	26.5
Ipo9	Importin-9	9.28	4.27	1	2	72.7
Ttc7b	Tetratricopeptide repeat protein 7B	9.22	2.37	1	3	94.1

Napb	Beta-soluble NSF attachment protein	9.19	10.4	2	2	33.5
Kpna4	Importin subunit alpha-3	9.09	3.26	1	2	57.9
Map2	Microtubule-associated protein 2	8.98	6.65	1	2	49.2
Add2	Beta-adducin	8.93	7.12	4	4	77.8
Sh3gl2	Endophilin-A1	8.92	11.36	3	3	39.9
Capzb	Capping protein beta 3	8.87	27.78	1	2	10
Krt10	Keratin, type I cytoskeletal 10	8.76	6.42	3	3	57
Acsl6	Long-chain-fatty-acid--CoA ligase 6	8.63	6.96	2	3	39.8
Serpinh1	Serpin H1	8.57	6.24	1	3	46.5
Dclk1	Serine/threonine-protein kinase DCLK1	8.56	11.29	3	3	40.3
Cyfip2	Cytoplasmic FMR1-interacting protein 2	8.51	3.99	4	5	145.6
Hnrnpa2b1	Heterogeneous nuclear ribonucleoprotein A2/B1	8.5	12.02	3	3	35.9
Sv2a	Synaptic vesicle glycoprotein 2A	8.45	6.2	2	2	82.6
Picalm	Phosphatidylinositol-binding clathrin assembly protein	8.34	6.53	1	3	64.6
Slc25a4	ADP/ATP translocase 1	8.27	21.48	1	4	32.9
Pfkm	ATP-dependent 6-phosphofructokinase, muscle type	8.22	11.54	6	6	85.2
Plcxd3	PI-PLC X domain-containing protein 3	8.19	8.72	1	2	36.3
Slc25a3	Phosphate carrier protein, mitochondrial	8.16	7.56	2	3	39.6
Atp2a2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	7.97	5.61	4	4	109.7
Krt42	Keratin, type I cytoskeletal 42	7.91	7.3	1	2	50.1
Atp1b1	Sodium/potassium-transporting ATPase subunit beta-1	7.84	12.5	3	4	35.2
Eif4g1	Eukaryotic translation initiation factor 4 gamma 1	7.82	4.8	1	3	54.4
Osbpl9	Oxysterol-binding protein-related protein 9	7.67	4.15	2	3	70.1
Vsnl1	Visinin-like protein 1	7.6	14.14	2	2	22.1
Gpr37l1	G protein-coupled receptor 37-like 1	7.59	4.89	1	2	38.3
Rph3a	Rabphilin-3A	7.58	6.31	2	2	75.4
Krt14	Keratin, type I cytoskeletal 14	7.56	7.02	1	2	52.8
Psmc4	26S protease regulatory subunit 6B	7.53	16.24	2	2	35.2
Gnb1	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	7.36	8.53	2	2	37.4

Actr10	Actin-related protein 10	7.3	4.32	1	2	46.2
Eif4b	Eukaryotic translation initiation factor 4B	7.28	6.2	1	2	46.5
Aak1	AP2-associated protein kinase 1	7.22	5.53	2	2	103.3
Pcmt1	Protein-L-isoaspartate(D-aspartate) O-methyltransferase (PIMT)	7.19	15.82	2	2	18.9
Nup155	Nuclear pore complex protein Nup155	7.18	2.72	1	3	88.8
Rab10	Ras-related protein Rab-10	7.01	16.5	2	3	22.5
Actn1	Alpha actinin 1a	7	2.48	2	2	102.7
Sv2b	Synaptic vesicle glycoprotein 2B	7	10.93	1	2	20
Pygm	Alpha-1,4 glucan phosphorylase	6.96	7.03	1	4	87.5
Map2k4	Dual specificity mitogen-activated protein kinase kinase 4	6.94	8.56	2	3	44.1
Dazap1	DAZ-associated protein 1	6.93	9.36	1	2	19.4
Arhgef2	Rho guanine nucleotide exchange factor 2	6.93	3.03	1	4	108.5
Pdia6	Protein disulfide-isomerase A6	6.88	6.14	1	2	42.9
Prdx2	Peroxiredoxin-2	6.87	22.73	3	3	21.8
Ighv1-31	Immunoglobulin heavy chain	6.86	33.67	1	3	11
Ddx5	Probable ATP-dependent RNA helicase DDX5	6.82	13.79	3	4	46
Synj1	Synaptojanin-1	6.75	2.37	2	2	144.5
Krt79	Keratin, type II cytoskeletal 79	6.73	2.26	1	2	57.5
Cul5	Cullin-5	6.61	4.77	1	2	65.5
Ptk2b	Protein-tyrosine kinase 2-beta	6.59	4.03	2	2	111
Atad3	ATPase family AAA domain-containing protein 3	6.44	3.72	2	2	66.7
Arhgap44	Rho GTPase-activating protein 44	6.41	2.33	1	2	88.9
Fus	RNA-binding protein FUS	6.4	19.23	2	3	14
Atad1	ATPase family AAA domain-containing protein 1	6.34	8.21	1	2	22.1
Aldh6a1	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	6.3	7.44	2	3	49.6
Prps1l3	Ribose-phosphate pyrophosphokinase 1	6.27	5.98	1	2	27.5
Uggt1	UDP-glucose:glycoprotein glucosyltransferase 1	6.27	2	2	2	176.3
Aldh2	Aldehyde dehydrogenase, mitochondrial	6.27	5.78	2	2	56.5

