

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Exploring the heart of MyBP-C

Pearce, Amy

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Exploring the heart of MyBP-C

By

Amy Pearce

**Thesis submitted for the degree of Doctor of
Philosophy**

King's College London

April 2021

Impact statement

The Covid-19 pandemic led to a complete closure of laboratory facilities exactly at the time when I was expecting to carry out important experiments for my thesis. From July, access to the laboratory was possible again, but with severe restrictions. Although I was generously awarded a paid three month extension from the British heart foundation, I still only had a limited amount of time in the laboratory to carry out experiments. Therefore, it was necessary to prioritise finishing experiments which had already been started in order to reach appropriate n numbers and achieve as complete a narrative as possible. Because of this there was not time to complete experiments which had been started or planned just before the restrictions came into effect.

Planned experiments included carrying out co-sedimentation assays with the central domains of cardiac myosin binding protein C (cMyBP-C) and myosin to see whether the central domains were able to interact with myosin. This could have possibly given more information about whether the effect seen on ATPase rates in the presence of the central domains were due to interactions with the thin or thick filament systems.

Recombinant production and purification of central domain fragments containing hypertrophic cardiomyopathy (HCM) linked missense mutations were also planned however due to the timescale involved in the optimisation of this procedure it was not possible to carry this out. These protein fragments would have been studied utilising techniques such as differential scanning fluorimetry to determine whether the mutations have an effect on domain stability. Furthermore, they would have been used in co-sedimentation assays in order to determine whether the mutations could be interfering with potential interaction partners.

These experiments would have provided more information about the functionality of the central domains as well as given further insight into the pathogenic mechanisms by which HCM-linked missense mutations cause disease.

Acknowledgements

Thank you to my supervisors Elisabeth Ehler and Thomas Kampourakis for all of their help and guidance throughout the years. Thank you to all of the members of the Ehler, Kampourakis and Gautel groups for your kindness and help. Thank you especially to Mark Holt for help with the STED microscopy and analysis and Luke Smith for helping me to understand Y2H and always being helpful and encouraging.

A special thank you to the people without whom I would never have made it through: Jess Chandler for hacking up a bovine heart with me. Tom Randall for always being a sounding board and a reassuring presence. Caraugh Albany for reminding me to not stress. Sarah Brown for the sanity saving coffee sessions. Beth Ormrod for pretty much always being by my side and listening to every single one of my rants. Sarah Kendall for having my back at every single occasion. Ailbhe O'Brien for too many reasons to articulate.

Finally, thank you to my family for all of their support. Lianne for helping me figure out adobe and pretending to understand when I talked about science. A special thank you to my mumma who had to actually read all of this and without whom there would be far less punctuation.

Abstract

Mutations in the sarcomeric protein cardiac myosin binding protein C (cMyBP-C) are one of the most common causes of hereditary hypertrophic cardiomyopathy (HCM). cMyBP-C is made up of 11 immunoglobulin-like or fibronectin-like domains, designated C0-C10, and has been found to interact with myosin, actin and titin. The N-terminus contains binding sites for both actin and myosin, suggesting it plays a role in the regulation of actomyosin cross-bridge cycling kinetics, which is the major role identified for MyBP-C. The C-terminus of the protein is thought to target MyBP-C to the thick filament backbone due to the fact that it contains binding sites for both myosin and titin. The function of the central domains (C3-C6) has not yet been elucidated; however the largest number of HCM-linked mutations are found in the C6 domain, followed by the C5 domain, intimating that this region does have an important functional role. The aims of this project were therefore:

1. To elucidate the structure and functional role of the central domains of cMyBP-C.
2. To understand the mechanism of disease of HCM-linked missense mutations in the C3C6 region.

To gain a better understanding of the role of the central domains a yeast two hybrid (Y2H) assay was carried out to identify putative novel binding partners. Recombinantly expressed fragments of the central domains were also used in biochemical assays in order to determine whether this region could be involved in interactions with actin. Constructs containing HCM-linked missense mutations within the central domain region were also generated and western blot analysis and immunofluorescence microscopy was utilised in order to gain insights into possible mechanisms of disease.

Whilst the results of these studies identified no novel binding partners via the Y2H screen, there was some evidence of an interaction between the central domains and actin. Furthermore, biochemical assays suggested that the central region could have a role in the regulation of actomyosin cross-bridge cycling kinetics. Experiments utilising HCM-linked missense mutation containing constructs highlighted the fact that there seems to be no unified disease mechanism by which these mutations lead to HCM. Instead it appears that these mutations lead to HCM via different pathways.

Further study of a wider range of missense mutations should therefore be carried out in order to inform on the function of specific regions of the protein, as well as allow for classification of different mutations depending on their pathogenesis to create more personalised methods of treatment.

Abbreviations

ACM – arrhythmogenic cardiomyopathy

ADP- adenosine-5'-diphosphate

AP – alkaline phosphatase

ATP – adenosine-5'-triphosphate

AV – adenovirus

BioID – biotin identification

BLAST – basic local alignment search tool

BSA – bovine serum albumin

CaMKII – Ca²⁺-dependent calmodulin kinase II

CK2 – casein kinase 2

cMyBP-C – cardiac myosin binding protein C

cMyBP-C^{+/-} - cardiac myosin binding protein C heterozygous knock-out

cMyBP-C^{-/-} - cardiac myosin binding protein C homozygous knock-out

CV – column volume

DAPI - 4',6-Diamidino-2-Phenylindole Dihydrochloride

DBSS-K – Dulbecco's balanced salt solution lacking potassium

DCM – dilated cardiomyopathy

dNTP – deoxynucleoside-5'-triphosphate

DMEM – Dulbecco's modified Eagle medium

DRX – disordered relaxed state

DSF – differential scanning fluorimetry

DTT - dithiothreitol

ECL – enhanced chemiluminescence

EDTA - ethylenediaminetetraacetic acid

EGTA - ethylene glycol tetraacetic acid

EHT – engineered heart tissue

ELC – essential light chains

ELIPA - enzyme linked inorganic phosphate assay

EM – electron microscopy

F-actin – filamentous actin

FBS – foetal bovine serum

FHL1 – four and a half LIM protein 1

FN - fibronectin

FT – flow through

G-actin – globular actin

GB – gold buffer

GFP – green fluorescent protein

GSK3 β – glycogen synthase kinase 3 β

HCM – hypertrophic cardiomyopathy

HF – heart failure

hiPSC-CMs – human induced pluripotent stem cell-derived cardiomyocytes

HMM – heavy meromyosin

HRP – horseradish peroxidase

ICM – ischaemic cardiomyopathy

IEC – ion exchange chromatography

IFU – infectious unit

Ig – immunoglobulin

IHM – interacting heads motif

IPTG - Isopropyl β -D-1-thiogalactopyranoside

KI – knock-in

KO – knock-out

L – lysate fraction

LB – Lennox broth

LMM – light meromyosin

LS – low salt

M199 – media 199

MALLS – multi-angle laser light scattering

MESG - 2-amino-6-mercapto-7-methylpurine riboside

MHC – myosin heavy chain

MLC – myosin light chain

MLCK – myosin light chain kinase

mRNA – messenger ribonucleic acid

MST – microscale thermophoresis

MyBP-C – myosin binding protein C

Myosin S1 – myosin subfragment 1

Myosin S2 – myosin subfragment 2

NaAC – sodium acetate

NMD – nonsense-mediated mRNA decay

NRC – neonatal rat cardiomyocytes

NTF – native thin filaments

PBS – phosphate buffered saline

PCR – polymerase chain reaction

PFA - paraformaldehyde

P_i – inorganic phosphate

PKA – protein kinase A

PKC – protein kinase C

PKD – protein kinase D

PNP – purine nucleotide phosphorylase

PP1 – protein phosphatase 1

PP2 – protein phosphatase 2

Pro-Ala – proline-alanine

PTM – post-translational modifications

qPCR – quantitative polymerase chain reaction

RLC – regulatory light chain

RPM – rotations per minute

RSK – ribosomal S6 kinase

RT – room temperature

SAXS – small-angle X-ray scattering

SDM – site directed mutagenesis

SD media – synthetic dropout media

SDS-PAGE - sodium dodecyl sulfate polyacrylamide gel electrophoresis

SHaRE – sarcomeric human cardiomyopathy registry

SOC medium – super optimal broth with catabolite repression medium

SPOC – spontaneous oscillatory contractions

SRX – super-relaxed state

STED – stimulated emission depletion microscopy

STORM - stochastic optical reconstruction microscopy

SUMO – small ubiquitin-related modifier

TB – terrific broth

TnC – troponin C

TnI – troponin I

TnT – troponin T

UPS – ubiquitin-proteasome system

WT – wild-type

Y2H – yeast two hybrid

YPD – yeast peptone dextrose

Contents

Acknowledgements.....	4
Abstract.....	5
Abbreviations.....	7
1. Introduction	20
1.1. Overview of the cardiac muscle.....	20
1.1.1. The Sarcomere	20
1.1.2. The thick filament	22
1.1.2.1. Myosin.....	22
1.1.2.2. Myosin Binding Proteins	23
1.1.3. The thin filament.....	24
1.1.3.1. Actin	25
1.1.3.2. Tropomyosin	25
1.1.3.3. Troponin complex	25
1.1.4. The elastic filament.....	26
1.1.5. Muscle contraction	27
1.2. Myosin binding protein C.....	34
1.2.1. MyBP-C structure	35
1.2.2. Localisation of MyBP-C in the sarcomere	37
1.2.3. Binding partners of MyBP-C.....	39
1.2.4. Function of MyBP-C	40
1.2.5. Post-translational modifications of MyBP-C.....	46
1.2.5.1. Phosphorylation of MyBP-C	47
1.2.5.2. Other types of post translational modifications	50
1.3. Cardiomyopathy.....	54
1.3.1. Hypertrophic Cardiomyopathy	54
1.3.2. HCM causing mutations	56
1.3.3. Truncation vs. missense mutations	57
1.3.4. Models of HCM causing cMyBP-C mutations	59
1.3.5. Current therapies	66
1.4. Current knowledge on the central domains of cMyBP-C	69
1.4.1. Structure of the central domains.....	69
1.4.2. Possible binding partners of the central domains	70
1.4.3. Possible role for the central domains	71

1.5. Aims and objectives	72
2. Materials and Methods.....	74
2.1. Buffers and Media.....	74
2.2. Antibodies	76
2.3. Generation of eukaryotic expression construct.....	77
2.4. Site directed mutagenesis of C3C6 constructs	80
2.5. COS-1 cells.....	82
2.5.1. Passage.....	82
2.5.2. Transfection	82
2.5.3. Immunofluorescence staining.....	82
2.5.4. Western blot analysis.....	83
2.6. Neonatal rat cardiomyocytes (NRC)	84
2.6.1. Isolation and culture	84
2.6.2. Transfection	84
2.6.3. Immunofluorescence staining.....	84
2.7.1. Confocal microscopy and image analysis.....	86
2.7.2. STED microscopy	86
2.7.2.1. Immunostaining	86
2.7.2.2. STED imaging & analysis	87
2.8. Full length cMyBP-C adenovirus	90
2.8.1. AV propagation	90
2.8.2. SDM of pENTR vector.....	92
2.8.3. Titration of AV.....	93
2.8.4. Determination of adenoviral concentration and infection duration	94
2.8.5. Infection of NRCs.....	94
2.9. Protein expression & purification	96
2.9.1. Protein expression optimisation	96
2.10.1. Native thin filament preparation	98
2.10.2. ATPase assay to check NTF activity.....	99
2.10.3. ATPase assay with C3C5.....	100
2.11. NTF co-sedimentation.....	101
2.12. Yeast Two Hybrid (Y2H) assay.....	102
2.12.1. Lithium acetate/carrier DNA transformation	102
2.12.2. Autoactivation check	103

2.12.3. LexA fusion expression check	103
2.12.4. Y2H assay screen.....	104
3. Characterisation of the wild-type central domains of cardiac myosin binding protein C	106
3.1. Localisation of the central domains of cMyBP-C	107
3.2. Production and purification of central domain fragments	110
3.2.1. Production of the C3C6 domain fragment.....	110
3.2.2. Optimisation of production and purification of the C3C5 domain fragment	112
3.3. Effect of the central domains on actomyosin cross-bridge cycling kinetics	115
3.3.1. Production of native thin filaments	115
3.3.2. Effect of the central C3C5 domains on the ATPase rate.....	116
3.4. Summary	119
4 – Possible binding partners of the central domains of cardiac myosin binding protein C	120
4.1. Yeast two hybrid assay screen to identify putative binding partners of cMyBP-C's central domains.....	121
4.1.1. Generation of bait constructs for the yeast two hybrid assay screen	121
4.1.2. Hits from the yeast two hybrid assay screen.....	124
4.1.2.1. Hits from the Y2H assay screen with the C4 bait construct.....	124
4.1.2.2. Hits from the Y2H assay screen with the C5 bait construct.....	125
4.1.2.3. Hits from the Y2H assay screen with the C3C6 bait construct	126
4.1.3. Validation of putative interaction partners identified by the yeast two hybrid screen	129
4.2. Possible interaction between the central domains and actin	131
4.2.1. Ratiometric analysis of co-localisation of the C3C6 domains and actin	131
4.2.2. Co-sedimentation of the central domains and native thin filaments.....	133
4.3. Summary	135
5 – Hypertrophic cardiomyopathy-linked mutations in the central domains of cardiac myosin binding protein C	136
5.1. HCM-linked mutant C3C6 domain containing pEGFP vectors	137
5.1.1. E542Q.....	142
5.1.2. G596R.....	144
5.1.3. N755K.....	146
5.1.4. W792R.....	149
5.1.5. R820Q.....	151
5.2.1. Protein expression levels of HCM-linked mutant pEGFP C3C6 constructs	153
5.2.2. Cellular localisation of pEGFP constructs containing HCM-linked mutations	154

5.2.3. Nuclear localisation of the mutant pEGFP C3C6 constructs	160
5.2.4. Effect of HCM-linked mutations on co-localisation of the central domains with actin.....	161
5.2.5. Possible myofibrillar disarray in cells expressing mutant C3C6 domain constructs	164
5.2.6. Summary of findings	173
5.2.7. Limitations of the experiments utilising the pEGFP vectors.....	174
5.3.1. Full length cardiac myosin binding protein C adenoviral constructs containing HCM-linked mutations	176
5.3.2. Optimisation of adenoviral infection	176
5.3.3. Titration of mutant adenoviral constructs.....	178
5.3.4. Protein expression levels of mutant constructs	180
5.3.5. Effect of HCM-linked mutants on localisation of cMyBP-C.....	182
5.3.5.1. Effect of HCM-linked mutants on overall cellular localisation of cMyBP-C.....	182
5.3.5.2. Effect of HCM-linked mutants on localisation of cMyBP-C at the myofibrillar level	190
5.4. Summary	203
6. Discussion.....	207
6.1. The central domains on their own show a diffuse, cytoplasmic localisation in cells	208
6.2. Potential instability of the C3C6 domains.....	209
6.3. Y2H assay screens failed to identify putative binding partners.....	211
6.4. Actin is a potential interaction partner of the central domains	213
6.5. The central domains may have a role in the regulation of actomyosin cross-bridge cycling kinetics	215
6.6. Missense mutations in the central domains lead to Hypertrophic Cardiomyopathy via several different mechanisms	217
6.7. Limitations of the study	228
6.8. Future directions.....	230
6.9. Final conclusions	234

Table of Figures and Tables

Figure 1 – Basic structure of the cardiac sarcomere.....	21
Figure 2 – Basic outline of the myosin molecule	23
Figure 3 - Basic outline of the thin filament structure.....	24
Figure 4 – Scheme of the excitation contraction coupling in the cardiac muscle	28
Figure 5 – Scheme of regulation of muscle contraction	29
Figure 6 – Structure of cMyBP-C.....	35
Figure 7 - Proposed models of the 3-dimensional sarcomeric organisation of Myosin binding protein C	38
Figure 8 – Putative post translational modifications of cardiac myosin binding protein C.....	47
Figure 9 - Characterisation of Hypertrophic Cardiomyopathy.....	54
Figure 10 - Myofibrillar disarray seen in Hypertrophic cardiomyopathy.....	55
Figure 11 – HCM-linked cMyBP-C mutations.....	56
Table 1 – Table of Buffers and Medias used during experimentations	75
Table 2 – Table of Antibodies used for western blot analysis and immunofluorescence staining	76
Figure 12 – Plasmid map of the pEGFP-C2 vector	77
Table 3 – Thermocycler settings used for amplification of C3C6 domains of cMyBP-C from adult human heart cDNA.....	78
Table 4 – Table of primer sequences used for site directed mutagenesis of C3C6 constructs to generate HCM-linked mutant constructs.	80
Table 5 – Thermocycler settings used for site directed mutagenesis	81
Figure 14– Plasmid map of the pAd vector containing the human full length cMyBP-C.....	91
Figure 15 – Plasmid map of the pLexA vector	102
Figure 16 – Validation of the GFP-tagged C3C6 domain containing construct.....	107
Figure 17 – Localisation of the GFP-tagged wildtype C3C6 containing plasmid.....	109
Figure – Localisation of C3C6 construct.....	109
Figure 18 – Optimisation of C3C6 protein induction	111
Figure 19 – Optimisation of C3C5 protein induction & purification	113
Figure 20 – Validation of native thin filament preparation	116
Figure 21 – Effect of C3C5 domains on ATPase activity.....	117
Figure 22 – Generation of bait constructs for the yeast two hybrid assay.....	123
Table 6 – Hits from the yeast two hybrid assay with the C4 bait plasmid.....	125
Table 7 - Hits from the yeast two hybrid assay with the C5 bait plasmid	126
Table 8- Hits from the yeast two hybrid assay with the C3C6 bait plasmid	128
Figure 23 – Forced cross of potential cardiac myosin binding protein C interactors identified in the yeast two hybrid assay.....	129
Figure 24 – Co-localisation of the central domains and F-actin	132
Figure 25 – Co-sedimentation of native thin filaments with the central domains of cardiac myosin binding protein C.....	134
Figure 26 – Hypertrophic cardiomyopathy linked missense mutations of the central domains of cardiac myosin binding protein C	137
Figure 27 – Sequence alignment of the C3 domain of myosin binding protein C	138
Figure 28 – Sequence alignment of the C4 domain of myosin binding protein C	139
Figure 29 – Sequence alignment of the C5 domain of myosin binding protein C	140
Figure 30 – Sequence alignment of the C6 domain of myosin binding C	141

Figure 31 – Structural representation of the E542Q mutation in the C3 domain of cardiac myosin binding protein C.....	143
Figure 32 – Structural representation of the G596R mutation in the C4 domain of myosin binding protein C	145
Figure 33 – Structural representation of the N755K mutation in the C5 domain of cardiac myosin binding protein C.....	146
Figure 34 – Structural representation of the W792R mutation in the C6 domain of cardiac myosin binding protein C.....	149
Figure 35 – Structural representation of the R820Q mutation in the C6 domain of cardiac myosin binding protein C.....	151
Figure 36 – Expression levels of GFP-tagged mutant C3C6 containing constructs.....	154
Figure 37 – Localisation of GFP-tagged C3C6 mutant constructs in COS-1 cells	156
Figure 37 cont.– Localisation of GFP-tagged C3C6 mutant constructs in COS-1 cells	157
Figure 38 – Localisation of GFP-tagged C3C6 mutant constructs in neonatal rat cardiomyocytes....	158
Figure 38 cont. – Localisation of GFP-tagged C3C6 mutant constructs in neonatal rat cardiomyocytes	159
Figure 39 – Nuclear localisation of GFP-tagged mutant C3C6 constructs in COS-1 cells and neonatal rat cardiomyocytes	161
Figure 40 – Co-localisation of the central domains and F-actin	163
Figure 41 – Myofibrillar disarray in neonatal rat cardiomyocytes expressing mutant GFP-tagged C3C6 containing plasmids	165
Figure 42 – Intensity profile plots from STED images of NRCs expressing mutant C3C6 domain constructs.....	169
Figure 43 – Intensity profile plots from STED images of NRCs expressing mutant C3C6 domain constructs.....	170
Figure 44 – Intensity profile plots from STED images of NRCs expressing mutant C3C6 domain constructs.....	171
Figure 45 – Sarcomeric organisation of NRCs expressing mutant C3C6 domain constructs.....	172
Table 9 – Summary of results from utilisation of the mutant GFP-tagged C3C6 domain mutant containing plasmids	173
Figure 46 – Optimisation of adenoviral dilution and infection duration	178
Figure 47 – Titration of full-length cardiac myosin binding protein C mutant adenoviral constructs	179
Figure 48– Protein expression levels of mutant full length-cardiac myosin binding protein C adenoviral constructs.....	181
Figure 49 – Localisation of full length wild type cardiac myosin binding protein C in neonatal rat cardiomyocytes.....	184
Figure 50 – Localisation of full length E542Q mutant cardiac myosin binding protein C in neonatal rat cardiomyocytes.....	185
Figure 51 – Localisation of full length G596R mutant cardiac myosin binding protein C in neonatal rat cardiomyocytes.....	186
Figure 52 – Localisation of full length N755K mutant cardiac myosin binding protein C in neonatal rat cardiomyocytes.....	187
Figure 53 – Localisation of full length W792R mutant cardiac myosin binding protein C in neonatal rat cardiomyocytes.....	188

Figure 54 – Analysis of localisation of full length cardiac myosin binding protein C mutant proteins	189
Figure 55 – Myofibrillar localisation of the wildtype full-length cMyBP-C	192
Figure 56 – Myofibrillar localisation of the E542Q mutant containing full-length cMyBP-C.....	193
Figure 57 – Myofibrillar localisation of the G596R mutant containing full-length cMyBP-C.....	194
Figure 58 – Myofibrillar localisation of the N755K mutant containing full-length cMyBP-C.....	195
Figure 59 – Myofibrillar localisation of the W792R mutant containing full-length cMyBP-C	196
Figure 60– Myofibrillar localisation of the R820Q mutant containing full-length cMyBP-C	197
Figure 61 - Intensity profile plots from STED images of NRCs expressing mutant full-length cMyBP-C	200
Figure 62 - Intensity profile plots from STED images of NRCs expressing mutant full-length cMyBP-C	201
Figure 63 – Sarcomeric organisation of NRCs expressing mutant full length cMyBP-C	202
Figure 64 – Summary of potential pathogenic mechanisms by which missense mutations in the central domains may lead to Hypertrophic Cardiomyopathy	223

1. Introduction

1.1. Overview of the cardiac muscle

The muscles in the body can be largely divided into smooth muscle, which is mainly found in the walls of hollow organs, and striated muscle. Striated muscle can be further separated into skeletal muscle, either slow or fast twitch, and cardiac muscle. Skeletal muscle is attached to the bones by tendons and is responsible for voluntary movement, whilst cardiac muscle is found solely in the heart (Hill and Olson, 2012).

The contractile tissue of the heart is made up of cardiomyocytes. The cardiomyocyte consists mainly of mitochondria, 1-2 nuclei and a parallel arrangement of myofibrils which run along the cell longitudinally and are subdivided transversely into sarcomeres (Sequeira et al., 2014). The cardiomyocytes are surrounded by a membrane called the sarcolemma, with the sarcolemma and the sarcomere linked by components called costameres. The intermediate filament protein desmin acts to connect the sarcomere to organelles as well as maintaining the structural integrity of the cell. Unlike skeletal muscle, where each muscle cell is isolated from each other, cardiomyocytes are connected at the ends via specialised contact structures, termed intercalated discs. The intercalated discs allow the transport of molecules and electrical signals which are integral for the fine-tuned synchronisation of muscle contraction in the heart (Henderson et al., 2017).

1.1.1. The Sarcomere

The sarcomere, the structure of which is outlined in figure 1, is the basic contractile unit which is defined as the region between two Z-discs, which confer the lateral boundaries of the sarcomere. The sarcomere consists mainly of three filaments: the myosin containing thick filaments, actin containing thin filaments and the titin elastic filaments, which will be

discussed in further detail below. The Z-disc and its associated proteins play a structural role, anchoring the thin filaments via the actin crosslinker α -actinin (Maruyama and Ebashi, 1965) and also acting as an anchor for the N-terminus of titin (Young et al., 1998). The regions on either side of the Z-disc are the I-bands which are bare of any myosin containing thick filaments, whilst the A-band is the region which extends the entire length of the thick filaments (Henderson et al., 2017). At the centre of the A-band is the M-band, which is integral to the stabilisation of the thick filament during contraction, as the thick filaments cross-link into the M-band via myomesin containing M-bridges. The repeating myomesin domains are capable of stretching 2.5x from a folded up state during contraction and relaxation and thus stabilise the M-band by acting like an elastic band (Pinotsis et al., 2012).

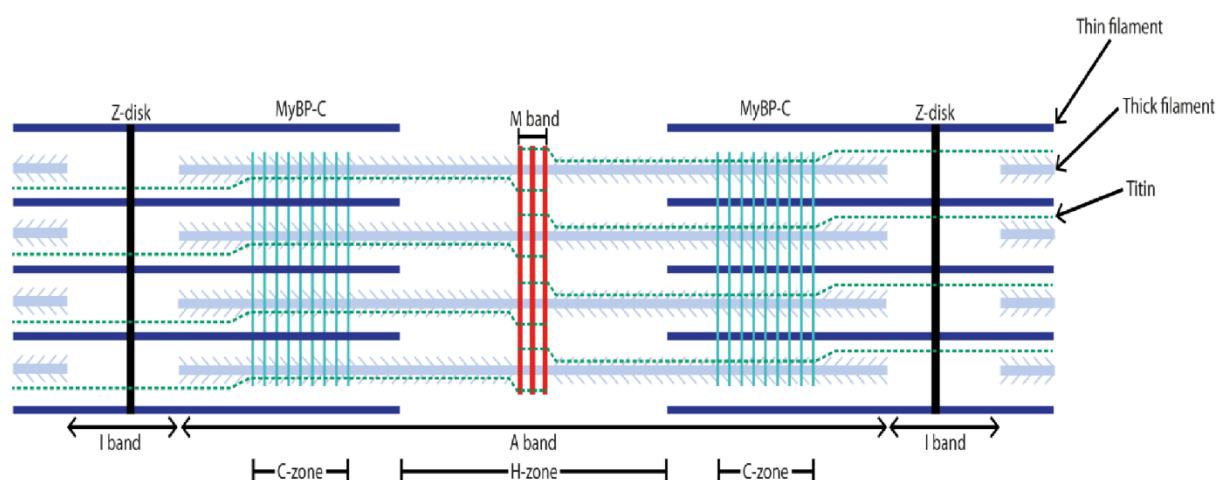


Figure 1 – Basic structure of the cardiac sarcomere

Outline of the sarcomere, the basic contractile unit of the myofibrils. The sarcomere is defined as the region between two Z-discs (black) and consists mainly of three filament systems: the thin filaments (dark blue), the thick filaments (pale blue) and the elastic titin filaments (green). MyBP-C (cyan) is located in the inner two thirds of the A-band in two groups, separated by the bare H-zone, as nine stripes 43 nm apart in the C-zone. At the centre of the A-band is the M-band (red), which acts to stabilise the thick filament during contraction.

1.1.2. The thick filament

1.1.2.1. Myosin

The thick filaments consist of myosin dimers arranged as bipolar filaments which interdigitate with the thin filaments. The myosin tail domains form the thick filament backbone whilst the myosin heads form a quasi-helical array on its surface (Reviewed in (Ehler, 2016)). There are 11 classes of myosins found in humans, although not all form filaments (Foth et al., 2006), each being composed of one or two heavy chains and several light chains (Ruppel and Spudich, 1996).

Myosin class II; the basic structure of which is shown in Figure 2, are the class found in the sarcomeres and consist of two myosin heavy chains (MHC), two myosin regulatory chains (RLC) and two myosin essential light chains (ELC) (Weeds and Lowey, 1971). The MHC consist of a head region which is responsible for binding filamentous-actin (F-actin) and adenosine-5'-triphosphate (ATP), the neck region which transduces force and the variable tail region which mediates interactions of myosin and filament formation (Henderson et al., 2017, Sellers, 2000). Proteolysis of class II myosins by trypsin originally produced two fragments which were named heavy meromyosin (HMM) and light meromyosin (LMM) (Szent-Gyorgyi, 1953). The HMM region was subsequently found to be further cleaved into subfragment 1 (S1), which is the head domain, and subfragment 2 (S2) which is the region between the head domain and the LMM (Mueller and Perry, 1962). The myosin light chains (MLC), regulate the motor function of the MHC via binding to either the myosin lever arm or the 'light chain domain' close to the head-rod junction (Henderson et al., 2017).

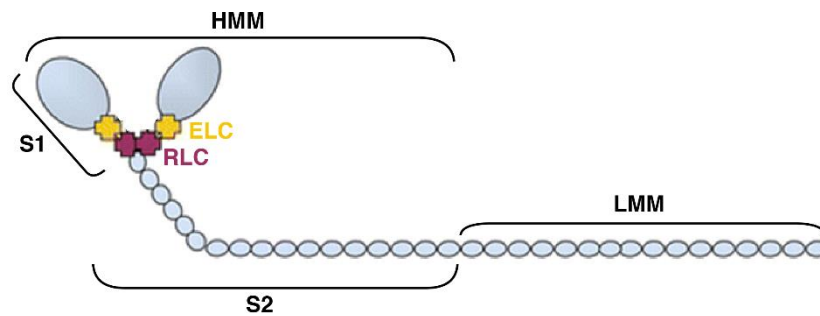


Figure 2 – Basic outline of the myosin molecule

Outline of the myosin class II molecule, dimers of which make up the basis of the thick filaments in the sarcomeres of striated muscle. The filaments are made up of two myosin heavy chains, dimers of which arrange as bipolar filaments, two regulatory light chains (RLC; purple) and two essential light chains (ELC; yellow). The MHC can be further divided into the heavy meromyosin (HMM) region which consists of the S1 or head region which is responsible for binding actin and ATP, the S2 which contains the neck region which transduces force and the light meromyosin (LMM) region. Figure adapted from (Pfuhl and Gautel, 2012)

1.1.2.2. Myosin Binding Proteins

Myosin binding proteins are a component of the thick filament which are located in the C-zone, in the inner two-thirds of the A-band, at a 43 nm periodicity (Bennett et al., 1986). The Myosin binding protein family consists of two proteins: myosin binding protein C (MyBP-C), which will be extensively discussed later, and myosin binding protein H, which is predominantly expressed in the Purkinje fibres in the heart (Alyonycheva et al., 1997). These proteins interact with the thick, thin and elastic filaments and are thought to be involved in filament assembly as well as in regulating contraction (Starr and Offer, 1978, Moos et al., 1978, Squire et al., 2003, Furst et al., 1992, Kampourakis et al., 2014).

1.1.3. The thin filament

The thin filaments are made up of pearlstring-like strands of F-actin molecules associated with tropomyosin and the troponin complex (Tardiff, 2011), outlined in figure 3, and are anchored via alpha-actinin at the Z-disc (Ribeiro Ede et al., 2014).

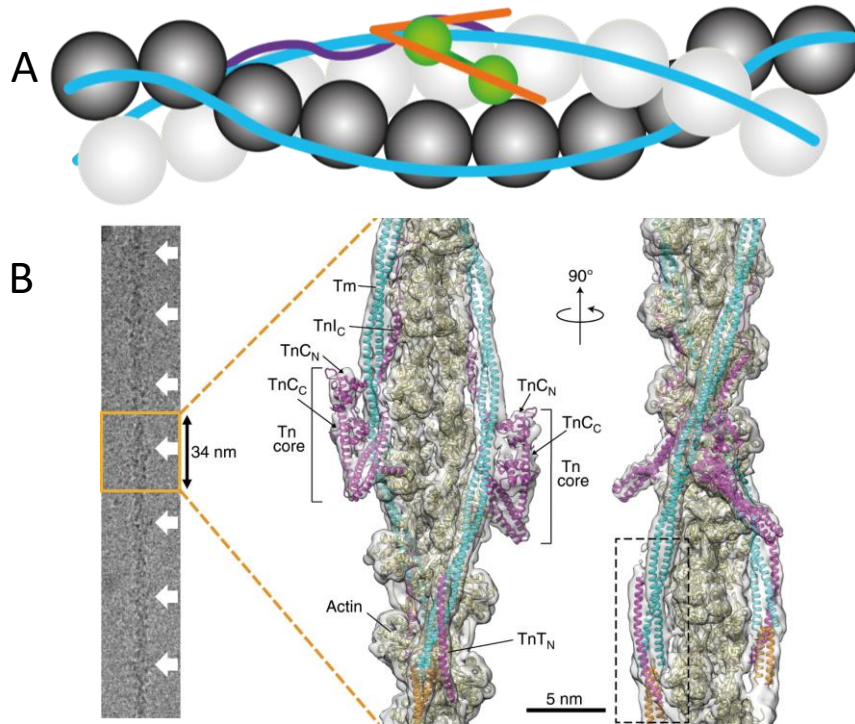


Figure 3 - Basic outline of the thin filament structure

(A) Cartoon of the basic structure of the thin filaments. The thin filaments are mainly made up of pearlstring-like strands of globular-actin (G-actin, MW = 42 kDa) molecules (grey circles). Dimers of tropomyosin (MW = 33kDa; blue) lie along the length of the thin filaments and forms a complex with troponin, which consists of three subunits: troponin C (TnC, MW = 18kDa; green), troponin I (TnI, MW= 23kDa; orange) and troponin T (TnT, MW = 34 kDa; purple), to regulate the interaction between the thin and thick filaments to control actomyosin crossbridge cycling. TnC undergoes a conformational change upon binding of calcium, which allows it to interact with TnI. TnI has inhibitory properties and has been shown to block actomyosin ATPase activity in vitro. Binding of TnC to TnI allows the dissociation of TnI from actin, which shifts tropomyosin and allows myosin to weakly bind to actin. The main role of TnT is thought to be to act as an anchor for the troponin complex to tropomyosin. (B) cryoEM image of the thin filament and a model fitted onto the cryoEM density map from two different viewpoints. 12 actin subunits are shown with a pair of tropomyosin coiled coils and a pair of the troponin complex. (Actin = beige; Troponin = purple; Tropomyosin = light blue and orange. Adapted from (Yamada et al., 2020)

1.1.3.1. Actin

Actin is responsible for about 20% of the entire mass of striated muscle and is the major component of the thin filaments. In solution actin is globular (G-actin); individual molecules which interact to form filamentous polymers (F-actin) which appear as two long-pitch helically entwined pearl strings (Hanson and Lowy, 1963). These polymers are polar; with barbed or positive ends which insert into the Z-disc and where new actin monomers are incorporated and pointed or minus ends which run into the M-band (Ehler, 2018, Huxley, 1963, Henderson et al., 2017).

1.1.3.2. Tropomyosin

Tropomyosin (MW = 33 kDa) is made up of two α -helical chains; dimers of tropomyosin lie head to tail along the length of the thin filaments and interact with the positively charged grooves of F-actin (von der Ecken et al., 2015), with each tropomyosin molecule binding 7 actin monomers (Ebashi, 1980). Tropomyosin forms a complex with troponins through which it regulates the interactions between the thin and thick filaments and thus acts to control actomyosin cross-bridge cycling (Parry and Squire, 1973).

1.1.3.3. Troponin complex

As stated above the troponin complex acts with tropomyosin to regulate actomyosin cross-bridge cycling. The troponin complex consists of three different subunits: Troponin C (TnC; MW = 18 kDa), Troponin I (TnI; MW = 23 kDa) and troponin T (TnT; MW = 34 kDa). TnC undergoes a conformational change upon the binding of calcium, which allows it to interact with TnI (McKay et al., 1997). TnI has inhibitory properties and has been shown to block actomyosin ATPase activity in vitro. Binding of TnC to TnI allows the dissociation of TnI from actin, which shifts tropomyosin and allows myosin to weakly bind to actin (Leavis and

Gergely, 1984). The main role of TnT was thought to be to act as an anchor for the troponin complex to tropomyosin (Franklin et al., 2012, Tanokura et al., 1983), however it is thought to also have a role in regulating actomyosin ATPase activity, calcium sensitivity and force generation in the sarcomere (Willott et al., 2010, Potter et al., 1995).

1.1.4. The elastic filament

The elastic filament is made up of titin (also called connectin) (Maruyama et al., 1976, Wang et al., 1979), which is the largest known protein in the human genome (Chauveau et al., 2014). It stretches half of the sarcomere from the Z-disc to the M-band (Fürost et al., 1988) and as such is important for the stability of the sarcomere. Titin has multiple different binding partners spanning along the length of the molecule.

Titin's N-terminus anchors the protein into the Z-disc and also contains binding sites for multiple proteins including actin (Linke et al., 1997), α -actinin (Young et al., 1998), tropomyosin (Raynaud et al., 2004) and obscurin (Young et al., 2001) suggesting an involvement in supporting the structure of the sarcomere. Furthermore, it is thought to contain binding sites for calpain, suggesting it also plays a role in mechanosensing, as calpain is thought to be involved in the sarcomere stretch response, although it is predominantly expressed in skeletal muscle (Ojima et al., 2007, Raynaud et al., 2005).

The region of titin corresponding to the I band contains the PEVK region, which consists of almost 70% proline (P), glutamic acid (E), valine (V) and lysine (K) residues, and is the main source of the elastic properties of titin (Labeit and Kolmerer, 1995). As such this region is one of the key components for passive tension, which is the lengthening of muscle in the absence of contractile force and is necessary for response to any changes in the mechanical demand (Trombitas et al., 2000).

The A-band region of titin links titin to the thick filament via binding sites for both MyBP-C and the myosin tail domains (Soteriou et al., 1993, Houmeida et al., 1995). This region of titin is composed of super repeats of seven fibronectin type III and four immunoglobulin like domains (Soteriou et al., 1993), with eleven of these super repeats corresponding to the thick filament repeats which is thought to confer the periodicity of myosin and MyBP-C in the C-zone (Freiburg and Gautel, 1996, Bennett et al., 2020). The M-band region of titin contains the titin kinase domain and is involved in signalling (Lange et al., 2005) as well as interacting with myomesin, stabilising titin in the thick filament (Obermann et al., 1995).

1.1.5. Muscle contraction

The 'sliding filament theory' was first put forward in 1954 in order to hypothesise how muscle contraction occurs (Huxley and Niedergerke, 1954, Huxley and Hanson, 1954). This theory states that contraction is driven by the sliding of the thin filaments along the thick filaments. Depolarisation of the cardiac cell membrane causes activation of L-type voltage operated calcium channels in the T-tubules leading to an influx of Ca^{2+} . Ca^{2+} then binds to and activates ryanodine receptor channels in the sarcoplasmic reticulum leading to a calcium-induced calcium release causing an increase in cytosolic Ca^{2+} (Eisner et al., 2003) (Figure 4).

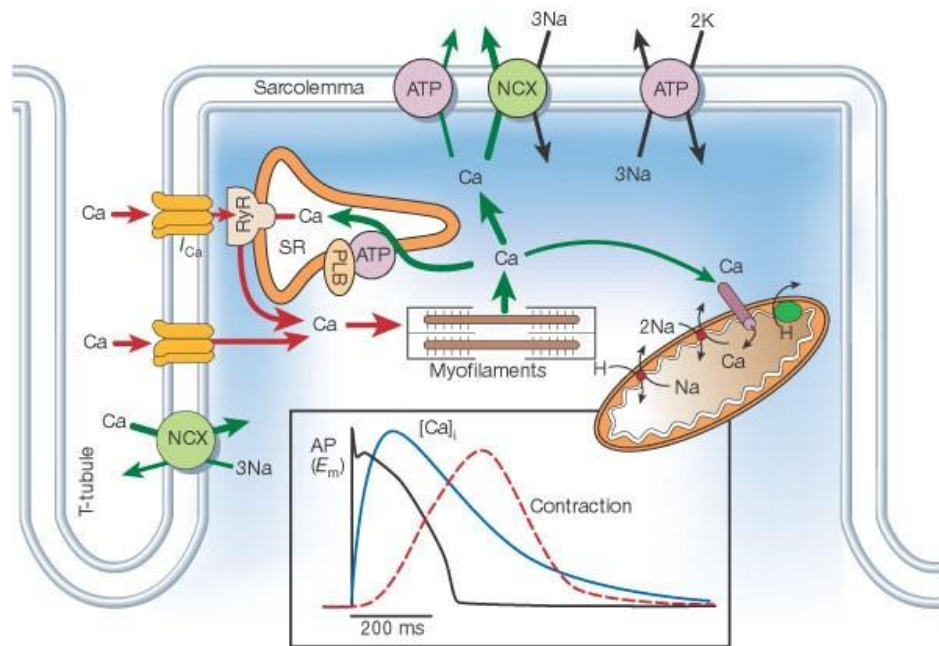


Figure 4 – Scheme of the excitation contraction coupling in the cardiac muscle

Overview of the excitation contraction coupling which occurs in the cardiomyocytes. Following depolarisation of the cardiac membrane, L-type voltage operated calcium channels in the T-tubules are activated. This leads to an influx of Ca^{2+} which bind to ryanodine receptors (RyR) on the sarcoplasmic reticulum (SR) causing calcium-induced calcium release and an increase in cytosolic Ca^{2+} . This Ca^{2+} then binds to the tropomyosin-troponin complex leading to conformational changes which allows formation of actomyosin crossbridges. From (Bers, 2002).

Ca^{2+} then binds troponin C leading to conformational changes in troponin I and troponin T and ultimately tropomyosin. This causes a transition from a 'blocked' to 'closed' state of tropomyosin exposing myosin-binding sites on actin. The myosin head, with adenosine diphosphate (ADP) and inorganic phosphate (P_i), can then interact with the exposed actin binding sites causing a change to the 'open' state of tropomyosin (Boussouf and Geeves, 2007) and exposing neighbouring myosin-binding sites allowing cooperative binding of further myosin heads (Craig and Lehman, 2001, Trybus and Taylor, 1980). This is followed by release of P_i and the 'working stroke', causing shortening of the sarcomere and/or force

development. ATP then binds to the myosin head causing detachment from actin, the ATP is then hydrolysed by myosin, allowing the cycle to restart (Lymn and Taylor, 1971, Huxley, 1957) (Figure 5).

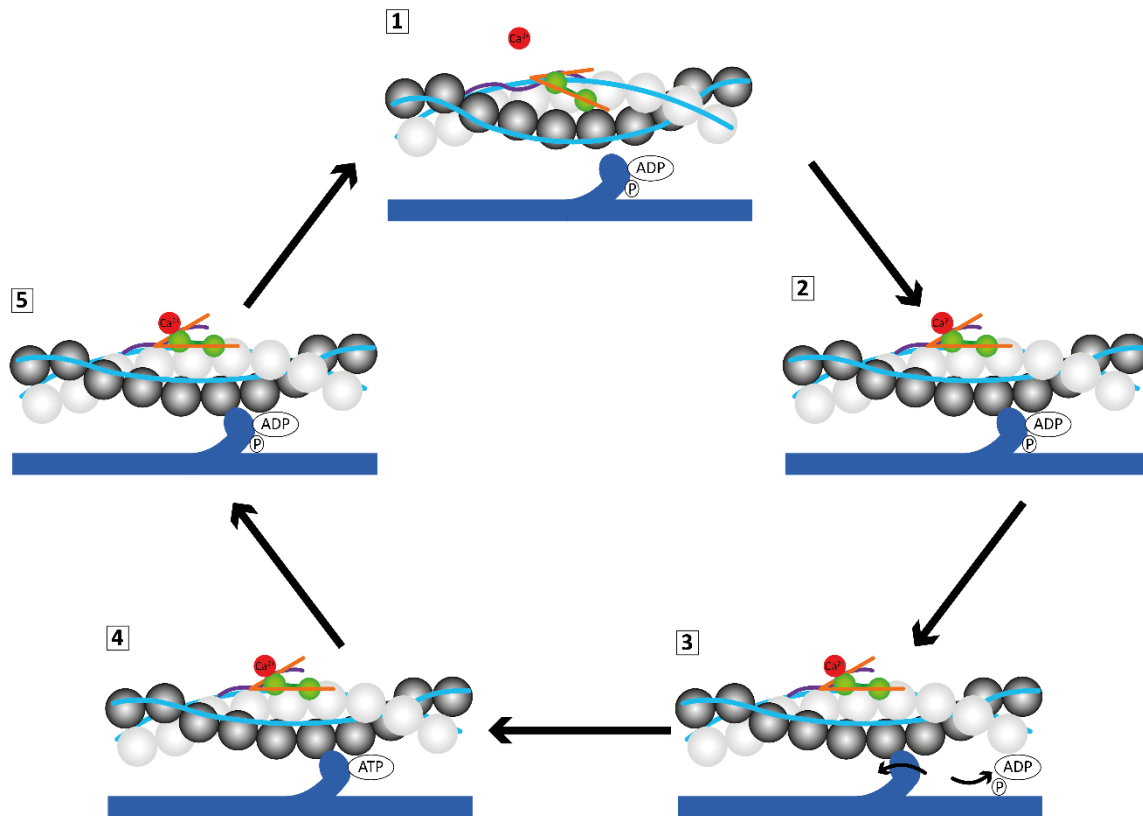


Figure 5 – Basic scheme of the process of muscle contraction

Myosin binding sites on actin (white & grey circles) are blocked by tropomyosin (light blue lines); myosin heads are sequestered to the surface of the thick filament (dark blue). Following depolarisation of the cardiac cell membrane there is an influx of Ca^{2+} (1). Calcium binds to troponin C (green) causing conformational changes in troponin I (orange) and troponin T (purple). This causes a conformational change to tropomyosin and exposes the myosin binding sites on actin allowing the myosin heads to bind to actin along with adenosine diphosphate (ADP) and inorganic phosphate (Pi) (2). The 'working stroke' then occurs causing shortening of the sarcomere (3). Adenosine triphosphate (ATP) then binds to the myosin head causing detachment from actin (4), the ATP is hydrolysed by myosin (5), allowing the cycle to restart.

Activation of the thin filament, as mentioned above, occurs via calcium binding to troponin leading to a structural change in tropomyosin. However, the thin filaments are not activated fully by calcium alone; myosin is also required for full activation. It has been shown that binding of myosin to the thin filament propagates further myosin binding by generating local fully active regions or “regulatory units”. Experiments suggest that two myosin heads are required in order to activate the regulatory unit on the thin filament, following which further myosins can bind to a total of 11 myosins per regulatory unit. These units have been shown to fuse to create larger regions of activation and it is thought that with sufficient levels of both calcium and myosin present, these regions would grow and lead to the entire thin filament being activated. Once myosin and calcium can no longer sustain the thin filament in its active state, activation was seen to “catastrophically collapse”, deactivating the thin filament (Desai et al., 2015).

Whilst the fact that the thin filaments are not able to be fully activated by only calcium but also require myosin is generally accepted, further detail about the regulation of the thin filaments has proved more controversial. Theories on the regulation of the thin filaments can broadly be separated into two different models: the two-state model and the three-state model.

The two-state model was originally put forward by Hill et al. in 1980 (Hill et al., 1980) and was subsequently further developed by Chalovich and Eisenberg (Chalovich and Eisenberg, 1982). This model suggested that the thin filament could exist in two states: an inactivated state with weaker binding affinities for myosin or an activated state with tighter myosin binding affinity (Hill et al., 1980). This model has more recently been added to by Heeley et al. (Heeley et al., 2019) who have carried out experiments measuring the rate of release of

P_i from myosin crossbridges in the presence of actin. These experiments found that the release of P_i was accelerated by actin and that the degree of acceleration was dependent on both the presence of Ca^{2+} and myosin heads being tightly bound. This led to the group putting forward a model in which actin is either active and able to promote the release of P_i from myosin or inactive and unable to promote P_i release, with both Ca^{2+} and myosin being able to bind both these states of actin and rebalance the equilibrium toward the active state (Heeley et al., 2019).

The three-state model of thin filament regulation was put forward by McKillop and Geeves in 1993 (McKillop and Geeves, 1993) and states the thin filament can be in three different states: the blocked state, the closed or calcium-induced state and the open or myosin induced state. In the blocked state the thin filament is unable to bind to the myosin head, with the majority of the myosin binding sites on actin being blocked by tropomyosin. When the thin filaments are in the closed state they can only bind the myosin head weakly, with Ca^{2+} removing the interaction between TnI and the actin-tropomyosin complex, allowing tropomyosin to revert to its most favourable binding sites on actin and exposing the majority of the myosin binding sites on actin. In this state the binding of myosin to actin is limited as complete binding is blocked by tropomyosin. In the open state, the myosin head can both bind and undergo isomerisation to a more strongly bound rigor-like conformation (McKillop and Geeves, 1993, Geeves et al., 2019).

Whilst the experiments by Heeley et al. showed that strong binding of the myosin heads alone was unable to fully activate the thin filament contrary to the three-state model (Heeley et al., 2019), the two-state model has a lack of structural support. The three-state model on the other hand is supported by multiple structural studies which have identified

the existence of three structural states of the thin filaments (Lorenz et al., 1995, Vibert et al., 1997, Poole et al., 2006, Bershtitsky et al., 2017).

Whilst the above focuses on the activation state of the thin filament in regulating muscle contraction, it is also important to consider the regulation of the thick filament. As there is no protein equivalent to tropomyosin which is capable of blocking the binding of all of the myosin heads to actin, it is hypothesised that the “off” state of the thick filament is achieved via a so called interacting heads motif (IHM) in which two myosin heads form an asymmetric structure with one head, termed ‘blocked’, sterically blocked from binding to actin by the other head, termed the ‘free’ head as its actin-binding activity is not blocked but its ATPase activity is inhibited. This IHM structure has been shown to be present in native thick filaments of tarantula striated muscle (Woodhead et al., 2005) and mouse cardiac muscle (Zoghbi et al., 2008), as well as in isolated myosin molecules from several different species and different types of striated muscle (Jung et al., 2008a, Jung et al., 2008b).

Multiple protein-protein interactions of myosin with itself and other proteins (intramolecular head-head and head-tail, intermolecular interactions, MyBP-C and titin) are thought to form a helical lattice stabilising the myosin heads on the thick filament surface (Irving, 2017). In this conformation the myosin heads are unable to bind to actin or hydrolyse ATP (Hooijman et al., 2011, Stewart et al., 2010). Because of this the IHM has been proposed to be the structural state of the so called super-relaxed state (SRX) of muscle, which is a biochemical state where the ATPase state of the myosin heads is suppressed and there is a lower population of heads capable of generating force (Hooijman et al., 2011, McNamara et al., 2016, Naber et al., 2011). Disruption of the SRX state leads to formation of the disordered relaxed state (DRX) where the myosin heads are released from the thick

filament backbone and are available to bind actin (Myburgh et al., 1995, McNamara et al., 2017). The IHM state is known to be destabilised upon the phosphorylation of the RLC which forms part of the myosin head (Alamo et al., 2008), with studies showing that RLC phosphorylation enhances active force, increases calcium sensitivity and alters the thick filament structure with myosin head domains becoming more perpendicular to the thick filament axis, leaving the helical folded off state. This structural change to the myosin heads was also shown to occur in unphosphorylated myosin head domains when neighbouring head domains were phosphorylated, suggesting that signalling occurs between the head domains (Kampourakis et al., 2016).

Another way in which it is thought that the thick filament is regulated is via mechanosensing, as it has been shown that the thick filaments can be switched on by mechanical stress independent of the calcium state (Fusi et al., 2016, Linari et al., 2015).

It is very probable that the regulation of the thin and thick filaments are interconnected and one such protein which could act as a mediator between the two systems is MyBP-C (Irving, 2017). As will be discussed in further detail later, the N-terminus of MyBP-C is capable of binding both the thin and thick filaments (Luther et al., 2011, Shaffer et al., 2009), and it is thought that binding to the thin filament stabilises them in their 'on' state, whilst binding to the thick filament stabilises it in the 'off' state (Kampourakis et al., 2014).