Pdha1	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	6.15	8.97	2	2	43.2
EZR	Ezrin	6.01	4.44	2	2	69.4
Sucla2	Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial	6.01	7.13	2	2	50.1
Vps52	Vacuolar protein sorting-associated protein 52 homolog	5.95	6.36	2	3	82
Clasp2	CLIP-associating protein 2	5.92	3.67	2	2	119.4
Atp5f1	ATP synthase F(0) complex subunit B1, mitochondrial	5.87	5.86	1	2	28.9
Syt1	Synaptotagmin-1	5.85	8.31	3	3	47.4
Opa1	Dynamin-like 120 kDa protein, mitochondrial	5.81	14.19	2	3	33.9
Caprin1	Caprin-1	5.79	6.53	1	2	37.6
Gstp2	Glutathione S-transferase P 2	5.77	19.15	2	2	21.1
Ctnnb1	Catenin beta-1	5.75	12.07	1	2	19.1
Prkce	Protein kinase C epsilon type	5.69	3.39	1	3	83.5
Atp5o	ATP synthase subunit O, mitochondrial	5.56	11.74	2	2	23.3
Pdcd6ip	Programmed cell death 6-interacting protein	5.52	1.96	1	2	96
Fh	Fumarate hydratase, mitochondrial	5.5	6.71	2	2	54.3
Stx1b	Syntaxin-1B	5.47	9.72	2	2	33.2
Acsf2	Acyl-CoA synthetase family member 2, mitochondrial	5.35	4.72	1	1	67.9
Trim21	E3 ubiquitin-protein ligase TRIM21	5.35	5.19	2	2	53.3
Anxa7	Annexin A7	5.31	6.02	2	2	44.4
Nceh1	Neutral cholesterol ester hydrolase 1	5.24	6.86	1	1	45.7
Cct3	T-complex protein 1 subunit gamma	5.23	4.73	2	2	56.5
Lphn3	Latrophilin-3	5.15	2.52	1	2	97.1
Phgdh	D-3-phosphoglycerate dehydrogenase	5.12	10.32	4	4	56.5
Pcbp2	Poly(rC)-binding protein 2	5.11	10.43	2	2	29.6
Ppp1cc	Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	5.01	5.26	1	1	37
Ndufs1	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial	4.85	4.13	1	1	79.6
Sec23a	Protein transport protein Sec23A	4.82	3.67	1	5	82.9
Hnrnpd	Heterogeneous nuclear ribonucleoprotein D0	4.75	20.56	2	2	12.1