1.2. Myosin binding protein C

MyBP-C was first detected in 1973 as an impurity in preparations of skeletal muscle myosin (Offer et al., 1973), which suggested it had myosin binding properties. However, it was subsequently found to also bind to actin and titin (Freiburg and Gautel, 1996, Moos, 1981, Moos et al., 1978, Yamamoto, 1986).

Three isoforms of MyBP-C have been identified: slow skeletal, encoded by MYBPC1 on chromosome 12, fast skeletal, encoded by MYBPC2 on chromosome 19, and cardiac MyBP-C (cMyBP-C), encoded by MYBPC3 on chromosome 11 (Weber et al., 1993, Carrier et al., 1997). All three isoforms consist of seven immunoglobulin (Ig)-like domains and three fibronectin (FN)-like domains, designated C1-C10, a proline-alanine (Pro-Ala) rich region before the C1 domain and a linker region between the C1 and C2 domains termed the M-motif.

The cardiac isoform (cMyBP-C), which is shown in figure 6, shows high amounts of similarity with the skeletal isoforms (54.4% identity with fast skeletal and 52.4% with slow skeletal) (Weber et al., 1993, Yasuda et al., 1995). However, it does differ from the skeletal isoforms with an additional Ig-like domain at the N-terminus, termed C0, a nine residue insertion in the M-motif including four putative phosphorylation sites and an additional 28 amino acids in the C5 domain (Freiburg and Gautel, 1996, Gautel et al., 1995, Yasuda et al., 1995).

1.2.1. MyBP-C structure

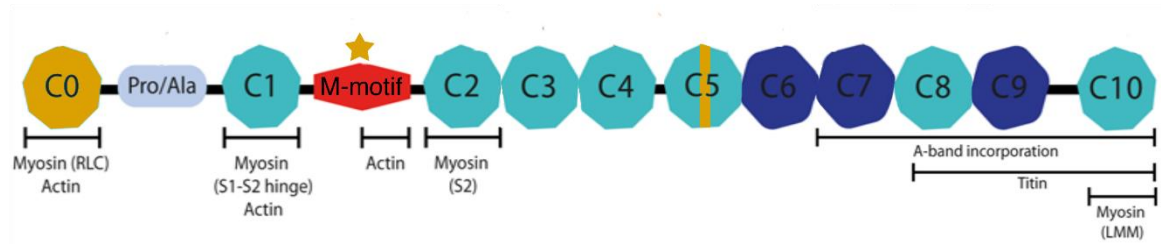


Figure 6 – Structure of cMyBP-C

Three isoforms of Myosin binding protein C have been identified: slow skeletal, fast skeletal and cardiac. All three isoforms consist of the same basic structure of immunoglobulin-like (octagon) or fibronectin-like (hexagon) domains termed C1-C10, a proline-alanine rich domain and an M-motif. The cardiac isoform differs from the skeletal isoforms with an extra N-terminal domain, termed C0, a 28 amino acid insert in the C5 domain and an additional 9 residues in the M-motif, including putative phosphorylation sites (shown in gold). The N-terminus is thought to contain binding sites for the myosin regulatory light chain, the S1-S2 hinge and the S2 fragment as well as actin. The C-terminus binds light meromyosin and titin, as well as being important for the proper incorporation into the sarcomeric A-band.

So far structural studies of the full length of the cMyBP-C protein have not been carried out, due to the size and flexibility of the protein however, rotary shadowing electron microscopy has been able to elucidate information about the overall structure of the cMyBP-C molecules (Previs et al., 2016). These experiments showed that some molecules exhibited a V-like conformation, which has also previously been seen for the skeletal isoform of MyBP-C (Bennett et al., 1985) and chicken cMyBP-C (Hartzell and Sale, 1985), whilst other molecules showed a second hinge point. These hinge points are thought to be located in the M-motif and near to the C5 domain, either between the C4 and C5 domains or the C5 and C6 domains. The former is hypothesised to have a role in altering the interaction of the N-terminal domains with actin or myosin, whilst the latter may be required for extension of cMyBP-C away from the thick filament (Previs et al., 2016).

The structure of individual domains C0, C1, C2, C5 and parts of the M-motif have been resolved (Ratti et al., 2011, Ababou et al., 2008, Ababou et al., 2007, Idowu et al., 2003, Howarth et al., 2012, Govada et al., 2008), as well as studies being carried out using small-angle X-ray scattering (SAXS) to analyse the N-terminal C0-C2 domain organisation (Jeffries et al., 2008, Nadvi et al., 2016) as well as various different fragments spanning the C0-C4 domains and the C5-C7 domains (Jeffries et al., 2011).

These structural studies have confirmed N-terminal myosin binding sites, with C0 proven to bind the myosin RLC (Ratti et al., 2011), whilst a binding site for C1 was found in the immediate vicinity of the S1-S2 hinge (Ababou et al., 2008). Furthermore, the C2 domain was found to have a similar structure to the protein telokin, which is a myosin binding fragment of the myosin light chain kinase and also binds myosin S2 with low affinity but high specificity (Ababou et al., 2007).

Structural studies have also indicated that the M-motif consists of two subdomains, a largely disordered and flexible N-terminal region containing three of the putative phosphorylation domains and an ordered and folded C-terminal region which possibly contains the actin binding site (Howarth et al., 2012).

Molecular dynamics simulations have also been carried out on the C1 domain revealing the formation of a potential binding pocket when MyBP-C undergoes mechanical force such as it would be subjected to during muscle contraction. This raises the possibility that other domains could also contain binding pockets that are only formed under mechanical strain seen in situ (Tiberti et al., 2019).

Structural analysis of the C5 domain of cMyBP-C is of particular interest, due to the cardiac specific inserts found in this domain. Studies have shown that two inserts are present in the cardiac isoform: one in the linker region between the C4 and C5 domain, which is thought to be required for correct folding of the C5 domain, whilst the second is an elongation of the CD loop of the C5 domain, which was found to be unstructured and may explain the decreased stability of the C5 domain compared to other domains of the protein (Idowu et al., 2003).

1.2.2. Localisation of MyBP-C in the sarcomere

MyBP-C is localised to the inner two thirds of the A-band of the sarcomere, termed the C-zone. MyBP-C appears as seven or nine transverse stripes, for the skeletal and cardiac isoforms respectively, spaced roughly 43nm apart. This periodicity indicates it localises to each third level of myosin heads (Craig and Offer, 1976) with this limit possibly being set by the presence of titin; as it is thought that two out of three of the myosin binding sites of MyBP-C in the C-zone are blocked by the 6 titin molecules located within each half thick filament (Liversage et al., 2001, Bennett et al., 2020, Freiburg and Gautel, 1996). Various models for the three dimensional sarcomeric organisation of MyBP-C have been proposed, with the two main models being a linear or collar-like arrangement which are shown in figure 7.

The collar model arose from yeast two-hybrid (Y2H) studies that indicated an interaction between the C5 and C8 and C7 and C10 domains, implying an interaction between C5C6C7 and C8C9C10 in a staggered parallel arrangement that encircles the thick filament backbone (Moolman-Smook et al., 2002).

On the other hand, both X-ray pattern modelling (Squire et al., 2003) and electron microscopy (EM) of negatively stained cardiac thick filaments (Al-Khayat et al., 2013, Zoghbi et al., 2008) suggests the C-terminal domains (C8-C10) are arranged linearly along the thick filament, supporting the linear model. Furthermore, results from a study using domain-specific immune-EM were also inconsistent with the collar model (Lee et al., 2015).

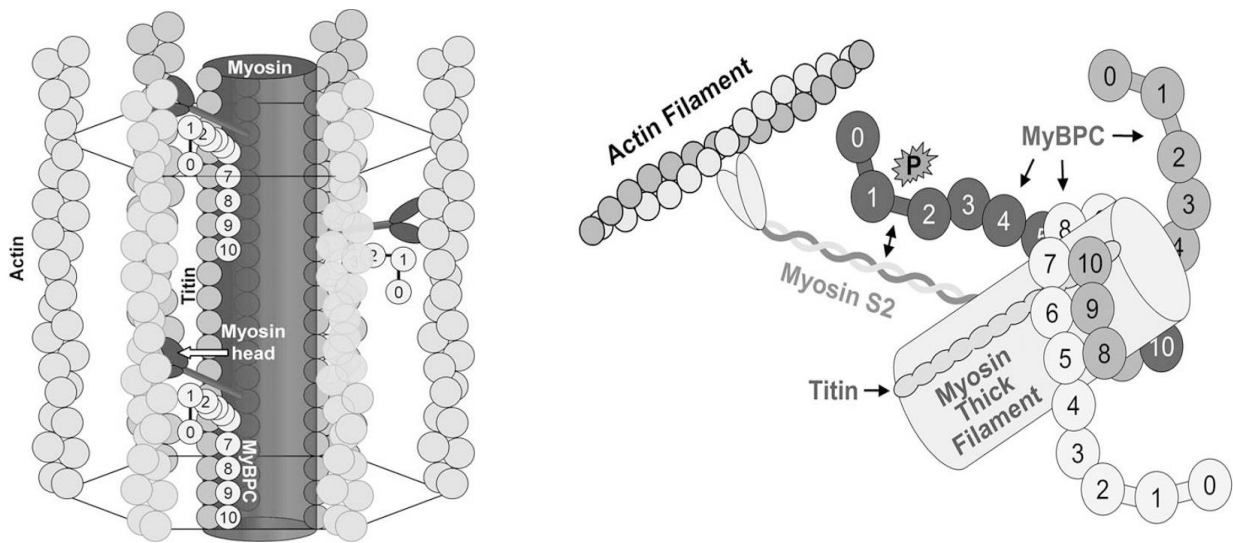


Figure 7 - Proposed models of the 3-dimensional sarcomeric organisation of Myosin binding protein C

Two different 3-dimensional sarcomeric organisations of myosin binding protein C (MyBP-C) have been proposed: the linear conformation (left) and the collar conformation (right). The linear conformation, based on studies utilising X-ray pattern modelling and electron microscopy, states that the four C-terminal domains (C7-C10) lay linearly along the thick filament backbone. The collar model is based on yeast two hybrid studies that found an interaction between the C5 and C8 domains and the C7 and C10 domains leading to the hypothesis that the C5-C7 domains bind to the C8-C10 domains in a staggered parallel manner to encircle the thick filament. Figure modified from (Moolman-Smook et al., 2002, Squire et al., 2003) by (Oakley et al., 2004)

1.2.3. Binding partners of MyBP-C

As the name suggests, MyBP-C binds to myosin, with binding sites for the LMM in the C10 domain (Okagaki et al., 1993), and sites for the myosin S2 and myosin RLC in the N-terminal C0-C2 region (Gruen and Gautel, 1999, Ratti et al., 2011). MyBP-C has also more recently been shown to bind to S1 myosin, with microscale thermophoresis experiments showing that MyBP-C bound a shortened version of S1 which did not contain the RLC. The precise domains involved in this interaction were not fully elucidated, however they did show that binding was 2x weaker for a fragment containing only the C0-C2 domains of MyBP-C compared to the full length protein, suggesting binding was not solely due to the N-terminal domains (Nag et al., 2017).

The C-terminus of MyBP-C (C8-C10) has also been shown to contain a binding domain for titin (Freiburg and Gautel, 1996). The interaction with titin, along with binding to the LMM, is thought to be responsible for anchoring MyBP-C to the backbone of the thick filament (Okagaki et al., 1993, Freiburg and Gautel, 1996).

It was subsequently found that MyBP-C was also able to bind actin, with studies showing an interaction with F-actin, isolated thin filaments and the I-bands of sarcomeres (Moos, 1981, Moos et al., 1978, Yamamoto, 1986), with possible binding sites in the M-motif (Bezold et al., 2013), the C0 domain (Kulikovskaya et al., 2003a, Lu et al., 2011, Orlova et al., 2011, Whitten et al., 2008), the Pro-Ala rich region (Squire et al., 2003) and the C1 domain (Bhuiyan et al., 2012, Shaffer et al., 2009). This binding of MyBP-C to actin has been shown to occur not only in vitro but also in intact skeletal muscle cells under resting conditions utilising electron tomography (Luther et al., 2011).

Calmodulin has also been found to interact with MyBP-C and it is thought to regulate binding and unbinding of MyBP-C to the myosin S2 fragment, either by initiating phosphorylation of MyBP-C via Ca^{2+} -dependent calmodulin kinase II (CaMKII) or phosphorylation of the RLC by myosin light chain kinase (MLCK) (Lu et al., 2012).

MyBP-C is also known to bind four and a half LIM protein 1 (FHL1) and is required for incorporation of FHL1 into the thick filament. (McGrath et al., 2006). FHL1 has been shown to activate downstream signalling pathways and acts as a mechano-sensor in the physiological regulation of the response to hypertrophic stress (Sheikh et al., 2008), implicating MyBP-C may also have a mechanosensory role.

The formin Fhod3 has been shown to strongly bind cMyBP-C, and this interaction has been shown to be required for the localisation of Fhod3 to the C-zone (Matsuyama et al., 2018). Fhod3 is known to have a key role in regulating the assembly of actin (Ushijima et al., 2018, Iskratsch et al., 2010).

1.2.4. Function of MyBP-C

cMyBP-C is thought to be key for normal cardiac muscle contraction, with studies showing its involvement in cardiac stress pathways (Sadayappan and de Tombe, 2012), whilst mutations in the gene which encodes cMyBP-C (MYBPC3) cause hypertrophic cardiomyopathy (HCM) (Harris et al., 2011). Studies carried out using knock-out (KO) mouse models, which will be discussed more extensively later, also show impairment in cardiac function (Harris et al., 2002), with these KO mice being viable but showing an abnormal sarcomere structure and ultimately developing HCM at about three months of age (Carrier et al., 2004).

MyBP-C is thought to be involved in myofilament assembly, as it has been shown that when purified myosin polymerises in vitro with MyBP-C the thick filaments are seen to be slightly longer and more regular than in the absence of MyBP-C (Koretz, 1979, Davis, 1988). Furthermore, when MyBP-C is present the concentration of myosin required for polymerisation is decreased suggesting MyBP-C may regulate the assembly of the thick filaments (Davis, 1988). Truncated forms of MyBP-C that lack the myosin binding domain have also been shown to inhibit myofibrillogenesis (Gilbert et al., 1996). Lastly, the timing of the appearance of MyBP-C during myofibrillogenesis, which has been shown to occur relatively late (Schultheiss et al., 1990, Rhee et al., 1994), first being detected as nascent or non-striated myofibrils transform into mature myofibrils, indicate it may have a role in the assembly of the myofilaments (Rhee et al., 1994, Lin et al., 1994, Gilbert et al., 1996).

These experiments were, however, carried out using in vitro methods, and as mentioned KO mouse models have shown that the presence of cMyBP-C is not necessary for viability (Harris et al., 2002, Carrier et al., 2004). Therefore, the role of cMyBP-C in myofibrillogenesis in vivo may be of less significance than these findings suggest.

The major known role for MyBP-C is its regulation of cross-bridge cycling; by maintaining the normal structures of the myosin and actin filaments, regulating the myosin and actin interaction rate and modulating the myosin ATPase activity (Ackermann and Kontogianni-Konstantopoulos, 2011). However, the molecular details of cMyBP-C's regulatory role during the cardiac cycle are still being debated.

It has been shown that the N-terminal region of MyBP-C is able to bind both the thin and thick filaments with equal affinity (Shaffer et al., 2009, Luther et al., 2011) and this ability to interact with both systems would allow it to be key in regulating the actomyosin cross-

bridge cycling. Multiple studies have shown that the N-terminal domains of MyBP-C increase calcium sensitivity of tension and filament sliding speed (i.e. activate the thin filament) when Ca^{2+} levels are sub-maximal, whilst when the filaments are maximally activated the N-terminus leads to a decrease in sliding velocity (Harris et al., 2004, Herron et al., 2006, Razumova et al., 2008, Razumova et al., 2006b). The precise regions within the N-terminus that are necessary for these effects are however under debate, with one study showing the C1 and M domains were effective at activating force in the absence of Ca^{2+} (Razumova et al., 2008) whilst another study suggested that the pro-ala rich region was the critical region for activation (Herron et al., 2006). These differences could be due to species differences that occur in the sequences of the N-terminal region, as the mouse and human sequences were utilised in the different studies respectively (Razumova et al., 2008, Herron et al., 2006).

The binding of MyBP-C to the S2 fragment of myosin has been suggested to lead to a restriction of cross-bridge formation by inhibiting the outward movement of myosin head domains towards the thin filament (Starr and Offer, 1978), whilst another study states that MyBP-C connects the S2 fragment to the LMM acting as a drag on the cross-bridge cycling (Hofmann et al., 1991). Furthermore, the N-terminal domains of cMyBP-C have been shown to have an inhibitory effect on myosin ATPase activation by actin at high concentrations, whilst at low levels they activate the ATPase activity (Belknap et al., 2014). MyBP-C is also thought to have a role in controlling the number of myosin heads available as it has been found, amongst other factors including the phosphorylation of the RLC, to move myosin from a so called SRX state where the myosin heads are bound to the thick filament backbone in a suppressed ATPase “off” state with a lower population of heads capable of

generating force (Hooijman et al., 2011, McNamara et al., 2016, Naber et al., 2011), into the DRX state where the myosin heads are not bound to the thick filament backbone and are available to weakly bind actin (Myburgh et al., 1995, McNamara et al., 2017). One way in which MyBP-C may achieve this is through its interaction with myosin S1, as it has been proposed that binding of a protein, possibly MyBP-C, to a relatively flat surface of the myosin head termed the myosin mesa could act to hold the myosin heads in their inactive form (Spudich, 2015).

Other studies have used structural evidence to propose a model of MyBP-C binding to the thin filament in positions which overlap with binding sites for the myosin heads and sterically clash with tropomyosin binding when in the low Ca^{2+} state (Whitten et al., 2008, Mun et al., 2011). Further studies using cMyBP-C have shown that these effects are dependent on the specific domains, with the longer C0-C3 fragment being able to displace tropomyosin and increase the thin filaments sensitivity to Ca^{2+} whilst shorter C0C1 and C0C1 with the first 17 residues of the M-motif were only capable of binding the thin filament without having any effect on Ca^{2+} sensitivity (Mun et al., 2014, Inchingolo et al., 2019). It has also been suggested that this binding of the N-terminus of cMyBP-C to the thin filaments occurs in two different manners depending on the calcium state. At low Ca^{2+} concentrations, it has been shown that the N-terminus binds weakly and diffuses from actin monomer to monomer until specific binding sites are found which allow for cMyBP-C to bind tightly allowing it to shift tropomyosin to its closed position to activate the thin filaments. At high concentrations of Ca^{2+} , when calcium binding to troponin has already shifted tropomyosin to its closed position, it is thought that the N-terminus can compete for myosin binding via tight binding to the thin filament (Inchingolo et al., 2019).

The studies mentioned above suggest the possible role MyBP-C may have on either the thick filament or thin filament in isolation. However the fact that, as mentioned before, the N-terminus is capable of binding both thick and thin filaments with equal affinity (Shaffer et al., 2009, Luther et al., 2011) would suggest it may have a regulatory effect on both systems. One study has shown that when it binds to the thin filaments it causes stabilisation of the thin filaments 'on' state, whilst binding to the thick filament stabilised it in its 'off' state, in which the myosin heads are folded back and unavailable to interact with actin (Kampourakis et al., 2014). Structural models posit that the N-terminus of MyBP-C extends away from the thick filament towards the thin filament (Squire et al., 2004, Moolman-Smook et al., 2002). This along with data from SAXS which shows the C0-M-C1-C2 fragment has a sufficient dimension to span the interfilament distance, presents evidence that MyBP-C is capable of interacting with both the thin and thick filaments and thus modulate actomyosin cross-bridge cycling (Jeffries et al., 2008).

A more recent study which utilised the high resolution microscopy technique stochastic optical reconstruction microscopy (STORM) suggested that the N-terminus of MyBP-C was biased toward binding to the thin filament in both relaxed and active muscle preparations and that it could bind to the myosin heads but only when the head was in close proximity to the thin filament. The data indicated that under conditions were the myosin heads appeared to be stabilised along the myosin filament surface as little as 10% of N-terminal MyBP-C was bound to the thick filament. This suggests that if MyBP-C is involved in regulating the SRX state possibly by stabilising the IHM, either a very small fraction of the protein is involved or other regions of the protein, such as the central domains, have a role in this stabilisation (Rahmanseresht et al., 2021).

Although a lot has been elucidated about the role MyBP-C plays, there is still a long way to go in order to fully understand its intricate function. Limitations of findings include the fact that many studies utilise isolated fragments and therefore do not take into account the stoichiometry and spatial constraints of the sarcomere. Furthermore, the entire process seems to be an intricate web including multiple different protein interactions which may have to occur in specific order or under specific conditions, therefore it is very difficult to obtain a clear overall picture of MyBP-C's regulatory role.

One method which attempts to overcome the limitation of the lack of spatial constraints seen in many in vitro methods is the recently developed “cut and paste” mouse model by the Harris group (Napierski et al., 2020) in which the COC7 domains can be removed in situ and replaced with either recombinantly expressed COC7, phosphorylated COC7 or subsets of the domains. This study showed that, in line with other experiments, removal of the N-terminal domains led to a decrease in myofilament calcium sensitivity (Belknap et al., 2014, Razumova et al., 2008, Razumova et al., 2006a) and an acceleration of cross-bridge cycling kinetics (Korte et al., 2003, Stelzer et al., 2006). This was reversed following addition of recombinant COC7, but the effect on the calcium sensitivity of the myofilaments was blunted when phosphorylated COC7 was added, again corroborating previous data, which will be described more fully later, suggesting phosphorylation of MyBP-C desensitises the myofilaments to calcium (Gresham and Stelzer, 2016, Colson et al., 2012).

The study also discovered that absence of the N-terminal domains led to sustained spontaneous oscillatory contractions (SPOC), which are defined as alternating cycles of slow sarcomeric shortening followed by rapid relaxation and sarcomeric re-lengthening in the absence of other cyclic changes such as release of calcium from the sarcoplasmic reticulum

(Wolfe et al., 2011). These are thought to be advantageous under periods of increased ionotropic drive in order to increase energetic efficiency. However, dysregulation could contribute to the dysfunction and arrhythmogenesis seen in HCM patients with cMyBP-C haploinsufficiency (Napieriski et al., 2020). This study went on to see whether the first four domains (C0C2) of cMyBP-C would be sufficient to reverse the functional effects seen following removal of C0C7. It was seen that the C0C2 increased the calcium sensitivity and eliminated SPOC but did not reverse the acceleration of crossbridge cycling kinetics. This suggests that whilst the C0C2 domains are sufficient to confer most of the effects of cMyBP-C on force, the C3C7 domains may also be involved in the regulatory role of the protein on cross-bridge kinetics (Napieriski et al., 2020).

It is thought that this new technique could be a useful tool to fill the gap between in vitro methods which lack the stoichiometry and spatial constraints of the sarcomere, and in vivo or ex vivo approaches which are time consuming and costly (Napieriski et al., 2020).

1.2.5. Post-translational modifications of MyBP-C

The human genome consists of about 25,000 genes (Consortium, 2004), however the proteome is much larger consisting of over a million different polypeptides (Jensen, 2004). This increase in diversity is due to both alternative messenger ribonucleic acid (mRNA) splicing and post-translational modifications (PTMs). Over 200 different types of PTM have been identified, with various effects on protein function, the scope of which is widened due to a probable interaction between different types of PTM (Brooks and Gu, 2003, Hunter, 2007). cMyBP-C is thought to undergo various different types of PTMs, some of the known putative sites of which are shown in figure 8.

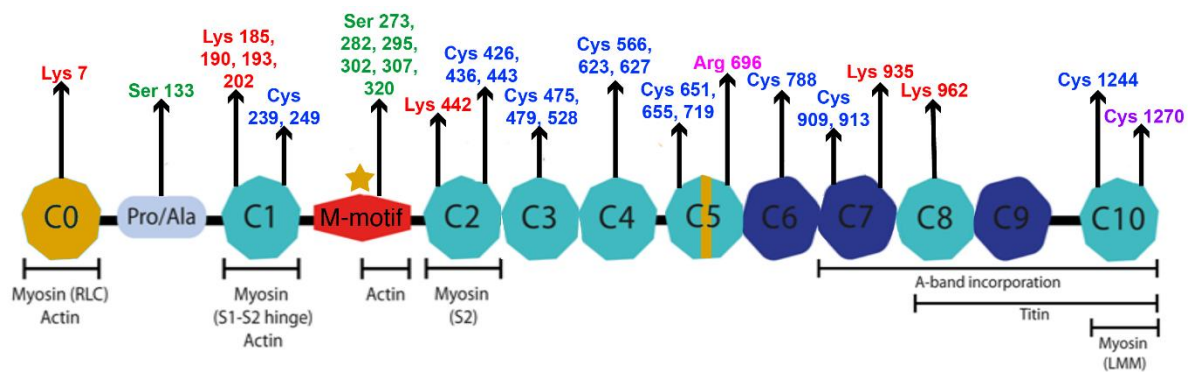


Figure 8 – Putative post translational modifications of cardiac myosin binding protein C

Figure showing the locations of putative sites of different types of post translational modifications of cardiac myosin binding protein C. Red = acetylation, green = phosphorylation, blue = s-glutathiolation, pink = citrullination, purple = s-nitrosylation.

1.2.5.1. Phosphorylation of MyBP-C

The post translational modification of MyBP-C via phosphorylation has been shown to be key in regulating the function of the protein. Phosphorylation is the covalent attachment of a phosphate group by protein kinases to either serine, threonine or tyrosine residues (Humphrey et al., 2015). Three putative phosphorylation domains have been identified in the M-motif of cMyBP-C; Ser273, Ser282 and Ser302 based on the human sequence (Gautel et al., 1995), which are highly conserved between species (James and Robbins, 2011). Further possible phosphorylatable sites in the M-motif (Ser307, Ser295, Ser315, Ser320) (Shaffer, 2009, Yuan et al., 2008) as well as a site in the Pro-Ala rich region (Ser133) (Kuster et al., 2013) have also been found. In all, it is thought that cMyBP-C may contain up to 17 sites that could be phosphorylated (Kooij et al., 2013).

Phosphorylation of cMyBP-C is thought to be necessary for normal cardiac function, with evidence that phosphorylation has a direct effect on the sarcomeric organisation and the

heart's contractile properties (Barefield and Sadayappan, 2010). cMyBP-C phosphorylation has been shown to be decreased both in mouse models of ischemia-reperfusion injury, hypertrophy and heart failure (HF) and human patients with atrial fibrillation, HCM and HF (Barefield and Sadayappan, 2010, Sadayappan et al., 2006). Dephosphorylation of cMyBP-C is associated with its degradation and can lead to disruption of thick filaments, a decrease in actomyosin cross-bridge formation and overall depression in contractility (James and Robbins, 2011).

Phosphorylation of cMyBP-C leads to desensitisation of myofilaments to Ca^{2+} and regulates the actomyosin cross-bridge kinetics (Kunst et al., 2000). It is thought that when cMyBP-C is phosphorylated binding to the S2 fragment of myosin is abolished (Gruen et al., 1999), allowing the thick filament to interact with thin filament proteins including actin and α -tropomyosin (Ge et al., 2009). This allows the myosin heads to extend to the thin filaments resulting in strong interactions between actin and myosin (Schlossarek et al., 2011). However, this phosphorylation-dependent interaction has not yet been shown to occur in vivo.

The different serines appear to show either a preference or specificity for different kinases. The canonical kinase for phosphorylation of cMyBP-C is protein kinase A (PKA) which phosphorylates Ser273, Ser282, Ser302 and Ser307, which are all in the region which binds the S2 segment of myosin (Barefield and Sadayappan, 2010). Protein kinase C (PKC) however only phosphorylates the Ser273 and Ser302 residues, and can further activate protein kinase D (PKD) (Avkiran et al., 2008) which is also thought to phosphorylate cMyBP-C at Ser302 (Bardswell et al., 2010). CaMKII is also able to phosphorylate multiple residues at Ser273, Ser282, Ser302, Ser307 (Squire et al., 2003). Ser282 has not only been found to

be phosphorylated by PKA and CaMKII but also by casein kinase 2 (CK2) and ribosomal S6 kinase (RSK) (Cuello et al., 2011, Gupta and Robbins, 2014). Lastly, Ser133 which is found in the Pro-Ala rich region of cMyBP-C is thought to be phosphorylated by the glycogen synthase kinase 3 β (GSK3 β) (Kuster et al., 2013).

It is thought that as the different phosphorylation sites are substrates for different kinases, cMyBP-C acts as an integrator of various different signalling pathways in order to regulate contractile function. For example, phosphorylation via PKA occurs via β -adrenergic signalling pathways whilst phosphorylation by PKC and PKD show a link with α -adrenergic stimulation (Jensen et al., 2011, Janssen et al., 2018). Furthermore, it has been shown that the phosphorylation state of cMyBP-C (i.e. non-phosphorylated, mono-phosphorylated, bis-phosphorylated, etc.) governs its regulatory function on the cardiac myofilament (Ponnam et al., 2019). Studies have indicated that PKA mono-phosphorylation weakens the interaction between the N-terminus of cMyBP-C with myosin S2, with phosphorylation at Ser282 being the most important target site for regulating this interaction, whilst bis-phosphorylation eliminates the interaction completely, releasing the myosin heads from the thick filament backbone and thus activating the thick filament (Ponnam et al., 2019, Kensler et al., 2017, Colson et al., 2012, Colson et al., 2008). This then allows binding of cMyBP-C to the thin filament stabilising it in its on state. However, the effects of this increased binding to the thin filament is thought to be counteracted by the fact that phosphorylation of the N-terminal regions of cMyBP-C by PKA inhibits its ability to activate the thin filament causing a decrease in calcium sensitivity and leading to relaxation (Colson et al., 2012, Gresham and Stelzer, 2016). Phosphorylation by PKA at Ser282 is thought to be the main regulator of this decrease in activation of the thin filament and this mono-phosphorylation has been shown

to reduce the activating effect by 30%, whilst bis-phosphorylation further reduces the effect and tris-phosphorylation abolishes the activating effect (Ponnam et al., 2019).

Furthermore, there is evidence that phosphorylation of Ser282 may be necessary to allow for phosphorylation of Ser273 and Ser302 by PKC (Mohamed et al., 1998, Schlender and Bean, 1991, McClellan et al., 2001), indicating it may act as a switch rendering other phosphorylatable residues more or less accessible to kinases (James and Robbins, 2011, Sadayappan et al., 2011).

Much less is known about the effects of phosphatases on cMyBP-C, with no known specific phosphatase shown to significantly dephosphorylate cMyBP-C. However, in vitro studies have shown protein phosphatase 1 (PP1), which dephosphorylates many target proteins of PKA, dephosphorylates purified cMyBP-C by 30-40% (Schlender et al., 1987), with similar findings seen in skinned donor heart tissue (Zaremba et al., 2007) and skinned mouse cardiomyocytes (Yang et al., 2008). It has also been shown to lead to a decrease in tris- and tetrakis-phosphorylated and an increase in unphosphorylated cMyBP-C in human heart muscle samples (Copeland et al., 2010). Protein phosphatase 2 (PP2) has also been shown to dephosphorylate cMyBP-C to a lesser degree in skinned cardiomyocytes, whilst alkaline phosphatase (AP) dephosphorylates cMyBP-C in human myocardium samples (Zaremba et al., 2007).

1.2.5.2. Other types of post translational modifications

Although phosphorylation is the most studied PTM affecting cMyBP-C, other types of PTM have also been found to have a role in the function of the protein.

Acetylation is the addition of an acetyl group from acetyl-CoA onto the N-terminus of a protein or to lysine residues, which is catalysed by acetyltransferases (Duan and Walther, 2015). Eight possible acetylation sites have been identified in cMyBP-C (Lys7, Lys185, Lys190, Lys193, Lys202, Lys442, Lys935, Lys962), five of which are located in the C0-C1 region (Ge et al., 2009). It is uncertain whether acetylation of cMyBP-C is upregulated in disease states or if it is functionally important, however a proteolytic degradation fragment of cMyBP-C has been found to be heavily acetylated, indicating that acetylation could possibly lead to a promotion of proteolysis and thus decrease the stability of cMyBP-C (Govindan et al., 2012).

Another form of PTM is citrullination, which is the conversion of an arginine to a citrulline residue (György et al., 2006) and often leads to a loss of protein structure and protein-protein interactions (Fert-Bober and Sokolove, 2014). Little information regarding citrullination of cMyBP-C exists, however Arg696 in the C5 domain has been found to be citrullinated in both healthy and diseased heart tissue (Fert-Bober and Sokolove, 2014), although the effect this modification has on cMyBP-C function has not been ascertained.

Oxidation can also lead to various different forms of PTMs, some of which are thought to affect cMyBP-C function.

Three sites have been found in the murine cMyBP-C: Cys479, Cys627 and Cys655 which can be a target for S-glutathionylation; the addition of glutathione groups to cysteine residues (Xiong et al., 2011), which has been reported to lead to an increase in the Ca^{2+} sensitivity of myofilaments (Patel et al., 2013). Studies utilising human heart samples have identified 15 cysteines that could be S-glutathionylated, with Cys249, Cys426, Cys436, Cys443, Cys475, Cys566, Cys651 and Cys719 all showing increased levels of S-glutathionylation in heart

samples from patients with dilated cardiomyopathy (DCM) or ischaemic cardiomyopathy (ICM) compared with healthy donor hearts (Stathopoulou et al., 2016). S-glutathionylation is known to cause changes in protein function via altering the structural conformation of the protein or modifying protein-protein interactions (Janssen-Heininger et al., 2013). S-glutathionylation of cMyBP-C leads to a decrease in cross-bridge cycling kinetics, which is thought to be due to S-glutathionylation of targets within the C1-M-C2 domain affecting the interaction of cMyBP-C with actin and the S2 fragment and RLC of myosin. Furthermore, it is also thought that there may be cross-talk between S-glutathionylation and phosphorylation, with S-glutathionylation of cMyBP-C at Cys249 preventing phosphorylation by protein kinases possibly by altering the conformation of the C1-M-C2 domains (Stathopoulou et al., 2016).

cMyBP-C is also thought to be S-nitrosylated, which is the covalent attachment of a nitric oxide to sulfhydryl residues of a protein (Duan and Walther, 2015), at the Cys1270 residue (Kohr et al., 2011). A study identified that S-nitrosylation of sarcomeric proteins led to a decrease in calcium sensitivity, a decreased maximal isometric force and prolonged relaxation. The study further identified cMyBP-C as being susceptible to S-nitrosylation (Figueiredo-Freitas et al., 2015).

Carbonylation is a type of irreversible oxidative modification that attaches a carbonyl group to various different amino acid residues (Cai and Yan, 2013), which occurs under conditions of prolonged oxidative stress and has been shown to modify cMyBP-C by impairing its ability to bind to actin (Aryal et al., 2014). It has been hypothesised that this modification could contribute to the cardiac dysfunction that has been observed during the oxidative stress caused by chemotherapy (Aryal et al., 2014).

Another PTM that has been found to target cMyBP-C is sumoylation, which is the reversible addition of the small ubiquitin-related modifier (SUMO) by SUMO-specific enzymes that is known to affect protein stability and localisation (Flotho and Melchior, 2013). Studies have identified sumoylation targets in the N-terminal domain of cMyBP-C and have shown that this modification has an effect on the half-life of the protein (Gupta and Robbins, 2014).

Whilst phosphorylation has been shown to be key in the function of cMyBP-C, it is clear that more work needs to be carried out to fully understand the role of phosphorylation, as well as understand any effects that other types of PTMs may be having on the protein. Furthermore, the possible interaction thought to occur between S-glutathiolation and phosphorylation (Stathopoulou et al., 2016) opens the possibility that there could be a much more intricate web of PTMs affecting the functionality of cMyBP-C.

1.3. Cardiomyopathy

Cardiomyopathies are myocardial disorders which are not solely due to coronary artery disease or abnormal loading conditions (Lopes and Elliott, 2014). There are three major groups: HCM which is found in 1:500 of the population, although more recent studies estimate the prevalence to be 1:200 (Semsarian et al., 2015), DCM found in 1:2500 of the population and arrhythmogenic cardiomyopathy (ACM) which affects 1:5000 of the population (Sabater-Molina et al., 2018).

1.3.1. Hypertrophic Cardiomyopathy

HCM was first described in 1958, with cases of asymmetrical hypertrophy or muscular hamartoma in the hearts of young people who had suffered sudden death (Teare, 1958). HCM is the most commonly inherited cardiovascular disease and the leading cause of sudden cardiac death in the young, which can occur without any prior symptoms of the disease (Frey et al., 2011). The disease is characterised by the thickening of the ventricular and/or septal wall which leads to a decrease in chamber size and thus blood volume at filling (Seidman and Seidman, 2001), a diagram of which can be seen in figure 9.

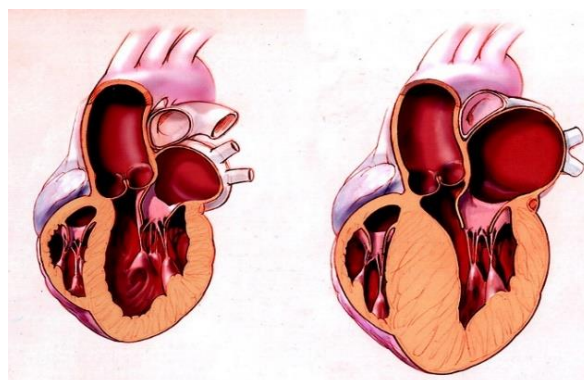


Figure 9 - Characterisation of Hypertrophic Cardiomyopathy

Diagram depicting the difference between a healthy heart, (left) and that of a patient with hypertrophic cardiomyopathy HCM (right). HCM is characterised as the thickening of the ventricular and/or septal walls leading to a decreased chamber size. Image from (Nishimura et al., 2003).

Typical phenotypic presentations of the disease, some of which are shown in figure 10, include left ventricular hypertrophy, myocardial hypercontractility, myofibrillar disarray and an increase in interstitial space with fibrosis formation, all of which can lead to diastolic dysfunction which commonly occurs in HCM patients (Davies and McKenna, 1995, Hughes, 2004). The mortality rate for patients with HCM is thought to be around 0.9-1.5% per year, although this is thought to have been reduced to 0.5% with the development of new therapeutics (Antunes and Scudeler, 2020). The leading cause of death from HCM is sudden cardiac death, which is thought to occur in about 16% of patients, mainly those under the age of 30 (Ho et al., 2018).

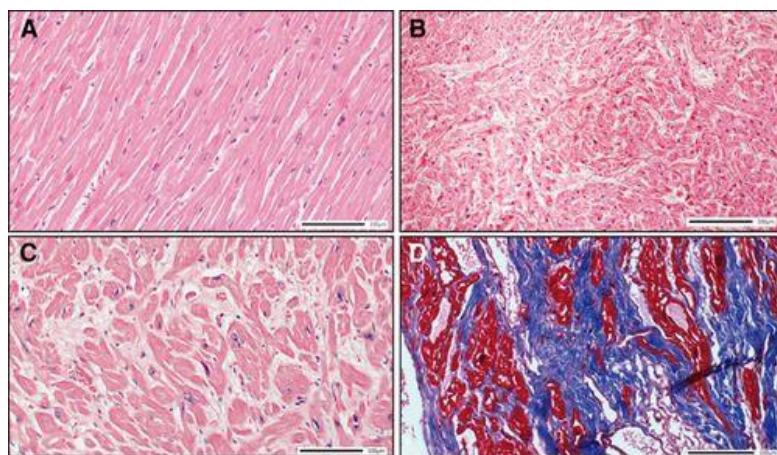


Figure 10 - Myofibrillar disarray seen in Hypertrophic cardiomyopathy

Images showing the myofibrillar disarray and interstitial fibrosis typically seen in patients with hypertrophic cardiomyopathy. (A) shows a haematoxylin and eosin (H&E) stained myocardial section from a healthy heart showing normal myocardial architecture. (B) shows myocardial section from the heart of an HCM patient stained with H&E showing disorganised myocardial architecture (4x magnification). (C) shows a 20x magnification of the H&E stained HCM myocardial section showing myocyte disarray. (D) shows a myocardial section from an HCM patient stained with Masson trichrome showing interstitial fibrosis (blue). Figure from (Marian and Braunwald, 2017)

1.3.2. HCM causing mutations

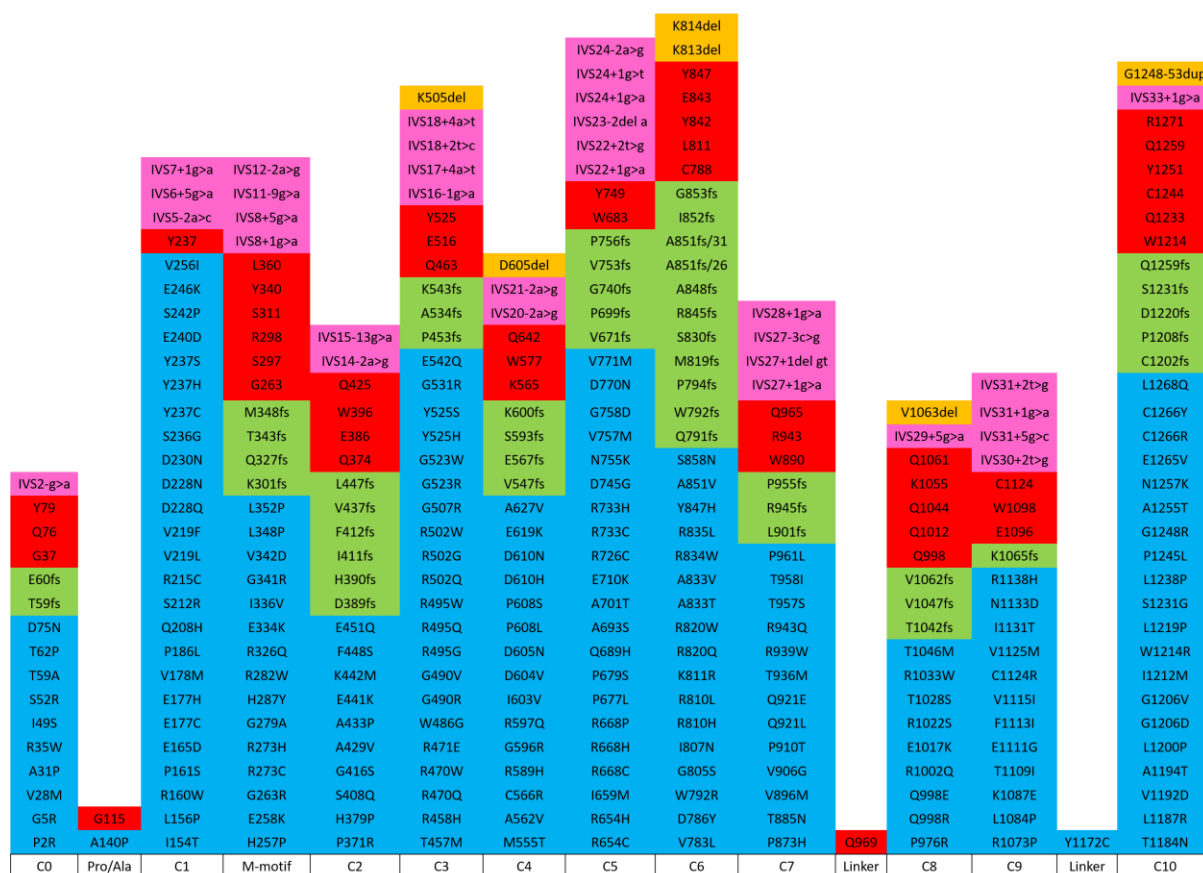


Figure 11 – HCM-linked cMyBP-C mutations

Figure showing mutations in each of the domains of the gene encoding cardiac MyBP-C, MYBPC3, linked to cases of HCM. The highest number of clinically linked variants occur in domain C6, followed by the C5 domain. Sequence variations include missense mutations causing single amino acid substitutions (blue), insertions/deletions predicted to cause reading frameshifts (green), nonsense mutations predicted to code premature stop codons (red), splice site mutations (pink) and in frame insertion/deletions (yellow) Adapted from Harris et al. supplemented by data from Wang et al. (Harris et al., 2011, Wang et al., 2018b)

Around 70% of HCM cases are caused by a mutation (Richard et al., 2003), which mainly occurs in one of the cardiac sarcomere proteins (Lopes et al., 2013), although mutations have also been found in genes of proteins involved in calcium homeostasis (Richard et al., 2003). The second most frequently mutated gene is MYBPC3, which encodes for cMyBP-C,

accounting for ~40% of HCM cases (Carrier et al., 2015). These mutations lead to a wide range of clinical manifestations, with more severe phenotypes associated with truncation mutations compared with missense mutations (Erdmann et al., 2001). Mutations have been found along the full length of the protein, as can be seen in figure 11, with the C6 domain containing the largest number of clinically linked mutations of all of the domains, followed by C5 (Harris et al., 2011), suggesting that the central domains have an important role in MyBP-C function.

1.3.3. Truncation vs. missense mutations

Almost 60% of cMyBP-C mutations are truncation mutations, and it was first thought that the disease mechanism was through a dominant negative mutation or 'poison polypeptide' mechanism. This was due to experiments showing truncated protein lacking the C-terminal domains, including the binding sites for myosin and titin, leading to mislocalisation of the protein in the sarcomeres (Yang et al., 1998). However, as no truncated protein has been detected in human samples of patients with HCM (Rottbauer et al., 1997, Moolman et al., 2000, van Dijk et al., 2009, Marston et al., 2009) it is possible that haploinsufficiency is the mechanism of disease, with cell surveillance mechanisms leading to elimination of the truncated protein. This theory is corroborated by experiments using human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) from a HCM patient showing ~50% lower levels of MYBPC3 mRNA and cMyBP-C protein (Prondzynski et al., 2017).

Possible cell surveillance mechanisms include nonsense-mediated mRNA decay (NMD), an evolutionary conserved mechanism found in all eukaryotes, which detects and eliminates mRNA transcripts containing premature stop codons (Garneau et al., 2007), which ~ 2/3rds of MYBPC3 mutations have been found to generate (Schlossarek et al., 2011). Another form

of cell surveillance is the ubiquitin-proteasome system (UPS), which prevents the accumulation of damaged, misfolded and mutant proteins (Schlossarek et al., 2011).

The remaining 40% of cMyBP-C mutations are missense mutations, with haploinsufficiency proposed as a possible disease mechanism. However, it has been shown that an A31P missense mutation leads to a stable MyBP-C protein that is incorporated into the sarcomere, which has led to the 'poison polypeptide' theory as a possible disease-causing mechanism for MYBPC3 missense mutations (van Dijk et al., 2016). This could occur through the missense mutation disrupting the function of the protein either by altering the proper domain folding thus affecting the structure of the protein or disrupting the protein's interactions with other proteins (Harris et al., 2011).

Although generally patients with missense mutations have a less severe phenotype than those with truncating mutations (Erdmann et al., 2001), there is still a high degree of heterogeneity of disease pathogenesis seen in patients carrying different types of mutations in the MYBPC3 gene. Therefore, although it is possible that generally similar types of mutations may behave in similar ways, it is likely that there is not a single disease mechanism that can be applied for all HCM-linked mutants.

A more recent study utilising the Sarcomeric Human Cardiomyopathy Registry (SHaRe), however, has suggested that the percentage of HCM patients carrying truncating mutations is slightly higher at ~71%. Furthermore, analysis utilising this database suggested there was no difference in the disease phenotype when comparing truncating and non-truncating mutations, with a slight increase in paediatric diagnoses in the non-truncating group. This study also showed that truncating mutations were spread evenly across all of the domains of cMyBP-C, whilst non-truncating mutants appeared to cluster in the C3, C6 and C10

domains, which could suggest that these domains are key regions in the functioning of the protein. The SHaRe database is mainly comprised of those of European ancestry, therefore this could explain the differences seen in this study compared with previous data (Helms et al., 2020b).

1.3.4. Models of HCM causing cMyBP-C mutations

Much has been elucidated about the function of cMyBP-C through animal models. Although animal models of disease are often not perfect recapitulations of human disease, they do have the benefit of showing what is occurring in situ with all of the constraints of the sarcomere and environmental factors in effect. Multiple models of cMyBP-C linked HCM have either been identified or created in order to better understand the function of cMyBP-C as well as how mutations in this protein may lead to disease.

In 1999, Kittleson et al. identified the first naturally occurring animal model of familial HCM, in the Maine coon breed of cats, which showed symptoms similar to that of the human disease (Kittleson et al., 1999). A later study showed that 34% of the Maine coon breed carried a MYBPC3 mutation (Fries et al., 2008), with Meurs et al. identifying that all Maine coon cats they studied with HCM were carrying an A31P missense mutation of cMyBP-C (Meurs et al., 2005). Another cat breed, Ragdoll, has also been associated with having HCM, with an R820W mutation in cMyBP-C being identified in cats of this breed with HCM (Meurs et al., 2007).

The first animal model created to study HCM caused by mutations in cMyBP-C utilised mice expressing cMyBP-C lacking both the titin and myosin binding domains, leading to a C-terminally truncated protein (Yang et al., 1998). The truncated protein was shown to make up 40-60% of the total protein and showed normal stability. However, immunofluorescence

staining showed diffuse staining and patchy striations indicating the protein was not correctly incorporating into the sarcomere, therefore it would only weakly be associated with the contractile apparatus. The animals exhibited no apparent heart enlargement or myocardium wall thickening, whilst skinned fibre samples taken from affected animals were found to have an increased calcium sensitivity of force development indicating subtle alterations were occurring in force production. The study also included a small sample of older animals, which exhibited a high degree of myocyte disarray as well as abnormal nuclei (Yang et al., 1998), whilst a subsequent study utilising the model indicated that affected animals had a significantly compromised exercise capacity, with an increased mortality during exercise (Yang et al., 2001). This can be seen as recapitulating the human disease, as often patients do not exhibit symptoms until later in life, furthermore HCM is known to be the leading cause of sudden cardiac death in young athletes (Frey et al., 2011).

The same group subsequently engineered a transgenic construct lacking the C-terminal 80 amino acids of cMyBP-C; therefore lacking the myosin binding domain (Yang et al., 1999). Very little truncated protein was detected in the transgenic mice, consistent with findings from endomyocardial biopsies from patients with HCM in which no truncated protein could be detected (Rottbauer et al., 1997, Moolman et al., 2000, van Dijk et al., 2009, Marston et al., 2009). Transgenic mice showed sarcomeres to be consistently out of register and mitochondrial abnormalities were detected, whilst histology showed mild hypertrophy but no remarkable changes in the overall morphology of the cardiomyocytes. Mechanical measurements on skinned fibres were in agreement with the studies utilising truncated protein lacking both the titin and myosin binding sites, in that they showed impairment in fibre mechanics. These transgenic mice also showed significant impairment in exercise

capacity, with functional defects seen under cardiac stress. Older animals at 12 months old showed ~20% increase in heart to body weight ratio and showed a mild but significant hypertrophy.

The first animal model utilising a KO of cMyBP-C was generated by Harris et al. in 2002 by inappropriate splicing of exons 2 and 11 leading to a reading frame shift and introduction of a neo resistance cassette (Harris et al., 2002). Both RT-PCR and Western blot analysis detected no product for homozygous KO mice (cMyBP-C^{-/-}), whilst heterozygous KO animals (cMyBP-C^{+/-}) showed the same levels of protein as wildtype (WT). It was also seen that no other isoforms of MyBP-C were upregulated in order to compensate for this loss. Hypertrophy was seen in cMyBP-C^{-/-} animals as early as three weeks old, with adult animals showing a significant increase in heart:body weight ratio as well as an increase in left ventricular diameter, anterior and posterior wall thickness, whilst molecular markers of hypertrophy were also seen to be upregulated in cMyBP-C^{-/-} but not in cMyBP-C^{+/-}. Histology revealed KO mice showed myocyte disarray and fibrosis consistent with HCM, with misalignment of the myofibrils and mitochondrial abnormalities seen. KO animals also showed a decrease in systolic function as well as a decrease in myofilament Ca²⁺ sensitivity.

Subsequent studies utilised this KO mouse model in order to gain a better understanding into the role of MyBP-C in the sarcomere. One study identified a decrease in LV functional capacity and an attenuation of increase in myocardial contractility following β -adrenergic stimulation in animals lacking cMyBP-C, implicating that cMyBP-C has a role in compensatory adaptations to pressure overload that are required for maintaining ventricular function (Brickson et al., 2007).

Use of this KO model also showed that in the absence of MyBP-C myocytes displayed an increase in power output due to an increase in loaded shortening rates, which was thought to be due to faster crossbridge cycling at each given load. This suggests a model in which MyBP-C functions to constrain the interaction between actin and myosin, therefore limiting loaded shortening velocity and thus power output (Korte et al., 2003).

This KO model also allowed studies to identify the effect that lack of MyBP-C would have on sarcomeric structure, with electron microscopy and image analysis showing that KO thick filaments had the same length, diameter and appearance as WT thick filaments (Kensler and Harris, 2008). Although the M-band was seen to have the same five-line structure as WT, in the KO the outer two lines were very weak (Luther et al., 2008). Furthermore, although the KO retained similar relaxed cross-bridge arrangements to WT, the thick filaments were seen to be more easily disrupted and regions of disorder were more common (Kensler and Harris, 2008). Lastly, synchrotron low angle X-ray diffraction showed radial displacement of crossbridges away from the thick filament in KO relaxed myocardium (Colson et al., 2007). This is in line with the hypothesis that MyBP-C acts to tether myosin crossbridges closer to the thick filament backbone, acting to reduce the likelihood of binding to actin and limiting activation of the thin filament (Hofmann et al., 1991, Calaghan et al., 2000).

A subsequent cMyBP-C KO mouse model was created by Carrier et al. by replacing the transcription site and exons 1-2 with a neo resistance gene, which caused total transcriptional inactivation of cMyBP-C gene expression (Carrier et al., 2004). No mRNA was amplified using qRT-PCR and no protein was detected by western blot analysis for the homozygous KO mice, whilst heterozygous animals showed a 25% decrease in protein levels at 10-11 months. KO animals showed upregulated mRNA for markers of hypertrophy,

asymmetrical septal hypertrophy, impaired relaxation, myocardial disarray, increased interstitial fibrosis and calcification in fibrotic areas (Carrier et al., 2004). These animals showed mild hypertrophy at six weeks which became more pronounced by 30 weeks old (Pohlmann et al., 2007). Studies indicated that the probability of myosin being weakly bound to actin were higher in KO animals as well as the crossbridge cycling velocity and force generated being lower. This suggests cMyBP-C prevents weak and inefficient binding of myosin crossbridges to actin (Lecarpentier et al., 2008).

These KO models make it clear that whilst lack of cMyBP-C is not embryonic lethal and does not interfere with the formation of the sarcomeres during embryogenesis, it is required for maintaining cardiac function and the organisation of the sarcomeres (Carrier et al., 2015).

Other models have used knock-in (KI) approaches in order to study the function of cMyBP-C, the first of which attempted to model a G>A transition on the last nucleotide of exon 6, which is associated with human HCM with a severe phenotype and bad prognosis. This mutation was seen to lead to a missense mutation (E264K) as well as a nonsense and insertion/deletion mutation, highlighting the heterogeneity of mutations in cMyBP-C (Vignier et al., 2009).

KI mouse models have also been utilised in order to elucidate the specific function given by unique elements found in the cardiac isoform of the protein. For instance, a study generated KI mice with cMyBP-C that lacked the C1 domain and the linker between C0 and C1 creating a protein with a similar size to that of the skeletal isoforms. The protein incorporated into the sarcomere and there were no ultrastructural changes seen to the heart, however skinned muscle fibres from homozygous KI animals showed an increased

Ca²⁺ sensitivity indicating the possibility that the extra N-terminal domain regulates force at the crossbridge level (Witt et al., 2001).

KI mice have also been used in order to study the effect of phosphorylation to the function of cMyBP-C by generating transgenic mice with known phosphorylation sites (Ser273, 282, 302) mutated. Mice with phosphorylation sites mutated to non-phosphorylatable alanines led to animals with depressed cardiac contractility, altered sarcomeric structure and an upregulation of transcripts associated with hypertrophy (Sadayappan et al., 2005), highlighting the importance of phosphorylation of cMyBP-C for normal cardiac function.

Although animal models have been very useful in elucidating the functional role of MyBP-C, they do not come without limitations. One of which is the fact that there may be species differences. For example there is a species difference in the myosin isoform expressed, with the adult mouse heart expressing mainly α -MHC whilst in humans β -MHC predominates. It has been shown that phosphorylation of cMyBP-C has different effects depending on which MHC isoform the filaments contain (Sadayappan et al., 2006, Weisberg and Winegrad, 1996), therefore it is possible that other observations could be similarly skewed due to species differences in the sarcomeric proteins.

Another issue with animal models is the fact that HCM-linked cMyBP-C mutations have often been seen to have a poor genotype-phenotype correlation, implicating the disease expression could be influenced by other factors. One way in which these issues have been resolved is via the use of patient derived induced pluripotent stem cell derived cardiomyocytes (iPSCMs) or engineered heart tissue (EHT) generated from iPSCMs isolated from patients with HCM. In 2013, Stohr et al. tested whether the EHT model would be viable by generating KI-EHTs carrying an E264K mutation that had been previously modelled in

mice (Vignier et al., 2009), which was found to recapitulate the findings obtained from studies using this mouse model (Stöhr et al., 2013), indicating that the use of EHTs could be a viable avenue to explore.

A recent study utilising patient and genome engineered iPSCMs identified the effects of several premature termination codon causing mutations in cMyBP-C in the early development stage. They showed that these mutations have a common pathomechanism of loss of function due to a reduction in synthesis of MyBP-C caused by uncompensated reduction in MYBPC3 mRNA. They also identified that iPSCMs were capable of compensatory slowing of degradation of MyBP-C, which is not seen in heart tissue from patients, which they suggest is due to this compensatory capacity being lost by the time HCM develops (Helms et al., 2020a). Interventions to reverse this loss of compensation could therefore be a viable target for patients with these types of mutations.

iPSCMs do however have the drawback of exhibiting an immature phenotype (Yang et al., 2014) which could mean certain proteins that are involved in the pathogenesis of the disease are either not present or expressed at different levels to that seen in the mature heart. Furthermore, HCM generally occurs later in life with little to no symptoms seen prior to onset (Frey et al., 2011), therefore it is possible that little information about the disease pathogenesis could be seen in this immature model. Whilst EHTs show a more mature phenotype, their complexity also brings difficulties in the form of issues with reproducibility (Eschenhagen et al., 2015). Although the use of iPSCMs and EHTs in modelling the function of cMyBP-C and the pathogenesis of HCM-linked mutations is not without flaws, this method is still relatively new and does have many promising aspects.

The heterogeneity seen in the pathogenesis and disease progression of HCM seen in patients not only with different mutations in cMyBP-C, but also patients carrying the same mutation makes it clear that the way in which mutations in the MYBPC3 gene lead to HCM varies from individual to individual (Wang et al., 2018a, Fananapazir and Epstein, 1994). Therefore, being able to generate an iPSCM model from an individual patient could lead to beneficial insights into not only the mechanism by which different mutations lead to disease but could also lead to the ability to test and develop more individualised treatments.

1.3.5. Current therapies

Current therapies for HCM mainly focus on symptomatic alleviation, such as beta-blockers and calcium channel blockers. These therapies, however, do not reverse the cardiac hypertrophy which is the major determinant of mortality and morbidity in HCM patients (Schlossarek et al., 2011). Furthermore, these drugs have very little evidence justifying their use, with no randomized trials demonstrating their efficacy in the treatment of HCM (Spoladore et al., 2020).

Patients with severe symptoms may require surgical intervention to relieve the obstruction in the left ventricle, such as myectomy or alcohol ablation (Sabater-Molina et al., 2018). Newer less invasive techniques have also been developed in order to reduce obstruction such as radiofrequency ablation (Lawrenz et al., 2011) and high-intensity focused ultrasound (Nazer et al., 2014).

New alternative therapies are now being developed in order to target the underlying mechanism of pathogenesis instead of only alleviating symptoms.

One novel therapeutic which has been developed is the drug Mavacamten, which targets hypercontractility which is thought to be involved in the pathogenesis of HCM, as it is an allosteric modulator of cardiac β -myosin which causes reversible inhibition of actomyosin crossbridge cycling (Kawas et al., 2017, Anderson et al., 2018). Use of the drug in mice showed suppression of hypertrophy, cardiomyocyte disarray and expression of profibrotic genes (Green et al., 2016), implicating that it attenuates the HCM phenotype. Subsequent phase IIa clinical trials showed a decrease in post-exercise LV outflow tract obstruction as well as improvements in exercise capacity in patients with obstructive HCM and it was shown to be generally well tolerated with few side effects (Heitner et al., 2019).

As it is believed that mutated cMyBP-C may be degraded by one of the cell surveillance systems, such as NMD or UPS (Schlossarek et al., 2011), one possible therapeutic approach could be to use inhibitors of NMD or UPS to increase levels of the mutant cMyBP-C in the hope that the mutant protein could be functional. Studies utilising UPS inhibitors have been shown to decrease hypertrophy in hypertensive rats (Meiners et al., 2008), whilst in an isoprenaline-induced cardiac hypertrophy murine model and a mouse model of pressure overload regression of inhibition of cardiac hypertrophy was achieved (Stansfield et al., 2008, Hedhli et al., 2008).

Another possible treatment could be the replacement of cMyBP-C by gene therapy, with a study using hiPSC-CMs from a HCM patient showing gene replacement restored cMyBP-C levels to 81% compared to control levels (Prondzynski et al., 2017). However, if there were still mutant peptides present at low levels it is possible that they could act as 'poison polypeptides' (Schlossarek et al., 2011).

Exon skipping is thought to be a novel avenue of therapy in genetic disorders and data has shown that modified antisense oligonucleotides can induce exon skipping leading to the production of novel cMyBP-C (Gedicke et al., 2010).

Prevention of germline transmission of MYBPC3 mutations is a desirable avenue of treatment, as HCM typically presents later in life and thus mutations are likely to be transmitted to the next generation. A recent study has identified utilisation of the CRISPR-Cas9 genome editing technique is capable of correcting a heterozygous MYBPC3 mutation in human preimplantation embryos (Ma et al., 2017). However, this would require the precise HCM-linked mutation that the parent is carrying to be known and could also be ineffective if other unidentified compound mutations are present.

1.4. Current knowledge on the central domains of cMyBP-C

1.4.1. Structure of the central domains

The structure of the central domains of cMyBP-C has not been fully elucidated, with no structural analysis obtained for the C6 domain, however NMR studies have been carried out for the C3 domain (Zhang et al., 2014) whilst the C4 domain has been resolved for the slow skeletal isoform of the protein (Niraula et al., 2008).

More thorough experimentation elucidating the structure of the C5 domain of cMyBP-C has been carried out and has allowed insight into the structure of the insertions unique to the cardiac isoform, revealing two inserts. The first is a linker region between the C4 and C5 domains which has also been shown to be integral to the fold of the C5 domain. The second insert is an elongation of the CD loop which is unstructured and thought to be responsible for the lower stability of the C5 domain, compared to the other domains (Idowu et al., 2003).

Further importance of understanding the structure of the central domains was highlighted by structural studies into HCM-linked mutations in the central domains that elucidated that these mutations could cause disruption to the structural integrity of individual domains. This implication, that mutations leading to aberrations in the correct structural folding of individual central domains could be causative of HCM, highlights the probability that these central domains do have a role in the proper function of MyBP-C in its role of regulating sarcomere dynamics (Nadvi et al., 2016).

1.4.2. Possible binding partners of the central domains

As mentioned earlier, previous yeast two hybrid studies have suggested that the C5 domain can bind to the C8 domain (Moolman-Smook et al., 2002) and thus could have a role in the localisation of the protein. The structural organisation of MyBP-C in the sarcomere is still being debated, however the results of a more recent study utilising domain-specific immune-EM were inconsistent with this collar model (Lee et al., 2015).

Studies in the slow skeletal isoform of MyBP-C have suggested an interaction between the C6-C10 region of MyBP-C and muscle creatine kinase (Chen et al., 2011), which is a protein that is thought to have a key role in the monitoring and localised production of ATP in the sarcomere (Henderson et al., 2017). It is thought that as the C-terminal region of MyBP-C also interacts with myosin, it could act as an adaptor between the two proteins and thus facilitate efficient ATP metabolism (Wang et al., 2018b). However, it is not known whether this interaction also occurs for the cardiac isoform of the protein. Furthermore, although muscle creatine kinase is a cytosolic protein, about 5-10% has been found to localise to the M-band (Wallimann et al., 1992), therefore it is possible that this localisation would not support an interaction with MyBP-C in the C-zone.

It has also been suggested that the central domains may be involved in MyBP-C's interaction with actin, with a study suggesting that the C5-C10 region is able to bind actin specifically and relatively strongly. However as this same region also contains binding sites for both titin and myosin there would not be the capacity for all three to bind simultaneously, although it is possible that the C5-C10 region of the protein could cycle through these different binding partners (Rybakova et al., 2011). Studies by Inchingolo et al. have also suggested the C3 domain may be involved in the interaction of the N-terminus of cMyBP-C with the thin

filament and was required to displace tropomyosin and sensitise the thin filament to Ca^{2+} whilst shorter C0C1 and C0C1 with the first 17 residues of the M-motif could bind but had no effect sensitisation of the thin filament (Inchingolo et al., 2019). However, this experiment did not use a C0C2 fragment therefore it does not definitively implicate the C3 domain in this action, as the C2 domain could be sufficient for this effect.

1.4.3. Possible role for the central domains

The studies showing various possible binding sites in the central domains, along with the high number of HCM-linked missense mutations in this region suggest that these domains have a role in the protein's function, with studies suggesting that the central domains, or at least the most N-terminal C3 domain, could be involved in the role of MyBP-C in sensitising the thin filaments and regulating crossbridge cycling kinetics (Napieriski et al., 2020, Inchingolo et al., 2019). However, it is also possible that these domains act as a spacer which allows for the protein to span the interfilament distance and the N-terminus to bind the thick and thin filaments. Mutations in these domains could lead to modifications in domain structure and possibly alter the length or flexibility of said spacer and affect the ability of the N-terminus to cycle between the thin and thick filaments and thus its ability to regulate actomyosin cross-bridge cycling. On the other hand, it could be argued that as a high number of mutations appear to be tolerated in the central domains this may suggest that this region is less important for the overall stability and function of the protein. It is possible that there are fewer mutations found in the N-terminal and C-terminal domains because mutations in these regions are more likely to have a significant impact on the function of the protein which would not be tolerated.

1.5. Aims and objectives

Since its discovery more than 50 years ago MyBP-C has been found to be so much more than merely an impurity in preparations of muscle myosin, with a key role in regulating the actomyosin cross-bridge cycling being discovered, even if the exact mechanism is still being debated.

Whilst the extremities of the protein have clear functional and structural roles, the role of the central C3-C6 domains is yet to be elucidated. Furthermore, only the structure of the C5 domain has been resolved for the cardiac isoform (Idowu et al., 2003). This lack of information about the functional role and structure of these domains, along with the fact that the C6 domain contains the largest number of HCM-linked mutations, followed by the C5 domain (Harris et al., 2011), as well as the C5 domain containing cardiac specific inserts (van Dijk et al., 2014), which suggests it may have a cardiac specific function, makes further study of this region of key importance.

This led to the hypothesis that the central region of cMyBP-C has a role in the function of the protein and does not merely act as a spacer. Furthermore, HCM-linked missense mutations in these central domains may lead to disease by interfering with this functional role. To explore this hypothesis the following aims were defined:

1. To elucidate the structure and functional role of the central domains of cMyBP-C.
2. To understand the mechanism of disease of HCM-linked mutations in the C3C6 region.

In order to achieve these aims a yeast two hybrid assay was carried out to identify possible binding partners, whilst a recombinantly expressed fragment of the central domains was

expressed, purified and utilised in biochemical assays to determine its interaction with actin. Constructs, firstly green fluorescent protein (GFP)-tagged plasmids and subsequently GFP-tagged adenoviral vectors (multidomain or full-length respectively), with HCM linked mutations within the C3-C6 domain regions were also generated. Protein levels, cellular localisation and effect on the structure of the myofibrils were assessed using western blot analysis and immunofluorescence microscopy, including using super-resolution stimulated emission depletion (STED) microscopy, to gain a clearer understanding of the disease mechanism.

2. Materials and Methods

2.1. Buffers and Media

Name	Contents
ADS Buffer	100 mM NaCl, 20 mM HEPES, 1 mM NaH ₂ PO ₄ , 5 mM Glucose, 5 mM KCl, 1 mM MgSO ₄ , pH 7.35 in dH ₂ O
ATPase assay buffer	10 mM MOPS pH 7, 5 mM MgCl ₂ , 1 mM DTT
Blotting Buffer	25 mM Tris-Base pH 8.3, 192 mM Glycine, 0.01% SDS (w/v) and 20% Methanol (v/v)
Blue Wonder (Loading buffer)	3.7M urea, 134.6 mM Tris pH 6.8, 5.4% SDS (w/v), 2.3% NP-40 (w/v), 4.45% β-mercaptoethanol (w/v), 4% glycerol (v/v), 6 mg/mL bromophenol blue
Buffer 1 (NTF preparation)	100 mM KCl, 3 mM MgCl ₂ , 3 mM EDTA, 5 mM imidazole, 5 mM DTT, pH 7.5
Buffer A	20 mM HEPES pH 7.5, 300 mM NaCl, 1 mM DTT, 20 mM imidazole
Buffer B	20 mM HEPES pH 7.5, 300 mM NaCl, 1 mM DTT, 500 mM imidazole
COS-1 cell complete media	DMEM with 4 mM glutamine, 10% FBS (v/v), 1% penicillin/streptomycin (v/v)
DBSS-K (Dulbecco's balanced salt solution with potassium added)	116 mM NaCl, 1 mM NaH ₂ PO ₄ , 0.8 M MgSO ₄ , 32.1 mM NaHCO ₃ , 5.5 mM glucose, 1.8 mM CaCl ₂ , pH 7.2
Enzyme Solution	0.06150 g collagenase type 2 (300 U/mg), 0.0450 g Pancreatin in 75 mL ADS buffer
Gold Buffer	20 mM Tris, 155 mM NaCl, 2 mM EGTA, 2 mM MgCl ₂ pH 7.5
HEK 293 cell complete medium	DMEM with 4 mM glutamine, 10% FBS (v/v), 1% penicillin/streptomycin (v/v)
Lisbeth's Mounting Medium	15 ml 0.1 M Tris-HCl pH 9.5 and 35ml 100% glycerol with 50 mg/mL <i>n</i> -propyl-gallate
Low Salt Buffer	0.9% NaCl (w/v), 0.01% Tween-20 (v/v), 9 mM Tris Base pH 7.4
Lysis buffer (Y2H)	10 mM TrisHCl, 1% β-mercaptoethanol (w/v), 5 U lyticase, pH 8
Neonatal Rat Cardiomyocyte (NRC) Maintenance Medium	25% Medium 199 (v/v), 4% horse serum (v/v), 2% GlutaMax (v/v), 1% penicillin/streptomycin (v/v), 0.1 mM Phenylephrine (PE) and 10 μM cytosine arabinoside (AraC) in basal medium DMEM
Neonatal Rat Cardiomyocyte (NRC) Plating Medium	66% DMEM (v/v), 16.5% M199 (v/v), 10% horse serum (v/v), 5% fetal calf serum (v/v), 4 mM Glutamine, 1% penicillin/streptomycin (v/v)

Neonatal Rat cardiomyocyte (NRC) Transfection Medium	20% M199 (v/v), 75% DBSS-K (v/v), 4% horse serum (v/v), 4 mM glutamine
NTF co-sedimentation buffer	20 mM imidazole pH 7.4, 180 mM KCl, 1 mM MgCl ₂ , 1 mM EGTA pH 8, 1 mM DTT
Salt buffer (NTF preparation)	100 mM KCl, 3 mM MgCl ₂ , 10 mM imidazole, 5 mM DTT, pH 7.3
1x SDS running buffer	25mM Tris, 200mM glycine, 1% SDS
TEV protease dialysis buffer	25 mM HEPES, 200 mM NaCl, 1 mM MgCl ₂ , 1 mM DTT, 5% glycerol (v/v)
Thin filament storage buffer	50 mM KCl, 2 mM MgCl ₂ , 0.4 mM ATP, 0.5 mM DTT, 5 mM imidazole, pH7
Urea sample buffer	8 M urea, 20 mM Trisbase pH 7, 23 mM glycine, 0.2 mM EDTA, 5% glycerol (v/v), 10 mM DTT, 0.004% bromophenol blue (w/v)
X-gal buffer	60 mM NaH ₂ PO ₄ , 10 mM KCl, 1 mM MgSO ₄ , 38 mM β-mercaptoethanol, 0.33 mg/ml X-gal
Y2H transformation buffer	10 mM TrisHCl pH 8, 1 mM EDTA, 100 mM Lithium acetate, 40% PEG 3500 (v/v)

Table 1 – Table of Buffers and Media used during experimentations

2.2. Antibodies

Antibody	WB/IF	Dilution	Company
Primary antibody			
Rabbit α all-actin	WB	1:2000	Sigma Aldrich
Mouse α sarcomeric alpha-actinin	IF	1:500 (STED)	Sigma Aldrich
Mouse α GFP	WB	1:1000	Roche
Mouse α His	WB	1:2000	Novagen
Rabbit α LexA	WB	1:1000	Abcam
Mouse α Myomesin	IF	1:50	N/A (Grove et al., 1984)
Rabbit α MyBP-C	IF/WB	1:50 (STED)/ 1:1000 (WB)	N/A (Linke et al., 1999)
Secondary antibody/Counterstain			
Alexa Fluor 546 phalloidin	IF	1:100	Invitrogen
Alexa Fluor 633 phalloidin	IF	1:100	Invitrogen
Atto647N α Rabbit	IF	1:500 (STED)	Sigma Aldrich
Cy3 Goat α mouse IgG	IF	1:500/1:200 (STED)	Jackson Immunochemicals
4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI)	IF	1:100	Sigma Aldrich
GFP-Booster Alexa Fluor® 647	IF	1:500	Chromotek
GFP-Booster Atto 647N	IF	1:200 (STED)	Chromotek
HRP-Rabbit α Mouse	WB	1:5000	Dako
HRP-Goat α Rabbit	WB	1:10,000	Calbiochem

Table 2 – Table of Antibodies used for western blot analysis and immunofluorescence staining

2.3. Generation of eukaryotic expression construct

The C3C6 domains of cMyBP-C was amplified from adult human heart cDNA using forward [TTT CTC GAG TGA GCC CCC TGT GCT CAT CAC G] and reverse [TTT GGA TCC ACG CGT TTA ACC GAT AGG CAT GAA GGG] primers (Sigma Aldrich) and cloned into the XhoI and BamHI restriction sites of a pEGFP-C2 vector (Clontech; figure 12).

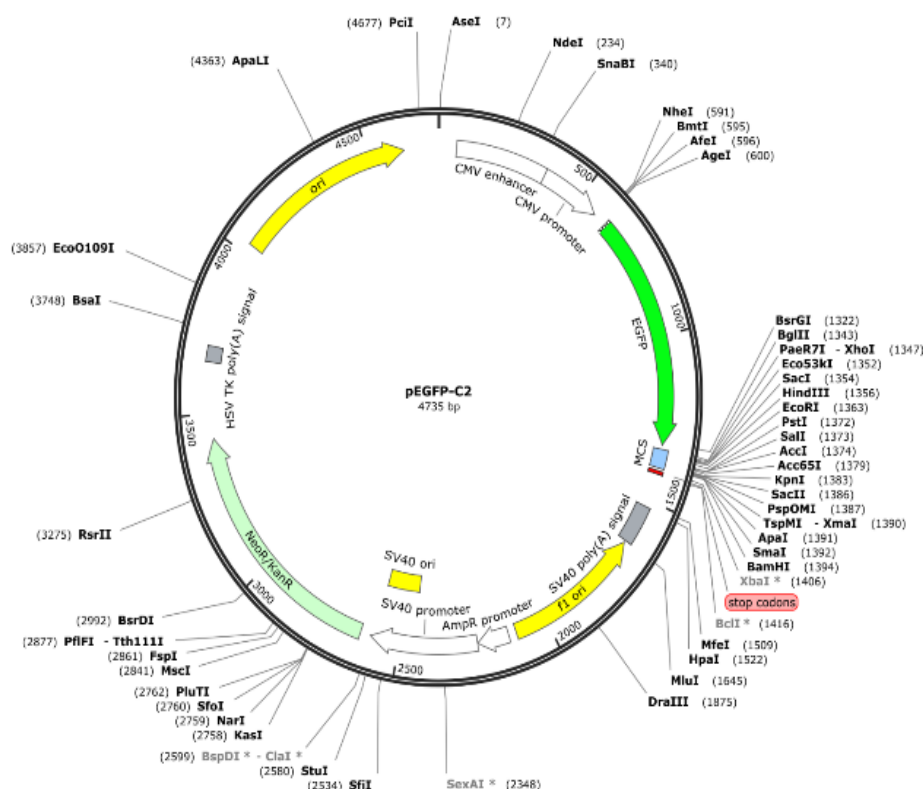


Figure 12 – Plasmid map of the pEGFP-C2 vector

The C3C6 domains of the human cardiac myosin binding protein C (cMyBP-C) was amplified out of a human heart cDNA library and cloned into a pEGFP-C2 vector using the XhoI and BamHI restriction sites. Vector map created with SnapGene.

MyBP-C cDNA library template was mixed with 10 μ M forward primer, 10 μ M reverse primer, Phusion DNA polymerase (NEB), Phusion DNA polymerase buffer (NEB), 10 mM deoxynucleoside triphosphate (dNTP; Qiagen) and H₂O. The mixture was put into a polymerase chain reaction (PCR) machine using the parameters outlined in table 3.

	Temperature (°C)	Time	
Initial denaturation	98 °C	30 sec	
Denaturation	98 °C	30 sec	} 30 Cycles
Annealing	65 °C	30 sec	
Elongation	65 °C	90 sec	
Final elongation	65 °C	10 min	

Table 3 – Thermocycler settings used for amplification of C3C6 domains of cMyBP-C from adult human heart cDNA

A 1% agarose gel with added GelRed nucleic acid gel stain (Biotium) was cast and the PCR product was loaded and run on the gel. The band was visualised using UV light and excised from the gel. DNA was extracted from the gel using a QIAquick gel extraction kit (Qiagen). Restriction digest reactions were set up for both the C3C6 insert and the pEGFP-C2 vector. The C3C6 insert or the pEGFP-C2 vector was mixed with Xho1 restriction enzyme (NEB), BamH1 restriction enzyme (NEB), Buffer 3 (10x; NEB), bovine serum albumin (BSA) and H₂O. The mixtures were incubated at 37 °C for two hours. The pEGFP-C2 vector was dephosphorylated with Antarctic phosphatase (NEB) and buffer 3 (10x; NEB) and incubated for one hour at 37 °C. The pEGFP-C2 vector was heat shocked at 65 °C for 10 mins. The vector and C3C6 insert were both purified using the QIAquick PCR purification kit (Qiagen). A ligation reaction was set up with a 1:3 insert:vector ratio. C3C6 insert, 1µg pEGFP-C2 vector, 1µl T4 DNA ligase (NEB) and 10x T4 DNA ligase buffer (NEB) were mixed together. A control reaction without the insert was also made. The reactions were incubated overnight at +4 °C.

The ligation reaction was added to Top10 Ultracompetent *E. coli* cells (ThermoFisher Scientific) and incubated on ice for 30 mins. The mixture was heat shocked at 42 °C for two minutes and 100x volume of super optimal broth with catabolite repression (SOC) medium (ThermoFisher Scientific) was added and incubated at 37 °C on a shaker for two hours. The mixture was plated onto a Lennox broth (LB; ThermoFisher Scientific)-agar plate containing kanamycin (50 µg/ml) and incubated overnight at 37 °C. A colony was selected from the plate and added to 3 ml LB broth containing kanamycin (100 µg/ml) and incubated at 37 °C in a shaking incubator overnight. The plasmid DNA was extracted using a QIAprep spin mini prep kit (Qiagen).

2.4. Site directed mutagenesis of C3C6 constructs

HCM-linked mutations within the C3C6 region were identified using the ClinVar database (Landrum et al., 2018) and primers were designed to generate these mutations (Table 4; Sigma-Aldrich).

Mutation	Primer sequence
W792R	GGCAGGCGGCTCCCGCTGTACTGTGCAG (Rev)
N755K	GGTCACAGTGAAGAAACCTGTGGGCGAGG (Fwd)
D610H	GTAGTCAGCCTCGTGGGCAGGTGTGACG (Rev)
G596R	GGTGTCCCACATCAGGCGGGTCCACAACTG (Rev)
E542Q	CTCCAGCTTCTTTTGCTGCACAATGAGC (Rev)
G531R	GCACTGTGCACTAGCCGGGGCCAGGCGCTGG (Fwd)
R817Q	CAGCCGCATCCACTGGTAGCTCTTCTTC (Rev)
R820Q	AGAGCTACCGGTGGATGCAGCTGAACTTCGACC (Fwd)
A833T	GATCATGCGCCGCGTTTCATGACTCAG (Rev)

Table 4 – Table of primer sequences used for site directed mutagenesis of C3C6 constructs to generate HCM-linked mutant constructs.

Either 10 μ M forward, 10 μ M reverse or no primer (control) was added to pEGFP-C2 with C3C6 insert, with 10 mM dNTP, 5x high fidelity buffer, Phusion DNA polymerase (NEB) and H₂O. The mixture was then put in a thermocycler using the parameters outlined in Table 5. DpnI (NEB) was added to the mixture and incubated at 37°C for 5 hrs.

	Temperature (°C)	Time	
Initial denaturation	98 °C	90 sec	
Denaturation	98 °C	30 sec	} 30 Cycles
Annealing	65 °C	30 sec	
Elongation	65 °C	30 sec/kb	
Final elongation	65 °C	10 min	

Table 5 – Thermocycler settings used for site directed mutagenesis

The PCR reaction was added to ultracompetent *E. coli* cells (ThermoFisher Scientific) and incubated on ice for 30 minutes. The mixture was heat shocked at 42°C for 30 seconds then put on ice for 1 minute. LB (ThermoFisher Scientific) media was added and the mixture was incubated at 37°C, 200 RPM for 1 hr. The mixture was spread onto LB agar (ThermoFisher Scientific) plates containing kanamycin (50 µg/ml) and incubated at 37°C overnight. Colonies were selected from the plates and LB media with kanamycin (50 µg/ml) was inoculated and put in a shaking incubator at 37°C, 200 RPM overnight. The cells were pelleted by centrifugation at 3000 rpm for 10 minutes, then plasmid DNA was extracted using a QIAprep spin mini prep kit (Qiagen). DNA was then sent off for sequencing in order to verify the correct mutation had occurred (Source Bioscience).

2.5. COS-1 cells

2.5.1. Passage

COS-1 cells (Gluzman, 1981) were taken from liquid nitrogen and thawed, then added to COS-1 cell complete medium (Dulbecco's modified eagle medium - high glucose (DMEM; Sigma-Aldrich), 10% foetal bovine serum (FBS), 1% penicillin/streptomycin). Cells were pelleted by centrifugation at 1200 RPM for 5 minutes. The supernatant was removed then the cell pellet was resuspended in 10 ml COS-1 cell complete medium and the cells were plated on 10 cm plates then put in an incubator at 37°C, 5% CO₂ for 24 hrs.

Cells were washed with phosphate buffered saline (PBS) then trypsin-ethylenediaminetetraacetic acid (EDTA) (Sigma-Aldrich) was added and incubated for 5 minutes. COS-1 cell complete media was added to the cells and centrifuged at 1000 RPM for 5 minutes. The supernatant was removed and the cell pellet was resuspended in media then plated onto 35mm plates and incubated at 37°C, 5% CO₂ for 30 minutes before transfection.

2.5.2. Transfection

1µg of DNA (pEGFP empty vector, pEGFP C3C6 WT or pEGFP C3C6 mutant) was added to 100µl of media 199 (M199; Sigma-Aldrich) then 3 µl of Escort IV (Sigma-Aldrich) was added and mixed, then incubated for 15 minutes at room temperature (RT) and the mixture was added dropwise to the cells. The cells were incubated at 37°C, 5% CO₂ for 48 hours.

2.5.3. Immunofluorescence staining

Cells were washed with PBS then were fixed with 4% paraformaldehyde (PFA) in PBS for 10 minutes. Cells were permeabilised with 0.2% TritonX-100 in PBS for 1 minute and washed 3x 5 minutes in PBS.

Cells were incubated in counterstain solution (4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI), Alexa Fluor 546 phalloidin; Table 2) diluted in 1% BSA/Gold buffer in a humid chamber on a rotating shaker at room temperature for 1 hour. Cells were then washed 3x 5 minutes in PBS, then cells were mounted using Lisbeth's mounting medium (0.1 M Tris-HCl/glycerol and 50 mg/mL *n*-propyl-gallate at pH 9.5).

2.5.4. Western blot analysis

Following transfection COS-1 cells were washed with PBS then 50 µl loading buffer (3.7M urea, 134.6 mM Tris pH 6.8, 5.4% SDS, 2.3% NP-40, 4.45% beta-mercaptoethanol, 4% glycerol, 6mg/ml bromophenol blue) was added and cells were scraped and transferred to an Eppendorf tube. The cell lysate was then boiled for 2 minutes.

Cell lysate samples were run on a precast 4-20% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel (Generon) then transferred onto a nitrocellulose membrane (GE life sciences) via wet transfer. Ponceau Red stain (Sigma) was used to verify transfer was successful.

The membrane was blocked for 1 hour on a rotator at room temperature (RT) in 5% skimmed milk (Sainsbury's) in low salt (LS) buffer. The membrane was then incubated in primary antibody solution (Mouse α GFP; Table 2) in 3% milk/LS for 1 hour at RT on a shaker. The membrane was washed 3x 5 mins in 3% milk/LS, then incubated in secondary antibody solution (horseradish peroxidase (HRP)-Rabbit α mouse; Table 2) on a shaker at RT for 1 hour. The membrane was washed 3x 5 mins in LS then incubated in Clarity™ western enhanced chemiluminescence solution (ECL, Biorad). The membrane was then developed using a BioRad gel imaging machine.

2.6. Neonatal rat cardiomyocytes (NRC)

2.6.1. Isolation and culture

Rat pups were sacrificed, a neck tie cut was performed and the heart removed. The hearts were rinsed with 1x ADS and minced. Neonatal rat cardiomyocytes (NRCs) were isolated via sequential digest by adding collagenase (Worthington) and pancreatin (Sigma-Aldrich) enzyme solution and incubating at 37 °C, 200 RPM for 15 minutes. After each sequential digest the tissue was triturated then the supernatant was removed and strained, pelleted at 800 RPM then resuspended in plating medium. The preparation of cardiomyocytes was enriched by pre-plating the cells in plating medium for 90 minutes at 37°C, 5% CO₂ in order to remove most remaining fibroblasts. The supernatant was then removed and plated on 35mm dishes which had been pre-coated with collagen (Advanced BioMatrix). The cells were then incubated at 37°C, 5% CO₂ overnight.

2.6.2. Transfection

Media on cells was changed to transfection media 2 hours prior to transfection. 1µg of DNA (pEGFP empty vector, pEGFP C3C6 WT or pEGFP C3C6 mutant) was added to 100µl of M199, 3 µl Escort III (Sigma-Aldrich) was added and the mixture was incubated for 15 minutes at room temperature; the mixture was then added dropwise to the media. After 5 hours the media was changed to maintenance media. The cells were incubated at 37°C, 5% CO₂ for 48 hours.

2.6.3. Immunofluorescence staining

Cells were washed with PBS then fixed with 4% PFA/PBS for 10 minutes. Cells were then permeabilised with 0.2% TritonX-100 in PBS for 5 minutes. Cells were then incubated in

primary antibody solution (Mouse α myomesin; Table 2) diluted in 1% BSA/GB in a humid chamber on a rotating shaker at RT for 1 hr. The cells were washed 3x 5 minutes in PBS, then incubated in secondary antibody/counterstain solution (DAPI, Cy3 α mouse, Alexa Fluor 633 phalloidin; Table 2) diluted in 1% BSA/GB in a humid chamber on a rotating shaker at RT for 1 hr. Cells were then washed 3x 5 minutes in PBS, then cells were mounted using Lisbeth's mounting medium.

2.7.1. Confocal microscopy and image analysis

Cells were imaged using a laser scanning SP5 II inverted confocal microscope (Leica microsystems, Wetzlar, DE), equipped with blue diode 405, argon, helium neon 456 and helium neon 633 lasers, using a 63x NA 1.4 oil immersion lens with a 2x zoom and Leica application suite Advanced Fluorescence software.

Figures were compiled using Adobe Photoshop software. Nuclear localisation of GFP-tagged constructs was analysed using Mathematica 11.3 software (Wolfram, Champaign IL), DAPI was used to make a binary mask to extract out nuclear pixels in the green channel and for those cells containing green in the nucleus the average fluorescence intensity per pixel was determined. (Image analysis in collaboration with Dr. Mark Holt).

2.7.2. STED microscopy

Resolution of images obtained via microscopy is constrained by the diffraction limit, which is the limit at which two objects cannot be resolved and is dependent on both the wavelength of the light and the aperture of the optics used (Abbe, 1873). Various techniques have been developed to overcome the diffraction limit and increase resolution, including stimulated emission depletion (STED) microscopy. STED works by using an excitation laser beam and a donut shaped depletion laser beam, which bleaches the peripheral fluorophores back to ground state. Therefore, only molecules in the centre of this ring emit fluorescence (Hell and Wichmann, 1994).

2.7.2.1. Immunostaining

NRCs were fixed, permeabilised and immunostained as described above with primary antibody (mouse α alpha-actinin, rabbit α MyBP-C; Table 2) and secondary

antibody/counterstain solution (Cy3 α mouse, Atto 647n α rabbit, GFP booster Atto 647n; Table 2). Cells were mounted using 70% glycerol/PBS according to manufacturer instructions (Abberior).

2.7.2.2. STED imaging & analysis

Cells were imaged using a STEDYCON (Abberior) attached to a Leica TCS SP5 confocal microscope equipped with 594 nm and 640 nm excitation lasers and a 775 nm depletion laser using a 100x NA 1.4 oil immersion lens. Images were recorded at a pixel size of 20 nm. Figures were compiled using Adobe Photoshop software.

Analysis of the images was carried out in collaboration with Dr Mark Holt, a general scheme of which can be seen in figure 13. For analysis the STED images along with their respective metadata-containing text files were imported into Wolfram Mathematica 12.2 (Champaign, IL). Images were then resized to 20 nm x 20 nm per pixel using information in the metadata file.

The images were then transformed using a Fast Fourier Transform (FFT) and converting to absolute magnitude to give a 2D spectrogram. The original image was multiplied by a Gaussian kernel with the same image dimensions to reduce edge artefacts. An inverse FFT was then applied to this to produce a 2D cepstrogram. This shows a linear series of bands emanating from the centre of the spectrogram with a spacing that is the same as the original image. The cepstrogram was then rotated so that the major axis of the signal was horizontal. The cepstrogram was then further transformed using a Radon Transform which shows distance of signal versus angle/direction. This has the benefit of smearing any background noise to give a cleaner signal. A line scan through the middle of the Radon Transform was then taken to visualise intensity as a profile plot. These profiles were used to

calculate the average sarcomere length (the major non-zero peak in the profiles, corresponding to alpha-actinin staining). These plots were then normalised to take account of slight variations in sarcomere length and the mean and standard deviations calculated for each mutant and control.

Images of cells displaying good MyBP-C localisation produced profiles that had two further intensity peaks corresponding to the MyBP-C signal positioned between the two major peaks in the profile. This was accentuated by removing the baseline signal using the Difference of Gaussians (DOG) approach (similar to background subtraction as commonly used in image processing). It was then possible to extract the intensity of the first MyBP-C peak as a measure of signal integrity. It was found that this peak was the best indicator for this.

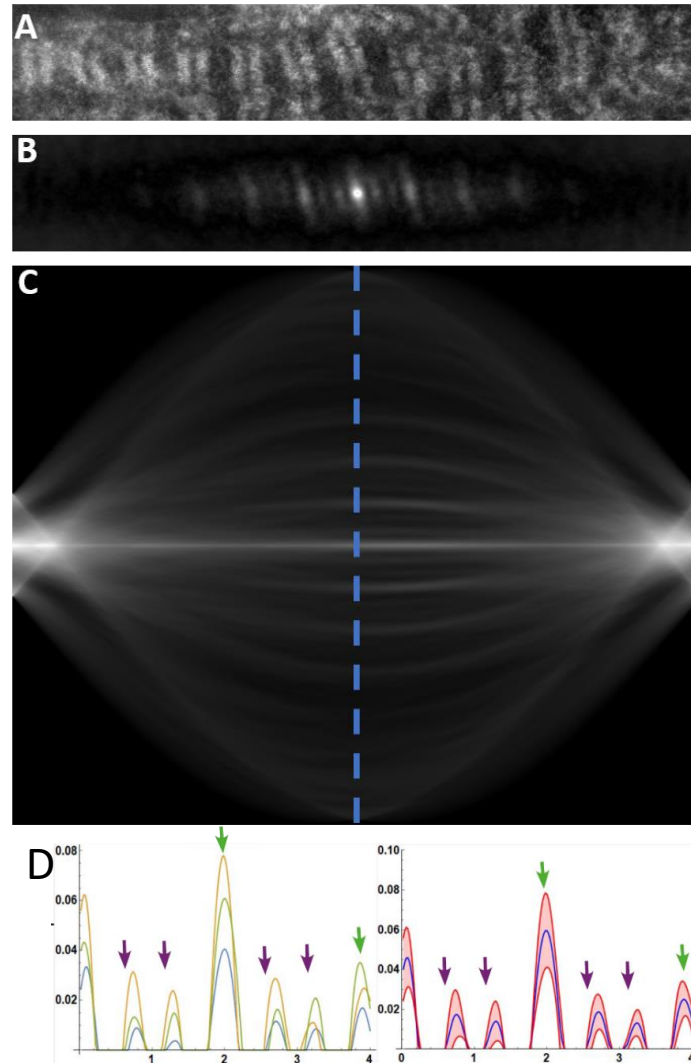


Figure 13 – Scheme for analysis of STED images

Diagram showing an example for the general scheme for analysis of images of neonatal rat cardiomyocytes images from STED microscopy. The original image (A) was transformed using a Fast Fourier Transform (FFT) to give a 2D spectrogram. An inverse FFT was then applied to give a 2D cepstrogram (B). The cepstrogram was then further transformed using a Radon Transform (C). A line scan was then taken through the middle of the Radon Transform (C; dashed line) to visualise intensity as a profile plot (D). The major non-zero peaks (green arrow) denote alpha-actinin signal and thus sarcomere length. The two intensity peaks positioned between the major peaks (purple arrows) correspond to the MyBP-C signal.

2.8. Full length cMyBP-C adenovirus

2.8.1. AV propagation

A pAd vector containing the full length human cMyBP-C tagged at the N-terminus with GFP was generated by Vectorbuilder (figure 14). The plasmid was added to ultracompetent *E. coli* cells (ThermoFisher Scientific) and incubated on ice for 30 minutes. The mixture was heat shocked at 42°C for 30 seconds then put on ice for 1 minute. LB (ThermoFisher Scientific) media was added and the mixture was incubated at 37°C, 200 RPM for 1 hr. The plasmid was streaked onto LB agar plates containing 100 µg/ml ampicillin; colonies were picked, grown up in 3ml LB broth containing 100 µg/ml ampicillin and the DNA was extracted by a mini-prep kit (Qiagen). The DNA was digested with 1µl Pac-1 restriction enzyme (NEB) per µg of DNA for 1 hour at 37°C in order to linearize the plasmid. The DNA was then ethanol precipitated, 1/10 volume of sodium acetate (NaAC; 2M, pH 5.2) was added to the Pac-1 digested DNA, then 3x volume of ethanol was added and was left to incubate overnight at RT. The mixture was centrifuged at 14K RPM for 30 minutes, the supernatant was discarded and the pellet was washed with 70% ethanol. The mixture was centrifuged again for 15 minutes, the supernatant was discarded and the DNA was resuspended in TE buffer.

HEK 293 cells were then transiently transfected with the pAd full-length cMyBP-C DNA. HEK 293 cells were cultured in complete medium (DMEM, 10% FBS, 1% Pen/strep) in 6 well plates until they reached 70% confluency. 1 µg of pAd DNA was added to 250 µl DMEM then p3000 reagent (ThermoFisher Scientific) was added. In a separate tube 3 µl lipofectamine 3000 (ThermoFisher Scientific) was added to 250 µl DMEM and incubated for 5 minutes at RT. The DNA and lipofectamine 3000 mixtures were added together and allowed to incubate

for 20 minutes at RT. The mixture was then added dropwise to the HEK 293 cells and incubated at 37°C, 5% CO₂ for 24 hours, after which media was changed to complete medium. 48 hours after transfection cells were trypsinised and replated in a 10 cm dish.

The infection was allowed to continue until 80% cytopathic effect was seen, adenovirus containing cells were harvested and then underwent 3x freeze-thaw cycles (-80°C for 30 minutes followed by 37 °C for 15 minutes). The cell lysate was then centrifuged at 3000 RPM for 15 minutes and the supernatant was taken and stored at -80°C. This crude lysate was then added to HEK cells and left for 48 hours until 80% cytopathic effect was seen. Adenoviral containing cells were harvested and underwent 3x freeze-thaw cycles (-80°C for 30 minutes followed by 37 °C for 15 minutes). The lysate was the centrifuged and supernatant was then taken and stored at -80 °C

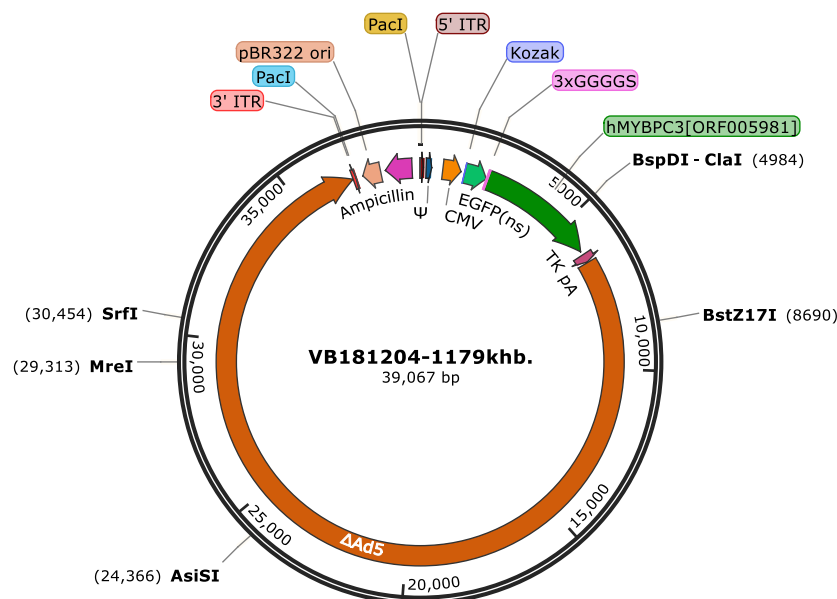


Figure 14– Plasmid map of the pAd vector containing the human full length cMyBP-C

A pAd vector containing the human full length cardiac myosin binding protein C was generated by Vectorbuilder. This was then propagated and amplified in HEK293 cells. Vector map created with SnapGene.

2.8.2. SDM of pENTR vector

The full length human cMyBP-C gene was amplified out of the pAd vector as described in 2.3. using forward [caattctctaaggaaataacttaaccATGGTGAGCAAGGGCGAG] and reverse [gtctagatatctcgacggatcctaTCACTGAGGCACTCGCAC] primers (Sigma Aldrich) and cloned into a pENTR vector (Thermo Fisher) using the Gibson assembly technique.

50 ng of pENTR vector was mixed with 100 ng of insert and 2x assembly mix (NEB) which was incubated at 50 °C for 15 minutes. The mixture was then added to ultracompetent E. coli cells (ThermoFisher Scientific) and incubated on ice for 30 minutes. The mixture was heat shocked at 42°C for 30 seconds then put on ice for 1 minute. 100x volume of LB (ThermoFisher Scientific) media was then added and the mixture was incubated at 37°C on a shaker for 1 hr. The mixture was spread onto LB agar (ThermoFisher Scientific) plates containing kanamycin (50 µg/ml) and incubated at 37°C overnight. Colonies were selected from the plates and 3 ml of LB media with kanamycin (50 µg/ml) was inoculated and put in a shaking incubator at 37°C overnight. The cells were pelleted by centrifugation at 3000 rpm for 10 minutes, then plasmid DNA was extracted using a QIAprep spin mini prep kit (Qiagen). DNA was then sent off for sequencing in order to verify the insert was correct (Genewiz).

Site directed mutagenesis was then carried out on the pENTR full-length cMyBP-C construct as outlined in 2.2. LR recombination was then carried out between the pENTR full-length cMyBP-C constructs with the desired mutations and the pAd vector. 100 ng of the entry clone was added to 300 ng of destination vector (donated by Dr Luke Smith) and TE buffer. 2 µl of LR Clonase II enzyme mix (Thermo Fisher) was then added and incubated at RT overnight. 1 µl of Proteinase K (Thermo Fisher) was added to the mixture and incubated at

37 °C for 10 minutes. The reaction mixture was then transformed into ultracompetent E. coli cells (Thermo Fisher) as previously described and plasmid DNA was extracted using the QIAprep spin mini prep kit (Qiagen).

Mutant pAd full-length cMyBP-C constructs were then propagated and amplified as outlined in 2.8.1.

2.8.3. Titration of AV

Adenovirus (AV) were titred using the AdEasy viral titer kit (Agilent Technologies). HEK293 cells were plated in 24 well plates at a density of 2.2×10^5 in complete medium. A 10-fold dilution series was carried out for each AV construct over dilutions 10^{-2} to 10^{-6} . Each dilution was then added in duplicate to the HEK293 cells and the cells were incubated at 37 °C, 5% CO₂ for 24 hours. The medium was aspirated and the plate was left for 10 minutes to dry. The cells were then fixed with ice-cold 100% methanol and incubated at -20 °C overnight.

The methanol was aspirated and the cells were washed 2x with 1x PBS with 1% BSA (PBS/BSA). Mouse anti-hexon antibody diluted 1:500 in PBS/BSA was then added and incubated at 37 °C for 1 hour. The antibody was aspirated and the cells were washed 2x with PBS/BSA, then HRP-conjugated goat anti-mouse antibody diluted 1:1000 in PBS/BSA was added and incubated at 37 °C for 1 hour. The antibody was aspirated and the cells washed 2x with PBS/BSA then DAB 1x working solution (1 part 10x DAB substrate, 9 parts peroxide buffer) was added and incubated at RT for 15 minutes.