Pura	Transcriptional activator protein Pur-alpha	4.74	7.17	1	1	34.9
Josd2	Josephin-2	4.74	13.7	1	1	16.1
Khsrp	Far upstream element-binding protein 2	4.73	3.21	2	2	76.7
Ldha	L-lactate dehydrogenase	4.73	7.17	1	1	26
Ehd3	EH domain-containing protein 3	4.67	5.05	1	1	60.8
Enpp6	Ectonucleotide pyrophosphatase/phosphodiesterase family member 6	4.65	25.33	1	1	8.8
Nudt21	Cleavage and polyadenylation specificity factor subunit 5	4.53	12.78	1	2	26.2
Ogt	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit	4.51	1.91	1	1	116.9
Iars2	Isoleucine--tRNA ligase, mitochondrial	4.49	6.43	1	1	62.4
Prkar1a	cAMP-dependent protein kinase type I-alpha regulatory subunit	4.34	9.34	1	1	28.8
Hbb	Beta-globin	4.34	25.85	3	3	15.7
Tom1l2	TOM1-like protein 2 (Target of Myb-like protein 2)	4.24	2.56	1	1	55.6
Dnaja2	DnaJ homolog subfamily A member 2	4.2	3.64	1	1	45.7
Lgi3	Leucine-rich repeat LGI family member 3	4.17	2.74	1	1	61.8
Krt79	Keratin, type II cytoskeletal 79	4.11	2.26	1	1	57.5
Ganab	Neutral alpha-glucosidase AB	4.01	3.39	1	1	106.8
Vat1L	Synaptic vesicle membrane protein VAT-1 homolog-like	3.95	3.89	1	1	36.8
SDH1-2	Succinate dehydrogenase [ubiquinone] flavoprotein subunit 2, mitochondrial	3.89	2.22	1	1	69.3
Mms19	MMS19 nucleotide excision repair protein homolog	3.88	9.79	1	4	30.9
Dpysl4	Dihydropyrimidinase-related protein 4	3.88	3.5	1	1	61.9
Stk39	STE20/SPS1-related proline-alanine-rich protein kinase	3.85	2.52	1	1	60.3
Ppp2r2a	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform	3.84	5.69	1	1	28.9
Ikbkap	Elongator complex protein 1	3.81	1.73	1	1	149.5
Trim2	Tripartite motif-containing protein 2	3.81	3.94	1	1	61.6
Rabgef1	Rab5 GDP/GTP exchange factor	3.79	3.87	1	1	56.8
Pfas	Phosphoribosylformylglycinamidine synthase	3.79	1.42	1	1	144.5

Dpp9	Dipeptidyl peptidase 9	3.79	2.44	1	1	97.9
Fam49a	Protein FAM49A	3.74	7.43	1	1	37.3
Aars	Alanine--tRNA ligase, cytoplasmic	3.74	4.15	1	1	49.6
Vamp3	Vesicle-associated membrane protein 3	3.63	16.5	1	1	11.5
Ctsb	Cathepsin B	3.62	9.28	1	1	20.9
Hk1	Hexokinase-1	3.61	2.24	1	1	74.3
Map2k1	Dual specificity mitogen-activated protein kinase kinase 1	3.59	5.33	1	1	33.4
Hpcal4	Hippocalcin-like protein 4	3.57	7.74	1	1	18
Eif5a	Eukaryotic translation initiation factor 5A-1	3.57	31.58	1	1	8.4
Krt2	Keratin, type II cytoskeletal 2 epidermal	3.56	1.98	1	1	70.9
Slc6a1	Sodium- and chloride-dependent GABA transporter 1	3.56	8.2	1	1	28.6
Snca	Alpha-synuclein	3.56	11.43	1	1	14.5
Trim3	Tripartite motif-containing protein 3	3.54	3.07	1	1	78.2
2210010C04 Rik	Protein 2210010C04Rik	3.5	4.86	1	1	26.4
Atp6v1h	V-type proton ATPase subunit H	3.48	3.31	1	1	55.8
Igkv3-7	Immunoglobulin kappa light chain	3.46	18	1	1	10.9
Xpo7	Exportin-7	3.46	1.47	1	1	123.7
Nptn	Neuroplastin	3.44	14.75	1	1	13.2
Arfip2	Arfaptin-2 (ADP-ribosylation factor-interacting protein 2)	3.43	6.45	1	1	37.7
Canx	Calnexin	3.4	2.71	1	1	67.2
Ppp3r1	Calcineurin subunit B type 1	3.39	16.47	1	1	19.3
Wdr7	WD repeat-containing protein 7	3.37	1.54	1	1	163.3
Prkcb	Protein kinase C beta type	3.27	1.79	1	1	76.7
Dnaja3	DnaJ homolog subfamily A member 3, mitochondrial	3.26	5.08	1	1	49.5
Mtch2	Mitochondrial carrier homolog 2	3.25	4.76	1	1	32.3
Ptbp1	Polypyrimidine tract-binding protein 1	3.23	7.02	1	1	32.3
Comtd1	Catechol O-methyltransferase domain-containing protein 1	3.21	9.92	1	2	28.9
Rnf14	E3 ubiquitin-protein ligase RNF14	3.18	6.13	1	1	40.7