The number of positive cells in 10 randomly chosen fields were counted using a 20x microscope objective. Infectious units (IFU) were then determined using the formula:

average number of positive cells per field x fields per well / volume of diluted virus used in each well (ml) x dilution factor.

2.8.4. Determination of adenoviral concentration and infection duration

COS-1 cells were cultured as described in 2.5.1. Cells were infected with WT AV full length cMyBP-C at dilutions ranging from 1:100 – 1:4. Cell viability was assessed using both visual observation to determine number of GFP-expressing cells versus amount of visible cell death and western blot analysis as described above to determine amount of protein produced (primary antibody – Rabbit α cMyBP-C; secondary antibody – HRP-Goat α Rabbit; Table 2).

Once the best dilution was chosen NRCs were infected with the WT AV construct and the infection was allowed to continue for 24 or 72 hours and cell viability was assessed using both visual observation and western blot analysis, as described above, to determine the best infection duration.

2.8.5. Infection of NRCs

NRCs were isolated as described in 2.3.1. The day following isolation the media on the cells was changed to maintenance media and infected with either WT or mutant adenovirus at a dilution of 1:100 for the WT construct. Mutant constructs were added at an equivalent dilution based on the IFU calculated via titration.

Cells were incubated for 24 hours at 37 °C, 5% CO₂ then either SDS samples were made for western blot analysis to determine protein expression levels (primary antibody - mouse α GFP or rabbit α MyBP-C, secondary antibody – HRP-rabbit α mouse or HRP-Goat α rabbit) or

cells were fixed and immunostained for confocal microscopy as previously described (conventional confocal microscopy: GFP booster Alexafluor 647, DAPI, Alexa Fluor 633 phalloidin (Table 2); STED microscopy as described in 2.6.1.).

2.9. Protein expression & purification

2.9.1. Protein expression optimisation

A pET-His-Tev plasmid containing the C3C6 insert (donated by Dr. Thomas Kampourakis) was transformed into BL21(DE3)RIPL E.Coli cells (Agilent) and plated on LB-Agar plates containing ampicillin (100 µg/ml). Colonies were picked and 20 ml LB containing ampicillin (100 µg/ml) and chloramphenicol (35 µg/ml) were transfected and grown up overnight in a shaking incubator at 37°C.

Cells were pelleted by centrifugation at 4500 RPM for 10 minutes then the pellet was resuspended in 1 L of LB containing ampicillin (100 µg/ml). The culture was incubated in a shaking incubator at 37 °C until reaching an OD600 of 0.6. Protein expression was then induced by adding 0.1 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG; VWR chemicals) and incubating in a shaking incubator at 37 °C for 3 hours. Cultures were then pelleted by centrifugation at 5000 RPM for 10 minutes.

Protein expression was checked by western blot using an anti-His antibody (Table 2). Cell pellet was then resuspended in BugBuster Master Mix (Millipore) with EDTA-free protease inhibitor (Roche) and 1 mM dithiothreitol (DTT) and left on a rotator for 20 minutes at RT. The lysate was then centrifuged at 15,000 g at 4 °C for 30 minutes and the supernatant was sterile filtered using a 0.45 µm filter.

Protein was then purified using a HisTrapFF column (GE life sciences). The column was equilibrated with 25 column volumes (CV) buffer A (20 mM HEPES pH 7.5, 300 mM NaCl, 1 mM DTT, 20 mM imidazole), then lysate (Fraction L) was loaded onto the column and the flow through (FT fraction) was collected. The column was then washed with 10 CV buffer A (Fraction W). Protein was eluted from the column with 10 CV of a 50:50 mixture of Buffer A

and Buffer B (20 mM HEPES pH 7.5, 300 mM NaCl, 1 mM DTT, 500 mM imidazole) (Fractions E1 and E2). Protein was then eluted using 10 CV of 100% buffer B (Fractions E3 and E4).

Blue Wonder loading buffer was then added to the fractions (L, FT, W, E1, E2, E3 and E4), the samples were boiled for 2 minutes then run on a 4-20% pre-cast SDS-PAGE gel (Generon). The gel was stained for protein using simply blue safe stain (Invitrogen).

Following low and impure yield of protein, optimisation of protein expression was carried out, using Terrific broth (TB; Sigma-Aldrich) instead of LB and inducing using either 0.1 mM or 1 mM IPTG at either 37°C for 3 hours or 18°C overnight. Due to multiple bands appearing following staining of the gel, indicating protein degradation, additional EDTA-free protease inhibitor was added to the lysis buffer and lysis was instead carried out at +4°C for 40 minutes.

Following these optimisation steps, yield was still seen to be very low and there was still evidence of protein degradation occurring. Therefore, a pET-His-Tev plasmid containing the C3C5 domains (donated by Dr Thomas Kampourakis) was expressed and purified using the same steps outlined above instead.

Protein containing fractions were pooled and TEV protease was added at a concentration of 1:50 to remove the His tag and was then dialysed against a solution of 25 mM HEPES, 200 mM NaCl, 1 mM MgCl₂, 1 mM DTT, 5% glycerol, overnight at +4 °C. Dialysed protein was then run again on the HisTrap FF column. The column was pre-equilibrated with buffer A, then digested protein was applied to the column and FT was collected. The column was then washed with 10 CV of buffer A (Fractions W1A & W1B), followed by 10 CV of buffer B (Fractions W2A & W2B).

2.10.1. Native thin filament preparation

Native thin filaments (NTFs) were isolated from bovine heart tissue. Bovine heart tissue was homogenised using a gentle MACS homogeniser in salt buffer (100 mM KCl, 3 mM MgCl₂, 10 mM imidazole, 5 mM DTT pH 7.3). Three volumes of 1:1 salt buffer:glycerol was added to the homogenised tissue and the mixture underwent gentle agitation at 4 °C for six hours. The glycerinated mixture was stored at -20 °C overnight.

An equal volume of salt buffer was added to the glycerinated tissue and the mixture was blended using a Sorvall Omnimix. Afterwards the mixture was centrifuged at 10,000 x g using a F12S rotor in a Sorvall evolution RC centrifuge for five minutes. The supernatant was discarded and the pellet was resuspended in 2 volumes of salt buffer supplemented with 0.6% TritonX-100. The mixture then underwent gentle agitation at 4 °C for 1 hour. This step was repeated a further three times. After the final centrifugation, the pellet was resuspended in 2 volumes of salt buffer without Triton X-100 and was centrifuged at 10,000 x g for 5 minutes. This wash step was repeated a further 3 times to ensure the removal of the Triton. The pellet was then resuspended in 4 volumes of Buffer 1 (100 mM KCl, 3 mM MgCl₂, 3 mM EDTA, 5 mM imidazole, 5 mM DTT, pH 7.5) and centrifuged at 10,000 x g for 5 minutes. The pellet was then resuspended in 0.5 volumes of buffer 1 and 5 mM of ATP was added. The mixture was then blended for 1 minute and centrifuged at 5,000 x g for 10 minutes.

The supernatant was taken and the pH was adjusted to pH= 5.5-6. The supernatant was then centrifuged consecutively using a F50L rotor in a Beckman optima LE 80K ultracentrifuge at 30,000 x g for 30 minutes, 50,000 x g for 30 minutes, 80,000 x g for 180 minutes. The supernatant was removed and the pellet was overlaid with thin filament

storage buffer (50 mM KCl, 2 mM MgCl₂, 0.4 mM ATP, 0.5 mM DTT, 5 mM imidazole, pH 7) and incubated on ice overnight. The pellet was resuspended in the thin filament storage buffer and incubated on ice for 30 minutes, then centrifuged at 30,000 x g for 30 minutes.

The supernatant was taken and undiluted, 1:3 and 1:6 dilutions of the sample were run on a 15% SDS gel which was subsequently stained with Coomassie to check the presence of the thin filament proteins.

2.10.2. ATPase assay to check NTF activity

ATPase activity of the NTFs was determined via performing an ATPase assay using the ATPase/GTPase enzyme linked inorganic phosphate assay (ELIPA) Biochem kit (Cytoskeleton Inc.). This assay utilises the absorbance shift from 330 to 360 nm which occurs when 2-amino-6-mercapto-7-methylpurine riboside (MESG) is catalytically converted by purine nucleotide phosphorylase (PNP) in the presence of inorganic phosphate (P_i). One molecule of P_i causes one molecule of MESG to be converted therefore absorbance at 360 nm is directly proportional to the amount of P_i generated by myosin ATPase

NTFs were dialysed against ATPase assay buffer (10 mM MOPS, 5 mM MgCl₂, 1 mM DTT) overnight. A phosphate standard curve was generated by adding inorganic phosphate at different concentrations ranging between 0-50 nM to ATPase assay buffer, 1 µM 2-amino-6-mercapto-7-methylpurine riboside (MESG) and 1 µM purine nucleotide phosphorylase (PNP) in a 96 well plate. The mixture was allowed to incubate for 15 minutes at RT and then the absorbance was read at 360 nm using a ClarioStar spectrophotometer plate reader (BMG Labtech).

2.10.3. ATPase assay with C3C5

NTFs and purified C3C5 were dialysed against ATPase assay buffer overnight. Different concentrations of purified C3C5 ranging between 0-100 μM were added to 10 μM NTFs, 0.02 μM Myosin S1, 1 μM MESG, 2 μM PNP and either 0.1 mM CaCl_2 (with calcium) or 1 mM ethylene glycol tetra-acetic acid (EGTA) (no calcium) in a 96 well plate. The absorbance was then read at 360 nm for a few cycles to achieve a baseline reading. 1 mM of ATP was then added and the run was continued and absorbance at 360 nm was read for another 60 cycles. Absorbance levels over time were plotted using GraphPad Prism and the slope was calculated for each concentration of purified protein.

2.11. NTF co-sedimentation

NTFs, purified C3C5 and purified C0C2 (donated by Dr Thomas Kampourakis) were dialysed against NTF co-sedimentation buffer (20 mM imidazole pH 7.4, 180 mM KCl, 1 mM MgCl₂, 1 mM EGTA pH 8, 1 mM DTT) overnight.

Purified C3C5 and C0C2 were added to 10 µM NTFs at either 25 µM or 50 µM concentration in an ultracentrifuge tube with NTF co-sedimentation buffer to a final volume of 50 µl and 1 mM ATP and 1 mM DTT was added. The samples were incubated at RT for 30 minutes then centrifuged at 100,000 x g for 1 hour in an F50L rotor using a Beckman optima LE 80K ultracentrifuge.

Urea sample buffer (8 M urea, 20 mM Trisbase pH 7, 23 mM glycine, 0.2 mM EDTA, 5% glycerol, 10 mM DTT, 0.004% bromophenol blue) was added to the supernatant and the pellet was resuspended in a 50:50 urea sample:co-sedimentation buffer. Samples were boiled for 5 minutes then run on a 4-20% gradient gel and stained with Coomassie.

2.12. Yeast Two Hybrid (Y2H) assay

2.12.1. Lithium acetate/carrier DNA transformation

The C3C6 fragment, as well as the individual C4 and C5 domains, were amplified from the pEGFP C3C6 vector using forward [TTTGAATTCTGGCGGACGCAGCCTACCAG] and reverse [TTTGTCTGACGCTGGGCCTGGACATGCCG] primers and cloned into an N-terminal LexA fusion protein (Takara; figure 15) using the EcoRI and Sall restriction sites as outlined in 2.1..

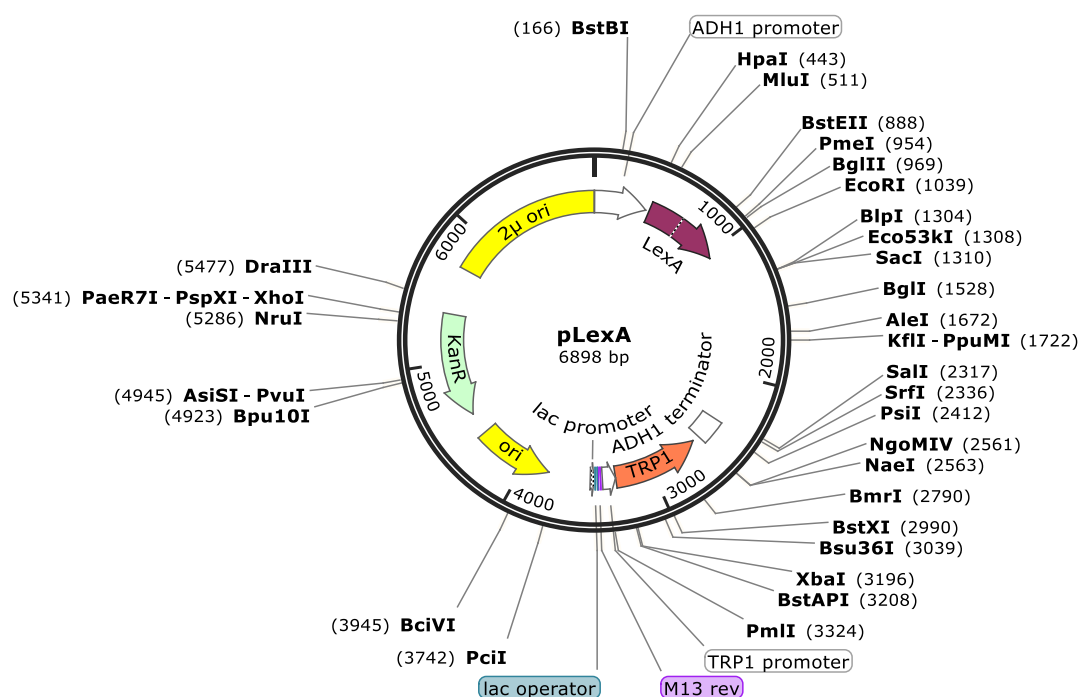


Figure 15 – Plasmid map of the pLexA vector

The C3C6 domains of cardiac myosin binding protein C was amplified out of the pEGFP C3C6 vector and cloned into the pLexA vector using the EcoRI and Sall restriction sites to create a yeast two hybrid bait plasmid. Vector map created with SnapGene.

Bait plasmid DNA was transformed into the L40 yeast strain by the lithium acetate/carrier DNA method. L40 yeast strain was streaked onto a YPD agar plate and incubated for 3 days at 30 °C. Cells were scraped from the plate and suspended in water then pelleted by centrifugation. The supernatant was removed and the cell pellet was again resuspended in

water and centrifuged. The cells were then mixed with 60 µg carrier DNA (Takara), which had been denatured at 98°C for 10 minutes, and 500 ng of bait plasmid DNA or an empty Gal4 construct to test for autoactivation. 500 µl of Transformation buffer (10 mM TrisHCl pH 8, 1 mM EDTA, 100mM Lithium acetate, 40% PEG 3500) was added and the mixture was incubated at 42 °C for 15 minutes. Cells were pelleted by centrifugation then washed 2x with water. Cells were then plated on synthetic defined media lacking tryptophan, histidine and leucine (SD-WHL) agar plate (Takara) and incubated at 30 °C for 3 days.

2.12.2. Autoactivation check

Following transformation of either bait plasmid with empty Gal4 construct or positive control DNA (pLexA KIN3 and pGadΔE, donated by Dr Luke Smith) in L40 yeast strain, colonies were picked and resuspended in water then spotted onto SD-WHL agar plates and incubated at 30 °C for 2-3 days.

Colonies were then picked and spotted onto a nitrocellulose membrane on a SD-WHL agar plate and incubated at 30 °C for 2-3 days. The nitrocellulose membrane was then freeze-thawed 3x in liquid nitrogen. The membrane was then placed on top of Whatman paper pre-soaked in X-gal buffer (60mM NaH₂PO₄, 10mM KCl, 1mM MgSO₄, 38mM β-mercaptoethanol, 0.33 mg/ml X-gal) and incubated at 30 °C until blue colouration was seen for the positive control.

2.12.3. LexA fusion expression check

Following transformation of bait plasmid DNA, cells were spread on SD lacking tryptophan (SD-W) agar plates (Takara) and incubated at 30 °C for 2-3 days. Colonies were picked and used to infect 3 ml of SD-W media (Takara) and incubated at 30 °C, in a shaking incubator

for 48 hours. Cells were pelleted by centrifugation, 4000 RPM at +4 °C for 10 minutes, then the pellet was resuspended in 100 µl of Y2H sample buffer (loading buffer, 100 mM PMSF, protease inhibitors).

Samples were boiled then run on a 4-20% precast gel (Generon) and wet transferred onto a nitrocellulose membrane. The membrane was blocked in 5% milk/LS buffer, incubated in primary antibody (Rabbit α LexA; Table 2) then secondary antibody (Goat α Rabbit; Table 2). The blot was developed using Clarity™ western ECL (Biorad) and imaged using a Biorad gel imaging machine.

2.12.4. Y2H assay screen

Following transformation of bait plasmid DNA, cells were spread on SD-W agar plates (Takara) and incubated for 2-3 days at 30 °C. Several colonies were then picked and inoculated in 3 ml of SD-W media (Takara) and incubated at 30 °C in a shaking incubator overnight. The culture was centrifuged at 4000 RPM for 3 minutes and the pellet was resuspended in enough yeast peptone dextrose (YPD; Takara) media to reach an OD600 of 0.2 then incubated at 30 °C in a shaking incubator until the culture reached an OD600 of 0.5. The culture was centrifuged at 4000 RPM for 3 minutes, the pellet was resuspended in YPD media then mate and plate human cardiac bait library (Takara) was added at a ratio of 2:1 bait:library. The bait and library mixture was centrifuged at 4000 RPM for 3 minutes, the pellet was resuspended in 200 µl YPD media, then spread on YPD agar plates and incubated at 30 °C for 5 hours.

Cells were scraped from the plates with YPD media and centrifuged at 4000 RPM for 3 minutes. The pellet was resuspended in 4 ml of water then plated on 10x SD-WHL agar plates and incubated at 30 °C until colonies appeared.

Colonies were selected and resuspended in water then inoculated in 3 ml of SD lacking tryptophan and leucine (SD-WL) media and incubated in a shaking incubator at 30 °C for 3 days. Cells were centrifuged at 4000 RPM for 5 minutes and the pellet was resuspended in 3 ml Y2H lysis buffer (10 mM Tris HCl, 1% β -mercaptoethanol, 5 U lyticase, pH 8.0) and incubated at 37 °C in a shaking incubator for 90 minutes. Lysate was freeze-thawed 3x in liquid nitrogen then DNA was extracted using the QIAprep spin miniprep kit (Qiagen). DNA was sent for sequencing (Genewiz) and potential binding partners were identified using a Blast database search.

DNA of any promising hits was then forced crossed with the bait plasmid, using the lithium acetate/carrier DNA transformation method described previously. Colonies were then picked and spotted onto nitrocellulose. The nitrocellulose membrane was freeze-thawed 3x in liquid nitrogen then the membrane was placed onto Whatmann paper that had been pre-soaked in X-gal buffer (see 2.12.2). If blue colouration occurred this indicated a true interaction was happening between the DNA from the Y2H screen hit and the bait plasmid.

3. Characterisation of the wild-type central domains of cardiac myosin binding protein C

Whilst many studies have been carried out to determine the structure and function of cMyBP-C, very little is understood about the functionality of the central domains of the protein. It is still unclear whether the central region merely functions as a spacer to ensure the N-terminal domains are in the correct proximity to actin and myosin or if they have a more functional role and contain binding sites for potential interaction partners.

In order to gain a clearer picture of the role of these central domains, both eukaryotic and bacterial constructs containing the central domains of the protein were utilised to study the subcellular localisation of the central domains, as well as elucidate possible functional roles of this region of the protein. These constructs contained only the central region of the protein so that localisation and functional effects could be observed without the influence of the surrounding N- and C-terminal domains of the protein.

3.1. Localisation of the central domains of cMyBP-C

To study the localisation of the central domains of cMyBP-C, a eukaryotic GFP-tagged C3C6 containing construct was generated by cloning the C3C6 region, amplified from human heart cDNA, into a pEGFP vector.

First this construct was transiently transfected into COS-1 cells to check expression of the construct, then samples were run on an SDS-PAGE gel and immunoblotted to validate the construct had the correct molecular weight of 73 kDa. As can be seen in figure 16, a single band was shown at just below 75 kDa for the GFP-tagged C3C6 domain protein, which is the expected molecular weight, with no band being seen in the negative control of non-transfected (NT) cells. EGFP alone showed a band at 27 kDa which is its expected molecular weight.

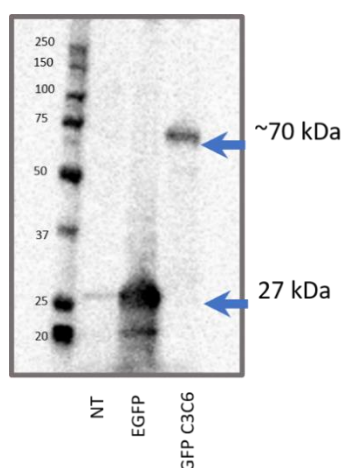


Figure 16 – Validation of the GFP-tagged C3C6 domain containing construct

In order to validate the GFP-tagged C3C6 domain containing construct which was to be utilised in further experiments, western blot analysis was used to determine the construct was the correct molecular weight. The construct, as well as a control of EGFP alone, was transiently transfected in COS-1 cells, samples were taken and run on an SDS-PAGE gel, transferred onto nitrocellulose and immunostained using an anti-GFP antibody. The results showed that the GFP-tagged C3C6 domain containing construct showed the expected molecular weight of 73 kDa, non-transfected (NT) showed no band at the expected molecular weight, whilst EGFP alone showed an expected molecular weight of 27 kDa.

This construct was then transiently transfected into both COS-1 cells and NRCs, then immunostaining and confocal microscopy was utilised to determine localisation in both of these cell types. As can be seen in figure 17B, in COS-1 cells C3C6 showed a diffuse cytoplasmic localisation, with no aggregation of the construct. The construct also showed little to no nuclear localisation (figure 17B; red arrow) when compared to EGFP alone, shown via counterstaining with DAPI, which did show a high level of nuclear localisation (figure 17A; red arrow). Furthermore, there did appear to be some co-localisation between the GFP-tagged C3C6 and F-actin at the periphery of the cells (Figure 17B; blue arrows), which was also seen with the EGFP alone but to a lesser extent (Figure 17A; blue arrows).

Similarly, in the NRCs C3C6 localised in a cytoplasmic, mostly diffuse manner and did not appear to aggregate. It also did not seem to localise to the nucleus in the NRCs. The EGFP alone construct was again seen to show high levels of nuclear localisation, as was seen in the COS-1 cells (Figure 17C, red arrow). Co-localisation with actin was also seen mainly in the myofibrils in the NRCs with the GFP-tagged C3C6 fragment (figure 17D, blue arrows).

From the results of these experiments it seems that the C3C6 domains are stable, as they did not aggregate. The fact that they showed a mainly diffuse cytoplasmic localisation indicates that these domains do not have a role in the incorporation of the protein into the sarcomere, which is as expected as this region does not contain any of the domains which have previously been found to carry out this role, i.e. domains C7-C10 (Okagaki et al., 1993, Freiburg and Gautel, 1996). The co-localisation seen between the GFP-tagged C3C6 domains with actin, which was either absent or much lower with EGFP alone, suggests the possibility that the central domains may be interacting with actin and thus could contain binding domains for actin. Although co-localisation does not necessarily mean that the two proteins

are interacting, merely that the two proteins are in close proximity. However, it did appear that the GFP alone construct did show some decoration of the myofibrils, although to a lower extent than the C3C6 containing construct, suggesting that the GFP tag could be responsible for any co-localisation seen with actin.

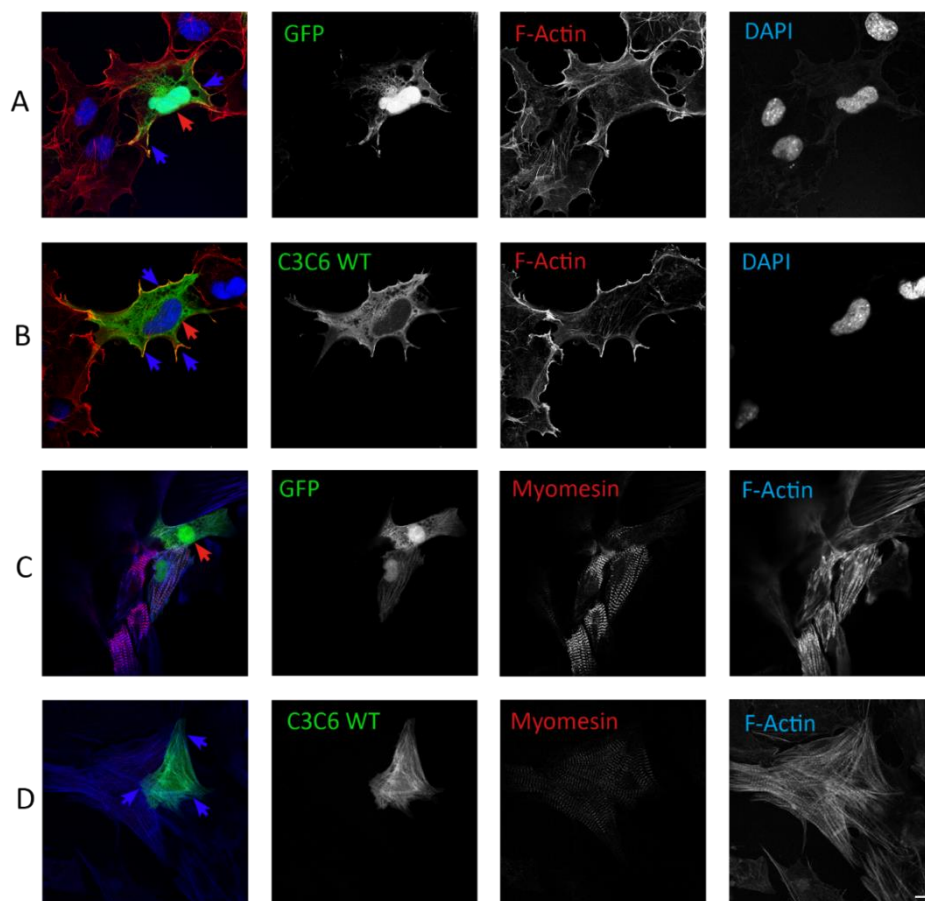


Figure 17 – Localisation of the GFP-tagged wildtype C3C6 containing plasmid

Confocal micrographs of transiently transfected COS-1 cells and NRC. A pEGFP empty vector or pEGFP vector containing the C3C6 domains of cMyBP-C was transiently transfected into COS-1 cells (A & B) and neonatal rat cardiomyocytes (NRC; C & D). Cells were then immunostained for F-actin (red in COS-1 cells; blue in NRCs), myomesin (red in NRCs) or DAPI (blue in COS-1 cells) to determine localisation of the C3C6 construct. The pEGFP C3C6 construct showed a diffuse cytoplasmic localisation in both COS-1 cells and NRCs and also showed some co-localisation with F-actin. The construct did not appear to aggregate suggesting it was not degraded and also did not appear to localise to the nucleus, as the EGFP alone was shown to do. Nuclear localisation indicated by red arrows; co-localisation between central domains and F-actin indicated by blue arrows. Scale bar = 10 μm

3.2. Production and purification of central domain fragments

3.2.1. Production of the C3C6 domain fragment

To study the stability and possible functional role of the central domains, protein production and purification of a His-tagged pET C3C6 construct was carried out in order to produce protein which could subsequently be used in functional studies.

The pET C3C6 domain containing construct was transformed into BL21 (DE3) RIPL E. coli cells and protein production was induced using different induction conditions (IPTG concentration at 0.1 mM or 1 mM; induction conditions of 2 hours at 37 °C or overnight at 18 °C). Cells were lysed and samples run on an SDS-PAGE gel, transferred onto nitrocellulose and probed using an anti-HIS tag antibody (Fig 18A) to determine whether soluble protein was being produced, as protein in the insoluble fraction tends to be aggregated protein that may form inclusion bodies of non-folded, inactive protein (Brondyk, 2009). The results indicated that the strongest band at the expected molecular weight of 50 kDa in the soluble fraction occurred under the conditions of 1 mM IPTG for 2 hours at 37 °C (figure 18A, Arrow).

Lysed protein was then purified on a HisTrap FF column and samples were taken and run on an SDS-PAGE gel which was stained with Coomassie (Fig 18B). A faint band was seen at the correct molecular weight of 50 kDa (figure 18B, arrow), however there was also a lower band seen at just below 40 kDa suggesting degradation of the protein product was occurring (Figure 18B, Asterisk).

Due to this the lysis step of the procedure was changed to occur at +4 °C instead of at RT and an increased amount of protease inhibitor was added to the lysis buffer. Samples were again purified on a HisTrap FF column and samples were run on an SDS PAGE gel which was

stained with Coomassie (Fig 18C). However, there did not appear to be any decrease in the levels of degradation of the protein seen, with the lower band still remaining (figure 18C, Asterisk), and the yield continued to be low (figure 18C, arrow).

Due to this a shorter pET construct containing just the C3C5 domains was trialled to see if a better yield could be achieved.

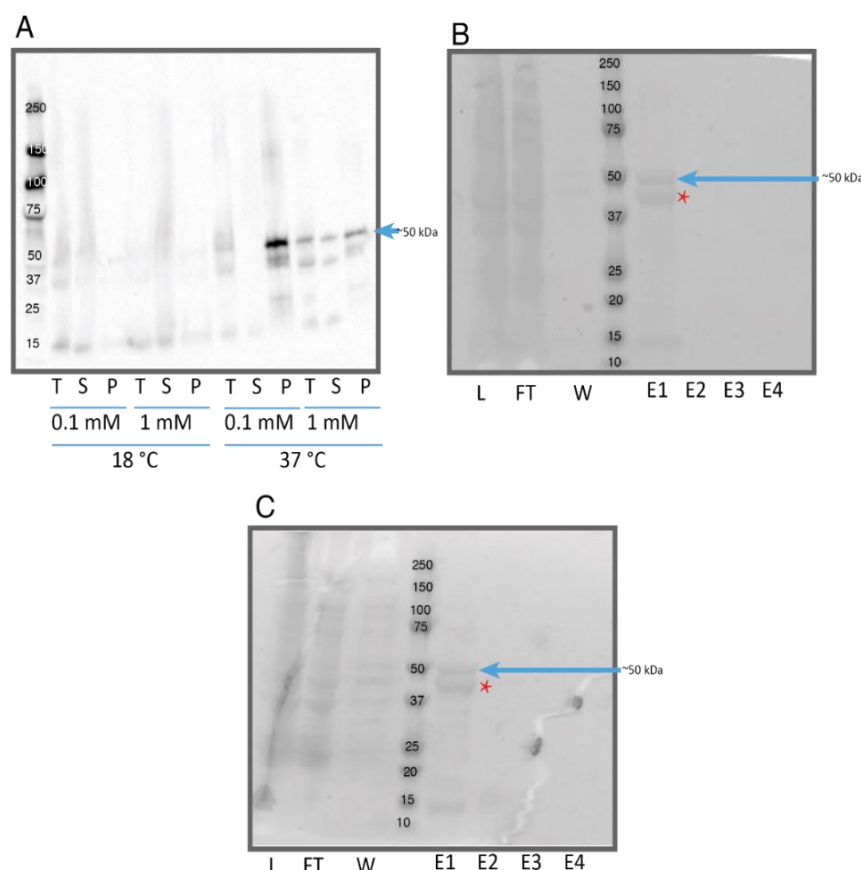


Figure 18 – Optimisation of C3C6 protein induction

Immunoblot and Coomassie stained SDS-PAGE gels of recombinantly expressed constructs. A pET C3C6 plasmid was transformed into BL21(DE3)RIPL E. coli cells and optimisation of protein induction was carried out. The results showed the strongest band at the correct molecular weight for the C3C6 fragment in the soluble fraction occurred following induction at 37 °C with 1 mM IPTG (A). Protein was subsequently purified using the HisTrapFF column. Samples were run on a precast 4-20 % SDS-PAGE gel and the gel was stained to detect protein. However, there was a fragment at a lower molecular weight indicating protein degradation (B; asterisk). To try to reduce fragmentation optimisation of the lysis steps were carried out, however, this did not decrease the amount of protein degradation seen (C; asterisk). Abbreviations: E=elution fraction; FT=flow through; L=lysate; P=pellet; S=soluble; T=total; W=wash

3.2.2. Optimisation of production and purification of the C3C5 domain fragment

To see whether use of a shorter fragment of the central domains of cMyBP-C would result in production of a higher protein yield, a HIS-tagged pET C3C5 domain containing construct was transformed into BL21 (DE3) RIPL E. coli cells and induction of protein production was carried out at either 37 °C for 2 hours or 18 °C overnight using either 0.1 mM or 1 mM of IPTG. Cells were lysed and samples were run on an SDS-PAGE gel, transferred onto nitrocellulose and probed using an anti-HIS tag antibody (Fig 19A). The results indicated that the conditions of 18 °C overnight produced the strongest band at the expected molecular weight of about 35 kDa in the soluble fraction, whilst there seemed to be little difference in yield when utilising 0.1 mM or 1 mM IPTG (Fig 19A, asterisk), therefore the conditions of 18 °C overnight with 0.1 mM IPTG were used in all further induction procedures.

The cells were lysed and purified on a HisTrap FF column, samples were run on an SDS PAGE gel which was stained with Coomassie and a large band was seen at the correct molecular weight in the elution fractions 1-4 (figure 19B, arrow), however there were many other bands seen. The His tag was removed, and the protein was once more purified on a HisTrap FF column, samples were run on an SDS PAGE gel which was stained with Coomassie following which a pure protein product was seen in the flow through (FT) fraction (Fig 19C, asterisk).

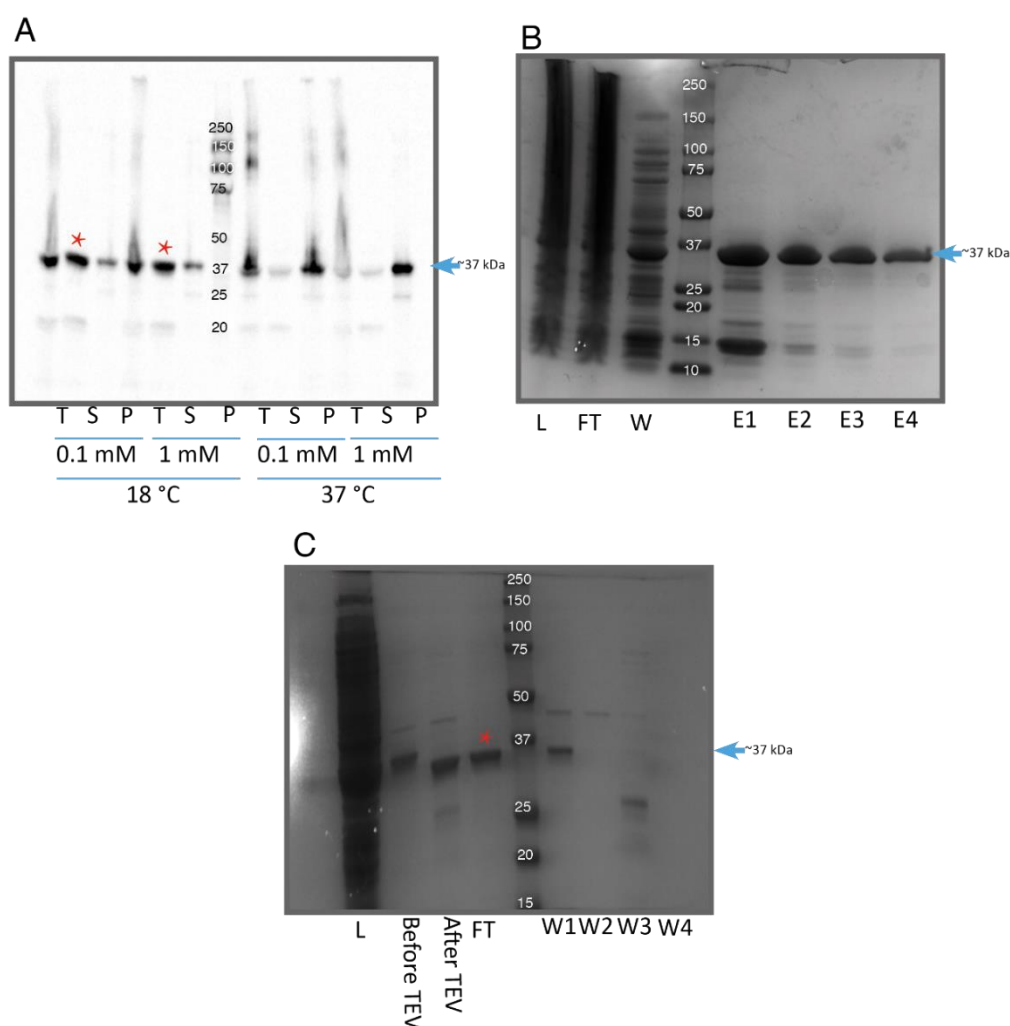


Figure 19 – Optimisation of C3C5 protein induction & purification

Immunoblot and Coomassie stained SDS-PAGE gels of recombinantly expressed constructs. A His-tagged pET C3C5 containing construct was transformed into BL21(DE3)RIPL E. coli cells and optimisation of protein induction was carried out. The results showed the strongest band at the correct molecular weight for the C3C5 fragment in the soluble fraction occurred following induction at 18 °C, with little difference seen between 0.1 mM and 1 mM IPTG (A; asterisks). Protein was then purified using the HisTrap FF column. Samples were run on an SDS-PAGE gel and the gel was stained to detect protein (B). The protein was then treated with TEV protease to remove the His tag from the protein. Protein was purified again on the HisTrap FF column and a pure band was produced at the correct molecular weight in the flow through (FT) fraction (C; asterisk). Abbreviations: FT=flow through; L=lysate; W=wash

These recombinant expression experiments of the central domains indicate that whilst both the C3C6 and the C3C5 domains showed soluble expression, the longer C3C6 fragment expressed at a much lower yield and showed signs of a degradation product, suggesting that it is less stable. Furthermore, the fact that the degradation product appeared to be at the correct molecular weight for a three domain fragment and the shorter C3C5 fragment did not show any degradation product suggests the possibility that the C6 domain was being cleaved off. Additionally, western blot analysis utilised an antibody against the N-terminal His tag for detection, therefore degradation must have been occurring at the C-terminus of the fragment. This could mean that the linker between the C5 and C6 domains is less stable and the C7-C10 domains are required to stabilise the C6 domain. However further experimentation such as mass spectrometry would be required to confirm this.

3.3. Effect of the central domains on actomyosin cross-bridge cycling kinetics

A major role for the N-terminal domains of cMyBP-C is in regulating the actomyosin crossbridge cycling kinetics. To determine whether the central domains also have an effect on crossbridge cycling the recombinantly expressed central domain fragment was utilised in an ATPase assay.

As stated earlier, following the working stroke during contraction, ATP binds to myosin to induce its detachment from actin and is then hydrolysed into ADP and inorganic phosphate (Huxley and Niedergerke, 1954, Huxley and Hanson, 1954, Huxley, 1957). The ATPase assay is able to measure levels of inorganic phosphate and thus the hydrolysis of ATP by myosin. Therefore, it can determine whether the presence of the central C3C5 domains has any effect on the rate of ATP hydrolysis by myosin and thus whether it has an effect on the actomyosin cross-bridge cycling kinetics. This assay also measures the activation state of the thin filament, as at low ATPase rates the thin filament is off, whilst at high rates the thin filament is on. Therefore, this assay may show whether the central domains could be affecting the regulatory state of the thin filaments.

3.3.1. Production of native thin filaments

NTFs, for use in the ATPase assay and the co-sedimentation assay, were isolated from bovine heart tissue as detailed in section 2.10.1. Following isolation of the NTFs samples were taken and run on an SDS-PAGE gel and Coomassie stained to ensure that all of the components of the thin filament were present at their respective expected molecular weights (actin – 42 kDa, tropomyosin – 29 kDa, TnT – 34.6 kDa, TnC – 18.4 kDa and TnI – 23.9 kDa), which can be seen in figure 20A. Calcium regulation of the isolated NTFs was then

checked via an ATPase assay, either in the presence of EGTA or calcium. The rate of ATP turnover by myosin was calculated and was found to be over 10x increased in the presence of calcium compared with EGTA (Figure 20B), indicating the NTFs were functionally intact.

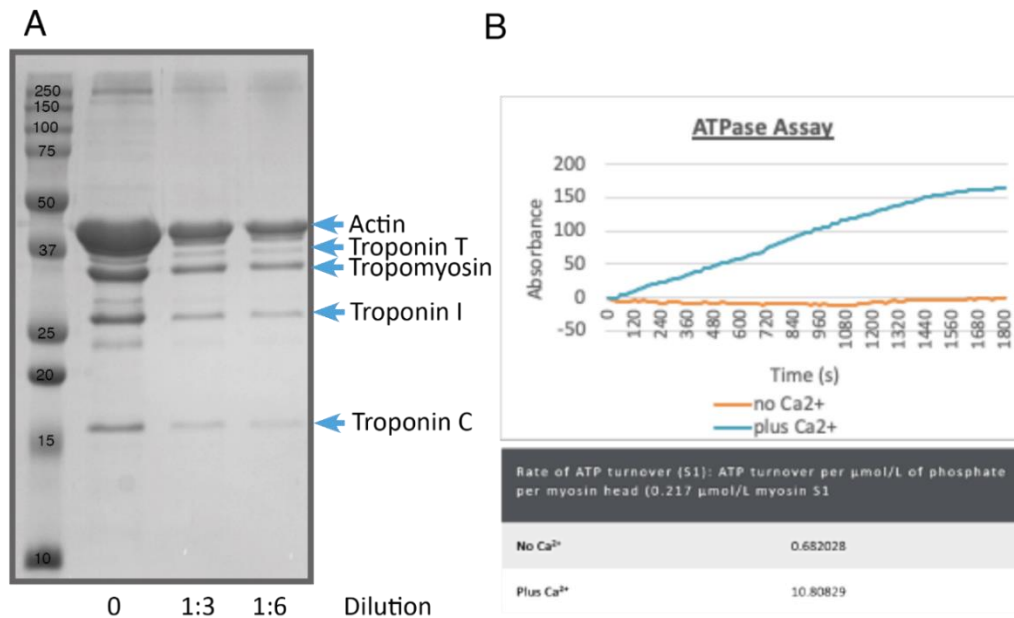


Figure 20 – Validation of native thin filament preparation

Native thin filaments (NTFs) were isolated from bovine heart tissue. Sample was run on an SDS-PAGE gel and Coomassie stained to ascertain all of the components of the thin filaments were present (A). Activity of the NTFs was then checked using an ATPase assay, with absorbance read at 360 nm by a spectrophotometer plate reader, which showed the ATPase rate was more than 10x higher in the presence of calcium compared to without calcium, indicating the NTFs were functioning correctly (B).

3.3.2. Effect of the central C3C5 domains on the ATPase rate

Following validation of the correct functioning of the NTFs, an ATPase assay was carried out, as described in section 2.12., utilising the purified recombinant C3C5 protein at increasing concentrations (between 0-100 μM) in either the presence or absence of calcium and the ATPase rate was calculated. Absorbance levels over time were plotted using GraphPad Prism and the slope was calculated for each concentration of purified protein.

In the absence of calcium, as can be seen in the graph in figure 21, the rate did not seem to increase and remained relatively constant no matter what concentration of the C3C5 fragment was added. In the presence of calcium, there was a trend to an increase in ATPase rate with increasing concentrations of C3C5, with a significant increase in rate seen at 50 and 70 μM compared to in the absence of the central domains ($p= 0.0042$ and 0.0219 respectively; one way ANOVA; $n=3$; figure 21).

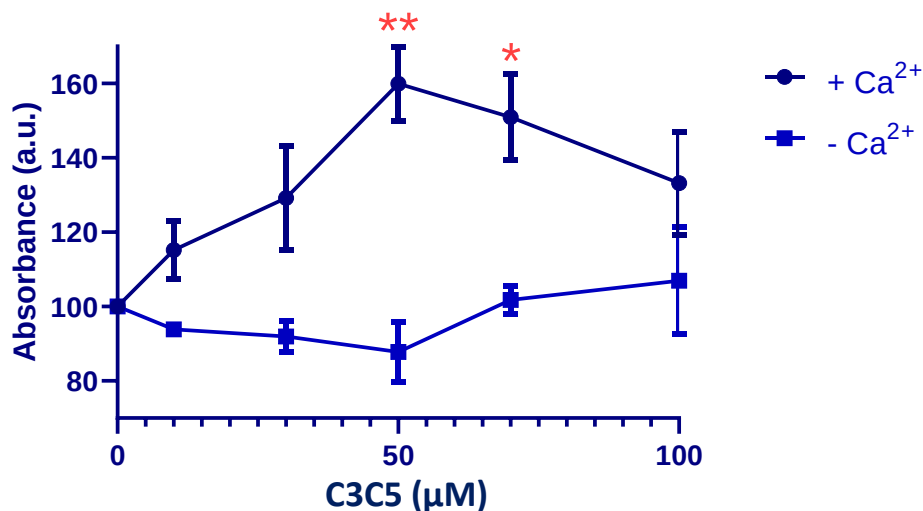


Figure 21 – Effect of C3C5 domains on ATPase activity

Purified recombinant C3C5 protein was added in increasing concentrations (0-100 μM) in an ATPase assay either in the presence or absence of calcium, absorbance was measured at 360 nm by a spectrophotometer plate reader and ATPase rates were calculated by plotting absorbance over time and calculating the slope. ATPase rates remained relatively constant in the absence of calcium, whilst in the presence of calcium, the rate showed a trend to an increase in ATPase rate with increasing concentrations of C3C5, with a significant increase in rate seen at 50 and 70 μM ($p= 0.0042$ and 0.0219 respectively; t-test; $n=3$).

This suggests that the central domains of cMyBP-C may also be involved in the regulation of the actomyosin crossbridge cycling kinetics and that this regulatory role is not solely carried out by the most N-terminal domains (C0-C2). It remains unclear whether this effect was

occurring due to the central domains of cMyBP-C altering the regulatory state of the thin filaments or via an interaction with myosin increasing its ability to bind actin.

3.4. Summary

The results of these experiments suggest that the central domains of cMyBP-C expressed in isolation show a diffuse cytoplasmic localisation but do show a possible co-localisation with actin, which will be explored further in chapter 4.

Recombinant expression of the central domains suggested that the C3C6 fragment was less stable than the shorter C3C5 fragment, with a possibility that the C6 domain was being cleaved off. This could mean that the linker between the C5 and C6 domains is unstable but could be stabilised in the presence of the C-terminal domains C7-C10.

Lastly, the ATPase assay indicated the possibility that the central domains could be involved in the role cMyBP-C has in regulating the actomyosin crossbridge cycling kinetics, as it was seen to significantly increase the ATPase rate at high concentrations of the C3C5 recombinant protein when in the presence of calcium. This potential effect could occur via interaction with either the thin or thick filament proteins. This is contrast to previous experiments which showed that at high concentrations the C0C2 domains inhibited ATP hydrolysis (Belknap et al., 2014). However, this difference could be due to the apparent differences in the binding affinities shown in the co-sedimentation assay, as the C3C5 domains appeared to bind to the NTFs with a much lower affinity when compared with the C0C2 domains.

4 – Possible binding partners of the central domains of cardiac myosin binding protein C

The high number of HCM-linked mutations found in the central domains, along with the cardiac specific insert in the C5 domain, suggests that these domains do have a functional role. Whilst it is possible that these domains merely act as a spacer, they could also have an active role in protein function via interaction with other proteins. There have been no confirmed binding partners found for the central domains of MyBP-C thus far; however, there have been studies indicating possible interactions with actin, creatin kinase and the C-terminus of cMyBP-C (Rybakova et al., 2011, Chen et al., 2011, Moolman-Smook et al., 2002). Due to this, yeast two hybrid assay screens were carried out in order to identify putative binding partners of the central domains of cMyBP-C.

The possible co-localisation seen previously between the central domains and actin was also explored further to see whether an interaction between the two was actually occurring or if the two proteins were merely in close proximity. This was achieved by carrying out ratiometric analysis on the confocal images shown in the previous chapter to quantify the levels of co-localisation between the GFP-tagged C3C6 domain containing construct and F-actin, in comparison with GFP alone. This possible interaction was further studied by carrying out a co-sedimentation assay using the purified C3C5 protein fragment and NTFs.

4.1. Yeast two hybrid assay screen to identify putative binding partners of cMyBP-C's central domains

As no binding partners for the central domains of MyBP-C have so far been confirmed, a yeast two hybrid assay screen was carried out to determine possible novel binding partners for this region.

The yeast two hybrid assay takes advantage of the fact that eukaryotic transcription factors have two sub-domains: an activating domain and a DNA binding domain. The screen involves a bait plasmid containing the protein of interest and a prey library, with each containing one of the two sub-domains of the transcription factor. If the bait and prey interact the transcription factor reconstitutes, which leads to downstream activation of a reporter gene (Fields and Song, 1989). In this case the reporter gene was a Histidine biosynthesis enzyme; therefore, only prey interacting with the bait plasmid would be able to grow on medium lacking Histidine.

4.1.1. Generation of bait constructs for the yeast two hybrid assay screen

A yeast two hybrid assay screen was carried out using a bait plasmid, with the transcription factor DNA-binding sub-domain of LexA, containing the central domains of cMyBP-C and a human cardiac prey library containing the Gal4 activating sub-domain. Bait constructs containing the C3C6 domains of cMyBP-C, as well as single domain bait plasmids for the C4 or C5 domains, were generated via enzyme restriction cloning.

Before carrying out the screens the bait plasmids had to first be validated to ensure that they could be expressed in the yeast strain as well as to test for autoactivation, as around

5% of proteins are thought to show some activating activity (Van Crielinge and Beyaert, 1999) and this could result in false positives in the screen.

All three bait plasmids were transformed into an L40 yeast strain to determine whether they expressed in the yeast cells. Following transformation into the yeast, the cells were lysed and samples were run on an SDS-PAGE gel, transferred onto nitrocellulose and probed using an anti-LexA antibody. The results showed bands at the expected molecular weights for all three bait plasmids (76 kDa for C3C6; about 40 kDa for C4 and C5, figure 22 A&B), showing that they were all able to express in the yeast strain. For the C4 and C5 bait plasmids there were additional bands at a lower molecular weight, indicating the possibility of degradation (figure 22B, arrow), however as there was a strong band at the correct molecular weight these bait plasmids were still utilised.

The bait plasmids were then checked for autoactivation to make sure that they did not cause a false positive by directly binding to the activation sub-domain of Gal4 alone. Each bait plasmid was co-transformed with an empty Gal4 containing plasmid, spotted onto nitrocellulose and an X-gal assay was performed. In this assay if the bait plasmid was able to bind directly to the Gal4 sub-domain this would lead to gene expression resulting in the synthesis of β -galactosidase. Following interaction of β -galactosidase with the X-gal substrate blue colouration would appear. As a LexA KIN3 bait plasmid had previously shown autoactivation with the empty Gal4 construct (previous experiments by Dr. Luke Smith, King's College London) this was used as a positive control. The results, as shown in figure 22C, showed no blue colouration for any of the bait plasmids, whilst the positive control did turn blue, indicating that none of the plasmids being tested were autoactivating.

All three of the bait plasmids generated were able to be expressed in the yeast strain and none showed any signs of autoactivation, therefore all three were used to carry out the yeast two hybrid screens.

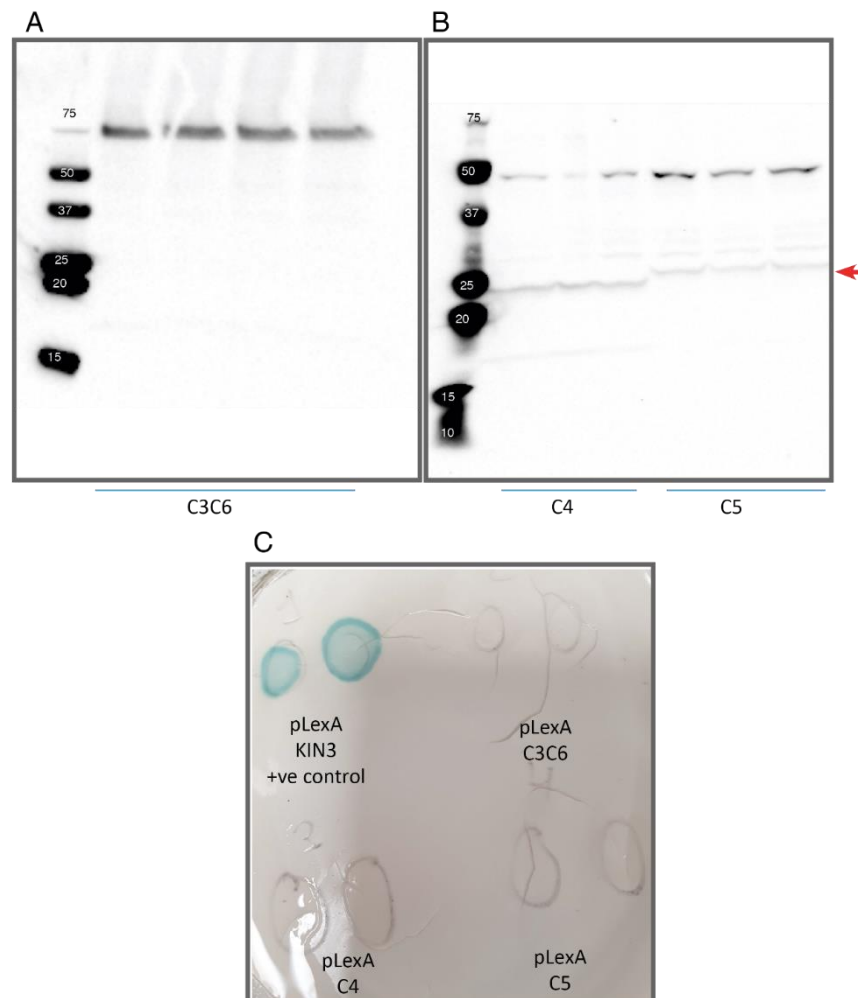


Figure 22 – Generation of bait constructs for the yeast two hybrid assay

Fragments from the central domains of cardiac myosin binding protein C (cMyBP-C) were cloned into an N-terminal LexA fusion protein. Constructs were transformed into L40 yeast strain and immunoblotting was used to ascertain expression of the C3C6 (A), C4 and C5 (B) bait constructs. Strong bands were seen at the expected molecular weight of 75 kDa for the 3 domain construct and 50 kDa for the single domain constructs. There were lower bands seen for the C4 and C5 domain contain bait constructs (B; arrow) which may suggest some degradation. Bait constructs were then co-transformed with an empty Gal4 construct to check for autoactivation via the X-gal assay (C). No blue colouration was seen for any of the constructs indicating no autoactivation was occurring. A pLex KIN3 construct which had previously shown autoactivation was used as a positive control.

4.1.2. Hits from the yeast two hybrid assay screen

Yeast two hybrid assay screens were carried out with each of the three bait plasmids containing domains from the central region of cMyBP-C, with a human cardiac prey library. The screen was plated on agar plates lacking Histidine, therefore any colonies which grew indicated an interaction between the bait and prey plasmids in order to reconstitute the transcription factor to lead to downstream activation of the Histidine biosynthesis enzyme. All colonies which grew were picked and DNA was extracted and sent for sequencing. A basic local alignment search tool (BLAST) database search then indicated any hits which were in frame with the Gal4 activating domain.

4.1.2.1. Hits from the Y2H assay screen with the C4 bait construct

70 colonies grew from the yeast two hybrid screen using the C4 bait plasmid; six of which were found to be in frame with the Gal4 domain (table 6). Of these the majority did not seem to be likely interactors consisting of : a hypothetical protein (CAE45949); Cytochrome C oxidase subunit 1 which is part of the mitochondrial electron transport chain; an unnamed protein (BAC85276.1); partitioning defective 3 homolog, an adaptor protein involved in asymmetrical cell division and cell polarisation (Tabuse et al., 1998) and membrane protein 11 isoform CRA-b which has a role in mitochondrial morphogenesis. Only one seemed to be of any interest: the ABHD18 isoform 9 (red in table 6). This is a protein that has an unknown function; however it has the highest expression levels in the heart (Fagerberg et al., 2014) as well as showing a striated appearance when stained for in cardiac muscle tissue, suggesting that it could possibly localise in the sarcomeres (from Human Protein Atlas available from <https://www.proteinatlas.org/ENSG00000164074-ABHD18/tissue/heart+muscle> (Uhlén et al., 2015)).

Protein	Function
C4	
Hypothetical protein CAE45949.1	N/A
Cytochrome C oxidase subunit 1	Part of the mitochondrial electron transport chain
Unnamed protein product BAC85276.1	N/A
ABHD18 isoform 9	Unknown (high expression in the heart) (Fagerberg et al., 2014)
Partitioning defective 3 homolog	Adapter protein involved in asymmetrical cell division and cell polarization processes (Tabuse et al., 1998)
membrane protein 11 isoform CRA-b	Role in mitochondrial morphogenesis (Rival et al., 2011)

Table 6 – Hits from the yeast two hybrid assay with the C4 bait plasmid

Following the yeast two hybrid assay screen DNA was extracted from colonies, sequenced and identified using a blast database search. Proteins which seemed likely candidates to be interactors of cardiac myosin binding protein C (cMyBP-C; red) underwent a forced cross with their respective bait construct to determine whether a true interaction was occurring.

4.1.2.2. Hits from the Y2H assay screen with the C5 bait construct

The yeast two hybrid screen utilising the bait plasmid containing the C5 domain of cMyBP-C grew 15 colonies, of which four were in frame with the Gal4 domain (table 7). The hits included: growth hormone-inducible transmembrane protein which is required for the mitochondrial tubular network and cristae organisation, as well as being involved in the apoptotic release of cytochrome c (Oka et al., 2008); B-cell growth factor which stimulates B-and T-cells (Yokota et al., 1986) and an unnamed protein product (BAB70813.1). The only

hit that appeared relevant to the cardiovascular system was a fragment of the titin cap (shown in red in table 7), which is the first 11 amino acids of the titin cap, also known as telethonin, which is thought to have a role in myofibrillar assembly (Mason et al., 1999).

Protein		Function
C5		
titin cap/telethonin fragment (first 11 AA)		Telethonin is required for myofibril assembly (Mason et al., 1999)
Growth hormone-inducible transmembrane protein		Required for the mitochondrial tubular network and cristae organization. Involved in apoptotic release of cytochrome c (Oka et al., 2008)
B-cell growth factor		B- and T-cell stimulation (Yokota et al., 1986)
Unnamed protein	product	N/A
BAB70813.1		

Table 7 - Hits from the yeast two hybrid assay with the C5 bait plasmid

Following the yeast two hybrid assay screen DNA was extracted from colonies, sequenced and identified using a blast database search. Proteins which seemed likely candidates to be interactors of cardiac myosin binding protein C (cMyBP-C; red) underwent a forced cross with their respective bait construct to determine whether a true interaction was occurring.

4.1.2.3. Hits from the Y2H assay screen with the C3C6 bait construct

Following the yeast two hybrid screen using the bait plasmid containing the C3C6 domains, 48 colonies grew. Of these, DNA extraction and analysis of the sequence identified only nine which were in frame with the Gal4 domain, which can be seen in table 8. The screen identified seven proteins which do not appear to be likely candidates for interaction with

MyBP-C: promyelocytic leukaemia zinc finger, which is involved in developmental and biological processes including spermatogenesis, hind limb formation, haematopoiesis and immune regulation (Suliman et al., 2012); cytochrome b, a component of the mitochondrial electron transport chain (Esposti et al., 1993); ferretin light polypeptide which is involved in iron storage (Harrison and Arosio, 1996); splicing factor arginine/serine-rich 3, isoform CRA_e which is involved in splicing regulation (Zhou and Fu, 2013); an unnamed protein product (BAG62403.1); centrosomal protein POC5, a protein involved in assembly of the centrosomes (Azimzadeh et al., 2009) and microtubule-associated proteins 1A/1B which are adaptor proteins that act to stabilise microtubules (Halpain and Dehmelt, 2006). Two did appear to have links to the cardiovascular system (red in table 8): atrial natriuretic peptide and beta myosin heavy chain. Atrial natriuretic peptide is involved in lowering blood pressure (Brenner et al., 1990), whilst beta myosin heavy chain forms part of type II myosin which is the isoform of myosin present in striated muscles and is responsible for contraction (Lodish et al., 2000).

Protein	Function
C3C6	
Atrial natriuretic peptide	Reduction in blood pressure (Brenner et al., 1990)
Promyelocytic leukemia zinc finger	major developmental and biological processes, such as spermatogenesis, hind limb formation, hematopoiesis, and immune regulation (Suliman et al., 2012)
Beta-myosin heavy chain	forms part of type II myosin (Lodish et al., 2000)
Cytochrome b	Component of the electron transport chain in mitochondria (Esposti et al., 1993)
Ferretin light polypeptide	Iron storage (Harrison and Arosio, 1996)
Splicing factor arginine/serine-rich 3, isoform CRA_e	Splicing regulation (Zhou and Fu, 2013)
Unnamed protein product BAG62403.1	N/A
Centrosomal protein POC5	Centrosomal assembly (Azimzadeh et al., 2009)
Microtubule-associated proteins 1A/1B	Stabilise microtubules, adaptor proteins (Halpain and Dehmelt, 2006)

Table 8- Hits from the yeast two hybrid assay with the C3C6 bait plasmid

Following the yeast two hybrid assay screen DNA was extracted from colonies, sequenced and identified using a blast database search. Proteins which seemed likely candidates to be interactors of cardiac myosin binding protein C (cMyBP-C; red) underwent a forced cross with their respective bait construct to determine whether a true interaction was occurring.

4.1.3. Validation of putative interaction partners identified by the yeast two hybrid screen

In order to ensure that the relevant potential hits shown above were reproducible and thus not merely false positives, forced crosses were performed between DNA extracted from colonies which grew during the screen and the corresponding bait plasmid. This was carried out by co-transfecting into yeast, spotting onto nitrocellulose and carrying out an X-gal assay, as described previously. The pLex-A KIN3 construct crossed with the empty Gal4 plasmid was again used as a positive control. If blue colouration was seen then this would have indicated a true interaction was occurring. However, as can be seen in figure 23, none of the forced crosses turned blue, whilst the positive control did turn blue, suggesting that all of these interactions were false positives.

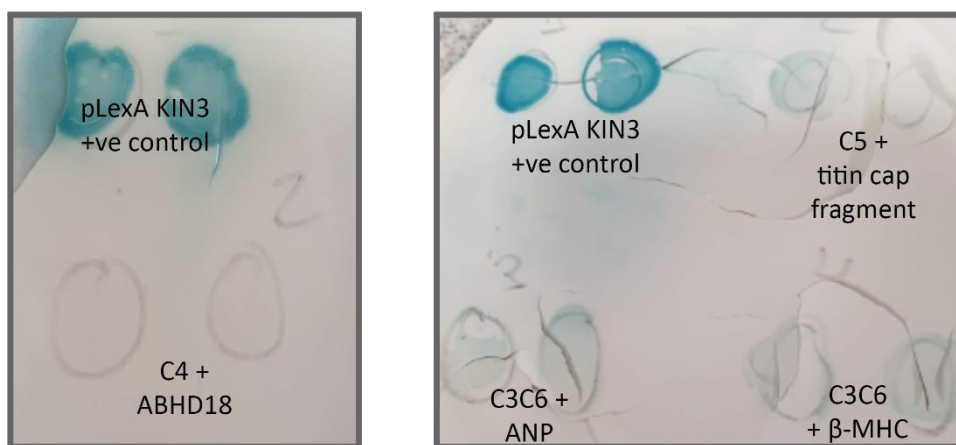


Figure 23 – Forced cross of potential cardiac myosin binding protein C interactors identified in the yeast two hybrid assay

Example of the forced cross tests carried out to determine whether potential interactors identified in the yeast two hybrid assays screens were true interactors. DNA from potential interactors identified in the yeast two hybrid assay screen were co-transfected into L40 yeast strain with their respective bait plasmid to determine whether a true interaction was occurring. Interaction was determined using the X-gal assay for each of the promising potential interactors. None of the potential binding partners turned blue indicating no true interactions were taking place. A pLex KIN3 construct which had previously shown autoactivation was used as a positive control.

The yeast two hybrid screen produced a high number of colonies, particularly the C3C6 and C4 containing bait plasmids, suggesting that there were no technical issues whilst carrying out the screen. However, the majority of the hits seen were either out of frame with the Gal4 domain or unrelated to the cardiovascular system or in a localisation which made it unlikely for interaction with MyBP-C to be feasible. Of those which did look promising none were positive during the forced-cross X-gal assay suggesting these hits were false positives.

4.2. Possible interaction between the central domains and actin

While no binding partners for the central domains of MyBP-C have been confirmed to date, past studies have indicated a possible interaction between the central domains of cMyBP-C and actin. One study indicated the C5-C10 region could bind specifically to actin (Rybakova et al., 2011), whilst another suggested the C3 domain might be involved in the N-terminus' interaction with the thin filament (Inchingolo et al., 2019). Furthermore, as was shown in section 3.1., confocal imaging indicated a possible co-localisation between the C3C6 domains and F-actin. Due to this, further quantification and experimentation was carried out to see if there was any interaction between the central domains and actin.

4.2.1. Ratiometric analysis of co-localisation of the C3C6 domains and actin

In order to determine whether the C3C6 domains were co-localising with F-actin, ratiometric analysis was carried out on confocal images of NRCs which had been transiently transfected with the GFP-tagged plasmid containing the C3C6 domains and immunostained with phalloidin to detect F-actin.

As shown in figure 24A the ratiometric analysis identifies areas with the GFP-tagged C3C6 domains alone as purple/dark blue, F-actin alone is red, whilst green indicates the presence of, indicating that the central domains and actin are localising in the same region which could suggest an interaction. The percentage of the cell which was green was calculated using ImageJ software.

The results, seen in the graph in figure 24B, showed a significant increase in the percentage of the cell showing the presence of both GFP and F-actin for the C3C6 domains when compared to GFP alone ($p=0.0009$; unpaired t-test; $n=3-4$). This suggests that the C3C6

domains are co-localising with actin. However, co-localisation does not necessarily mean an interaction is occurring, as the two proteins may just be in close proximity to each other.

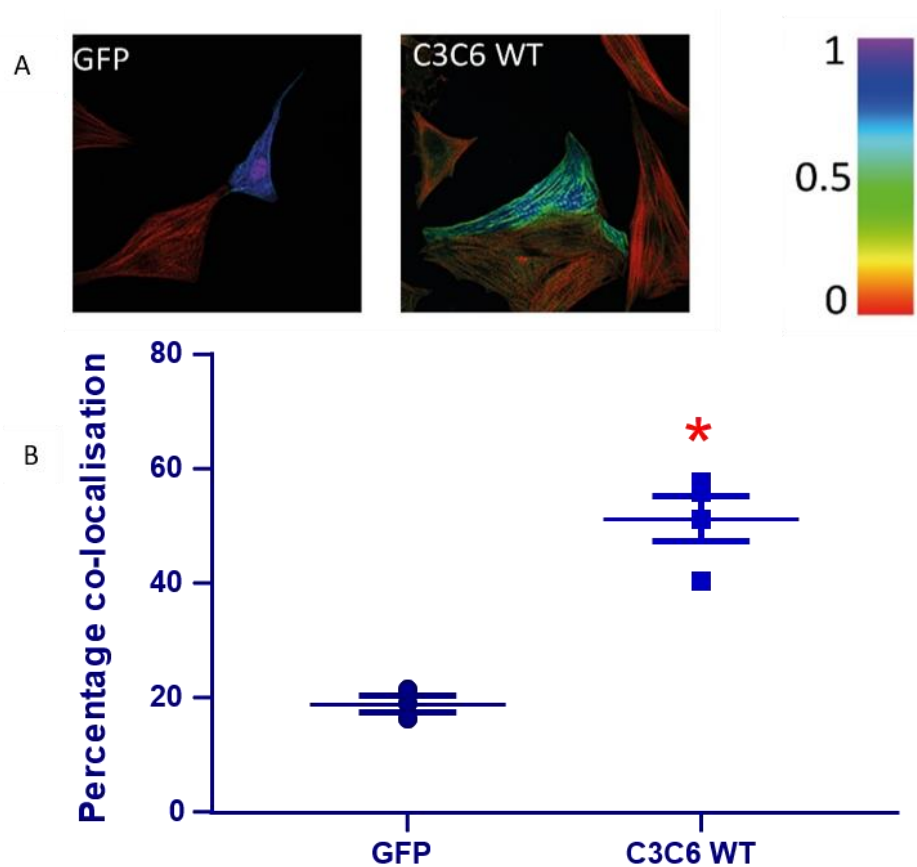


Figure 24 – Co-localisation of the central domains and F-actin

Images showing ratiometric analysis to determine levels of co-localisation between the central C3C6 domains of cardiac myosin binding protein C (cMyBP-C) and F-actin in neonatal rat cardiomyocytes (NRCs). NRCs were transiently transfected with either GFP alone or a GFP-tagged wildtype C3C6 containing plasmid, immunostained for F-actin and images were taken using a confocal microscope. Mathematica 11.3 software was then used to determine co-localisation, with red indicating F-actin alone, purple/dark blue indicating the GFP-tagged C3C6 containing construct alone and green indicating the presence of both (A). ImageJ was then utilised to determine the percentage of the cell which was green. The results indicated that the WT construct had a significant increase in co-localisation when compared to the empty pEGFP control vector ($p = 0.0009$, $n = 3-4$;

4.2.2. Co-sedimentation of the central domains and native thin filaments

The possible interaction between the central domains of cMyBP-C and actin was further studied using a co-sedimentation assay with NTFs. The co-sedimentation assay uses high speed ultracentrifugation of recombinant C3C5 protein and NTFs after an incubation period, following which anything that binds to the NTFs will co-sediment with the NTFs and would be present in the pellet fraction. C0C2 recombinant protein, which has previously been shown to co-sediment with actin (Shaffer et al., 2009), was also used as a positive control.

As can be seen in figure 25A, there was a band at the correct molecular weight for the C3C5 protein (35 kDa) in the pellet fraction (arrow), indicating that the C3C5 domains were co-sedimenting with the NTFs. However, densitometric analysis (figure 25B) showed that the co-sedimentation was far less than that of the C0C2 domains, indicating that the central domains may be interacting with the thin filaments, but at a much lower level than the N-terminal domains.

As can be seen in the lane for both C0C2 and C3C5 alone, there were bands indicating the fragments did partially sediment even in the absence of the NTFs. However, during the densitometric analysis this was subtracted from the band showing protein fragments with the NTFs in order to eliminate any protein that had just aggregated under the high speed centrifugation.

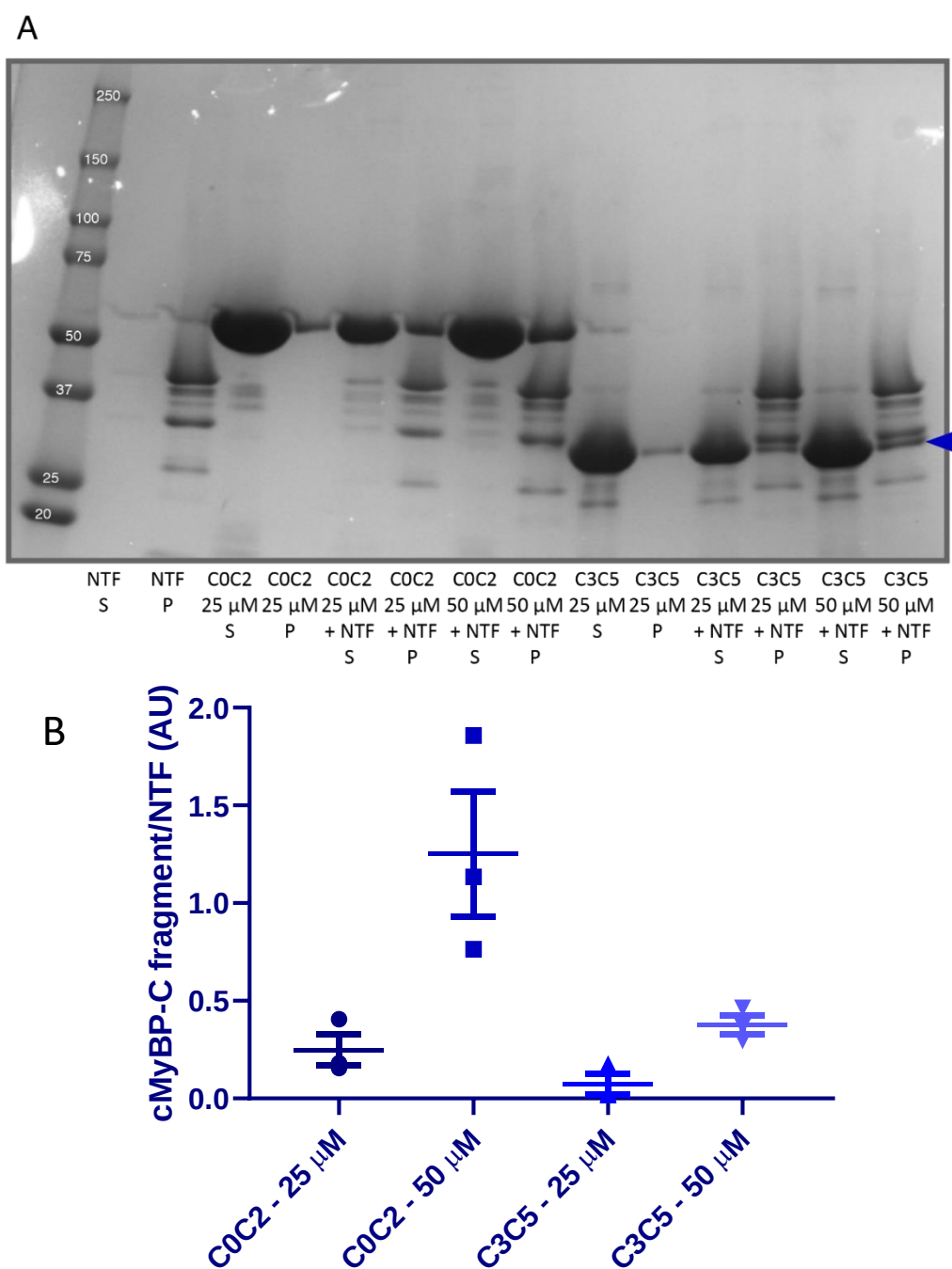


Figure 25 – Co-sedimentation of native thin filaments with the central domains of cardiac myosin binding protein C

Co-sedimentation of native thin filaments (NTFs) and fragments of cardiac myosin binding protein C (C3C5 and COC2 used as a control) was carried out. Samples were run on an SDS-PAGE gel and Coomassie stained to determine whether there was any interaction between the NTFs and the cMyBP-C fragments (A). The C3C5 fragment did appear to co-sediment with the NTFs (blue arrow), however densitometric analysis indicates that this co-sedimentation occurred at a much lower level than with the N-terminal COC2 fragment (B).

4.3. Summary

Different methods were utilised in order to identify putative novel interaction partners of the central domains of cMyBP-C. The Y2H assay screen did lead to a high number of colonies which suggested that no technical issues occurred during the screen itself; however the few hits which were in frame with Gal4 and were feasible interactors for MyBP-C were shown to be false positives.

Whilst the Y2H assay screens did not lead to any potential novel binding partners for the central domains of cMyBP-C, both ratiometric analysis of confocal images and co-sedimentation assays suggested that there could be an interaction between the central domains and actin. The co-sedimentation assay did however indicate that this interaction occurred at a much lower level than that of the more N-terminal C0C2 domains.

5 – Hypertrophic cardiomyopathy-linked mutations in the central domains of cardiac myosin binding protein C

The central domains contain a high number of HCM-linked mutations; with the highest number occurring in the C6 domain, followed by the C5 domain, as can be seen in figure 8, with a high proportion of these being due to missense mutations (~49 and 62% respectively; from data compiled from (Harris et al., 2011, Wang et al., 2018b)).

The pathogenesis of HCM-linked missense mutations in cMyBP-C is still under debate, with some studies suggesting they lead to an unstable protein which is degraded causing haploinsufficiency. Another study however indicates that the A31P missense mutation leads to a full-length protein capable of incorporating into the sarcomere (van Dijk et al., 2016) suggesting a ‘poison polypeptide’ mechanism of pathogenicity.

In order to further elucidate the mechanism by which these HCM-linked missense mutations cause disease, several HCM-linked missense mutations within the C3C6 region were generated in expression vectors and studied. Protein expression levels were determined in order to see whether the mutation was leading to haploinsufficiency via protein degradation, whilst localisation and possible effects on the sarcomeric ultrastructure were studied to see if the mutation could be acting as a ‘poison polypeptide’.

5.1. HCM-linked mutant C3C6 domain containing pEGFP vectors

In order to study the pathogenic mechanism of different HCM-linked missense mutations in the C3C6 region of cMyBP-C, several mutations were identified using the ClinVar database (Landrum et al., 2018), which are shown in figure 26. The mutations had a varying level of evidence of pathogenicity (from uncertain significance to pathogenic) but all mutations involved substitutions to a non-similar amino acid group, had an allelic frequency of less than 0.01% according to the gnomAD database (Karczewski et al., 2020) and were generally well conserved across different species and the different isoforms of MyBP-C (figures 27-30).

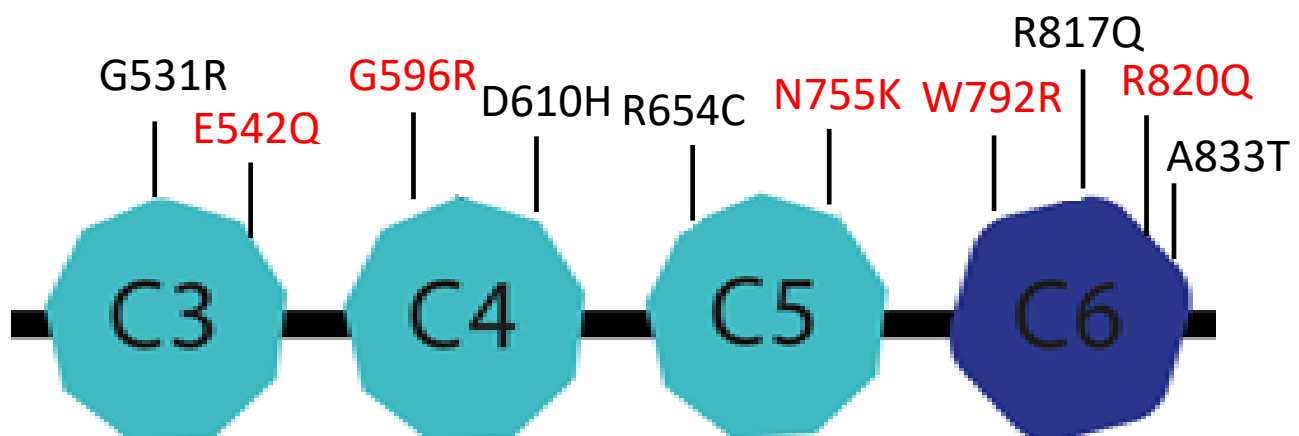


Figure 26 – Hypertrophic cardiomyopathy linked missense mutations of the central domains of cardiac myosin binding protein C

Figure showing the Hypertrophic cardiomyopathy (HCM) linked missense mutations found in the central domains of cardiac myosin binding protein C (cMyBP-C) utilised in this study. Missense mutations were identified using the ClinVar database, with all showing an allelic frequency of less than 0.01% according to the gnomAD database. These mutations were introduced into a GFP-tagged C3C6 containing plasmid in order to study whether the mutations had any effect on protein expression levels or localisation of the protein. Those mutants showing the most apparent phenotype in preliminary experiments (shown in red) were then taken forward into adenoviral vectors containing the full length cMyBP-C protein.

The mutations were then introduced into the GFP-tagged C3C6 domain containing plasmid described earlier using site directed mutagenesis. These mutant constructs were used to determine whether these mutations had any effect on protein expression, thus indicating haploinsufficiency, or whether they displayed any evidence of having a “poison polypeptide” effect such as mislocalisation of the protein or interference with protein-protein interactions. This was done by carrying out transient transfections in COS-1 cells and NRC which were analysed via immunoblotting, confocal microscopy and STED microscopy.

Following these experiments, those mutants showing the most interesting phenotype were taken forward into full-length adenoviral vectors. These were the E542Q, G596R, N755K, W792R and R820Q mutants, more information about which is detailed below.

5.1.1. E542Q

The negatively charged, polar and hydrophilic glutamic acid residue at the 542 position in the C3 domain is highly conserved throughout different species as well as in the fast skeletal isoform, however it is replaced by a non-polar leucine residue in the slow skeletal isoform of MyBP-C (Figure 27). The mutation of this glutamic acid to the polar, non-charged glutamine residue (E542Q; figure 31) was first described by Carrier et al. following DNA sequencing of familial HCM patients in a French population (Carrier et al., 1997). Both glutamic acid and glutamine are amino acids which prefer to be surface exposed (Betts and Russell, 2003), however glutamic acid has been shown to be much more abundantly found in α -helix structures and is known to stabilise coil-coil domains (Young and Ajami, 2000). Therefore, it is possible that substitutions of glutamic acid residues can cause structural destabilisation.

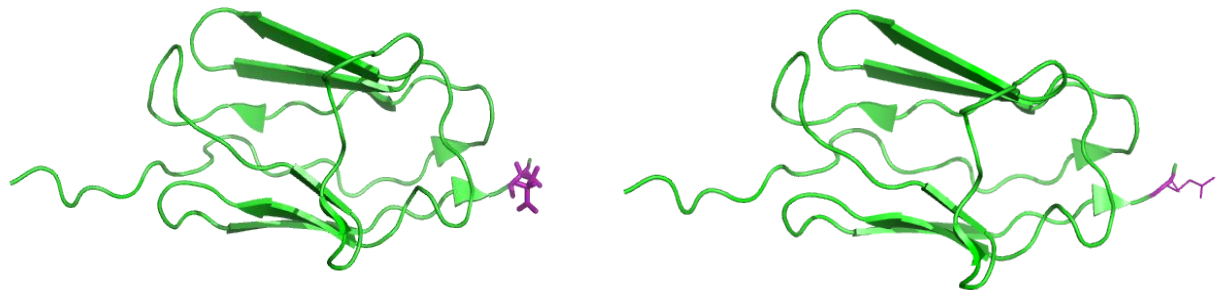


Figure 31 – Structural representation of the E542Q mutation in the C3 domain of cardiac myosin binding protein C

Representation of the C3 domain of cardiac myosin binding C (cMyBP-C) using Pymol software based on the NMR structural data of the human cMyBP-C C3 domain (Zhang et al., 2014). On the left the wild type C3 domain is shown with the negatively charged, polar and hydrophilic glutamic acid residue at the 542 position is highlighted in purple. In the HCM-linked E542Q mutation this residue is substituted with a polar, non-charged glutamine residue (highlighted in purple, right). Glutamic acid residues are frequently found in α -helices and are known to be involved in stabilising coiled-coils therefore substitutions of this amino acid could lead to structural instability.

A study investigated family members of an HCM patient carrying the E542Q mutation and identified an 83% penetrance of the mutation in this family, with 67% showing symptoms of HCM. The age of onset for symptoms of HCM was shown to be 36.2 (+/- 16) years of age (Charron et al., 1998). Another study found that carriers of this mutation, based on data from three different affected families, showed an age at diagnosis and an advanced wall thickening not above average for that of HCM as a whole, suggesting this particular mutation does not cause a particularly severe or early onset form of the disease (Rodríguez-García et al., 2010). The differences seen between the two studies could be due to factors such as ethnicity, although both were studying European cohorts, or there could have been compounding mutations leading to a more severe disease.

A subsequent study utilised transient transfection of a myc-tagged E542Q mutated cMyBP-C construct into foetal rat cardiomyocytes followed by immunofluorescence staining and confocal microscopy. This showed that this mutant protein was fully capable of

incorporating into the A-band of the sarcomeres, with little to no difference in expression levels or integration when compared with WT (Flavigny et al., 1999).

Furthermore, electron micrographs of cardiac samples from a myectomy patient carrying the E542Q mutation showed no loss of the sarcomeric structure and features, nor evidence of sarcomeric disarray, although histological analysis did show evidence of myofibrillar disarray (Vydyanath et al., 2012).

The previous literature therefore suggests that this mutation is pathogenic, leading to an average severity and age of onset of disease. Furthermore, studies indicate that this mutation leads to a full-length protein that is capable of incorporating correctly into the sarcomere, suggesting the mechanism of disease is not through haploinsufficiency.

5.1.2. G596R

The G596R mutation is found in the C4 domain and involves the substitution of a well conserved (figure 28), small, non-polar hydrophobic glycine residue for a larger, positively charged, polar, hydrophilic arginine (Figure 32). Meta-analysis of the effect different amino acid substitutions can have has shown that substitution of glycine is highly associated with pathological disease (Iqbal et al., 2020). Glycine is the smallest amino acid with a side chain consisting of a single hydrogen allowing for a high degree of conformational flexibility. Due to this glycine is able to reside in parts of structures which other amino acids are not able to, such as tight turns (Betts and Russell, 2003). As arginine is a much larger amino acid, if the G596 is involved in a tight turn the arginine would not allow for this and so could lead to structural destabilisation.



Figure 32 – Structural representation of the G596R mutation in the C4 domain of myosin binding protein C

The C4 domain of myosin binding protein C was made utilising Pymol software based on NMR structural data from the slow skeletal isoform (Niraula et al., 2008) as no structural data for the C4 domain of the cardiac isoform has been resolved thus far. On the left is the wild type C4 domain of MyBP-C with the small non-polar glycine residue at the 596 position. On the right, this residue is replaced with the larger, positively charged, polar, hydrophilic arginine residue. As glycine is the smallest amino acid it is able to reside in parts of structures which are not possible for other amino acids, such as tight turns. Due to this substitution of glycine can lead to destabilisation of protein structure.

The first mention of the G596R mutation being linked to HCM was by Gruner et al., in which sequencing of the DNA of a patient with apical HCM was found to carry the mutation (Gruner et al., 2011). This mutation was also seen in a 69 year old woman with elderly onset HCM and no family history of HCM (Arad et al., 2014) as well as from an autopsy of a 45 year old man following an unknown cause of death (Sanchez et al., 2016).

A study also used various modelling software programmes which make predictions as to whether or not a mutation is likely to be deleterious on the proteins structure, with all suggesting the mutation would be damaging (Sanchez et al., 2016).

5.1.3. N755K

At the 755 position in the C5 domain there is a small, polar, non-charged, hydrophilic asparagine residue which is conserved in several species as well as between the three isoforms of MyBP-C (Figure 29). In the HCM-linked N755K mutation this is replaced by a larger, positively charged, polar, hydrophobic lysine (Figure 33). The substitution to the larger positively charged residue could interfere with domain folding which could lead to protein destabilisation or may alter the ability of the protein to bind with potential interaction partners.

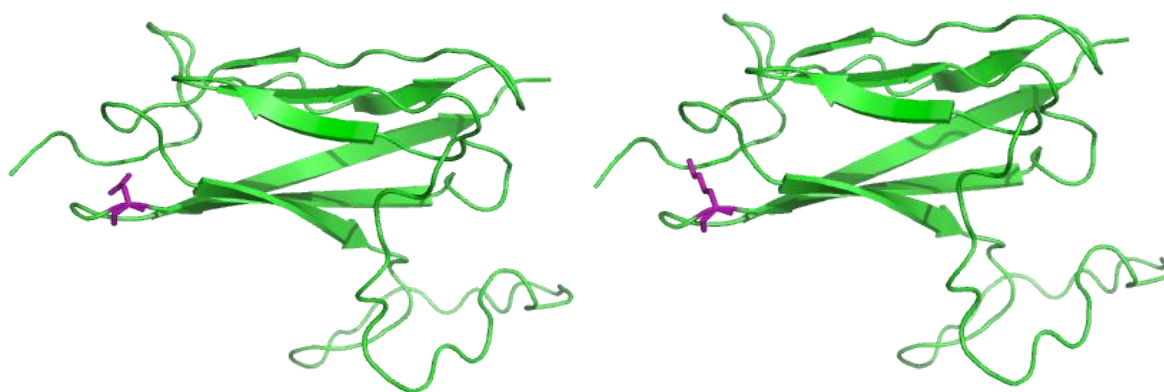


Figure 33 – Structural representation of the N755K mutation in the C5 domain of cardiac myosin binding protein C

The C5 domain of cardiac myosin binding protein C (cMyBP-C) was made utilising Pymol software based on NMR structural data (Idowu et al., 2003). On the left is the wild type C5 domain of cMyBP-C, with a small, polar, non-charged, hydrophilic asparagine residue at the 755 position. This is replaced in the N755K mutation by a larger, positively charged, polar, hydrophobic lysine (right). Structural studies suggest that this substitution to the much bulkier and charged lysine residue could interfere with the tight fold of the domain which could lead to destabilisation of the protein (Idowu et al., 2003).

The N755K mutation was first identified by Yu et al. following sequencing of a de novo HCM patient. They subsequently studied the family members of the patient and discovered that eight of the thirteen genetically tested were heterozygous for the N755K mutation and from this they discovered that in this family there was an 89% penetrance by the age of 40. The youngest person identified from this family with the mutation was diagnosed at the age of 11, whilst one of those identified as being a carrier died at the age of 19 of sudden cardiac death following exercise. They concluded that this mutation showed near complete penetrance with quite marked hypertrophy even in children (Yu et al., 1998a, Yu et al., 1998b).

As mentioned earlier, one theory as to how MyBP-C arranges on the thick filament is the collar model, in which there is an interaction between the C5C6C7 and C8C9C10 domains in a staggered parallel arrangement that encircles the thick filament backbone (Moolman-Smook et al., 2002). If this model is correct, N755 would be at the interface of the possible binding between the C5 and C8 domains. A study by Moolman-Smook et al. looked into the possibility that the N755K mutation could interfere with this interaction, and found that the binding of the mutant protein to the C8-C10 domains of MyBP-C was 10x weaker than the WT protein (Moolman-Smook et al., 2002).

Structural modelling and analysis was carried out for the N755K mutation by Idowu et al., with NMR analysis of the C5 domain with this N755K mutation showing a spectra typical of an unfolded protein, although there was still some evidence of some remaining structure. Furthermore, it was shown that this residue is located in a narrow space defined on either side by an FG hairpin and a BC turn. The residue is also surrounded by a number of conserved small residues, suggesting the necessity for tight constraints to accommodate the

fold of the domain. It was suggested that the substitution to a lysine residue in this mutation, which is a bulkier and charged amino acid, could disrupt this tight fold. The side chain of the N755 is also thought to be involved in H-bonding, which would be lost following substitution to the lysine residue (Idowu et al., 2003).

These studies suggest that this mutation leads to HCM with a relatively normal level of penetrance, however very early onset of the disease was seen, with one carrier showing symptoms of HCM at the age of 11 and then died of sudden cardiac death at the age of 19 (Yu et al., 1998b). This study did however only look at one family therefore it could be possible that the severity of disease seen could have been due to a compounding mutation or other factors within the family. Furthermore, structural studies and modelling suggest that the positioning of this mutation, close to the cardiac specific C5 insert as well as being an area requiring tight restraints in order to maintain the fold of the domain would likely have detrimental effects on the folding of domains and could thus have a destabilising effect on the protein. However, these structural experiments were carried out utilising the isolated C5 domain, and it is possible that in the full length protein the tight packing of the surrounding domains as well as of other proteins in the sarcomere could be able to stabilise the domain fold up to a point, however following mechanical stress it may not be able to maintain its partial structure and could collapse and could lose partial or complete function (Idowu et al., 2003). Additionally, the work by Moolman-Smook et al. suggests that the mechanism of disease for this mutant could be by interfering with potential interacting partners, in this case between the C5 and C8 domains which they posit is how MyBP-C arranges on the thick filament backbone (Moolman-Smook et al., 2002).

5.1.4. W792R

The W792R mutation involves the substitution of a highly conserved (Figure 30), aromatic, polar, hydrophobic tryptophan for a positively charged, hydrophilic arginine residue (Figure 34). The side chain of the W792 is predicted to angle toward the central hydrophobic region of the domain therefore the substitution to the positively charged arginine could possibly disrupt the proper folding of the domain.

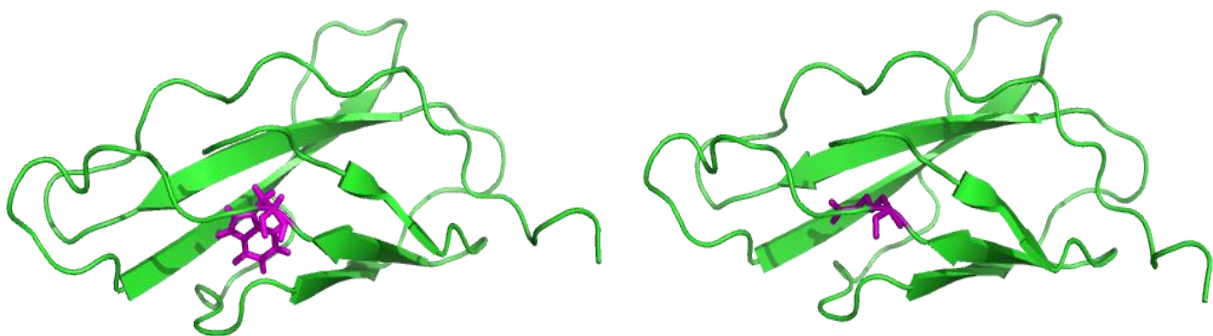


Figure 34 – Structural representation of the W792R mutation in the C6 domain of cardiac myosin binding protein C

As there is currently no structural data for the C6 domain for any of the isoforms of myosin binding protein C (MyBP-C), I-TASSER software was utilised to model the domain. This model was then visualised using Pymol software. On the left the aromatic, polar, hydrophobic tryptophan can be seen at the 792 position. In the W792R mutation this is replaced with a positively charged, hydrophilic arginine (right). The side chain of this tryptophan residue is thought to angle toward the hydrophobic core of the domain and the replacement of this amino acid with the positively charged arginine could therefore effect domain folding.

This mutation was first described in 2004 by van Driest et al. (Van Driest et al., 2004), following DNA sequencing of HCM patients at a tertiary referral centre for HCM. Subsequently a study identified the mutation in two unrelated female HCM patients, both had family history of HCM, were diagnosed with the disease at a relatively young age (26 &

27) and also had to undergo myectomy surgery at the age of 35 and 32 respectively. Immunohistochemistry on paraffin embedded tissue samples from these myectomy patients identified a small degree of disruption in the even distribution of cMyBP-C but were not significantly different from healthy samples (Theis et al., 2009).

A study by Smelter et al. suggested that protein expression levels of the W792R mutant was significantly lower, reaching no more than 30% of the WT levels. However, there was no difference in the mRNA levels indicating any degradation was occurring at the protein level. They also showed that the mutant protein appeared to properly localise to the C-zone and form doublets, although showing a significantly lowered fluorescence intensity, suggesting that the mutation does not affect localisation of the protein, but again does lead to degradation. The study also identified an increase in twitch amplitude in murine engineered cardiac tissue (mECTs) infected with mutant adenoviral construct which was comparable to that of mECTs with a cMyBP-C KO, suggesting that the accelerated contraction kinetics seen was due to a decrease in protein levels not via a dominant negative effect. Reduced expression of the mutant protein was also confirmed by a decreased bacterial expression, furthermore mass spectrometry showed a decrease in the full-length protein with an increase in levels of a proteolytic cleavage product, suggesting it is susceptible to digestion (Smelter et al., 2018).

These studies indicate that this mutation may lead to a fairly early onset and severe disease, as shown by the fact that two unrelated patients required surgical intervention at a young age. It is thought that this mutation leads to a decrease in the expression of the protein, possibly by destabilising the protein and making it more susceptible to degradation, leading to HCM through haploinsufficiency.

5.1.5. R820Q

The positively charged, polar and hydrophilic arginine residue at position 820 in the C6 domain is highly conserved throughout different species, however it is replaced by the similarly positively charged lysine in the fast skeletal isoform of MyBP-C (Figure 30). The R820Q mutation sees this residue replaced by a polar, non-charged glutamine (Figure 35). Arginine is known to often form salt bridges with other amino acids creating hydrogen bonds which can be important for protein stability. Substitution of this residue to a lysine which is unable to form these bonds could therefore affect the stability of the domain (Betts and Russell, 2003).

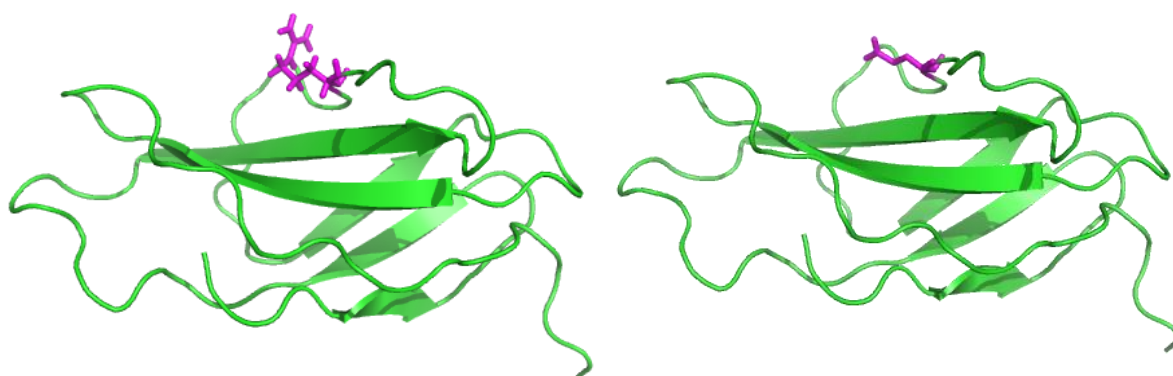


Figure 35 – Structural representation of the R820Q mutation in the C6 domain of cardiac myosin binding protein C

As there is currently no structural data for the C6 domain for any of the isoforms of myosin binding protein C (MyBP-C), I-TASSER software was utilised to model the domain. This model was then visualised using Pymol software. On the left can be seen the charged, polar and hydrophilic arginine at position 820. In the R820Q mutation this is replaced with a polar, non-charged glutamine (right). Arginine often forms salt bridges with other amino acids to create hydrogen bonds which can be important for protein stability. Replacement of this residue for a lysine, which is unable to form these bonds, could therefore destabilise the domain.

Structural studies carried out by Nadvi et al. identified that this was likely a surface exposed residue which would be unlikely to disrupt the domain fold, which they subsequently corroborated using biophysical techniques. They concluded that the mutation would be unlikely to interfere with the domain fold and thus would not affect the stability of the protein. Instead they suggested this substitution would more likely effect possible interactions, for example if the collar model for the arrangement of MyBP-C on the thick filament backbone is correct the R820 would be positioned at the interface with the C9 domain (Nadvi et al., 2016).

The R820Q mutation was first described by Konno et al. in 2003 (Konno et al., 2003) in which seven probands with HCM were identified as carrying the mutation, with no other mutations seen. The relatives of these HCM patients were then studied in order to determine whether they also carried the R820Q mutation, with seventeen of the twenty-four family members tested being shown to be carrying the mutation. This study showed that there was a penetrance of 40-50% in under 50 year olds and 70-100% in over 50 year olds. Furthermore, this study also showed that in 40% of over 70 year olds this mutation led to “burnt out” HCM, in which patients develop systolic dysfunction and left ventricular dilatation (Maron, 2002), which can be misdiagnosed as DCM.

These studies suggest that this mutation leads to a relatively late onset HCM, with possible transition into a “burnt out” phenotype. Furthermore, structural studies suggest that this mutation would not have an effect on protein stability leading to degradation of the protein but would more likely interfere with possible interactions with binding partners.

5.2.1. Protein expression levels of HCM-linked mutant pEGFP C3C6 constructs

To see whether any of the HCM-linked mutations led to a decrease in protein expression levels, and thus whether they could be leading to disease via haploinsufficiency, GFP-tagged plasmids containing the C3C6 region with the mutations shown in figure 26 were transiently transfected into COS-1 cells, samples were taken, run on an SDS-PAGE gel, transferred onto nitrocellulose and probed using an anti-GFP antibody. As is shown in figure 36A, a band was seen at the expected molecular weight of about 70 kDa for all of the mutant proteins, except for the R820Q mutant. The lower box shows an immunoblot against α -actin as a loading control indicating that this lack of a band for the R820Q mutant was not due to any differences in loading. Densitometry showed that none of the mutants had protein expression levels significantly different from the WT, apart from the R820Q mutation which led to a significant decrease in expression levels ($p=0.002$; one-way ANOVA; $n=3$; Figure 36B)

These results indicated that the missense mutations, except for the R820Q mutation, did not appear to lead to unstable protein. The R820Q mutation may therefore be causing disease via protein degradation leading to haploinsufficiency, however it is also possible that the lack of protein expression could have been due to a technical issue with the vector.

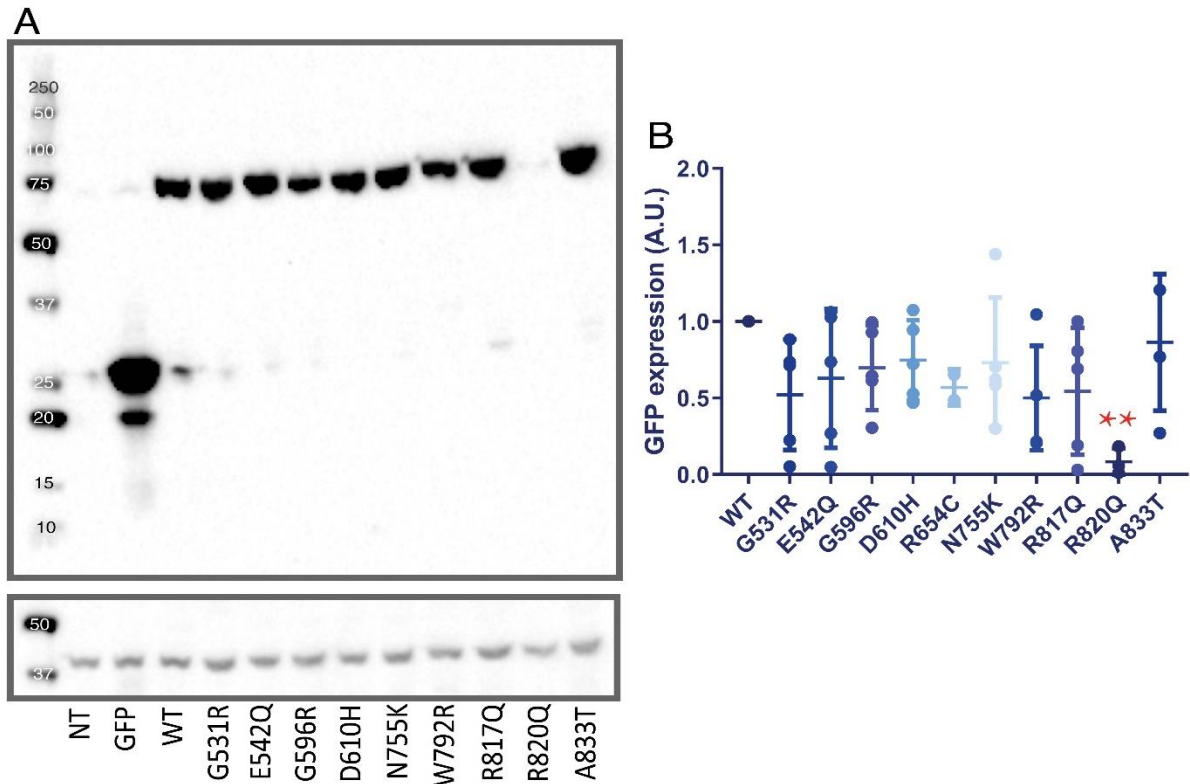


Figure 36 – Expression levels of GFP-tagged mutant C3C6 containing constructs

Immunoblot of COS-1 cells transiently transfected with GFP-tagged plasmids containing the mutant C3C6 domains using an α -GFP antibody. Cell samples were run on an SDS-PAGE gel and immunoblotted using an anti-GFP antibody (top panel) to determine protein expression levels of the mutant constructs. An anti-actin antibody was also used as a loading control (bottom panel). Bands at the expected molecular weight of 70 kDa were detected (A). Densitometry indicated the R820Q mutant showed significantly decreased expression levels ($p = 0.002$), whilst the rest of the mutations did not lead to any significant difference in levels of expression (B). ($n = 2-5$)

5.2.2. Cellular localisation of pEGFP constructs containing HCM-linked mutations

In order to determine whether the HCM-linked missense mutations could be causing disease by having an effect on the localisation of the protein, immunostaining and confocal microscopy was utilised. Wildtype and mutant C3C6 containing GFP-tagged plasmids were

transiently transfected into COS-1 cells and NRCs, immunostained for F-actin, DAPI and myomesin, and cellular localisation of the mutant C3C6 domains was determined using confocal microscopy.