Krt77	Keratin 77	3.16	4.95	1	1	25.8
mt-Nd4	NADH-ubiquinone oxidoreductase chain 4	3.14	7.14	1	1	28.5
Rgs7	Regulator of G-protein signaling 7	3.07	3.41	1	1	54.8
Diras2	GTP-binding protein Di-Ras2	3.07	5.53	1	1	22.5
Rpl23	60S ribosomal protein L23	3.06	33.33	1	1	6.2
Cryzl1	Quinone oxidoreductase-like protein 2	3.04	8.29	1	1	37.8
Ckb	Creatine kinase B-type	3.03	3.41	1	1	42.7
Gnaq	Guanine nucleotide-binding protein G(q) subunit alpha	3.03	5.29	1	1	42.1
Pde2a	cGMP-dependent 3',5'-cyclic phosphodiesterase	3.02	2.29	1	1	103.2
Necab2	N-terminal EF-hand calcium-binding protein 2	2.98	12.5	1	1	25.2
Hba	Alpha globin	2.98	25.86	1	1	6.2
Inpp4a	Type I inositol 3,4-bisphosphate 4-phosphatase	2.97	1.92	1	1	76.3
Ndufs7	NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial	2.93	12.05	1	1	24.7
Rtn1	Reticulon-1	2.92	4.36	2	2	83.5
Dsp	Desmoplakin (DP)	2.91	0.42	1	1	332.7
Rap1gap	Rap1 GTPase-activating protein 1	2.91	4.2	1	1	43.4
Psmd12	26S proteasome non-ATPase regulatory subunit 12	2.91	6.31	1	1	23.6
Dhx9	ATP-dependent RNA helicase A	2.88	2.29	1	1	59.2
Nsfl1c	NSFL1 cofactor p47 (p97 cofactor p47)	2.87	5.01	1	1	37.4
Atp6v1c1	V-type proton ATPase subunit C 1	2.86	4.19	1	1	43.8
Dync1li1	Cytoplasmic dynein 1 light intermediate chain 1	2.85	9.36	1	1	21
Actr3	Actin-related protein 3	2.85	5.09	1	1	24.1
Ywhaq	14-3-3 protein theta	2.84	4.9	1	1	27.8
Sept4	Septin-4	2.82	6.57	1	1	24.2
Syncrip	Heterogeneous nuclear ribonucleoprotein Q	2.81	6.63	1	1	21.6
Ctnnd2	Catenin delta-2	2.79	2.68	1	1	41.6
Mpst	3-mercaptopyruvate sulfurtransferase	2.73	4.38	1	1	33.1
Syngr1	Synaptogyrin-1	2.72	6.28	1	1	21.3

CPK21	Calcium-dependent protein kinase 21	2.71	5.08	1	1	59.9
4930544G11Rik	Testis specific expressed protein 5	2.69	12.44	1	1	21.6
Mgst3	Microsomal glutathione S-transferase 3	2.68	8.5	1	1	16.9
Cap1	Adenylyl cyclase-associated protein	2.66	11.59	1	1	17.9
Rufy3	Protein RUFY3	2.65	5.12	2	2	53
Cttn	Src substrate cortactin	2.64	2.55	1	1	57.1
Map1b	Microtubule-associated protein 1B	2.64	3.09	1	1	34.9
Hivep1	Zinc finger protein 40	2.62	3.42	1	1	92.9
Npepps	Puromycin-sensitive aminopeptidase	2.61	1.37	1	1	98.5
Lasp1	LIM and SH3 domain protein 1	2.6	23.21	1	1	6.5
Hadhb	Trifunctional enzyme subunit beta, mitochondrial	2.59	9.25	1	1	19
Syp	Synaptophysin	2.57	7.88	1	1	18.1
Abat	4-aminobutyrate aminotransferase, mitochondrial	2.55	6.47	1	1	26.2
Cct8	T-complex protein 1 subunit theta	2.54	9.09	1	1	14.5
Pacsin1	Protein kinase C and casein kinase substrate in neurons protein 1	2.54	5.22	1	1	50.5
Dagla	Sn1-specific diacylglycerol lipase alpha	2.51	1.63	1	1	115.2
Sec24b	Protein transport protein Sec24B	2.51	1.08	1	1	112
Vcp	Transitional endoplasmic reticulum ATPase	2.48	2.35	1	1	80.1
Capza2	F-actin-capping protein subunit alpha-2	2.47	5.75	1	1	20
Echs1	Enoyl-CoA hydratase, mitochondrial	2.46	5.76	1	1	20.9
Cct4	T-complex protein 1 subunit delta	2.46	5.7	2	2	54.8
Bdh1	3-hydroxybutyrate dehydrogenase	2.45	19.05	1	1	12.1
Padi2	Protein-arginine deiminase type-2	2.45	6.48	1	1	36.9
Gpd2	Glycerol-3-phosphate dehydrogenase, mitochondrial	2.39	1.24	1	1	80.9
Suclg1	Succinyl-CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial	2.38	4.62	1	1	36.1
Smap1	Stromal membrane-associated protein 1	2.37	3.38	1	1	37.8
Rab11b	Ras-related protein Rab-11B	2.34	10.4	1	1	14.3
Ncald	Neurocalcin-delta	2.33	4.6	1	1	20.2