The R820Q mutant did not express in either COS-1 cells or NRCs, whilst all other mutants did express, although at a low transfection rate. This suggests that the R820Q mutation led to an unstable protein that was degraded.

In the COS-1 cells all of the other mutant proteins appeared to have a similar diffuse cytoplasmic localisation to the WT (Figure 37). However, some looked like they had a more nuclear localisation which can be seen by the apparent co-localisation between the GFP and DAPI signals, when compared to WT, particularly G531R, E542Q, R817Q and A833T and therefore further quantification was subsequently carried out. Furthermore, there was also a similar co-localisation pattern seen between GFP and F-actin, as had been seen for the WT C3C6 domains described earlier. In the COS-1 cells there appeared to be some co-localisation between GFP-tagged C3C6 and F-actin at the periphery of the cells (figure 37).

In the NRCs the mutant proteins again appeared to have a similar diffuse cytoplasmic localisation to the WT (figure 38), as was seen in the COS-1 cells. The nuclear localisation of the mutant protein was less apparent in the NRCs, although the G596R mutant did show some evidence of nuclear localisation. Co-localisation with actin was also seen mainly in the myofibrils similar to the WT C3C6 domains (figure 38), although for some of the mutants, particularly G531R, G596R and A883T this was less obvious.

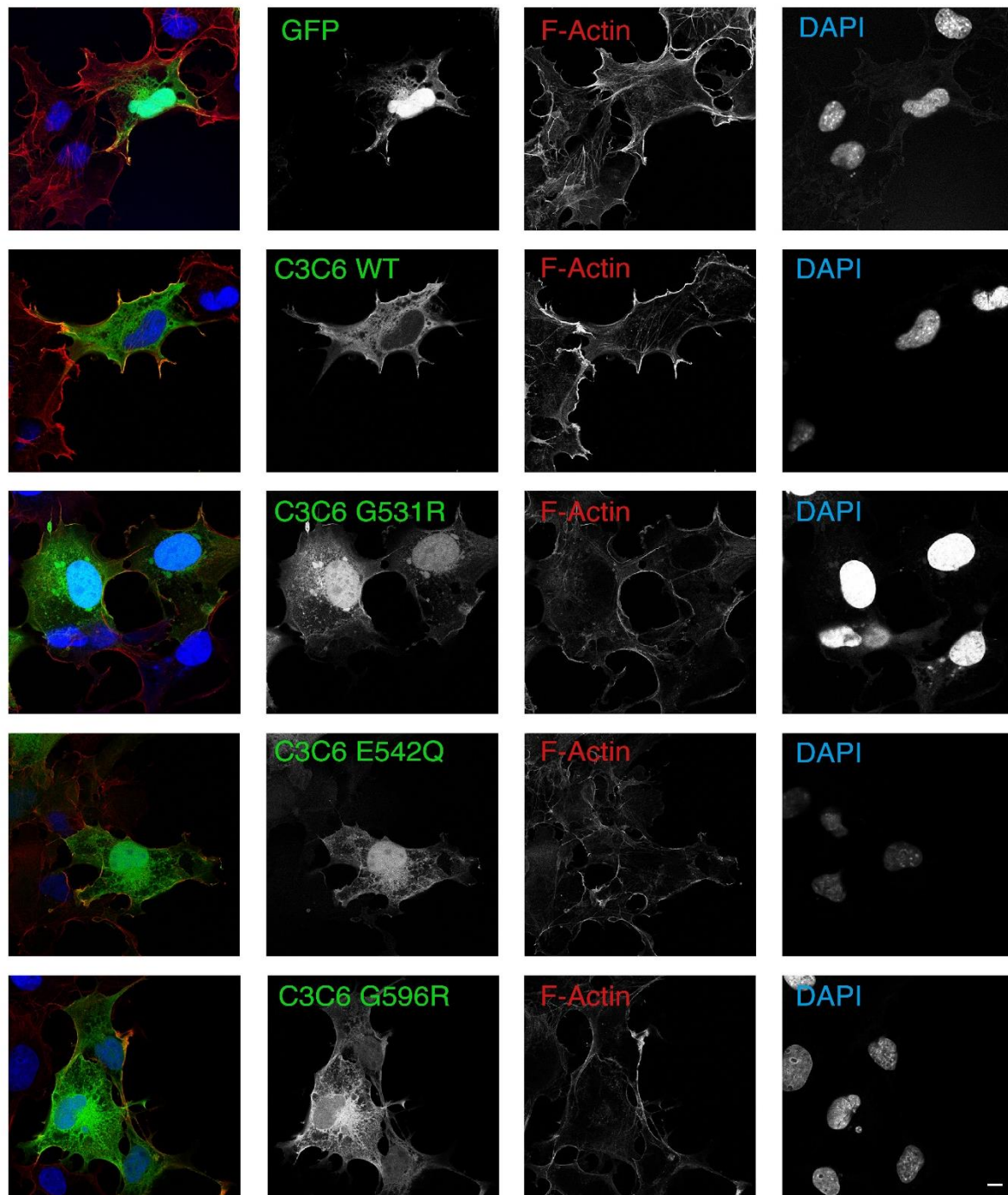


Figure 37 – Localisation of GFP-tagged C3C6 mutant constructs in COS-1 cells

Confocal microscopy images of immunostained COS-1 cells transiently transfected with GFP-tagged C3C6 mutant containing plasmids. Cells were immunostained for F-actin (red) and DAPI (blue) and confocal microscopy was used to determine localisation of the mutant constructs. All mutant constructs were expressed except for the R820Q mutant. All mutant constructs showed a diffuse cytoplasmic localisation, similar to the wild type (WT) construct. However, some mutant constructs did appear to have higher levels of nuclear localisation than the WT construct. Scale bar = 10 μ m

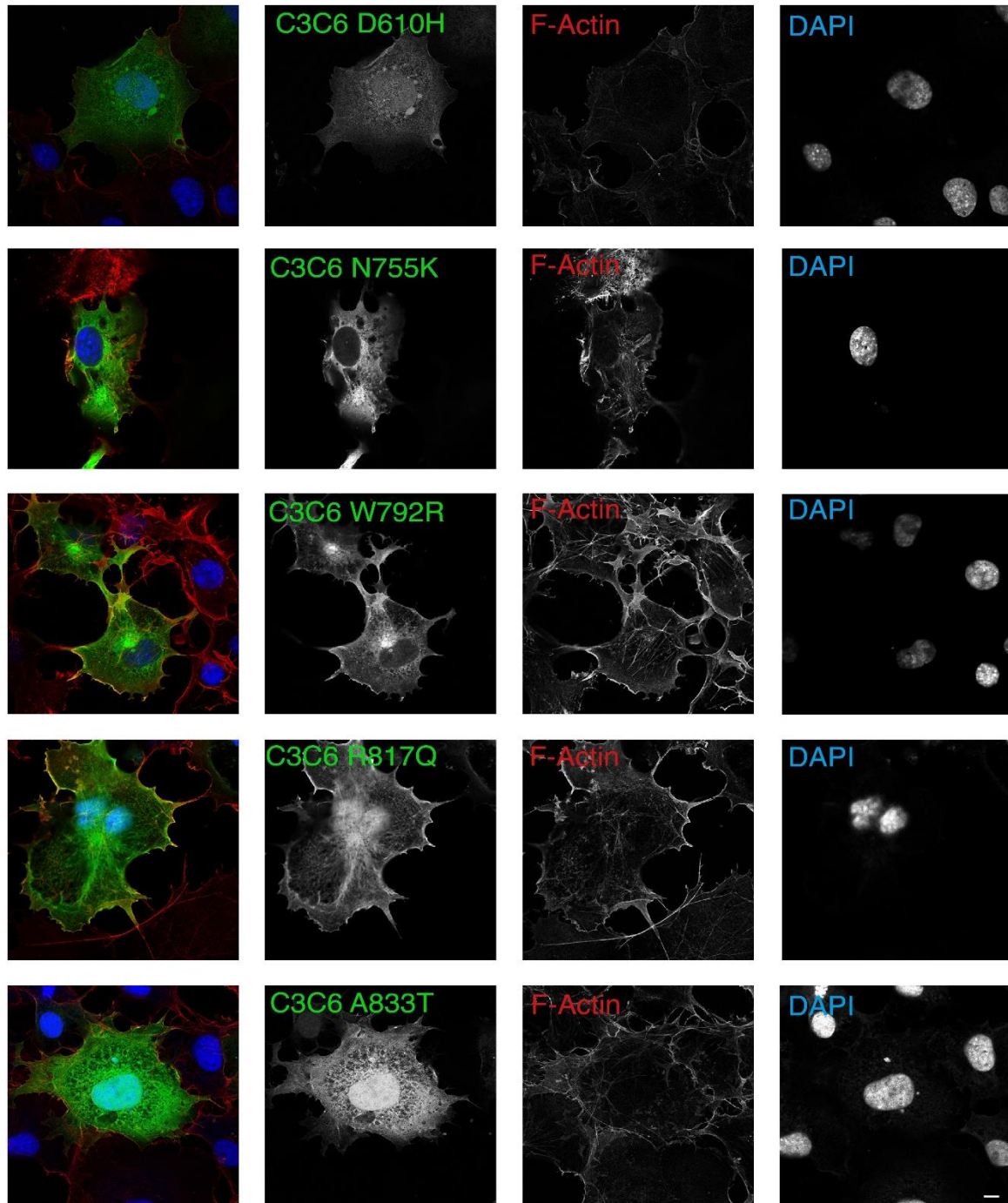


Figure 37 cont.– Localisation of GFP-tagged C3C6 mutant constructs in COS-1 cells

Confocal microscopy images of immunostained COS-1 cells transiently transfected with GFP-tagged C3C6 mutant constructs. Cells were immunostained for F-actin (red) and DAPI (blue) and confocal microscopy was used to determine localisation of the mutant constructs. All mutant constructs were expressed except for the R820Q mutant. All mutant constructs showed a diffuse cytoplasmic localisation, similar to the wild type (WT) construct. However, some mutant constructs did appear to have higher levels of nuclear localisation than the WT construct. Scale bar = 10 μ m

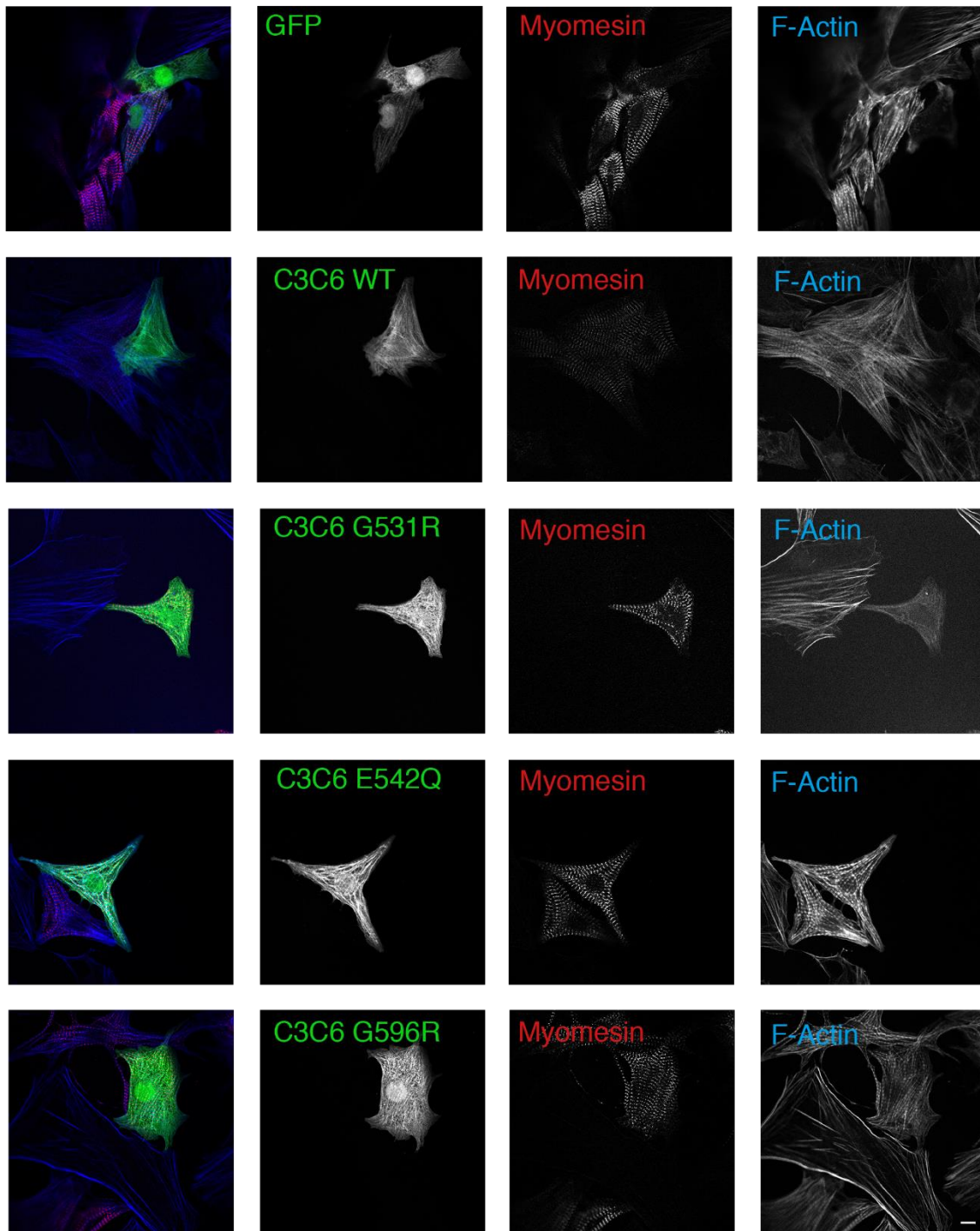


Figure 38 – Localisation of GFP-tagged C3C6 mutant constructs in neonatal rat cardiomyocytes

Confocal microscopy images of immunostained neonatal rat cardiomyocytes (NRCs) transiently transfected with GFP-tagged C3C6 mutant containing plasmids. Cells were immunostained for myomesin (red), F-Actin (blue) and DAPI (not shown) and confocal microscopy was used to determine localisation of the mutant constructs. All mutant constructs were expressed except for the R820Q mutant. All mutant constructs showed a diffuse cytoplasmic localisation, similar to the wild type (WT) construct. Scale bar = 10 μm

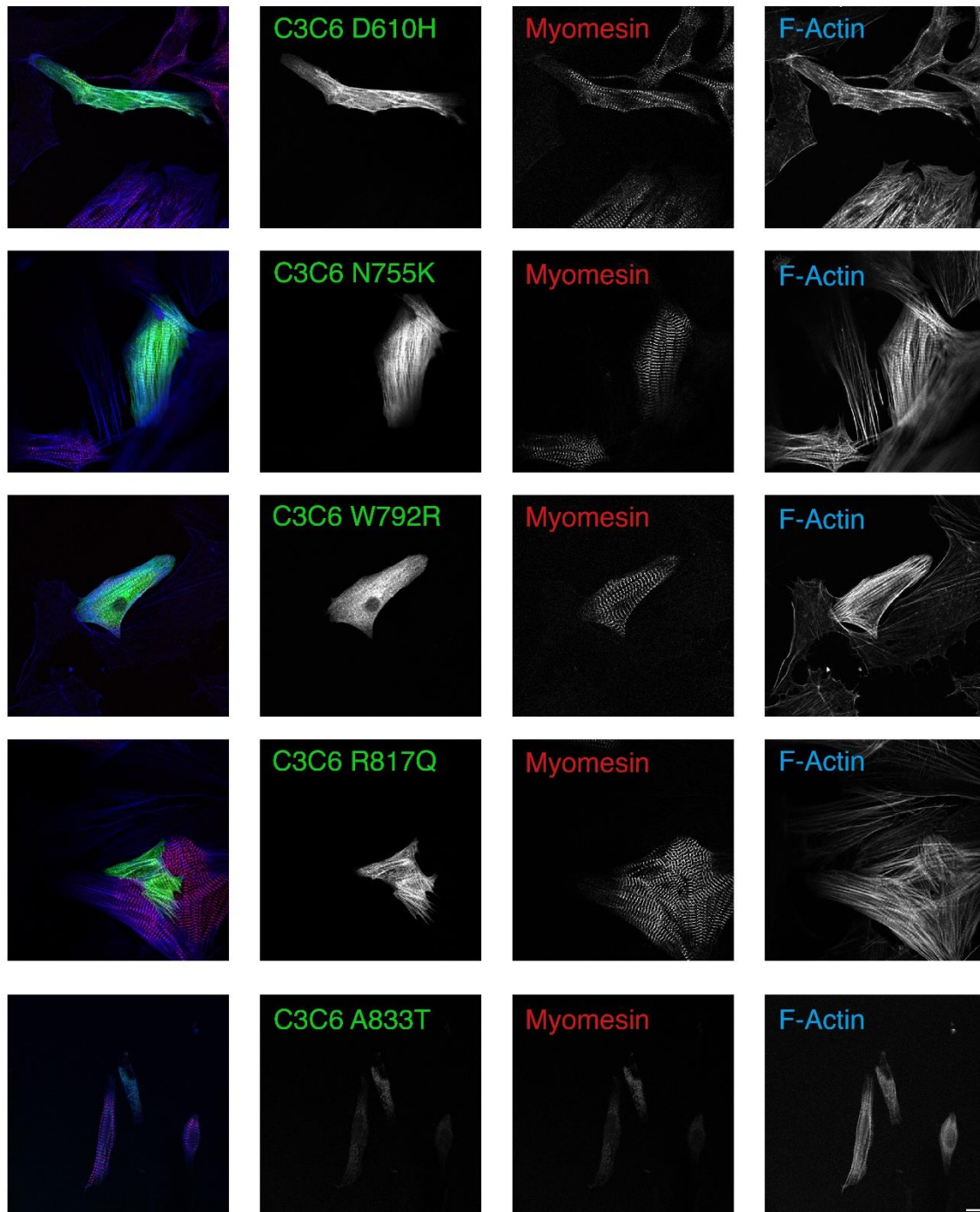


Figure 38 cont. – Localisation of GFP-tagged C3C6 mutant constructs in neonatal rat cardiomyocytes

Confocal microscopy images of immunostained neonatal rat cardiomyocytes (NRCs) transiently transfected with GFP-tagged C3C6 mutant containing plasmids. Cells were immunostained for myomesin (red), F-Actin (blue) and DAPI (not shown) and confocal microscopy was used to determine localisation of the mutant constructs. All mutant constructs were expressed except for the R820Q mutant. All mutant constructs showed a diffuse cytoplasmic localisation, similar to the wild type (WT) construct. Scale bar = 10 μm

5.2.3. Nuclear localisation of the mutant pEGFP C3C6 constructs

GFP-tagged proteins that fail to find a cytoplasmic binding partner often erroneously go to the nucleus (Seibel et al., 2007). However, confocal images showed that whilst GFP alone showed nuclear localisation, this was not the case for the WT C3C6 protein.

On the other hand, some of the mutant C3C6 domain proteins appeared to show some evidence of increased nuclear localisation compared to the WT C3C6 protein suggesting that some of the missense mutations could lead to mislocalisation of the protein. In order to further explore the possibility of the mutant plasmids leading to mislocalisation of the protein to the nucleus, Mathematica 11.3 software was utilised on the confocal microscopy images to measure nuclear localisation of the GFP-tagged constructs.

The results, shown in figure 39, indicated that there was a trend to an increase in nuclear localisation for the G531R, G596R and A833T mutants in the COS-1 cells compared to WT, whilst in NRCs only the G596R mutation appeared to cause a trend to an increase in nuclear localisation.

However, following quantification of nuclear localisation of the GFP-tagged mutant constructs none of the mutations were found to cause a significant mislocalisation of the protein to the nucleus compared with WT (one way ANOVA; n=2-4).

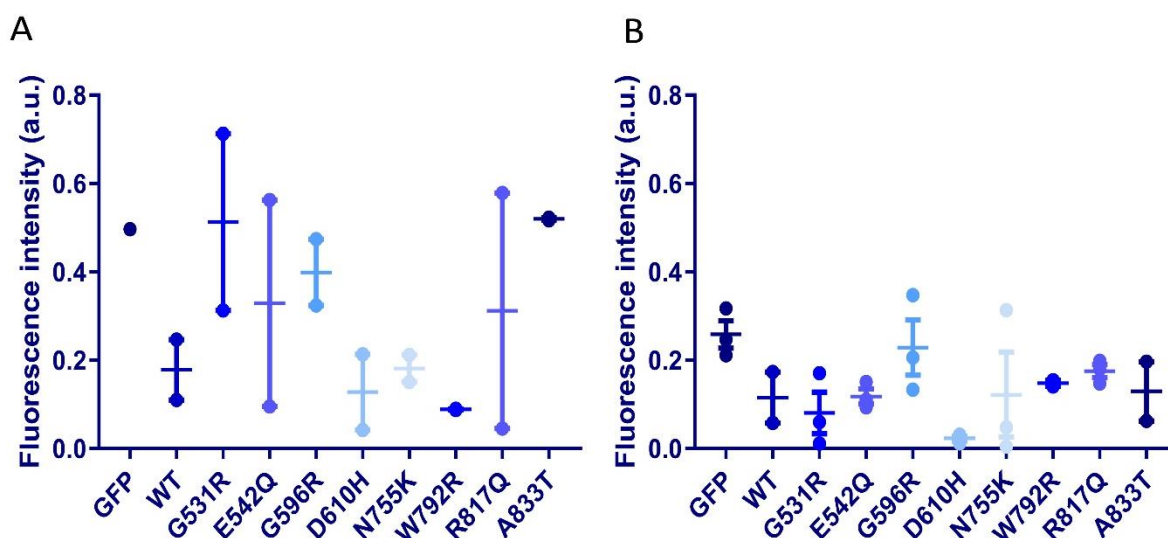


Figure 39 – Nuclear localisation of GFP-tagged mutant C3C6 constructs in COS-1 cells and neonatal rat cardiomyocytes

Figures showing quantification of nuclear localisation of mutant C3C6 GFP-tagged constructs in both COS-1 cells and neonatal rat cardiomyocytes (NRCs). Cells were transiently transfected with either WT or mutant C3C6 GFP-tagged constructs, immunostained for DAPI and confocal microscopy was used to determine subcellular localisation. Nuclear localisation was determined using Mathematica 11.3 software. The results indicated the WT construct showed a trend to a decrease in nuclear localisation compared with the pEGFP alone. In COS-1 cells the G531R, G596R and A833T mutant constructs showed a trend to an increase in nuclear localisation compared to the WT construct, whilst in NRCs only the G596R mutant construct showed a slight trend to an increase. (n= 2-3)

5.2.4. Effect of HCM-linked mutations on co-localisation of the central domains with actin

As described earlier, ratiometric analysis identified a significant increase in co-localisation of the GFP-tagged C3C6 WT plasmid with F-actin, indicating a possible interaction between the central domains of cMyBP-C and actin.

Confocal images also showed that there was co-localisation between F-actin and the GFP-tagged mutant C3C6 domain proteins, which for some of the mutations appeared less obvious. In order to determine whether any of the mutations could alter the levels of co-localisation with actin, ratiometric analysis was also carried out on confocal microscopy images for the GFP-tagged C3C6 mutants.

As mentioned earlier, ratiometric analysis identifies areas with the GFP-tagged C3C6 domains (WT or mutant) alone as purple/dark blue, F-actin alone is red, whilst green indicates the presence of both F-actin and the GFP-tagged C3C6, which could indicate co-localisation between the central domains and actin (figure 40A). The percentage of the cell which was green was determined via ImageJ software.

The results showed that the E542Q and G596R mutations caused a significant decrease in actin co-localisation when compared to WT ($p=0.0109$ and 0.0194 respectively; one way ANOVA; $n=1-5$; figure 40B), indicating that these mutations could be interfering with the possible interaction between the central domains and actin, whilst all of the other mutants appeared to have the same level of co-localisation with actin as the WT protein.

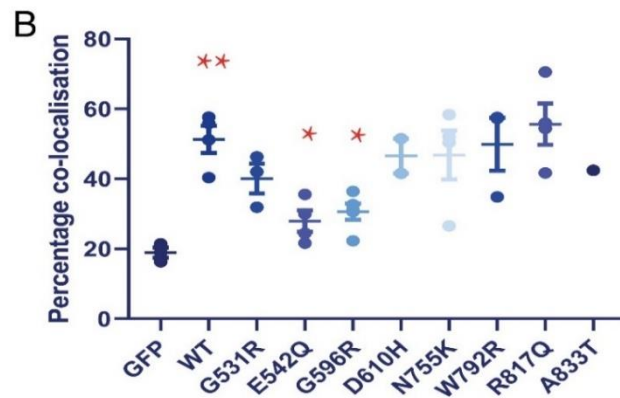
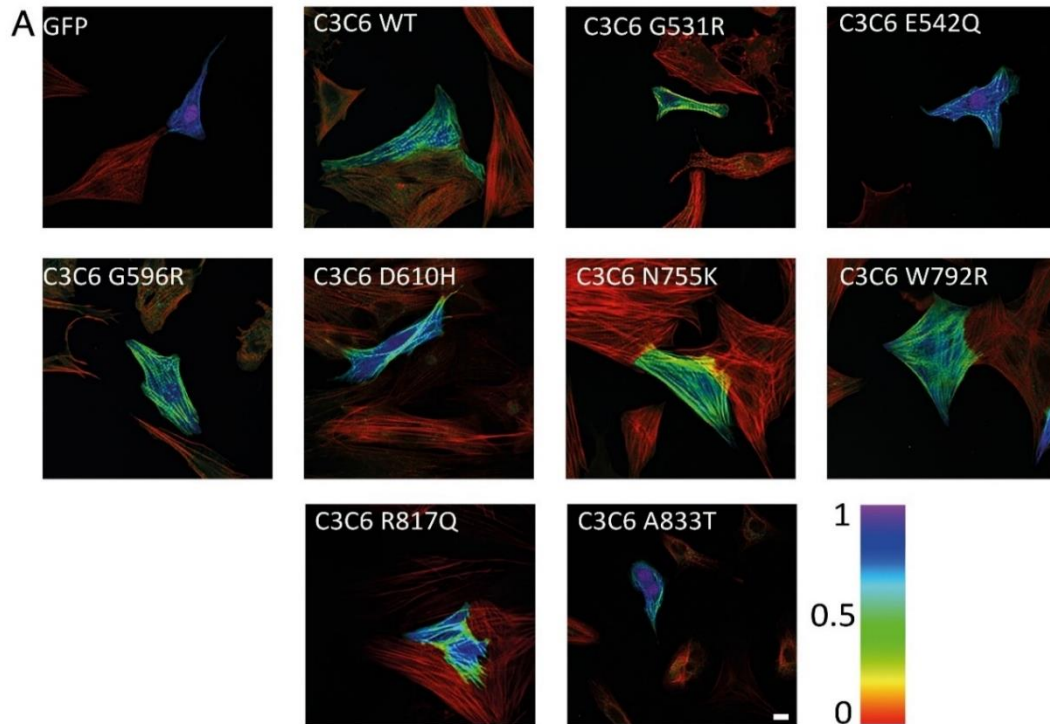


Figure 40 – Co-localisation of the central domains and F-actin

Images showing ratiometric analysis to determine levels of co-localisation between the GFP-tagged C3C6 constructs and F-actin in neonatal rat cardiomyocytes (NRCs). NRCs were transiently transfected with either wildtype (WT) of mutant GFP-tagged C3C6 containing plasmids, immunostained for F-actin and images were taken using a confocal microscope. Mathematica 11.3 software was then used to determine co-localisation, with red indicating F-actin alone, purple/dark blue indicating the pEGFP C3C6 construct alone and green indicating the presence of both (A). ImageJ was then utilised to determine the percentage of the cell which was green. The results indicated that the WT construct had a significant increase in co-localisation when compared to the EGFP alone ($p= 0.0009$) Furthermore, the E542Q and G596R mutants had a significant decrease in the percentage of co-localisation when compared to WT ($p= 0.0109$ and 0.0194 respectively) (B) ($n= 1-5$). Scale bar = $10\ \mu\text{m}$

5.2.5. Possible myofibrillar disarray in cells expressing mutant C3C6 domain constructs

In order to understand the impact of the HCM-linked missense mutations at the level of the myofibrils, high resolution stimulation emission depletion (STED) microscopy was utilised. NRCs were transiently transfected with the GFP-tagged mutant C3C6 plasmids and cultured for either 2 or 8 days. These were then immunostained for endogenous MyBP-C and a specialised STED secondary antibody and imaged using a STED microscope.

As an antibody against MyBP-C was used, which is raised against the C0 domain of the protein and therefore would not detect the C3C6 domain proteins, only endogenous cMyBP-C was being visualised. Images were taken of cells which were known to be expressing the GFP-tagged mutants, however any disarray seen cannot be directly attributed to the mutant protein but could still be an indirect impact on the myofibrils occurring due to the presence of the mutant protein.

As can be seen in figure 41, the NRCs transiently transfected with the N755K, W792R, R817Q and A833T mutant plasmids appeared to have an increase in myofibrillar disarray when compared with WT after 2 days of culturing, which was exacerbated following prolonged culture. The cells transiently transfected with the remaining mutations (G531R, E542Q, G596R), which showed relatively normal myofibrils at 2 days of culture, also appeared to have a worsening of myofibrillar disarray when the cells were cultured for 8 days.

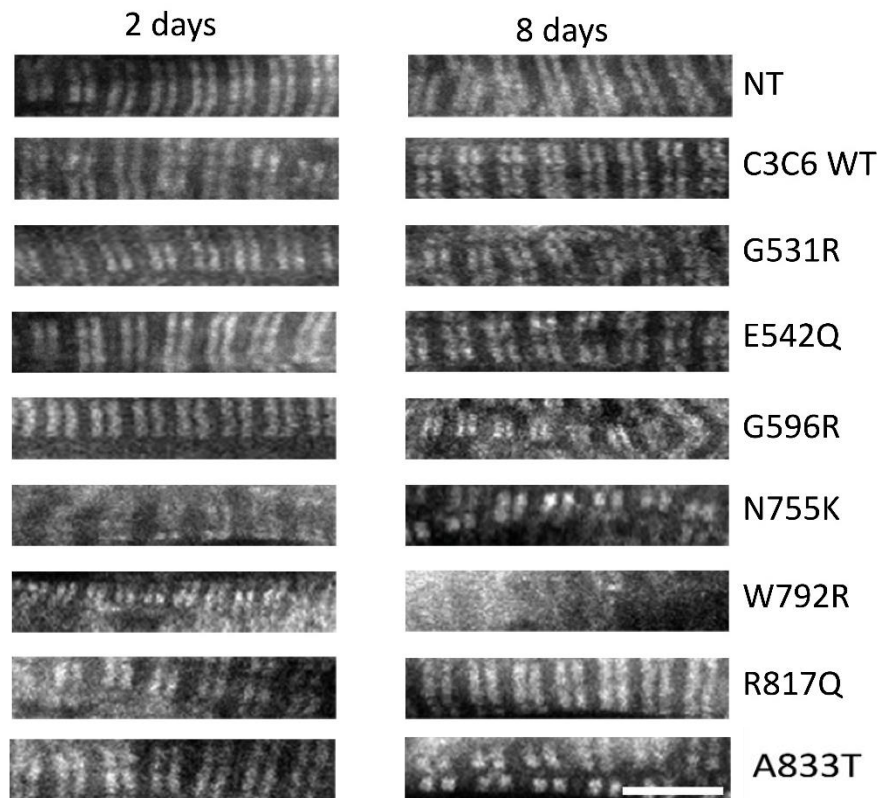


Figure 41 – Myofibrillar disarray in neonatal rat cardiomyocytes expressing mutant GFP-tagged C3C6 containing plasmids

STED microscopy images of immunostained neonatal rat cardiomyocytes transiently transfected with GFP-tagged C3C6 wildtype and mutant containing plasmids. NRCs were transiently transfected with the plasmids and cultured for either two or eight days and immunostained for cardiac myosin binding protein C. Cells which were known to be expressing the mutant protein were imaged using a STED microscope. Cells expressing mutant constructs appeared to have myofibrillar disarray, particularly N755K and W792R, which was exacerbated when the cells were cultured for longer. Scale bar = 5 μ m

Further analysis of the STED images allowed for a more in depth look at the possible myofibrillar disarray seen in the NRCs expressing the mutant constructs. Images were analysed as described in section 2.7.2.2., resulting in intensity profile plots, which can be seen in figures 42-44 for images from cells cultured either short term (ST; A) or long term (LT; B), with raw data for each individual image shown on the left whilst the plot of the averaged data is shown on the right (for averaged data the blue line represents mean value,

red indicates standard deviation; $n = 2-3$). A major peak at ~ 2 with a smaller secondary peak at ~ 4 denotes a sarcomere length at the expected $\sim 2 \mu\text{m}$ (green arrows). Two additional peaks positioned before each of the major peaks denote MyBP-C at the expected spacing (purple arrows). Strong, narrow peaks seen at each of these positions would indicate well organised sarcomeres. From these plots the sarcomere length could be extrapolated, as well as the intensity of the major sarcomeric peak and the major MyBP-C peak, graphs of which can be seen in figure 45.

As can be seen in figure 42 the analysis plots for both the NT and WT images showed well defined peaks for the images of the short term cultured cells suggesting a high level of sarcomeric organisation, although the peaks were seen to be less well defined in the NT long term cultured cells suggesting some disarray in the myofibrils following the longer culture. However, as can be seen in the plot showing the raw data, there did seem to be some variability between individual images for the NT, long term cultured cells.

The plots for both the G531R (figure 42) and E542Q (figure 43) mutant construct transfected cells showed well defined peaks similar to that seen for WT suggesting that these mutations do not have an effect on the sarcomeric organisation.

Cells transfected with the G596R mutant C3C6 construct (figure 43) showed slightly less well defined peaks than WT suggesting that this mutation could be causing minor sarcomeric disarray.

Plots from images of cells expressing both the R817Q and A833T mutant C3C6 domains (figure 44) showed plots very similar to WT for ST culture with strong well defined peaks. However, the plots for images of the LT cultured cells showed less well defined peaks

suggesting these mutations may cause disorganisation of the sarcomeres following longer culturing.

Finally, plots of images from cells expressing the N755K (figure 43) and W792R (figure 44) C3C6 mutant constructs showed less well defined peaks which were shown to worsen for cells which had undergone longer culture. This suggests that these mutations may cause myofibrillar disarray which is exacerbated by longer culture.

From the intensity profile plots sarcomere length was extrapolated, which is shown in figure 45A, which showed that the sarcomere length remained very similar regardless of whether cells were expressing WT or mutant C3C6 domain constructs, or if the cells were cultured for short or long term. However, as can be seen in figure 45B there was some variability in the standard deviation especially for the G596R and N755K as well as the G531R and W792R short term cultured C3C6 domain expressing cells, suggesting a possible increase in irregularity of sarcomere spacing.

The intensity of the sarcomere signal peaks (figure 45C) were generally similar, with low levels of variation shown by standard deviation (figure 45D). There was however a trend to a decrease across all mutants showing a decrease in intensity in LT cultured cells compared to ST, which was not apparent for WT, suggesting a possible increase in sarcomeric disarray following longer culture times. Although there were no significant differences in the sarcomere signal intensity, some mutants showed decreased intensity levels compared with WT, particularly N755K, W792R LT, R817Q, A833T LT, suggesting that these mutations may be causing disorganisation of the sarcomeres.

The intensity of the MyBP-C signal peak (figure 45E) also showed a trend to a decrease in intensity for LT cultured cells compared to ST for the mutants, again suggesting increased disorganisation of the sarcomeres following longer culture times. The standard deviation was seen to be relatively similar across the board, although G531R ST, G596R ST and R817Q ST did appear to have slightly higher levels of variability (Figure 45F). Whilst the intensity of the MyBP-C signal peak was shown to be similar to WT for the majority of the mutant C3C6 expressing cells, for the W792R C3C6 mutant intensity was shown to be decreased suggesting a possible misalignment of the sarcomeres caused by the presence of this mutation. The G531R, N755K and R817Q mutants also showed a decrease in signal intensity for the LT cultured cells suggesting possible disarray of the sarcomeres but only following the longer culture.

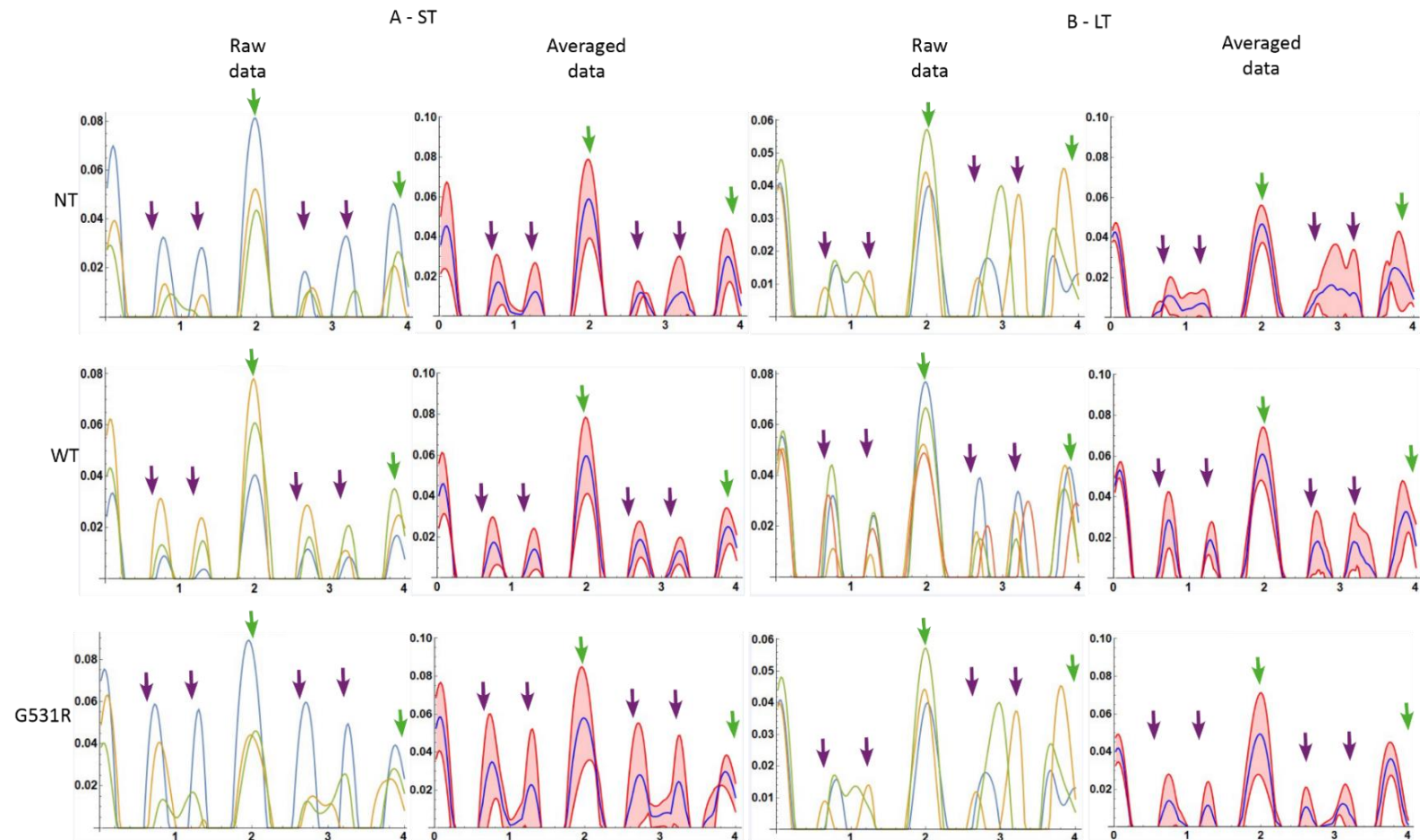


Figure 42 – Intensity profile plots from STED images of NRCs expressing mutant C3C6 domain constructs

Figure showing intensity profile plot analysis of STED images of neonatal rat cardiomyocytes (NRCs) that had either not been transfected (NT) or expressing wildtype (WT) or the G531R mutant C3C6 domain constructs following either short term (ST; A) or long term (LT; B) culture. Raw data showing plots for each individual image are shown on the left, with the averaged data shown on the right (blue line denotes mean intensity, red indicates standard deviation). A major peak at ~ 2 with a smaller secondary peak at ~ 4 denotes the sarcomere spacing, with 2 peaks prior to each major peak denoting the spacing of the MyBP-C stripes (purple arrows). All of the plots showed strong narrow peaks at each of these points indicating well organised sarcomeres, however the plot for the NT cells following LT culture did show less well defined peaks suggesting less well organised sarcomeres following a longer culture. (n=3)

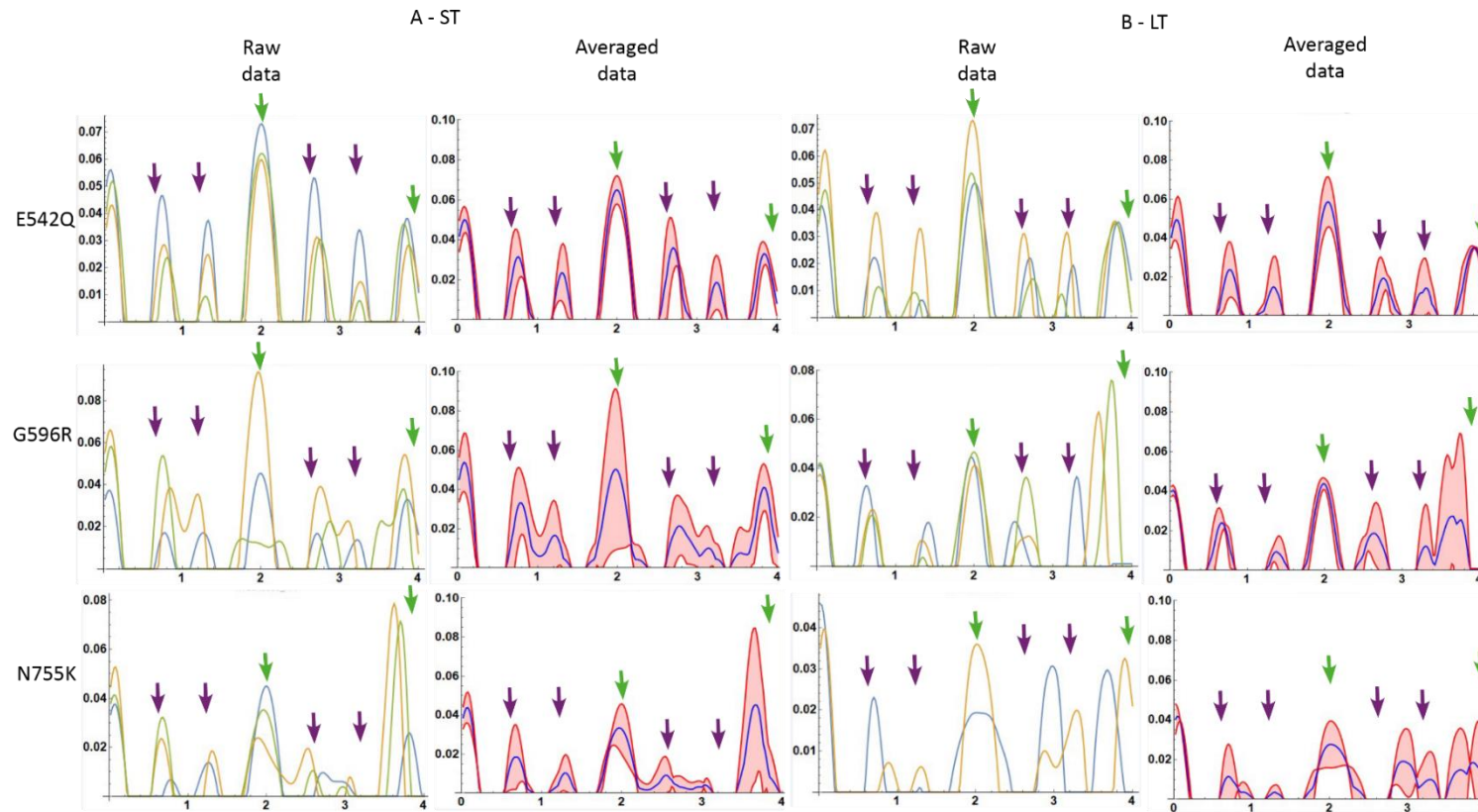


Figure 43 – Intensity profile plots from STED images of NRCs expressing mutant C3C6 domain constructs

Figure showing intensity profile plot analysis of STED images of neonatal rat cardiomyocytes (NRCs) expressing the E542Q, G596R and N755K mutant C3C6 domain constructs following either short term (ST; A) or long term (LT; B) culture. Raw data showing plots for each individual image are shown on the left, with the averaged data shown on the right (for averaged data blue line denotes mean intensity, red indicates standard deviation). A major peak at ~ 2 with a smaller secondary peak at ~ 4 denotes the sarcomere spacing, with 2 peaks prior to each major peak denoting the spacing of the MyBP-C stripes (purple arrows). The plots for the images of the E542Q C3C6 domain expressing cells showed strong narrow peaks similar to that of WT indicating well organised sarcomeres. Images from G596R C3C6 domain expressing cells produced plots similar to WT but with slightly less distinct peaks. The plots for the images from cells expressing the N755K C3C6 domain mutant also showed less distinct peaks, with the peaks for the LT plot showing even less distinct peaks, suggesting this mutation may lead to disorganisation of the myofibrils which worsened following longer culture. (n= 2-3)

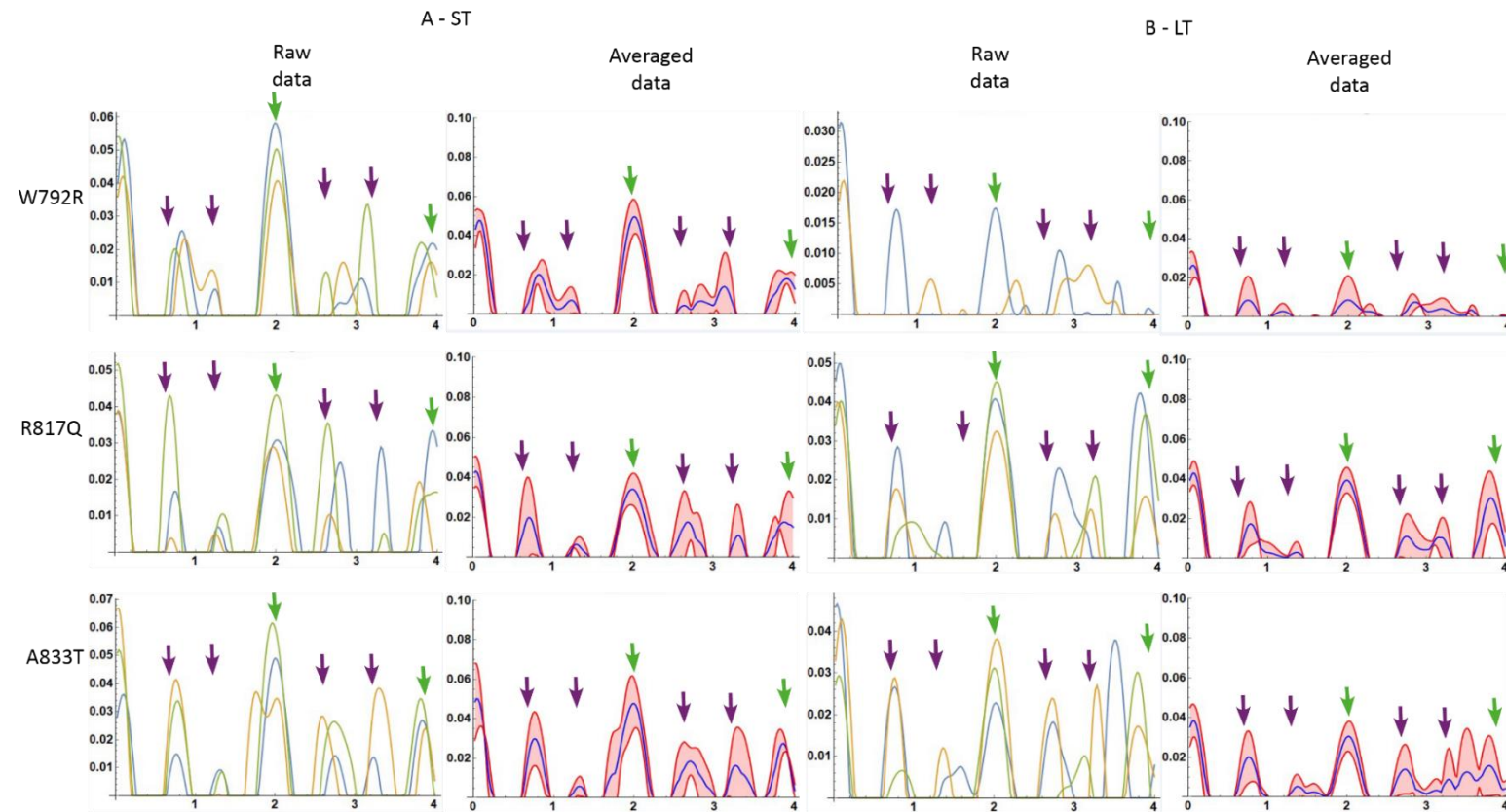


Figure 44 – Intensity profile plots from STED images of NRCs expressing mutant C3C6 domain constructs

Figure showing intensity profile plot analysis of STED images of neonatal rat cardiomyocytes (NRCs) expressing the W792R, R817Q and A833T mutant C3C6 domain constructs following either short term (ST; A) or long term (LT; B) culture. Raw data showing plots for each individual image are shown on the left, with the averaged data shown on the right (for averaged data blue line denotes mean intensity, red indicates standard deviation). A major peak at ~2 with a smaller secondary peak at ~4 denotes the sarcomere spacing, with 2 peaks prior to each major peak denoting the spacing of the MyBP-C stripes (purple arrows). The plots for the images of the W792R C3C6 domain expressing construct showed peaks which were less distinct than wildtype, with the plot for the LT images showing peaks which were much less distinct and of weaker intensity, suggesting that this mutation appears to lead to a lack of organisation of the sarcomeres which worsens following prolonged culture. The plots for images of cells expressing the R817Q and A833T C3C6 mutant domains showed relatively strong and distinct peaks similar to that of the WT however following longer culture the peaks were less distinct. (n= 2-3)

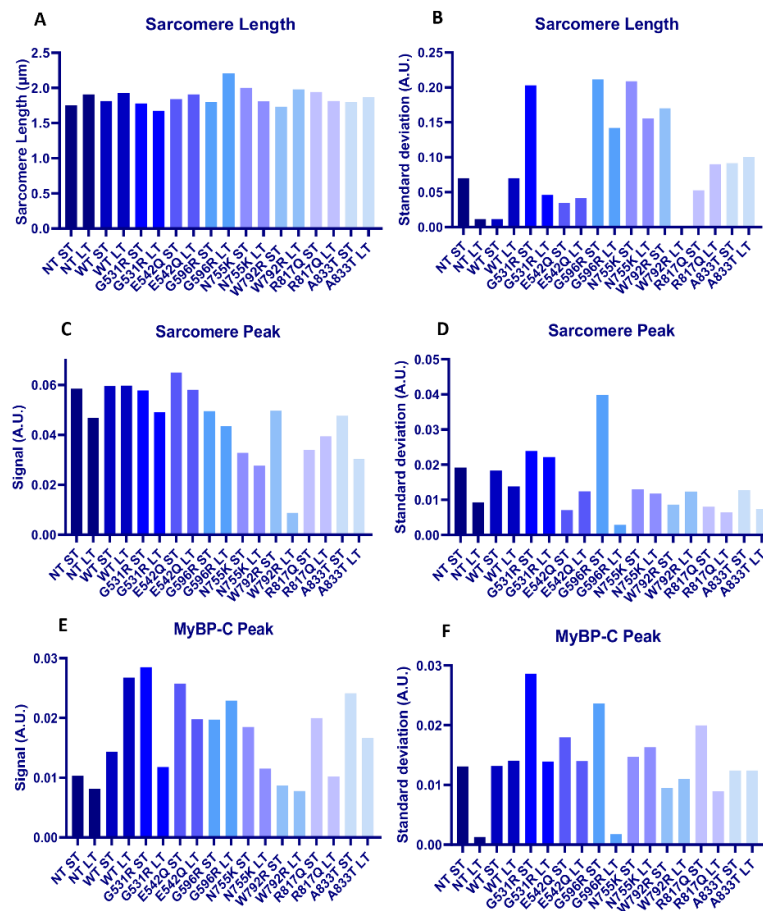


Figure 45 – Sarcomeric organisation of NRCs expressing mutant C3C6 domain constructs

Graphs showing various measures (sarcomere length, sarcomere peak intensity signal and MyBP-C peak signal intensity) indicating levels of sarcomeric organisation extrapolated from intensity profile plots generated from analysis of STED images. Sarcomere length (A) was shown to remain the same regardless of whether cells underwent short term (ST) or long term culture (LT) or whether cells were expressing wildtype (WT) or mutant C3C6 domain constructs. However, there was an increase seen in the standard deviation (B) in some cases, especially G531R ST, G596R (ST and LT), N755K (ST and LT) and W792R ST, suggesting a possible increase in variability of sarcomere spacing. The sarcomere peak intensity signal showed a trend to a decrease for cells expressing the R817Q, N755K, W792R (LT) and A833T (LT) mutant C3C6 domain constructs (C) suggesting these mutations may cause disorganisation of the sarcomeres. The standard deviation was seen to be relatively low suggesting a low amount of variability (D). MyBP-C peak intensity signal showed a trend to a decrease for cells expressing the W792R C3C6 domain mutant, as well as in the LT cultured cells expressing the G531R, N755K and R817Q C3C6 domain mutant constructs (E), suggesting that these mutations may cause disarray of the myofibrils, particularly following longer culture. Furthermore, there appeared to be a general trend of a decrease in signal intensity for both the sarcomere and MyBP-C peak for the LT culture cells compared to ST which was not seen in WT (C and E), suggesting the mutant proteins may be affecting sarcomeric organisation only following prolonged culture. (n= 2-3).