Rab2a	Ras-related protein Rab-2A	2.32	15.48	2	2	19.1
Krt15	Keratin, type I cytoskeletal 15	2.28	3.98	1	1	49.1
Aldoc	Fructose-bisphosphate aldolase C	2.28	4.41	1	1	39.4
Gtf2i	General transcription factor II-I	2.25	8.12	1	1	40.5
Pmpca	Mitochondrial-processing peptidase subunit alpha	2.24	3.85	1	1	48.4
Ipo5	Importin-5	2.24	2.51	1	1	90
Gnb5	Guanine nucleotide-binding protein subunit beta-5	2.22	9.31	1	1	22
Dctn1	Dynactin subunit 1	2.22	2.28	1	1	126.7
Mapt	Microtubule-associated protein tau	2.2	10.75	2	2	38.9
Ppp2cb	Serine/threonine-protein phosphatase 2A catalytic subunit beta isoform	2.2	3.96	1	1	32
Nap1l1	Nucleosome assembly protein 1-like 1	2.18	3.82	1	1	33.5
Sncb	Beta-synuclein	2.18	10.53	1	1	14
Mtfr1l	Mitochondrial fission regulator 1-like	2.17	11.88	1	1	22.4
Atl1	Atlastin-1	2.11	2.57	1	1	49.2
Krt10	Keratin, type I cytoskeletal 10	2.11	2.14	1	1	57
Vps51	Vacuolar protein sorting-associated protein 51 homolog	2.1	2.69	1	1	86.1
Igk12f2	Immunoglobulin kappa light chain	2.07	9.4	1	1	12.3
Immt	MICOS complex subunit Mic60	2.03	6.34	1	1	30.4
C2cd4cC2CD 4 family	C2 calcium-dependent domain-containing protein 4C	2.02	4.3	1	1	44.6
C1qc	Complement C1q subcomponent subunit C	2.02	7.32	1	3	26
cdk5	Cyclin-dependent kinase 5	2.02	4.97	1	1	18.5
Dlst	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	1.99	2.42	1	1	49
Pdhx	Pyruvate dehydrogenase protein X component, mitochondrial	1.94	1.8	1	1	54
Ighm	Immunoglobulin M	1.9	3.08	1	1	49.9
Rab18	Ras-related protein Rab-18	1.85	5.34	1	1	23
Atp5i	ATP synthase subunit e, mitochondrial	1.83	16.9	1	1	8.2

Prdx1	Peroxiredoxin-1	1.83	4.71	1	1	18.9
Plaa	Phospholipase A-2-activating protein	1.79	2.14	1	1	87.2
Pcbp1	Poly(rC)-binding protein 1 (Alpha-CP1)	1.76	3.93	1	1	37.5
Fubp1	Far upstream element (FUSE) binding protein 1	1.68	2.03	1	1	67.2
Tanc2	Protein TANC2	1.64	0.4	1	1	220.1
Usp5	Ubiquitin carboxyl-terminal hydrolase	1.62	1.56	1	1	93.3
Gpm6a	Neuronal membrane glycoprotein M6-a	1.6	4.12	1	1	29.8
Ran	GTP-binding nuclear protein Ran	0	6.48	1	1	24.4
Mpp2	MAGUK p55 subfamily member 2	0	3.62	1	1	61.5
Mycb	Protein B-Myc	0	21.76	1	3	18.5
Phyhipl	Phytanoyl-CoA hydroxylase-interacting protein-like	0	4.27	1	1	42.3
Rtn3	Reticulon-3	0	1.17	1	1	101.5
Rtn4	Reticulon-4	0	1.03	1	1	126.5
Igh	Immunoglobulin heavy chain	0	16.24	1	1	12.9
Prkcg	Protein kinase C gamma type	0	2.32	1	1	72.9
Gna14	Guanine nucleotide-binding protein subunit alpha-14	0	6.13	1	1	36.4
Bin1	Myc box-dependent-interacting protein 1	0	3.77	1	1	52.7
Rab5c	Ras-related protein Rab-5C	0	31.43	1	1	3.4
Plekhg5	Pleckstrin homology domain-containing family G member 5	0	3.94	1	1	115.7