5.2.6. Summary of findings

A summary of all of the findings from utilisation of the GFP-tagged mutant C3C6 domain containing vectors can be seen in table 9, as well as data about the pathogenicity of the mutation as stated in the ClinVar database (Landrum et al., 2018).

MUTATION	PROTEIN LEVELS	NUCLEAR LOCALISATION	ACTIN CO-LOCALISATION	DYSREG OF MYOFIBRILS (2 DAYS)	DYSREG OF MYOFIBRILS (8 DAYS)	EVIDENCE OF PATHOGENICITY (CLINVAR)
G531R		+			+	Conflicting views
E542Q			-		+	Pathogenic
G596R		++	-		+	Uncertain significance
D610H				N/A	N/A	Conflicting views
N755K				++	++	Pathogenic
W792R				+	+++	Pathogenic/ likely pathogenic
R817Q				+	+	Conflicting views
R820Q	Significant decrease	N/A	N/A	N/A	N/A	Pathogenic/ likely pathogenic
A833T		+		+	++	Conflicting views

Table 9 – Summary of results from utilisation of the mutant GFP-tagged C3C6 domain mutant containing plasmids

Table summarising the results gained from the experiments utilising the GFP-tagged C3C6 domain plasmids containing hypertrophic cardiomyopathy linked mutations, as well as the evidence of pathogenicity according to the ClinVar database.

The mutation which appeared to have the greatest effect on the protein was R820Q, which did not express in either the COS-1 cells or the NRC and showed a significant decrease in protein expression levels according to the western blot analysis. This could indicate the mutation is leading to an unstable protein or mRNA which is being degraded, however this is

contrary to previous structural studies which suggested that this mutation would not impede the domain folding and thus would not lead to protein degradation, but would be more likely to interfere with potential binding partners (Nadvi et al., 2016). It is possible that there was an issue with the R820Q pEGFP vector, or with its GFP tag, which either made it unstable or undetectable during the experiments, although the sequences of all of the plasmids were checked before being utilised for transient transfection.

Analysis of nuclear localisation of the GFP-tagged C3C6 mutant plasmids showed possible mislocalisation of the protein for the G531R, G596R and A833T mutants, whilst the E542Q and G596R mutation appeared to disrupt the possible interaction of the central domains with actin.

Lastly, STED microscopy of endogenous cMyBP-C indicated that the NRCs expressing all of the C3C6 domain mutation plasmids showed myofibrillar disarray following prolonged culturing, whilst the N755K, W792R, R817Q and A833T mutation expressing cells had disarray even after the 2 day culture.

These preliminary findings suggest that the R820Q mutation may lead to protein instability and degradation and therefore the mechanism of pathogenicity could be haploinsufficiency. The other mutations however appeared to lead to no differences in protein expression levels, indicating these mutations may act in a “poison polypeptide” manner and have some effect on the localisation of the protein or interfere with possible interaction partners.

5.2.7. Limitations of the experiments utilising the pEGFP vectors

The use of the pEGFP vectors, though relatively simple and quick, was not without issues. For example, the transfection rates were very low, even for the WT plasmid, therefore it is

possible that any phenotypic observations attributed to the mutations could just have been due to the fact that the only cells that were taking up the plasmid were cells that were not in the best of condition to begin with.

Another issue with this approach is the fact that only the C3C6 domains were being expressed, without the C-terminus which is known to be the region of the protein responsible for anchoring the protein into the thick filament backbone (Okagaki et al., 1993, Freiburg and Gautel, 1996) and thus these fragments would not incorporate into the A-band and so would not be completely physiologically relevant.

In order to remedy these issues, promising mutations were carried over into a full length human cMyBP-C adenoviral construct, which would not only be capable of properly incorporating into the sarcomere due to the inclusion of the C-terminus but should also have a much higher infection rate.

5.3.1. Full length cardiac myosin binding protein C adenoviral constructs containing HCM-linked mutations

In order to gain a more physiologically relevant understanding of any pathological effects the HCM-linked mutations were having, adenoviral constructs encoding for mutated full-length cMyBP-C were utilised. As the C-terminal domains required for incorporation of the protein into the thick filament backbone were present this would mean the protein would be able to localise in its correct stoichiometry.

A pAd construct with an N-terminal GFP tag and the full length human cMyBP-C was purchased from Vector Builder and propagated and amplified in HEK-293 cells to produce an adenoviral stock. The cMyBP-C region was then amplified out of the plasmid and cloned into a smaller pENTR vector using Gibson cloning in order to introduce mutations via site directed mutagenesis. Six HCM-linked mutations (E542Q, G596R, N755K, W792R, R817Q and R820Q) were chosen based on the preliminary findings from the experiments utilising the pEGFP vectors encoding for the shorter C3C6 constructs, as well as the evidence of pathogenicity from the ClinVar database (Landrum et al., 2018). All of the pENTR vectors were sequenced to verify that the mutation had been introduced.

The pENTR vectors then underwent LR recombination with the pAd vector, and these were propagated and amplified in the HEK-293 cells. All of the mutant pAd vectors were successfully amplified, except for the R817Q mutant.

5.3.2. Optimisation of adenoviral infection

In order to determine the optimal dilution to use the adenovirus in further experiments, COS-1 cells were infected with the WT adenovirus using a range of different dilutions (between 1:100 and 1:4) for 24 hours. Samples were taken and run on an SDS-PAGE gel,

transferred onto nitrocellulose and probed for MyBP-C. As can be seen in figure 46A, there was a band at the correct molecular weight of 150 kDa for cMyBP-C (arrow), which was not seen for the negative control of the spleen sample, indicating cMyBP-C protein was being produced. At the higher dilutions (1:4-1:20) there was a band at a lower molecular weight indicating possible degradation of the protein (asterisk). As there was a clear, single band at the correct molecular weight at the lowest dilution of 1:100 this was the dilution chosen for subsequent experiments.

The optimal duration of infection was also determined by infecting COS-1 cells for either 24 or 72 hours. As can be seen in figure 46B, there were multiple bands seen from the sample infected for 72 hours, indicating that degradation of the protein was occurring. When infection occurred for 24 hours, a single band at the correct molecular weight (150 kDa; arrow) was seen; therefore this infection duration was used for further experiments.

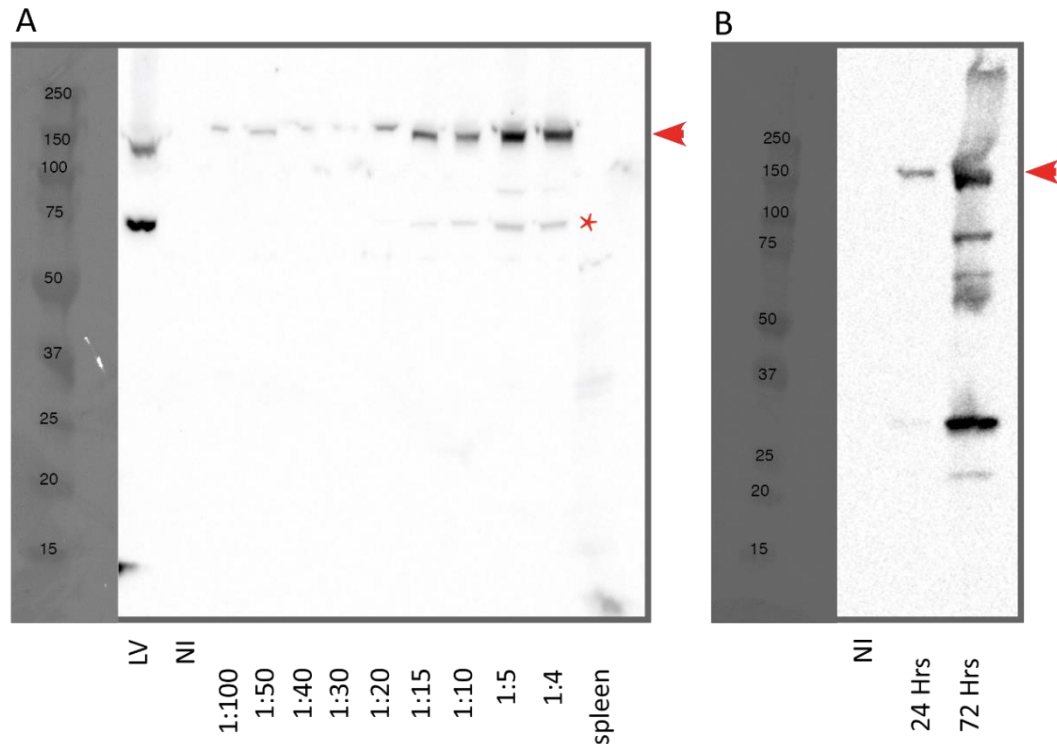


Figure 46 – Optimisation of adenoviral dilution and infection duration

Immunoblot of COS-1 cells infected with full length cardiac myosin binding protein C adenovirus at varying dilutions (A) and different infection durations (B). COS-1 cells were infected with increasing dilutions (1:100 – 1:4) of adenovirus and samples were run on an SDS-PAGE gel and immunostained for cMyBP-C. Results indicated that at higher dilutions (1:15-1:4) there were bands at lower molecular weight (asterisks) than the expected 150 kDa (arrow) suggesting protein degradation. As the lowest dilution of 1:100 produced a band at the correct molecular weight, this was used for all subsequent experiments. This dilution of adenovirus was subsequently used to infect COS-1 cells for 24 or 72 hours and samples were run on an SDS-PAGE gel and immunostained for cMyBP-C. samples from cells infected for 72 hours showed many bands at a lower molecular weight than the expected 150 kDa (arrow), indicating this longer infection period led to protein degradation, therefore 24 hour infection was used in all subsequent experiments.

5.3.3. Titration of mutant adenoviral constructs

To ensure that the WT and mutant adenoviruses were being used at the same titre, and therefore any differences seen in protein expression would be purely due to any effect the

mutations could be having on the stability of the protein, a titration experiment was carried out. This was done by infecting HEK-293 cells with a serial dilution of each adenovirus and using a titration kit to immunostain for the protein hexon, which is produced upon viral replication. Cells which were stained, and thus were infected with the adenovirus, were then counted and the viral titre was calculated. All of the adenoviruses showed hexon staining (Figure 47, arrows) indicating that all were capable of replicating and infecting the cells. Based on the titre calculations, in all of the following experiments the mutant adenoviruses were added at a titre equal to that of the WT when it was at a 1:100 dilution.

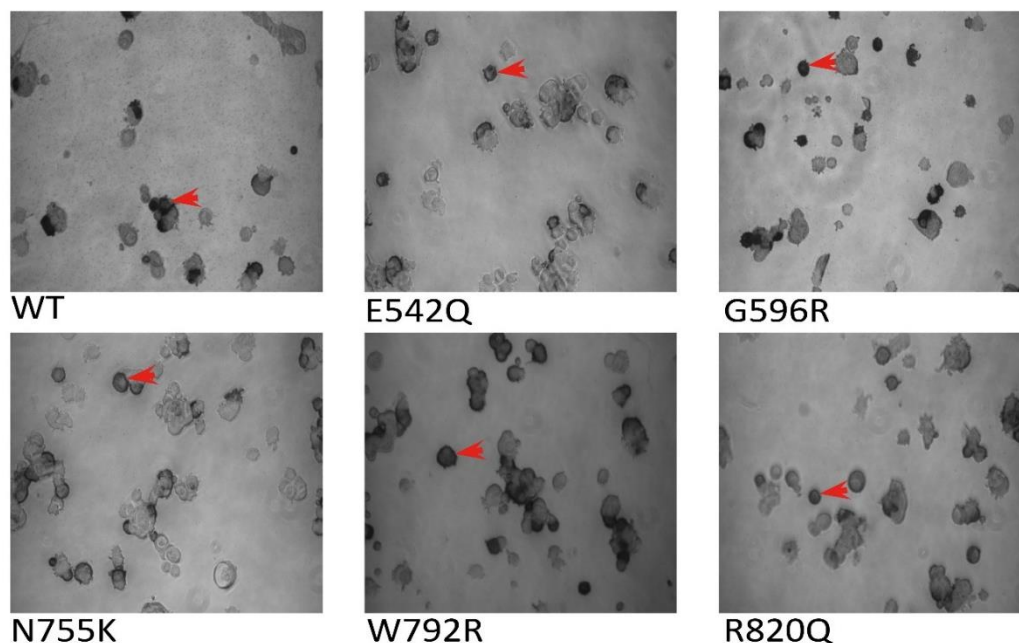


Figure 47 – Titration of full-length cardiac myosin binding protein C mutant adenoviral constructs

Images of titration experiment of full-length cardiac myosin protein C mutant adenoviral constructs. Neonatal rat cardiomyocytes were infected with the adenoviral constructs and a titration kit was used to determine the titre of each adenoviral construct. Cells were immunostained for the protein hexon, which is produced during viral replication. Cells which were infected by the adenovirus were therefore stained a dark brown colour (examples indicated by arrows). Number of infected cells were counted and used to determine the viral titre.

5.3.4. Protein expression levels of mutant constructs

In order to determine protein expression levels for all of the full-length mutant proteins, NRCs were infected with both WT and mutant adenovirus for 24 hours. Although as mentioned above, all of the adenoviruses were shown to be replicating and infecting cells, the R820Q mutant did not appear to express, with no GFP signal being seen. All of the other mutants did express and showed an infection rate similar to the WT adenovirus. Infected cells were then either processed into SDS-PAGE samples or fixed to carry out confocal microscopy.

SDS-PAGE samples were run on an SDS-PAGE gel, transferred to nitrocellulose and probed using an anti-GFP antibody, then an anti-MyBP-C antibody to determine a ratio between the endogenous and exogenous protein levels. Densitometry was carried out to determine protein expression levels, the results of which, figure 48A, showed that the N755K mutant showed a slight increase in expression levels, the W792R mutant had a slight decrease in protein expression and the R820Q mutant lead to a significant decrease in protein levels ($p=0.0197$; one-way ANOVA; $n=3$), whilst the other mutants did not show a difference in expression levels compared with WT.

The NRCs were also imaged using confocal microscopy and fluorescence intensity levels of GFP were measured using ImageJ. As can be seen in figure 48B, this showed that the W792R mutant lead to a significant decrease in levels of GFP fluorescence compared with WT ($p=0.0069$; one-way ANOVA; $n=3$).

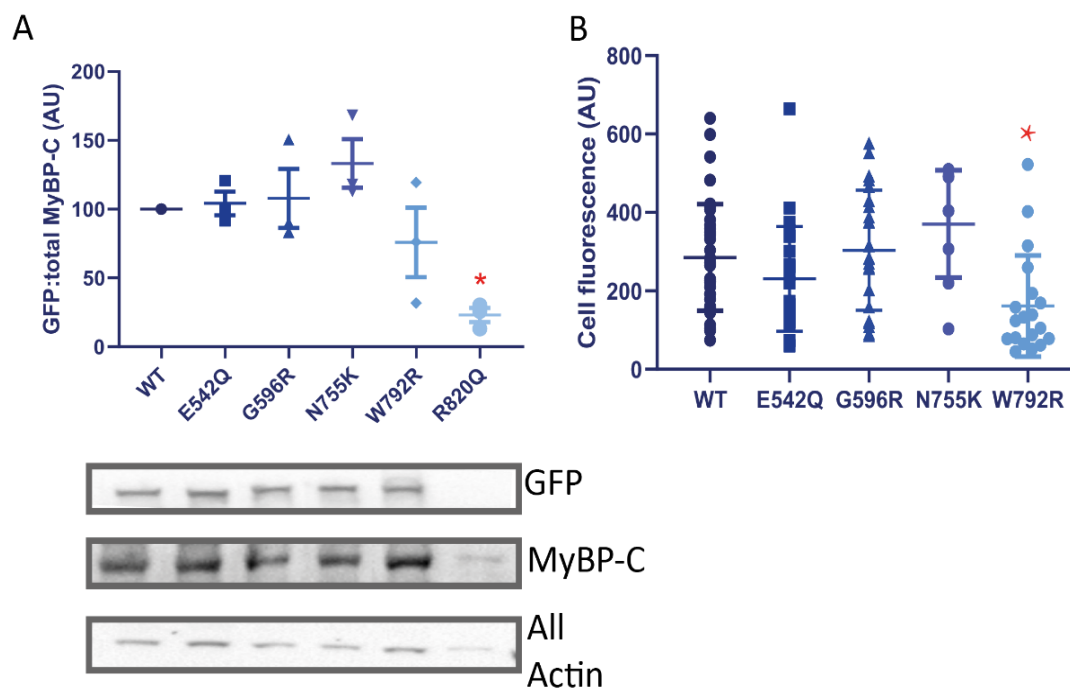


Figure 48– Protein expression levels of mutant full length-cardiac myosin binding protein C adenoviral constructs

Immunoblot and densitometric analysis of protein levels (A) of full-length cardiac myosin binding protein C adenoviral constructs containing hypertrophic cardiomyopathy linked point mutations. Neonatal rat cardiomyocytes (NRCs) were infected with adenoviral constructs for 24 hours, samples were taken and run on an SDS-PAGE gel, transferred onto a nitrocellulose membrane and probed with an anti-GFP, anti-MyBP-C and anti-all actin antibody (used as a loading control). Densitometry was used to determine protein expression levels. The results indicated that there was a significant decrease in protein levels of the R820Q mutant ($p=0.0042$), with a slight trend to a decrease for the W792R mutant. ($n=3$)

Fluorescence intensity data of full-length cardiac myosin binding protein C adenoviral constructs containing hypertrophic cardiomyopathy linked point mutations (B). NRCs were infected with adenoviral constructs for 24 hours, cells were fixed, immunostained and images were taken using a confocal microscopy. Fluorescence intensity levels of GFP were determined using ImageJ software. The results indicated a significant decrease in the fluorescence intensity for the W792R mutant ($p=0.0069$)

These experiments indicate that the E542Q and G596R mutations did not affect the protein expression levels and thus do not appear to cause destabilisation of the protein and lead to

disease through haploinsufficiency but may instead have some other effect on the functionality of the protein.

Western blot analysis showed that the N755K showed a slight increase in protein expression levels, whilst the W792R mutation led to a slight decrease in protein expression. Furthermore, the W792R mutation did lead to a significant decrease in the levels of expression based on fluorescence intensity data from confocal microscopy, which is in corroboration with a study carried out by Smelter et al. which also showed a decreased level of protein expression for the W792R mutant (Smelter et al., 2018).

Finally, the R820Q mutant did not appear to express, as was seen in the experiments utilising the GFP-tagged plasmid, implying that the lack of expression seen previously was probably not due to issues with the plasmid itself, but could be due to the mutation causing mRNA or protein instability leading to its degradation.

5.3.5. Effect of HCM-linked mutants on localisation of cMyBP-C

In order to determine whether the HCM linked full length mutant cMyBP-C could affect the localisation of the protein and thus be causing disease through a 'poison polypeptide' mechanism, NRCs were infected with the WT and mutant full-length cMyBP-C AV, fixed and immunostained and then imaged using both confocal and high resolution (STED) microscopy.

5.3.5.1. Effect of HCM-linked mutants on overall cellular localisation of cMyBP-C

The general cellular localisation of the adenoviral WT and mutant full-length cMyBP-C adenoviral constructs was studied utilising confocal microscopy. As can be seen in figures

49-53 using confocal microscopy, all of the constructs apart from R820Q expressed in the NRCs.

The GFP-tagged WT full-length cMyBP-C protein was seen to mainly localise in a striated pattern, although some cells did exhibit either a more diffuse cytosolic localisation of the GFP-tagged protein (Figure 49; teal arrows) or appeared to show overexpression of the protein (Figure 49; red arrows).

The E542Q mutant protein appeared to localise in the same striated pattern as the WT protein, with a few cells exhibiting a more diffuse localisation (Figure 50; teal arrows). The G596R mutant protein also showed a very similar cellular localisation pattern as the WT protein, with striated localisation seen in the majority of cells, although some cells did show evidence of diffuse localisation (Figure 51; teal arrows) or overexpression of the protein (Figure 51; red arrows).

Cells expressing the N755K protein appeared to more often show overexpression of the protein (figure 52; red arrows) or diffuse localisation of the protein (Figure 52; teal arrows) than either the WT protein or the E542Q and G596R proteins.

Lastly, in cells infected with the W792R mutant adenovirus there were a high number of cells which showed the protein localising in a diffuse manner as opposed to the striated pattern seen for the wildtype (figure 53, teal arrows). A few cells also appeared to show overexpression of the protein although this overexpression did not appear to be uniform across the cell but was more apparent in the periphery of the cells (figure 53; red arrows).

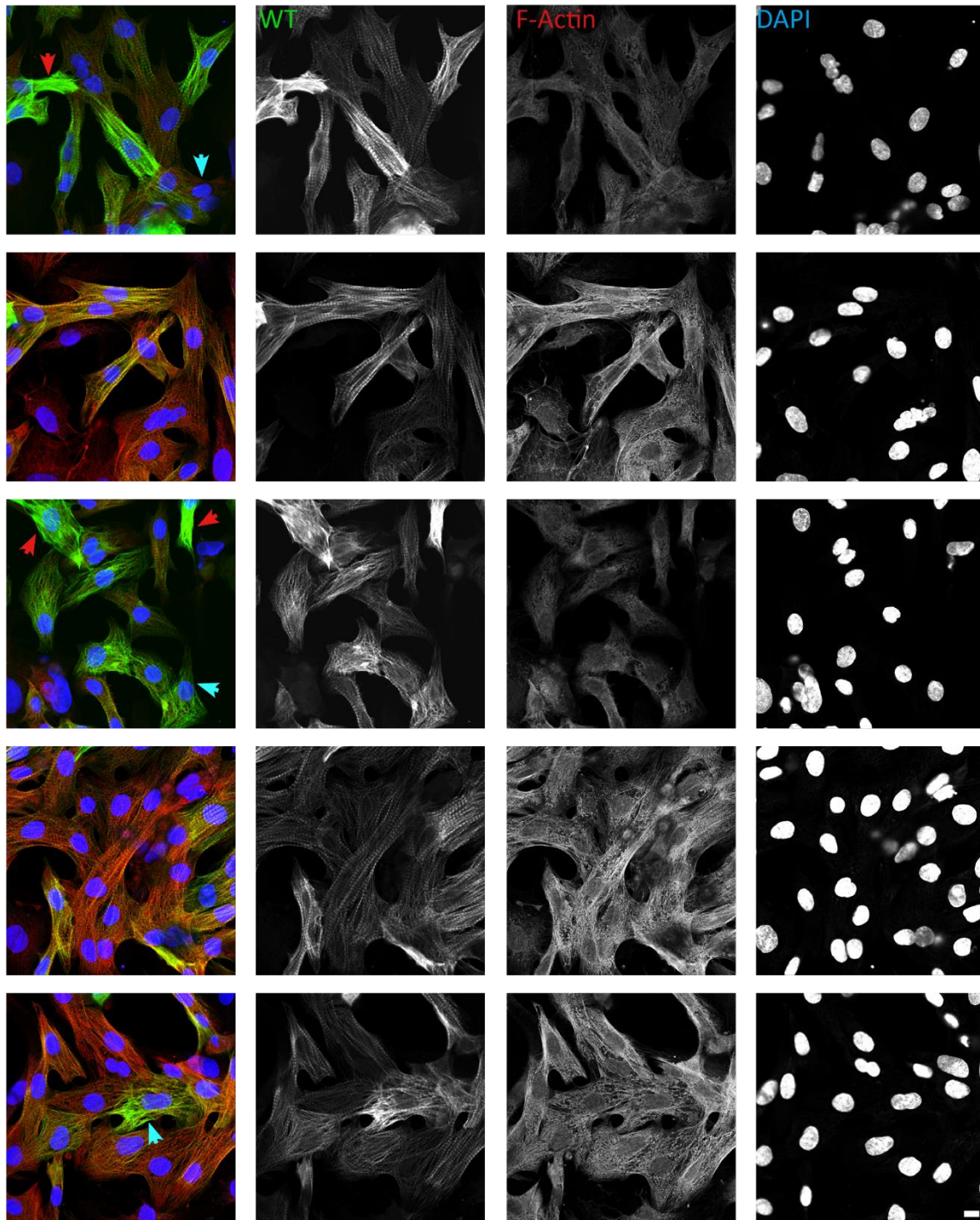


Figure 49 – Localisation of full length wild type cardiac myosin binding protein C in neonatal rat cardiomyocytes

Confocal images of neonatal rat cardiomyocytes (NRC) infected with GFP-tagged full length wildtype cardiac myosin binding protein C (cMyBP-C) adenovirus (AV; green). Cells were immunostained for F-actin (red) and DAPI (blue). The full-length protein was seen to mainly localise in a striated pattern. However, there was some evidence of diffuse expression (cyan arrow) and overexpression of the protein (red arrow) Scale bar = 10 μ m

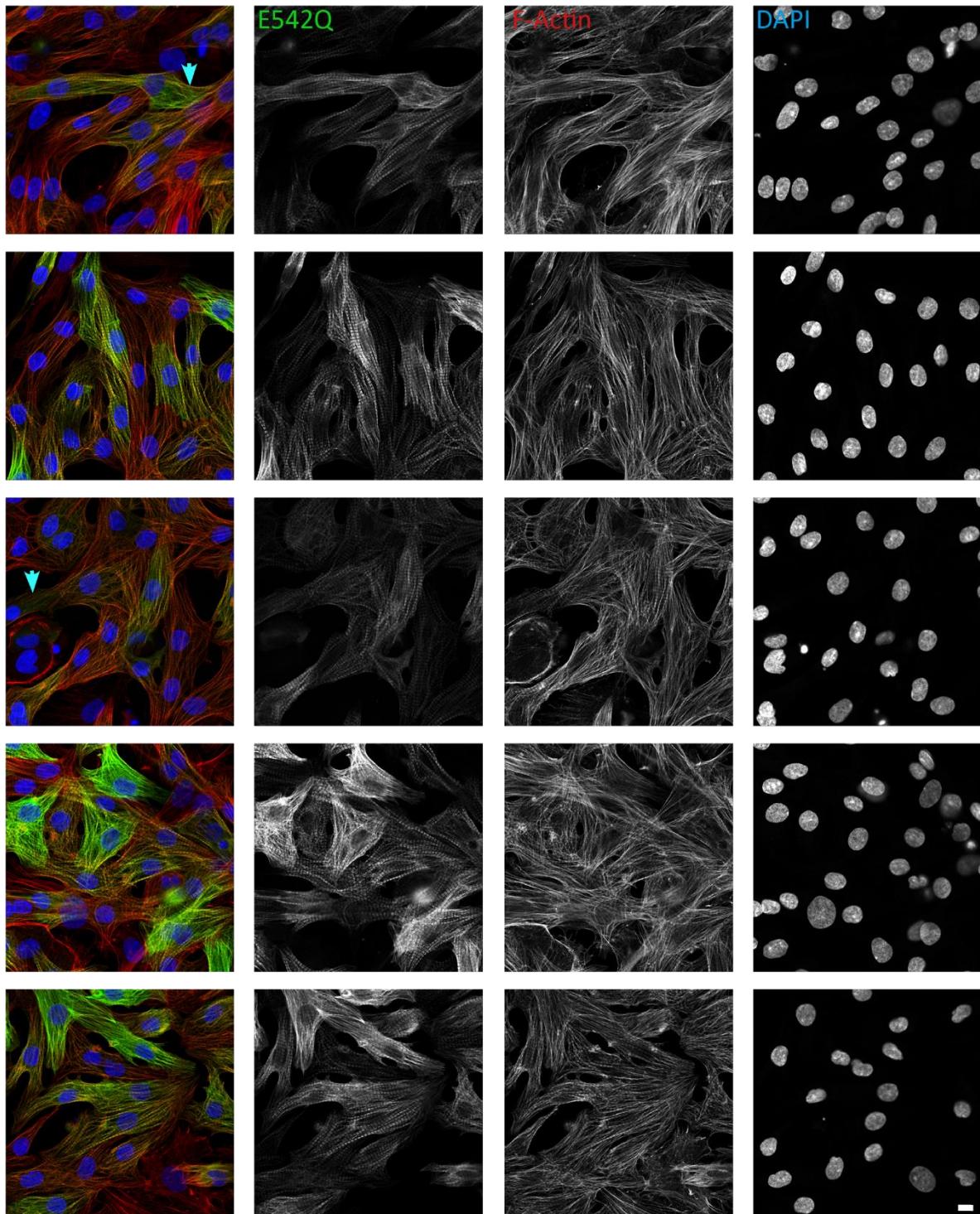


Figure 50 – Localisation of full length E542Q mutant cardiac myosin binding protein C in neonatal rat cardiomyocytes

Confocal images of neonatal rat cardiomyocytes (NRC) infected with GFP-tagged full length cardiac myosin binding protein C (cMyBP-C) adenovirus (AV; green) with the HCM-linked E542Q mutation. Cells were immunostained for F-actin (red) and DAPI (blue). The E542Q mutant protein was seen to show a very similar striated localisation as the wildtype protein. However, some cells show evidence of diffuse localisation (teal arrow) Scale bar = 10 μ m

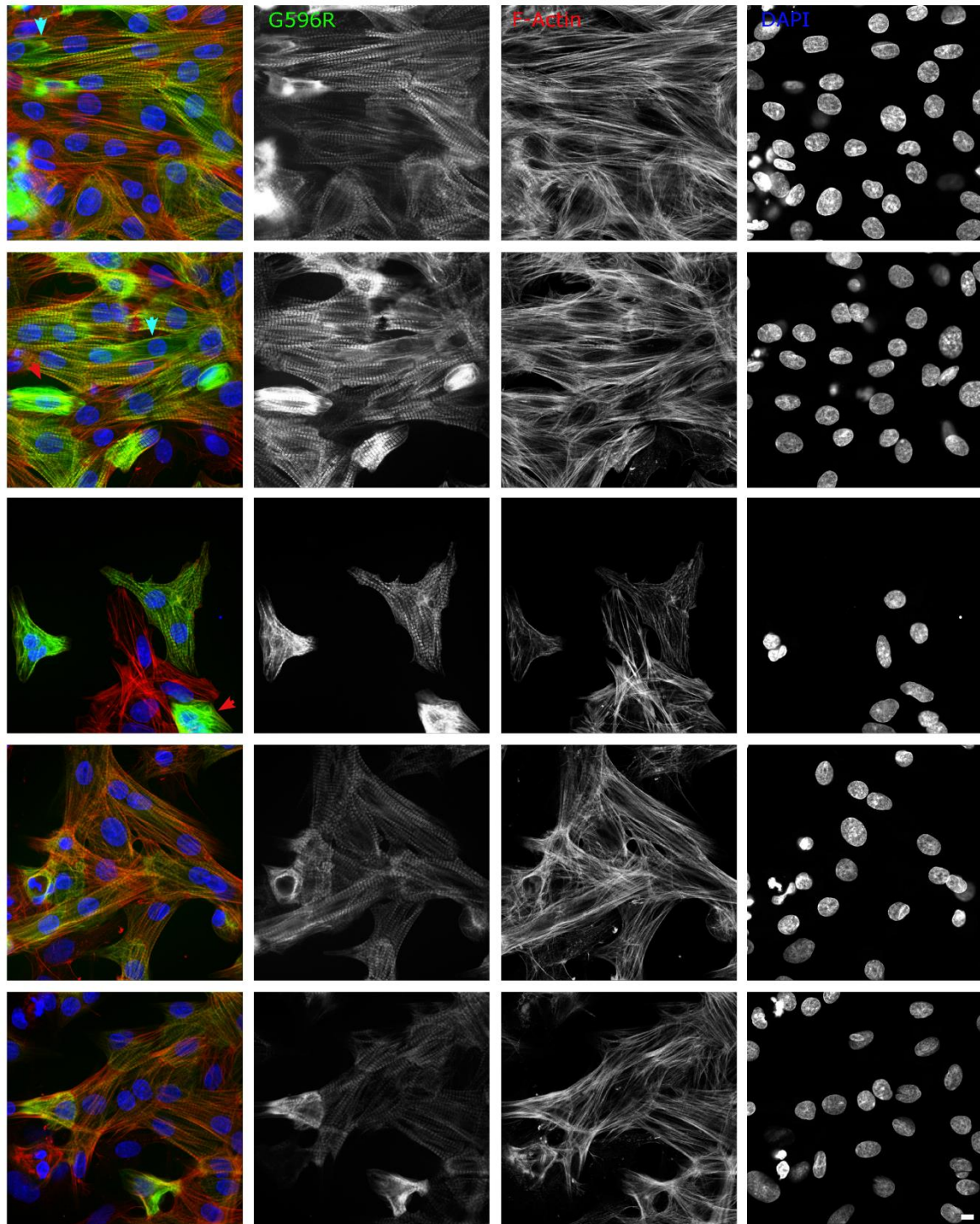


Figure 51 – Localisation of full length G596R mutant cardiac myosin binding protein C in neonatal rat cardiomyocytes

Confocal images of neonatal rat cardiomyocytes (NRC) infected with GFP-tagged full length cardiac myosin binding protein C (cMyBP-C) adenovirus (AV; green) with the HCM-linked G596R mutation. Cells were immunostained for F-actin (red) and DAPI (blue). The G596R mutant protein was seen to show a similar striated localisation as to the wildtype protein with some evidence of diffuse expression (cyan arrow) and overexpression of the protein (red arrow). Scale bar = 10 μ m

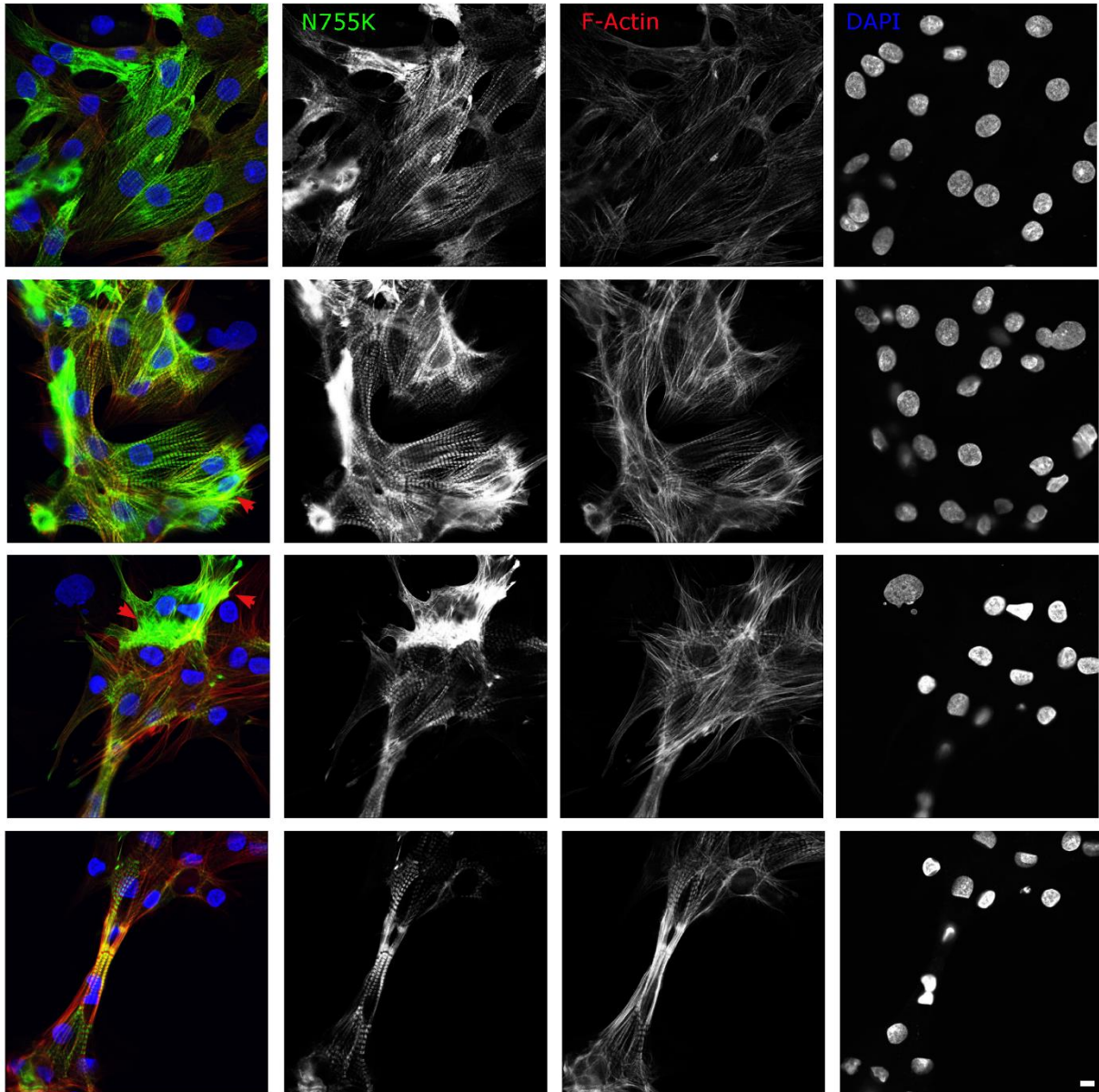


Figure 52 – Localisation of full length N755K mutant cardiac myosin binding protein C in neonatal rat cardiomyocytes

Confocal images of neonatal rat cardiomyocytes (NRC) infected with GFP-tagged full length cardiac myosin binding protein C (cMyBP-C) adenovirus (AV; green) with the HCM-linked N755K mutation. Cells were immunostained for F-actin (red) and DAPI (blue). The N755K mutant protein was seen in some instances to show the same striated localisation as the wildtype protein. However, some cells showed apparent overexpression of the protein (red arrow). Scale bar = 10 μ m.

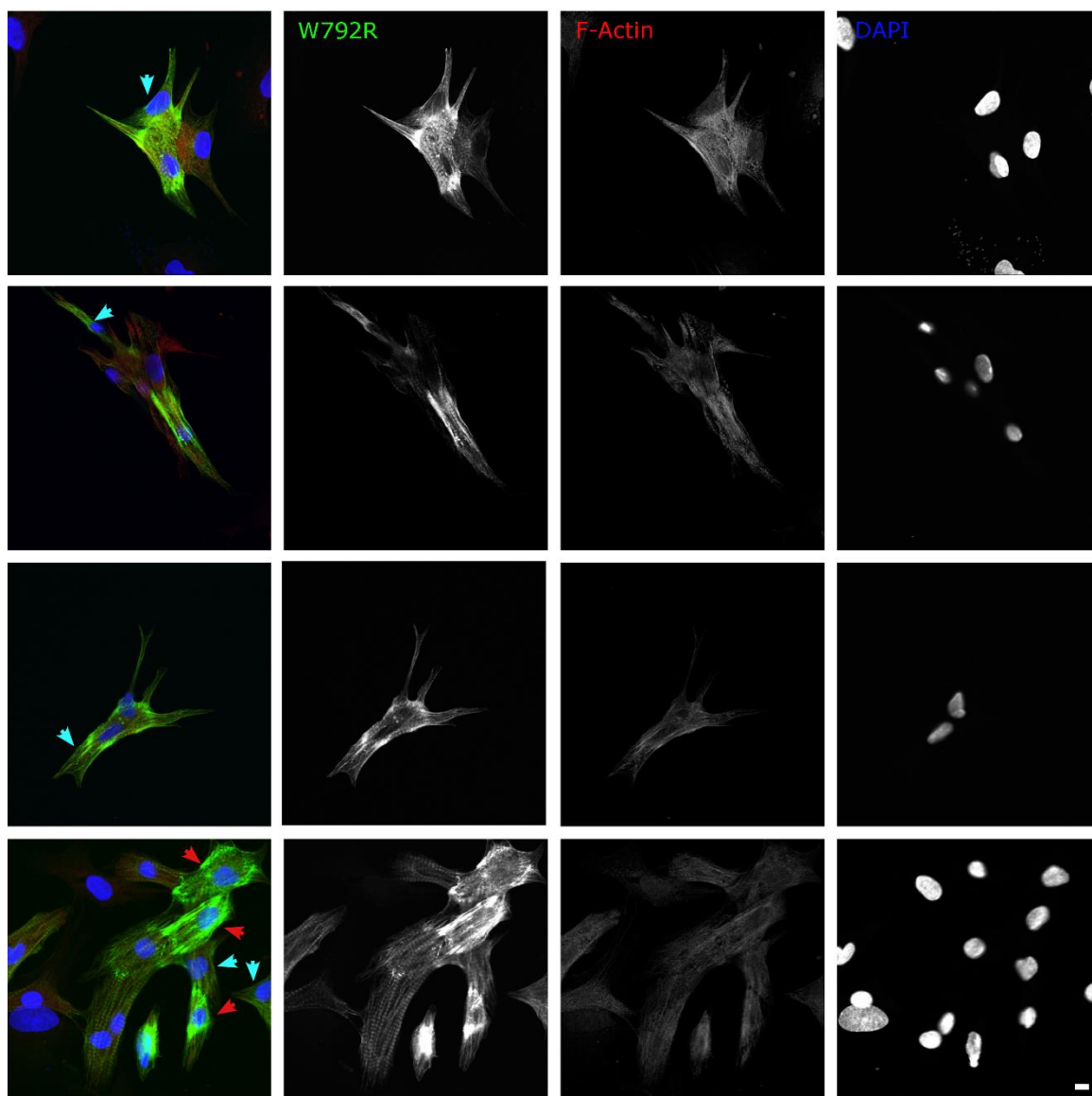


Figure 53 – Localisation of full length W792R mutant cardiac myosin binding protein C in neonatal rat cardiomyocytes

Confocal images of neonatal rat cardiomyocytes (NRC) infected with GFP-tagged full length cardiac myosin binding protein C (cMyBP-C) adenovirus (AV; green) with the HCM-linked W792R mutation. Cells were immunostained for F-actin (red) and DAPI (blue). The W792R mutant protein was seen in some instances to show the same striated localisation as the wildtype protein. However, some cells showed diffuse localisation of the protein. Scale bar = 10 μ m

The percentage of cells that showed overexpression, diffuse expression and normal striation were calculated from the confocal images. The E542Q and G596R mutant full-length adenoviral infected cells showed a similar localisation pattern to the WT, with very few cells exhibiting either diffuse expression or overexpression of the protein. The N755K mutation, however, led a trend to an increase in the number of cells overexpressing the protein, whilst the W792R mutant had a significant increase in the number of cells with a diffuse expression of the protein (Figure 54; $p < 0.0001$; one-way ANOVA; $n = 4-8$).

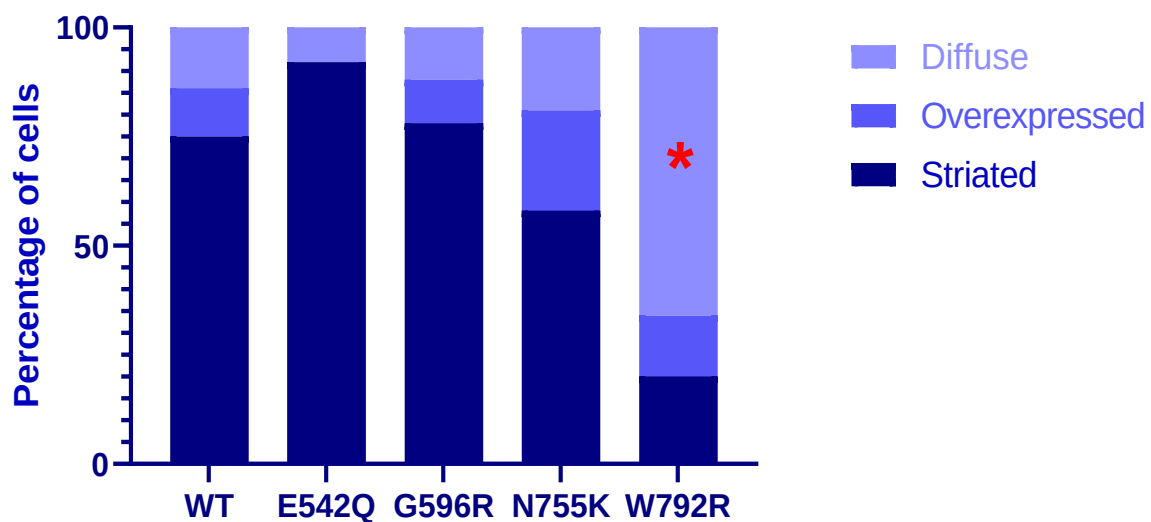


Figure 54 – Analysis of localisation of full length cardiac myosin binding protein C mutant proteins

Graph showing the expression patterns for each of the mutant full-length cardiac myosin binding C (cMyBP-C) adenoviral constructs in neonatal rat cardiomyocytes (NRC). Confocal images of NRC infected with wildtype and mutant full-length cMyBP-C adenovirus were analysed to determine the localisation of the protein. The protein either showed a striated localisation, indicating correct incorporation into the sarcomere, overexpression or diffuse localisation. The percentage of cells showing each of the three localisation patterns was calculated. The E542Q and G596R mutant proteins showed a similar localisation pattern as the wildtype protein. The N755K mutant protein however showed a trend to an increase in the percentage of cells showing an overexpression of the protein. The W792R mutant protein showed a significant increase in the percentage of cells which showed diffuse protein expression ($p = < 0.0001$; one-way ANOVA; $n = 4-8$)

5.3.5.2. Effect of HCM-linked mutants on localisation of cMyBP-C at the myofibrillar level

STED microscopy was utilised in order to determine whether the mutations were having any effect at the level of the myofibrils. NRCs were infected with the GFP-tagged full-length cMyBP-C adenoviral vectors (WT and mutants) for 24 hours. Cells were fixed and then immunostained using an anti-GFP nanobody as well as an antibody against the Z-disc protein α -actinin and were imaged using a STED microscope.

The STED microscopy showed that in the myofibrils of the NRC the wildtype GFP-tagged full-length protein was seen to localise in an expected well defined doublet pattern (Figure 55; magenta) in between the alpha-actinin stripes which denote the Z-discs of the sarcomere (Figure 55; green).

The GFP-tagged mutant full-length adenoviral constructs E542Q, G596R and N755K generally also showed a well-defined doublet pattern similar to that seen for WT, as seen in figures 56-58.

However, it was much more difficult to find cells that were expressing the GFP-tagged full length cMyBP-C W792R containing mutant adenovirus in the same striated doublet pattern as the WT or that of the other mutants. Even when the protein was seen to localise in a doublet pattern, these doublets appeared less distinct than those seen in the cells expressing the WT protein (Figure 59).

The adenoviral vector with the full-length cMyBP-C containing the R820Q mutation again did not show any expression as can be seen in figure 60A, there was however staining for α -actinin suggesting that this was not due to a technical fault during immunostaining or imaging. However, following a longer incubation time of 7 days, one cell was seen to express

the GFP-tagged protein as can be seen in figure 60B. However, no α -actinin staining was seen for this sample, although this could have been due to technical issues either during immunostaining or imaging.

This prolonged culture was also attempted for the other AV constructs, however after 7 days the GFP-tagged protein was overexpressed making visualisation of the myofibrils by STED impossible.