Appendix III: Pre-immune immunoglobulin IAP, C57BL/6J mice

Table of proteins identified in the pre-immune immunoglobulin IAP from digitonin-solubilised wild-type C57BL/6J mouse brain. Proteins are ranked by protein identity score, an output of the SEQUEST algorithm that provides the likelihood that the protein in question was genuinely present in the sample. Also provided are the percent of the protein covered by peptides identified in the sample (Coverage (%)), the number of distinct peptides identified for that protein in the sample (Number of unique peptides), the number of mass spectra corresponding to that protein (Spectral count), and the molecular weight of the protein in kilodaltons (Molecular weight (kDa)).

Gene symbol	Protein name	Protein identity score	Coverage (%)	Number of unique peptides	Spectral count	Molecular weight (kDa)
Tubb4b	Tubulin beta-4B chain	44.8	31.24	1	15	49.8
Tubb2a	Tubulin beta-2A chain	41.07	34.61	5	15	49.9
Tubb4a	Tubulin beta-4A chain	35.72	28.38	2	12	49.6
Tuba1a	Tubulin alpha-1A chain	31.6	22.39	5	11	50.1
Mb21d2	Protein MB21D2 (Mab-21 domain-containing protein 2)	31.39	21.79	9	10	55.6
Atp5a1	ATP synthase subunit alpha, mitochondrial	31.09	21.16	9	11	59.7
Krt14	Keratin, type I cytoskeletal 14	29.04	17.98	3	10	52.8
Atp5b	ATP synthase subunit beta, mitochondrial	28.72	25.9	9	11	56.3
Gapdh	Glyceraldehyde-3-phosphate dehydrogenase	25.09	21.62	5	8	35.8
Krt42	Keratin, type I cytoskeletal 42	24.19	14.82	2	8	50.1
Krt5	Keratin, type II cytoskeletal 5	22.97	14.31	3	9	61.7
Atp1a3	Sodium/potassium-transporting ATPase subunit alpha-3	20.73	11.35	7	7	111.6
Krt77	Keratin 77	20.61	7.53	1	11	61.3
Igh	Immunoglobulin H	19.98	14.88	3	5	52.2
Ighm	Immunoglobulin mu	19.62	13.88	5	6	49.9
Dpysl2	Dihydropyrimidinase-related protein 2	19.52	15.56	6	7	62.2
Krt1	Keratin, type II cytoskeletal 1	17.82	5.02	1	10	65.6

Krt14	Keratin, type I cytoskeletal 14	16.89	11.16	1	5	52.8
Hspa8	Heat shock cognate 71 kDa protein	16.77	14.47	6	6	60.8
Actb	Actin, cytoplasmic 1	16.6	25.87	2	8	41.7
Igkc	Immunoglobulin kappa light chain	16.37	23.4	1	6	25.8
Krt10	Keratin, type I cytoskeletal 10	15.92	7.66	4	6	57
Actb	Actin, cytoplasmic 1	15.37	22.4	1	7	41.7
Igkc	Immunoglobulin kappa light chain	15.26	23.74	1	6	24.1
Krt6a	Keratin, type II cytoskeletal 6A	14.92	12.3	1	6	59.3
Krt2	Keratin, type II cytoskeletal 2 epidermal	14.73	4.53	2	6	70.9
Krt42	Keratin, type I cytoskeletal 42	14.4	9.73	1	4	50.1
Tuba4a	Tubulin alpha-4A chain	14.15	10.71	1	5	49.9
Krt10	Keratin, type I cytoskeletal 10	12.96	9.45	5	7	57
Krt5	Keratin, type II cytoskeletal 5	12.66	5.52	3	4	61.7
Nefl	Neurofilament light polypeptide	12.59	10.59	3	4	57.8
Trim21	E3 ubiquitin-protein ligase TRIM21	12.11	9.52	4	5	53.3
4732456N 10Rik	Uncharacterised protein	11.82	7.62	1	5	58.2
Krt73	Keratin, type II cytoskeletal 73	11.72	4.27	2	3	58.9
Krt14	Keratin, type I cytoskeletal 14	11.49	35.48	1	3	10.7
Ighg2b	Immunoglobulin gamma heavy chain	11.48	14.85	4	5	44.2
Krt16	Keratin, type I cytoskeletal 16	10.85	6.61	1	4	51.6
C1qb	Complement C1q subcomponent subunit B	10.84	16.6	4	6	26.7
Plp1	Myelin proteolipid protein	10.23	17.33	4	4	30.1
Krt16	Keratin, type I cytoskeletal 16	9.9	10.23	1	4	51.6
Dlat	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	9.83	9.35	4	5	67.9
Krt79	Keratin, type II cytoskeletal 79	9.67	5.84	1	4	57.5
Krt15	Keratin, type I cytoskeletal 15	9.35	9.73	1	5	49.1
Krt76	Keratin, type II cytoskeletal 76	9.09	4.88	1	4	62.8

Cnp	2',3'-cyclic-nucleotide 3'-phosphodiesterase	8.25	11.25	4	4	44.6
Prss1	Trypsin-1	8.21	8.13	1	2	26.1
Krt2	Keratin, type II cytoskeletal 2 epidermal	8.15	4.95	3	3	70.9
IgK	Immunoglobulin kappa	8.05	16.22	1	4	12
C1qc	Complement C1q subcomponent subunit C	7.95	13.01	3	3	26
Syn1	Synapsin-1	7.59	5.81	2	3	74.1
Krt14	Keratin, type I cytoskeletal 14	7.4	33.33	1	2	10.7
Syn2	Synapsin-2	7.09	5.97	1	2	63.3
Ighg3	Immunoglobulin gamma-3	7.03	10.3	3	3	43.9
Ighg1	Immunoglobulin heavy chain	6.94	8.21	3	3	51
Ighv	Immunoglobulin heavy chain	6.86	23.89	2	3	12.7
Xpnpep1	Xaa-Pro aminopeptidase 1	6.85	9.47	2	3	31.6
Rab3d	Ras-related protein Rab-3D	6.67	20.75	2	3	12
Pcmt1	Protein-L-isoaspartate(D-aspartate) O-methyltransferase (PIMT)	6.6	17.39	2	2	22.4
C1qa	Complement C1q subcomponent subunit A	6.58	10.2	2	3	26
Prss1	Trypsin-1	6.54	8.13	1	3	26.1
Ina	Alpha-internexin	6.02	4.19	1	2	55.3
Ighv	Immunoglobulin heavy chain	5.93	10.92	1	2	13.5
Krt79	Keratin, type II cytoskeletal 79	5.87	4.14	1	2	57.5
Cfl1	Cofilin-1	5.72	15.06	2	2	18.5
Sirt2	NAD-dependent protein deacetylase sirtuin-2	5.56	7.12	2	2	39.4
Nefm	Neurofilament medium polypeptide	5.46	4.89	1	2	53.5
Acot13	Acyl-coenzyme A thioesterase 13	5.28	21.43	2	2	15.2
Hspa9	Stress-70 protein, mitochondrial	5.25	3.68	2	2	73.4
Igha	Immunoglobulin A heavy chain	4.96	26.67	1	2	9.7
Gnao1	Guanine nucleotide-binding protein G(o) subunit alpha	4.87	7.34	2	2	40.1
Igk	Immunoglobulin kappa	4.83	23.68	2	2	12.6
Atp5c1	ATP synthase subunit gamma, mitochondrial	4.81	14.94	2	2	16.8
Igkv	Immunoglobulin kappa light chain	4.78	21.5	1	2	11.6

Ighv	Immunoglobulin heavy chain	4.7	20	1	2	13.3
Igkc	Immunoglobulin kappa light chain	4.37	21.5	1	2	11.7
Eef1a1	Elongation factor 1-alpha 1	4.18	4.11	2	2	50.1
Tuba1b	Tubulin alpha-1B chain	3.96	15.15	1	1	10.9
Dsp	Desmoplakin (DP)	3.9	0.66	1	1	261.3
Sncb	Beta-synuclein	3.65	10.53	1	1	14
C1qa	Complement C1q subcomponent subunit A	3.64	6.12	1	1	26
Ighv	Immunoglobulin heavy chain	3.33	27.97	1	2	12.9
Sh3glb2	Endophilin-B2	3.28	3.48	1	1	41.8
Krt78	Keratin, type II cytoskeletal 78	3.21	2.85	1	1	54.7
Ubc	Polyubiquitin-C	3.17	23.88	1	1	22.6
Ighm	Immunoglobulin M heavy chain	3.17	13.83	1	1	10.6
Hsp90ab1	Heat shock protein HSP 90-beta	3.16	7.41	1	1	22.5
Nsf	Vesicle-fusing ATPase	3.14	1.75	1	1	82.6
Eno1	Alpha-enolase	2.99	26.87	1	1	7.4
Actb	Actin, cytoplasmic 1 (Beta-actin)	2.94	14.88	1	1	13.5
Camk2a	Calcium/calmodulin-dependent protein kinase type II subunit alpha	2.92	6.35	1	1	21.3
Plp1	Myelin proteolipid	2.88	5.35	1	1	26.5
Jup	Junction plakoglobin	2.84	1.61	1	1	81.7
Gnb1	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	2.79	10.19	1	1	12.3
Igkv5-39	Immunoglobulin heavy chain	2.62	6.11	2	2	27.8
Mbp	Myelin basic protein	2.57	8.2	1	1	13.8
Gapdh	Glyceraldehyde-3-phosphate dehydrogenase	2.54	9.72	1	1	15.5
Krt78	Keratin, type II cytoskeletal 78	2.54	2.44	1	1	54.7
Mbp	Myelin basic protein	2.35	9.84	1	1	13.8
Pdha1	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	2.2	3.33	1	1	43.2
C1qc	Complement C1q subcomponent subunit C	2.1	4.07	1	1	26
Arf3	ADP-ribosylation factor 3	2.04	22.45	1	1	5.3

Igkv	Immunoglobulin kappa light chain	1.94	11.4	1	1	12.4
Sugp1	SURP and G-patch domain-containing protein 1	1.6	1.71	1	1	72.6
Sdk2	Protein sidekick-2	0	1.47	1	1	239.8
Pemt	Phosphatidylethanolamine N-methyltransferase	0	10.16	1	1	14.2
Ccdc171	Coiled-coil domain-containing protein 171	0	1.13	1	2	152.2
Atp1b1	Sodium/potassium-transporting ATPase subunit beta-1	0	3.62	1	1	35.2
Arvcf	Armadillo repeat protein deleted in velo-cardio-facial syndrome homolog	0	0.94	1	1	105

Appendix IV: Papers and presentations arising from this thesis

Papers

Waite, A.J.*, **Carlisle, F.A.***, Chan, Y. M. & Blake, D. J. 2016. Myoclonus dystonia and muscular dystrophy: ε-sarcoglycan is part of the dystrophin-associated protein complex in brain. *Movement Disorders*. 31 (11):1694-1703 (* denotes joint first authors)

Xiao, J., Vemula, S.R., Xue, Y., Khan, M., **Carlisle, F.A.**, Waite, A.J., Blake, D. J., Dragatsis, I., Zhao, Y. & LeDoux, M. 2017. Role of major and brain-specific Sgce isoforms in the pathogenesis of myoclonus-dystonia syndrome. *Neurobiology of Disease*. 98: 52-65.

Carlisle, F.A., Waite, A.J. & Blake, D.J. 2016. [Manuscript in preparation] The human sarcoglycan genes produce alternatively spliced transcripts encoding isoforms that affect assembly and trafficking of sarcoglycan complexes.

Presentations (as presenting author)

Carlisle, F.A., Waite, A.J. & Blake, D.J. Alternative splicing generates sarcoglycan isoforms with altered intracellular trafficking and sarcoglycan complex assembly. Poster presentation at the EMBO Meeting, Birmingham UK, 5th-8th September 2015.

Carlisle, F.A., Waite, A.J., Chan, Y.M., Isles, A.R. & Blake, D.J. ε-sarcoglycan interacts with components of the dystrophin-associated glycoprotein complex and other membrane proteins in brain. Poster presentation at the Society for Neuroscience Annual Meeting 2015, Chicago IL, 17th-21st October 2015.

Carlisle, F.A., Waite, A.J., Chan, Y.M., Isles, A.R. & Blake, D.J. ε-sarcoglycan interacts with components of the dystrophin-associated glycoprotein complex and other proteins in brain. Poster presentation at the Genetics Society 2015 Autumn Meeting, London UK, 19th-20th November 2015.

Carlisle, F.A., Waite, A.J. & Blake, D.J. Alternative splicing generates sarcoglycan isoforms with altered trafficking and protein interactions. Oral presentation at the Southwest RNA Club Annual Meeting 2016, Bristol UK, 23rd May 2016. Award-winning presentation.