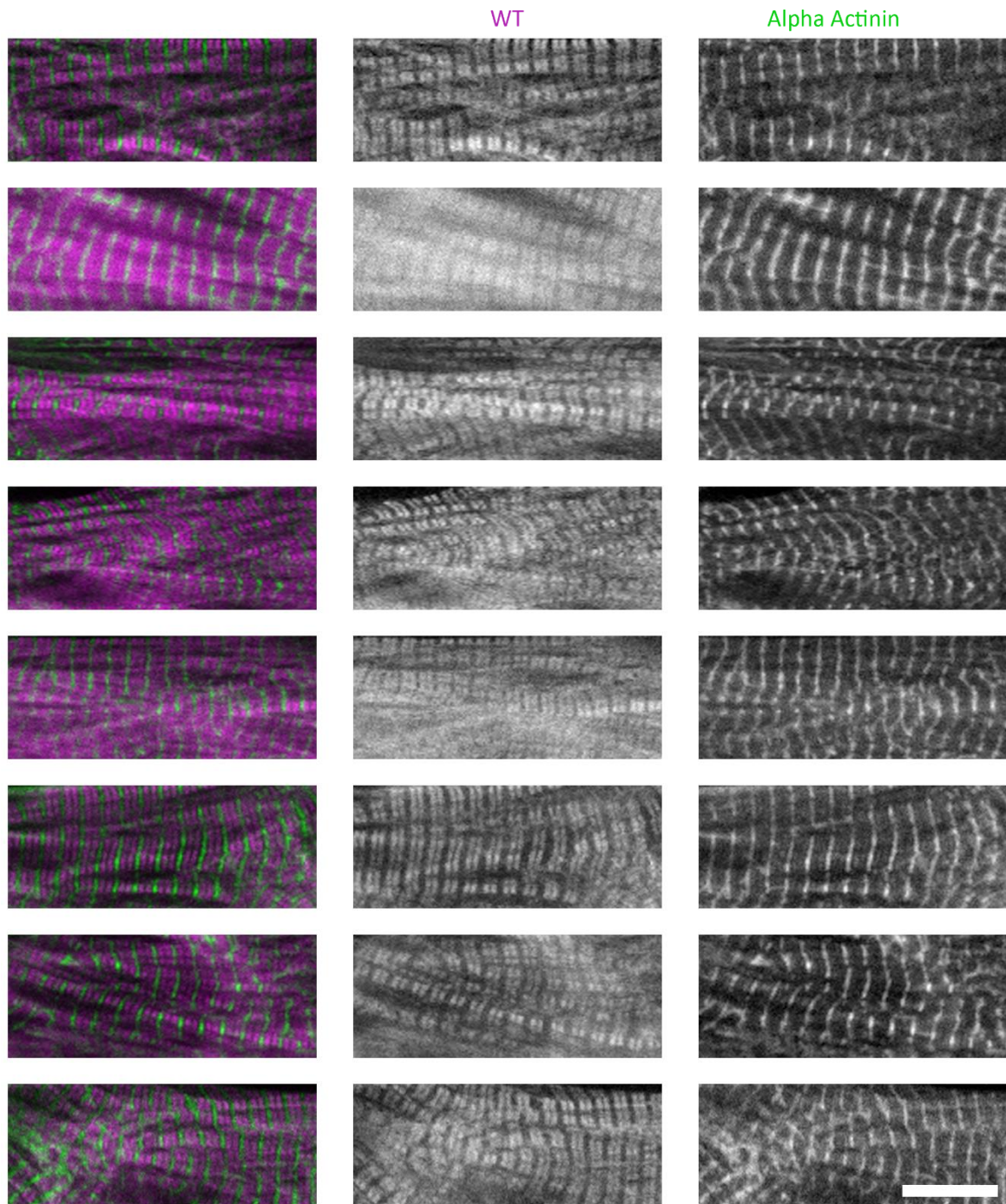


Figure 55 – Myofibrillar localisation of the wildtype full-length cMyBP-C
STED microscopy images of neonatal rat cardiomyocytes infected with the GFP-tagged wildtype full-length cMyBP-C adenovirus. Cells were immunostained with an anti-GFP nanobody (purple) and an anti- α actinin antibody (green). The wild-type protein was seen to localise in an expected doublet pattern. Scale bar = 10 μ m

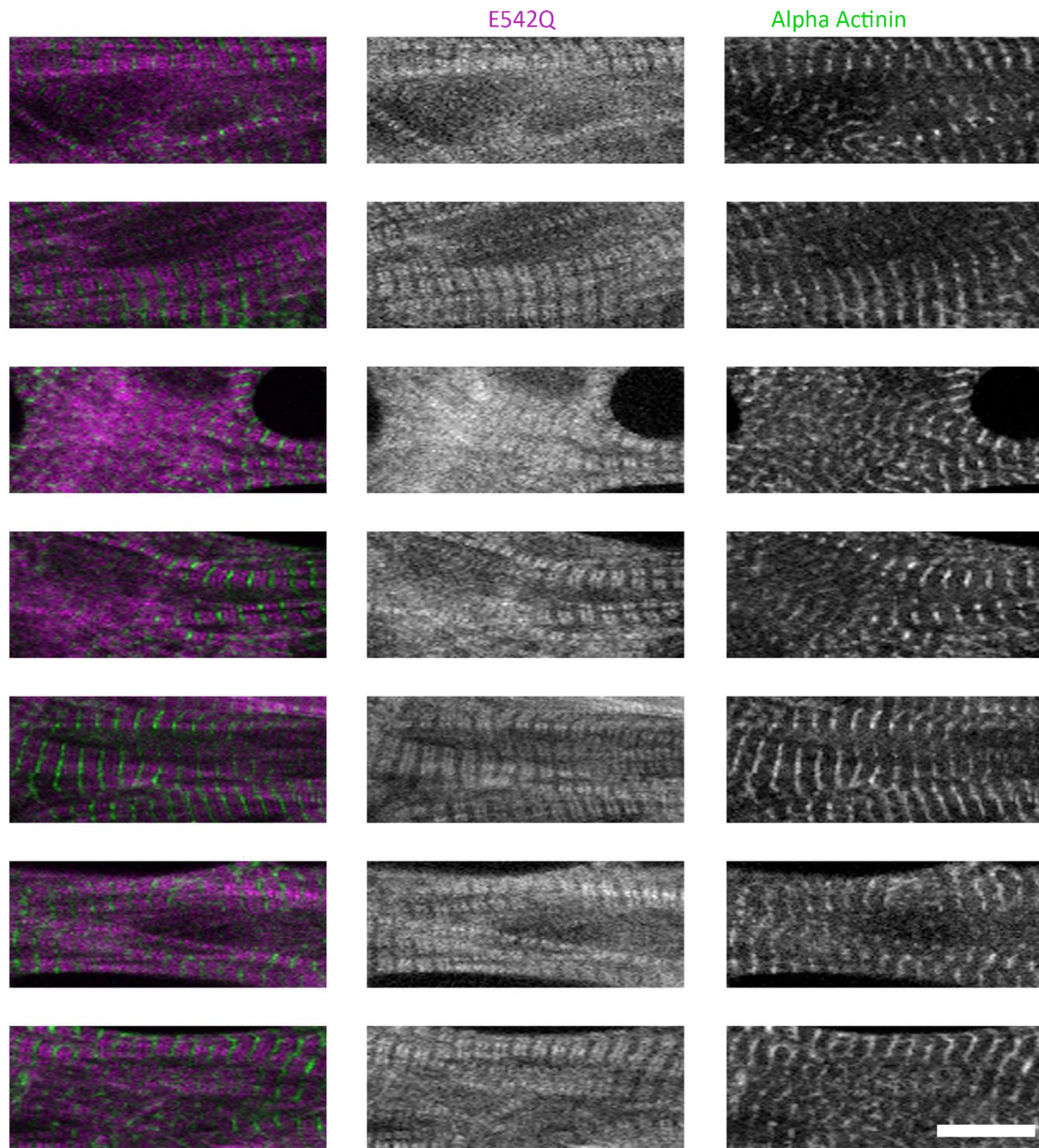


Figure 56 – Myofibrillar localisation of the E542Q mutant containing full-length cMyBP-C

STED microscopy images of neonatal rat cardiomyocytes infected with the GFP-tagged E542Q mutant full-length cMyBP-C adenovirus. Cells were immunostained with an anti-GFP nanobody (purple) and an anti-alpha actinin antibody (green). The E542Q mutant protein was seen to localise in a similar doublet pattern to the wildtype protein. Scale bar = 10 μ m

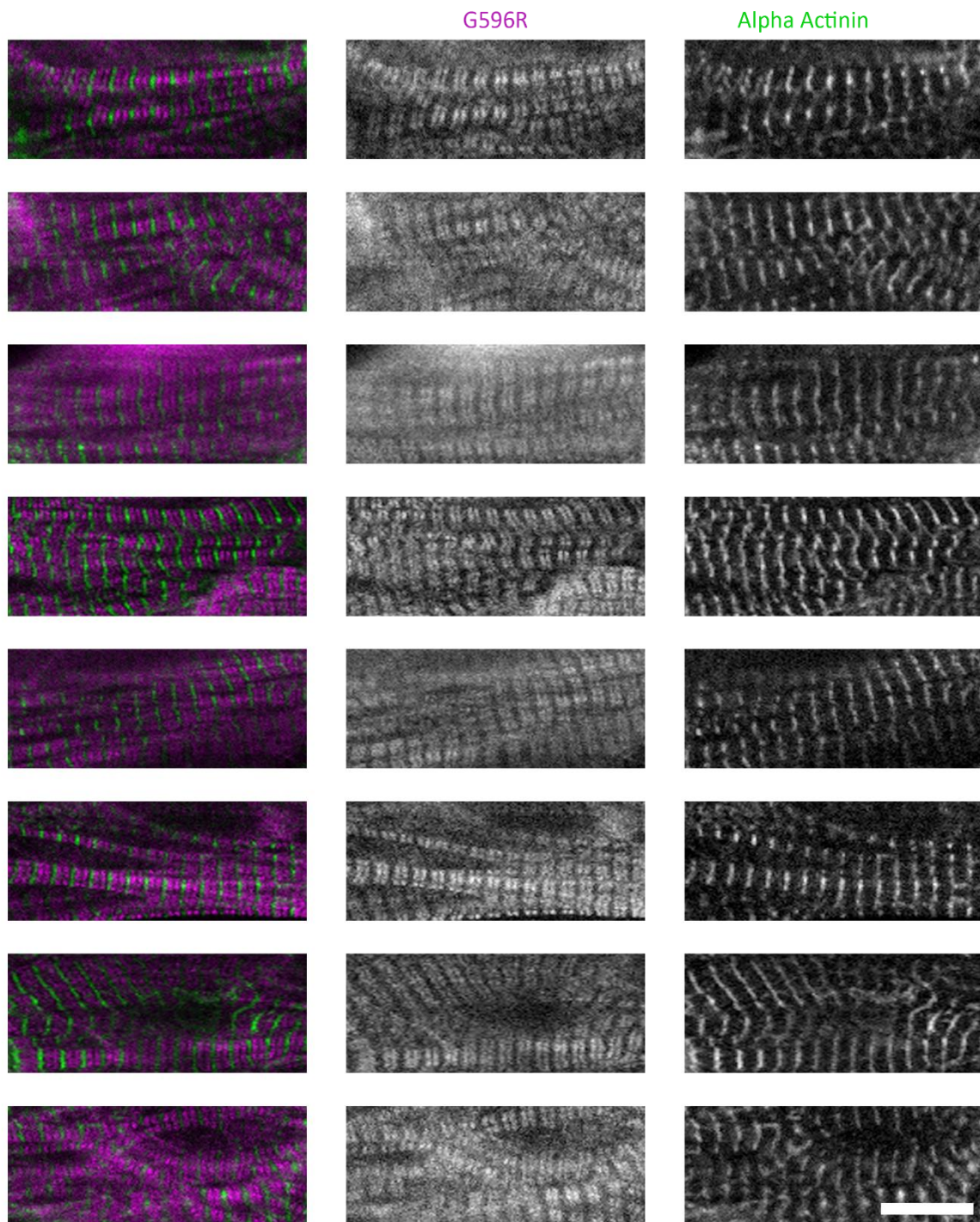


Figure 57 – Myofibrillar localisation of the G596R mutant containing full-length cMyBP-C

STED microscopy images of neonatal rat cardiomyocytes infected with the GFP-tagged G596R mutant full-length cMyBP-C adenovirus. Cells were immunostained with an anti-GFP nanobody (purple) and an anti-alpha actinin antibody (green). The G596R mutant protein was seen to localise in a similar doublet pattern to the wildtype protein. Scale bar = 10 μ m

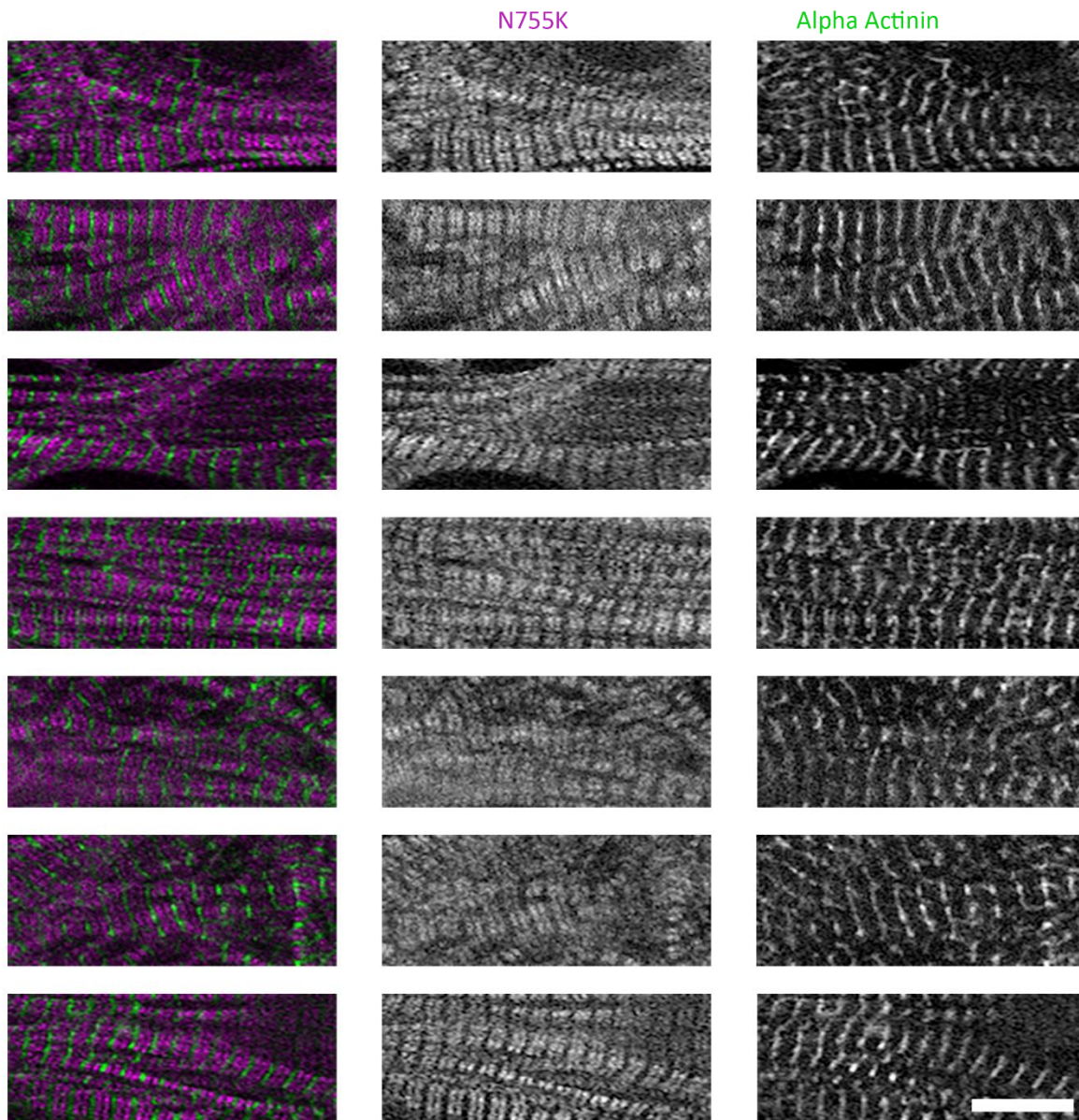


Figure 58 – Myofibrillar localisation of the N755K mutant containing full-length cMyBP-C

STED microscopy images of neonatal rat cardiomyocytes infected with the GFP-tagged N755K full-length cMyBP-C adenovirus. Cells were immunostained with an anti-GFP nanobody (purple) and an anti-alpha actinin antibody (green). The N755K mutant protein was seen to localise in a similar doublet pattern to the wildtype protein. Scale bar = 10 μ m

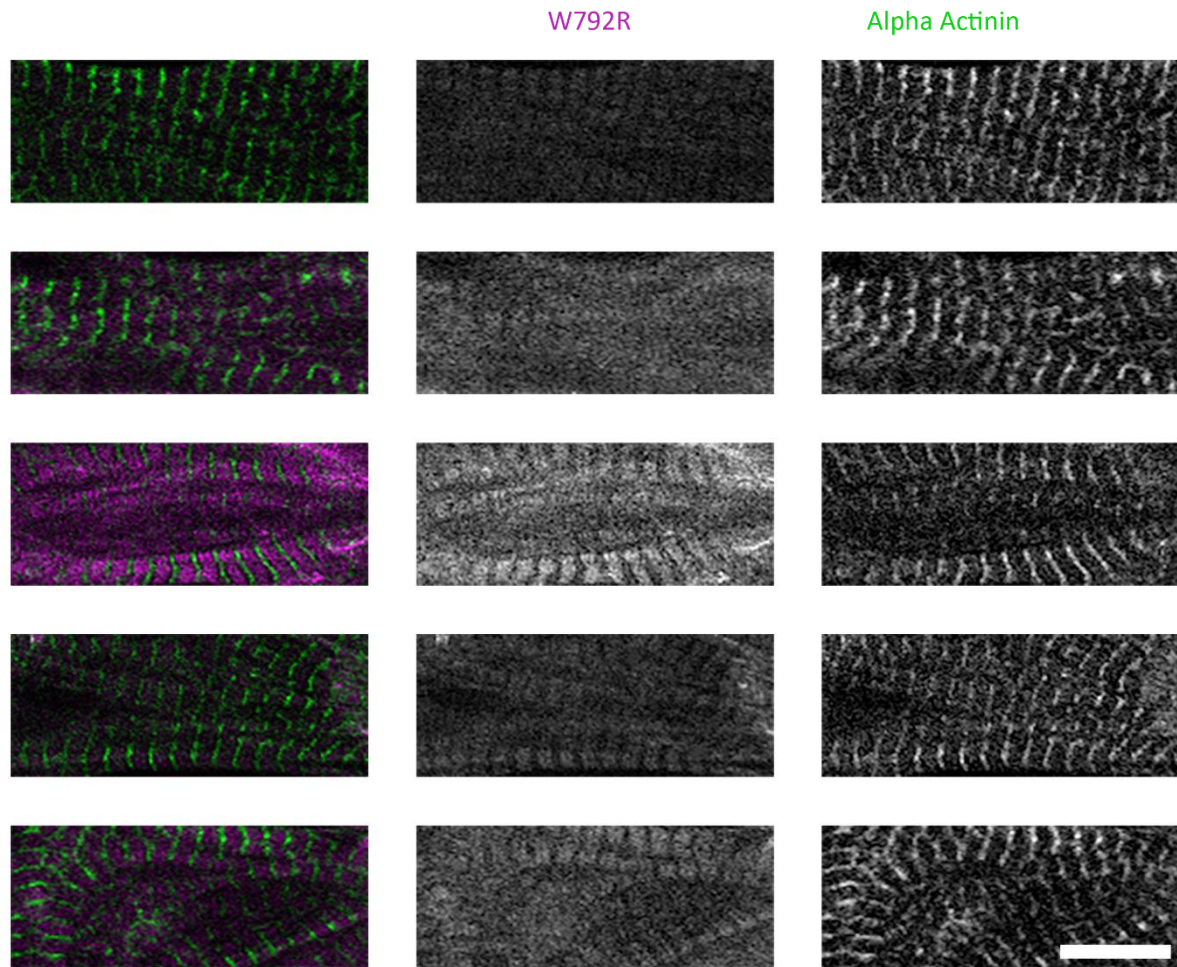


Figure 59 – Myofibrillar localisation of the W792R mutant containing full-length cMyBP-C

STED microscopy images of neonatal rat cardiomyocytes infected with the GFP-tagged W792R mutant full-length cMyBP-C adenovirus. Cells were immunostained with an anti-GFP nanobody (purple) and an anti-alpha actinin antibody (green). Far fewer cells were seen to show the doublet pattern seen with the wildtype protein and those that did have the doublet pattern were less distinct than those seen in cells expressing wildtype protein. Scale bar = 10 μ m

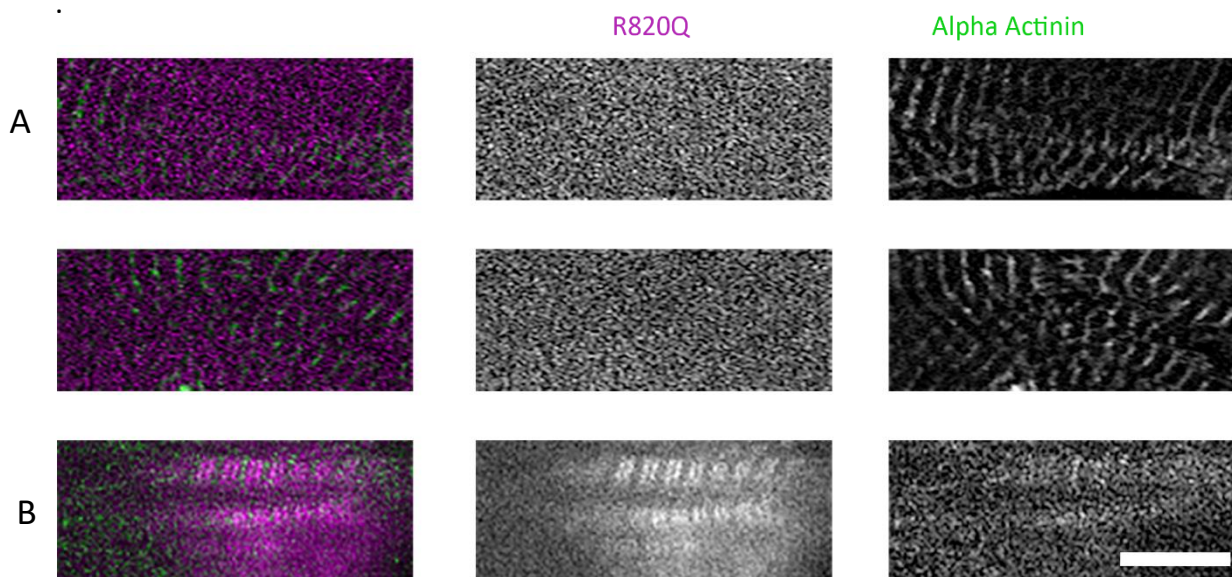


Figure 60– Myofibrillar localisation of the R820Q mutant containing full-length cMyBP-C

STED microscopy images of neonatal rat cardiomyocytes infected with the GFP-tagged R820Q mutant full-length cMyBP-C adenovirus. Cells were immunostained with an anti-GFP nanobody (purple) and an anti-alpha actinin antibody (green). The R820Q mutant protein was not seen to express in the NRCs (A). However, following a longer incubation of 7 days, one cell was seen to show the protein localising in a doublet pattern (B). Scale bar = 10 μ m

Analysis of these images, which was carried out as described in section 2.7.2.2., allowed further evaluation of the possible effects these mutant constructs could be having on the myofibrils. The analysis resulted in intensity profile plots, which are shown in figures 61 and 62, with raw data for each individual image shown on the left whilst the plot of the averaged data is shown on the right (for averaged data the blue line represents mean value, red indicates standard deviation; $n = 2-3$). A major peak at ~ 2 with a smaller secondary peak at ~ 4 denotes a sarcomere length at the expected 2 μ m (green arrows). Two additional peaks positioned before each of the major peaks denote MyBP-C at the expected spacing (purple arrows). Strong, narrow peaks seen at each of these positions would indicate well organised

sarcomeres. From these plots the sarcomere length could be extrapolated, as well as the intensity of the major sarcomeric peak and the major MyBP-C peak, graphs of which can be seen in figure 63.

As can be seen in figure 61, the plot for images of NI cells showed only well defined sarcomeric peaks, as would be expected as there should be no signal denoting the MyBP-C stripes for the NI cells, only alpha-actinin signal. Although there were some peaks in between the major sarcomeric peaks they did not appear to have the correct positioning to correspond to the MyBP-C stripes and were therefore more likely to be either shadows of the sarcomeric peak or background due to non-specific antibody binding.

The plot for images of cells expressing WT cMyBP-C protein (Figure 61) showed strong sarcomere peaks with slightly merged but still relatively well defined peaks for MyBP-C, indicating well organised sarcomeres.

The plots from images of the E542Q, G596R (Figure 61) and N755K (Figure 62) mutant protein expressing cells, appeared to be very similar to that of the WT although the peaks, particularly those signifying MyBP-C, were slightly less defined. This suggests these mutations are having no or very little effect on the sarcomeric organisation.

The plot from images of W792R mutant protein expressing cells (Figure 62) showed peaks which were far less prominent and well defined than those of the WT, suggesting that this mutation may lead to less well organised sarcomeres.

Lastly, the plot from images of cells expressing the R820Q mutant protein (Figure 62) showed either absence of expected peaks or broader and less well defined peaks. This would suggest that this mutation leads to a high level of myofibrillar disarray.

From these plots the sarcomere length as well as the signal intensity of the sarcomere and MyBP-C peaks were extrapolated, giving a measure of the level of organisation of the sarcomeres, which are shown in figure 63. Standard deviation was also plotted to show the level of variability between different cells.

As can be seen in figure 63A there was no difference in the sarcomere length with very little variability as shown in figure 63B, showing that the mutations did not appear to have an effect on the general sarcomeric spacing.

The signal intensity for the sarcomere peak also appeared similar (Figure 63C), again suggesting good sarcomeric organisation. However, there was a higher degree of variation for the W792R and R820Q mutant protein expressing cells as shown by a higher standard deviation (Figure 63D), suggesting higher levels of variability in organisation caused by these mutations.

Lastly, the signal intensity for the MyBP-C peak was shown to be much lower for the N755K, W792R and R820Q mutant protein expressing cells compared to WT (Figure 63E), with the standard deviation for all of these being very low suggesting (Figure 63F) this decrease in signal intensity was consistent. This suggests that these mutations lead to possible myofibrillar disarray and a lack of correct organisation.

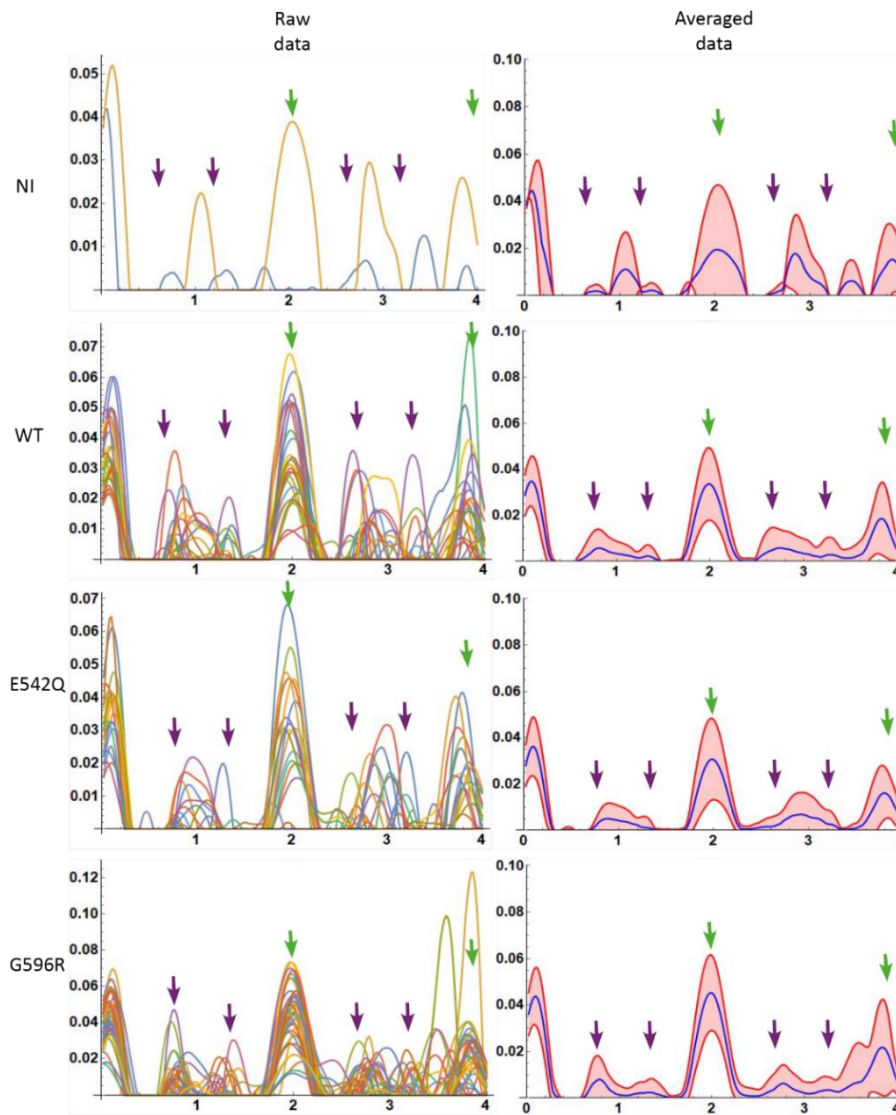


Figure 61 - Intensity profile plots from STED images of NRCs expressing mutant full-length cMyBP-C

Figure showing intensity profile plot analysis of STED images of neonatal rat cardiomyocytes (NRCs) that had not been infected (NI) or expressing GFP-tagged wildtype (WT) or E542Q or G596R mutant constructs. Plots for the raw data is shown on the left, with averaged data of the images shown on the right (for averaged data blue line denotes mean intensity, red indicates standard deviation). A major peak at ~ 2 with a smaller secondary peak at ~ 4 denotes the sarcomere spacing, with 2 peaks prior to each major peak denoting the spacing of the MyBP-C stripes (purple arrows). The plot for NI cells showed only well defined sarcomeric peaks, as would be expected, although there were some peaks in between these did not correspond with the correct spacing to be MyBP-C stripes but is more likely to be either shadows of the sarcomeric peaks or due to non-specific antibody binding. The plot for cells expressing the WT protein showed relatively strong and well defined peaks. Both of the plots for the E542Q and G596R mutant protein expressing cells also showed tracings similar to that of WT suggesting these mutations are not having a negative impact on sarcomeric organisation. (n=2-23)

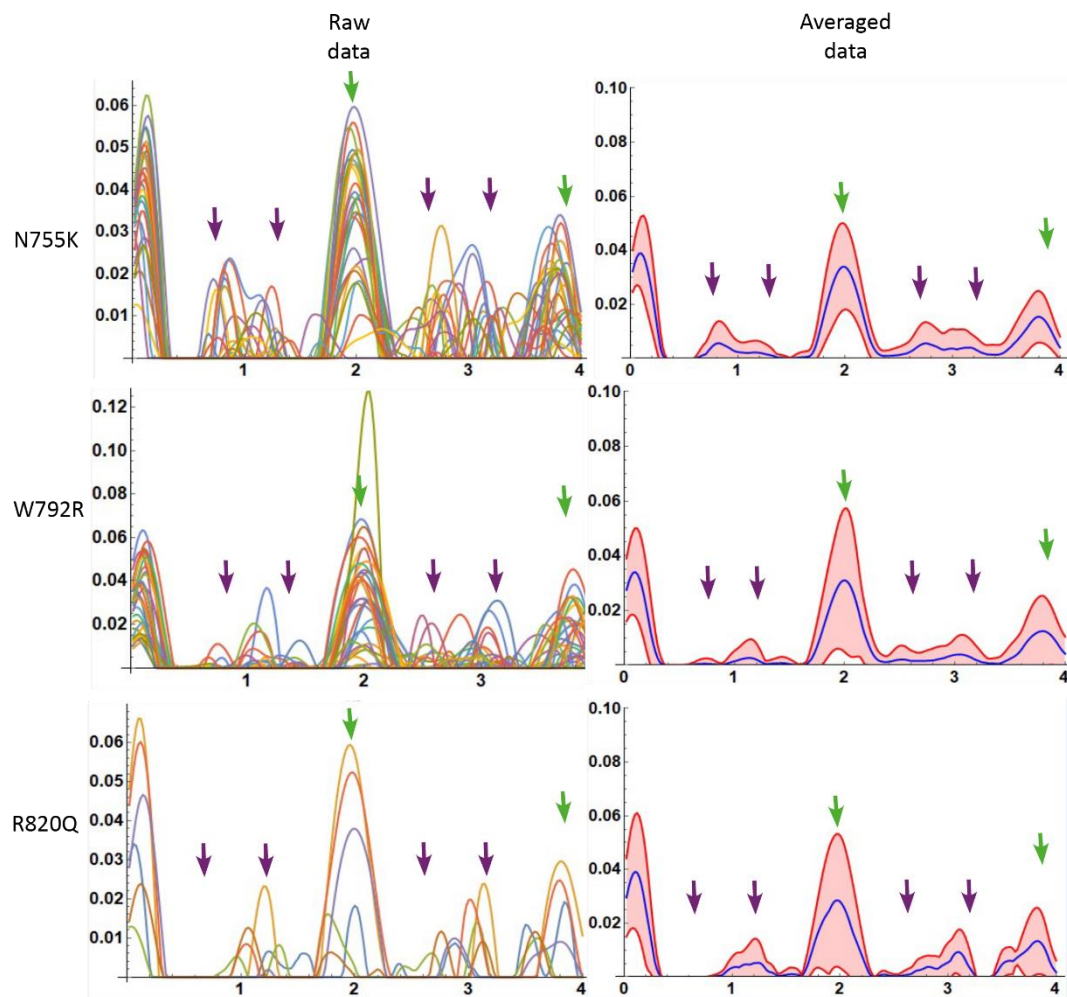


Figure 62 - Intensity profile plots from STED images of NRCs expressing mutant full-length cMyBP-C

Figure showing intensity profile plot analysis of STED images of neonatal rat cardiomyocytes (NRCs) expressing the N755K, W792R and R820Q mutant constructs. Plots for the raw data is shown on the left, with averaged data of the images shown on the right (for averaged data blue line denotes mean intensity, red indicates standard deviation). A major peak at ~ 2 with a smaller secondary peak at ~ 4 denotes the sarcomere spacing, with 2 peaks prior to each major peak denoting the spacing of the MyBP-C stripes (purple arrows). The plot for the N755K mutant protein expressing cells shows relatively strong and well defined peaks. The plot for cells expressing the W792R mutant protein showed less prominent and less well defined peaks whilst the R820Q mutant protein expressing cells showed a plot with either absent peaks or broader and less well defined peaks. This suggests that the W792R and R820Q mutations could be leading to disorganisation of the sarcomeres. (n= 4-18)

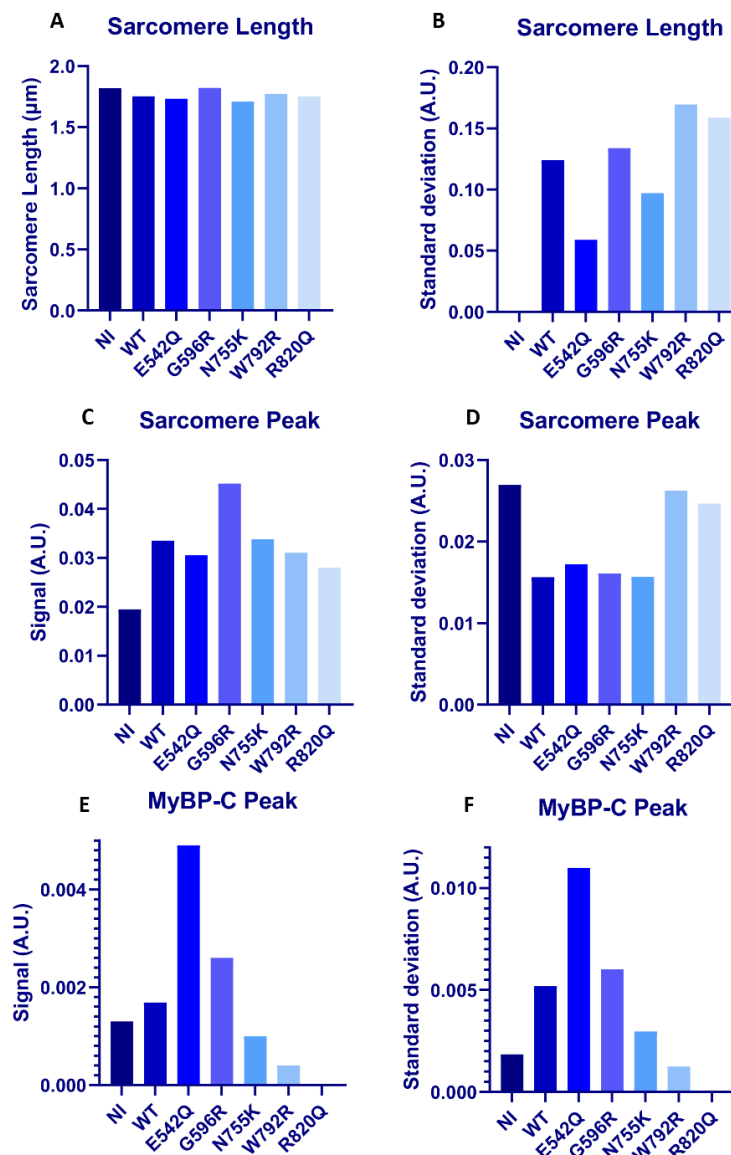


Figure 63 – Sarcomeric organisation of NRCs expressing mutant full length cMyBP-C

Graphs showing various measures (sarcomere length, sarcomere peak intensity signal and MyBP-C peak signal intensity) indicating levels of sarcomeric organisation extrapolated from intensity profile plots generated from analysis of STED images. There was seen to be very little difference in sarcomere length (A) with the standard deviation also remaining relatively similar (B) signalling a low level of variability suggesting none of the mutations had an impact on the general sarcomere spacing. The signal intensity for the sarcomere peak was also very similar (C) although there was a slight increase in standard deviation for the W792R and R820Q mutant expressing cells, suggesting a higher level of variability in sarcomere organisation for these mutants. The signal intensity for the MyBP-C peak was seen to be lower for the N755K, W792R and R820Q mutant expressing cells, with the standard deviation also being low suggesting this decrease in intensity was consistent, which may indicate that these mutations could lead to sarcomeric disorganisation. N= 2-23

5.4. Summary

Based on the observations from these experiments it seems clear that HCM-linked mutations of cMyBP-C may lead to disease via multiple different mechanisms.

The use of a GFP-tagged plasmid containing the C3C6 domains with the E542Q mutation showed no significant difference in the protein expression levels compared with WT, corroborated by the experiments using the adenovirus expressing the full length cMyBP-C with this mutation, and localisation was also seen to be the same as the WT plasmid. Furthermore, analysis of STED images showed no difference in sarcomeric organisation when compared with WT. There was, however, some evidence that the E542Q mutation decreased the co-localisation with actin seen for the WT C3C6 containing plasmid. This suggests that whilst the mutation may not have any effect on the stability of the protein its localisation or organisation of the sarcomeres, in agreement with the previous literature, it could interfere with the possible interaction between the central domains and actin.

In the experiments utilising the GFP-tagged plasmid encoding C3C6 with the G596R mutation, it was seen that there was no difference in protein expression levels, which was corroborated by the experiments utilising the full-length adenoviral vector containing this mutation. Analysis of STED images also showed that sarcomeric organisation was generally similar to WT suggesting that this mutation does not lead to disarray of the sarcomeres. However, it did appear to have a trend to an increase in nuclear localisation suggesting the mutation could be causing mislocalisation of the protein, as it is possible that if the protein cannot bind to its correct binding partner this could lead to the protein being mislocalised. It was also seen that the G596R mutation lead to a significant decrease in the possible interaction between the central domains of cMyBP-C and actin. Therefore, as with the

E542Q mutation, the G596R mutation may cause HCM through interfering with possible interaction partners of cMyBP-C.

The N755K mutation appeared to cause overexpression of the protein, with western blot of mutant full length adenovirus showing a slight increase in protein expression levels, whilst analysis of confocal images showed a trend to an increase in the percentage of cells that were overexpressing the GFP-tagged protein. It is possible that overexpression of the protein could lead to a potential increase in binding to interaction partners which could cause aberration of the functioning of the protein. Confocal microscopy indicated that the GFP tagged C3C6 mutant plasmid showed a localisation similar to WT and did not show any increase in nuclear localisation or any difference in co-localisation with F-actin compared to WT. STED microscopy of cells expressing the GFP-tagged C3C6 domain containing plasmid however did show evidence of myofibrillar disarray, which seemed to be exacerbated following longer culturing, which appeared to be corroborated by further analysis of the STED images. This could link with the theory put forward by Idowu et al. that the C5 domain with the N755K mutation may be stabilised by surrounding domains that could become destabilised following mechanical stress (Idowu et al., 2003).

The W792R mutant adenoviral construct showed a decrease in protein expression levels shown via western blot analysis, whilst fluorescence intensity data showed a significant decrease in expression levels, corroborating the results shown by Smelter et al. which showed a decrease in protein expression levels of about 70% (Smelter et al., 2018). STED imaging of myofibrils of cells expressing the GFP-tagged W792R mutant containing C3C6 plasmid appeared to show myofibrillar disarray which was worsened following prolonged culture, as was seen via analysis of STED images. Furthermore, the full-length adenoviral

vector containing this mutant showed an increase in diffuse expression of the protein seen both via conventional confocal microscopy and at the myofibrillar level utilising STED microscopy, suggesting that this mutation could cause a lack of incorporation into the sarcomere and protein degradation.

The mechanism by which the R820Q mutation could be causing disease appears to be via protein degradation or mRNA instability. No signal for the GFP-tagged C3C6 domain containing R820Q mutant was seen following infection of COS-1 cells or NRCs, whilst Western blot analysis showed that there was a significant decrease in protein expression levels. This suggested that either the protein was being degraded or there was an issue with the plasmid and/or the GFP tag. This mutation was therefore carried on into the adenoviral full-length cMyBP-C vector, however again there was no expression of the R820Q mutant seen in the HEK-293 cells, COS-1 cells or NRCs. During the propagation stages of the adenovirus the R820Q mutant appeared the same as the other adenoviral constructs and also during the titration experiment colouration was seen after staining for the protein hexon, which is produced on replication of adenovirus, suggesting that the adenovirus was replicating and that there was nothing wrong with the adenovirus. Furthermore, sequencing was carried out before both the GFP-tagged C3C6 domain plasmid and the full-length adenovirus were utilised in the experiments. Further evidence for the construct being correct comes from the anecdotal observation of one NRC being seen to express the R820Q construct and incorporating into doublets in the A-band following a 7-day culture. Analysis of these limited STED images suggested high levels of sarcomeric disarray. Therefore, it seems likely that this mutation leads to instability of the protein leading to its degradation.

These experiments highlight the heterogeneity in the mechanism of disease occurring for the different HCM-linked missense mutations of the central domains of cMyBP-C, and presumably those found in the other regions of the protein. Further experimentation into the mechanisms of pathogenicity of HCM-linked missense mutations should therefore be carried out in order to classify the different pathways and open up the possibility for more individualised treatment for HCM patients dependent on the causative mutation they are carrying.

6. Discussion

The overall aims of these experiments were to gain a better understanding of the structure and function of the central domains of cMyBP-C, and also explore the mechanism by which missense mutations within this region can lead to HCM. The importance of learning more about this understudied region is highlighted by the high number of HCM-linked mutations within these domains which make it likely that this region of the protein does have some kind of structural or functional role. It is however possible that these central domains could just function as a spacer to ensure the N-terminal domains are in their correct localisation and orientation, with the possibility that mutations in this region could have an effect on domain folding which could cause alterations in the protein structure that could affect the ability of the region to act as a spacer. Conversely, it is possible that the high number of mutations that appear to be tolerated within the central domains may indicate that this region is not important for the stability and function of the protein.

6.1. The central domains on their own show a diffuse, cytoplasmic localisation in cells

As was shown in both COS-1 cells and NRCs, the central C3C6 domains of MyBP-C localised in a mostly diffuse cytoplasmic manner, did not appear to mislocalise to the nucleus and did not show aggregation. Previous work has shown that the GFP-tag is prone to nuclear localisation which can lead to mislocalisation of GFP-tagged proteins to the nucleus (Seibel et al., 2007). However, the C3C6 domains did not appear to localise to the nucleus, suggesting that the GFP tag did not cause mislocalisation of the central domains in the wildtype. Furthermore, the fact that there did not appear to be any aggregation of the protein suggests that the protein was stable and was not targeted for degradation. The diffuse cytoplasmic localisation of the GFP-tagged C3C6 domains would be expected as this central region does not contain the regions that have previously been shown to be involved in incorporation of MyBP-C into the thick filament, i.e. the C7-C10 domains which are known to contain binding domains for both myosin and titin (Okagaki et al., 1993, Freiburg and Gautel, 1996, Gilbert et al., 1996).

These results suggest that the central domains are capable of being stably expressed in both types of eukaryotic cells as no evidence of aggregation was seen. Furthermore, this confirms that the central domains do not have a role in the localisation of the protein into the sarcomere, seen by the fact that expression of the protein was in a mostly diffuse, cytoplasmic manner.

6.2. Potential instability of the C3C6 domains

Recombinant expression of the multidomain fragments of the central domains of cMyBP-C in *E. coli* indicated that whilst both showed soluble expression, the longer C3C6 fragment produced a much lower yield than the C3C5 domains, with a higher level of degradation of the product being seen indicating that the C3C6 fragment was less stable than the shorter fragment.

Furthermore, the fact that the degradation product was at the correct molecular weight for a three-domain fragment suggests the possibility that the C6 domain is being cleaved off. This could mean that the linker between the C5 and C6 domains is less stable and requires the C7-C10 domains in order to stabilise it. The linker region between the C5 and C6 domains consists of small amino acids, which typically suggest a higher degree of flexibility (Huang and Nau, 2003), and so could support this theory.

In a study by Nadvi et al., a central domain construct was expressed spanning the C5-C7 domains, which did not show degradation (Nadvi et al., 2016), therefore it is possible that the C7 domain stabilises the link between C5 and C6. Furthermore, SAXS data has suggested that the C5C6 domain showed a significantly greater conformational flexibility than was seen for the comparably sized C3C4 domains, this was suggested to be important to provide a functionally important flexible central region (Jeffries et al., 2011). Although it is also possible that the increased flexibility measured was due to the disordered insert known to be present in the cardiac-specific insert found in the C5 domain (Jeffries et al., 2011).

However, if this data is correct and the C5C6 region does confer a functionally important flexible region this could suggest a reason as to why mutations in these domains may cause

HCM, as if these mutations interfere with the flexibility of the protein they could have a detrimental effect on the function of the protein.

On the other hand, the fact that the central C3C6 domains were shown to express in both COS-1 cells and NRCs with little evidence of aggregation or trafficking to the nucleus for protein degradation suggest that these domains are stable in mammalian cells. This difference could be due to the bacterial cells lacking something, such as muscle specific chaperones or proteins involved in PTMs, that is found in the eukaryotic cells which is required for the correct folding of the central domains and this is the reason for the decreased stability seen for the bacterially expressed protein. For example it has been found that the myosin heads require the chaperone protein Unc45 for correct folding to occur (reviewed in (Hutagalung et al., 2002)), therefore this could also be the case for MyBP-C. Furthermore, whilst much less is known about post-translational modifications of the central domains than the N-terminal region of cMyBP-C, as can be seen in figure 7, multiple putative sites have been suggested in the C3-C6 domains for PTMs. It is possible that PTM of residues in the central region is required for correct domain folding and as prokaryotes are known to have differing PTM capacity from eukaryotes (Sahdev et al., 2008), the PTMs necessary for folding of the protein may not be present in the *E. coli*.

6.3. Y2H assay screens failed to identify putative binding partners

The experiments described in this work utilised a Y2H assay screen in order to identify putative binding partners for the central domains. The Y2H technique is a popular tool used for identifying protein-protein interactions in an unbiased manner as it is relatively easy to carry out and inexpensive, being more accessible than mass spectrometry techniques as it does not require specialist or costly equipment. Mass spectrometry does however have the advantage of being able to determine all of the components of larger complexes whilst Y2H screens are only able to detect binary interactions (Brückner et al., 2009).

The Y2H assay screen does have its flaws, most notably the amount of false negatives and false positives produced, with studies suggesting that somewhere between 43-71% of interactions may be missed due to false negatives (Edwards et al., 2002) whilst It is thought that 44-91% of hits from Y2H screens are false positives (Mrowka et al., 2001).

False negatives can occur for a number of reasons, for example interactions involving membrane associated proteins or proteins localised to different sub-cellular compartments can be missed. Furthermore, the way in which interactors are bound to the bait or prey could cause steric hindrance interfering with any possible binding. The fact that the screen occurs in a yeast system could also mean that there is a lack of components required for PTMs of proteins which may be required for binding to occur (Brückner et al., 2009).

False positives may occur due to inappropriately high expression levels of the bait and prey or the proteins being localised in compartments which do not correspond with their natural cellular environment. Unspecific interactions can also occur with proteins which are not folded correctly, or with some proteins which are known to be “sticky” (Brückner et al., 2009).

The Y2H assay screen utilising bait plasmids containing the C3C6 domains as well as the individual C4 and C5 domains did not lead to the identification of any putative binding partners that could be confirmed.

The fact that a large number of colonies did grow, particularly in the screens utilising the C3C6 and C4 domain bait plasmids, suggests that there were no technical issues whilst carrying out the screen. All three bait plasmids were also shown to express well in the yeast cells, although western blot analysis of the expression levels did show the C4 and C5 domain containing bait constructs had some signs of degradation. Therefore, it is possible that these bait plasmids were not expressing correctly in the yeast which could have affected their ability to interact with putative binding partners within the screen, thus generating false negatives.

This screen was not able to recapitulate results from an earlier Y2H screen in which an interaction between the C5 domain and the C-terminal domains of MyBP-C was identified (Moolman-Smook et al., 2002). This could be due to the fact that a different expression system was used in this study, with the bait construct utilising the Gal4 activating domain being used in those studies whilst the assay in this study used LexA, and it has been shown that the use of different Y2H systems detects different subsets of interactions (Rajagopala et al., 2009, Stellberger et al., 2010).

The fact that no interaction partners were identified by the Y2H assay screens could suggest that the central domains act merely as a spacer. However, whilst the Y2H assay has its benefits, as mentioned above a major flaw of the technique is the high number of false negatives. Therefore, it is possible that the central domains do have interaction partners, but they were not identified by the screen due to the reasons listed above.

6.4. Actin is a potential interaction partner of the central domains

Whilst no putative interaction partners for the central domains of MyBP-C were identified by the Y2H assay screens, analysis of confocal images of NRCs expressing the GFP-tagged C3C6 domains identified partial co-localisation between the central domains and F-actin. This suggests that the central domains could bind to actin, however co-localisation does not necessarily confirm that an interaction is occurring between the two proteins.

The possibility that the central domains could be interacting with actin was further studied by utilising a co-sedimentation assay, in which purified recombinant C3C5 was incubated with NTFs and centrifuged at very high speeds. The recombinant C3C5 was seen to co-sediment with the NTFs suggesting that binding was occurring, however the co-sedimentation was at a much lower level than that seen for the C0C2 domains of cMyBP-C. This intimates that whilst the central domains may be capable of binding to actin, it is with a much lower affinity than the N-terminal domains.

Previous studies have carried out co-sedimentation assays utilising several different recombinant murine cMyBP-C fragments with either F-actin or NTFs. These experiments suggested that the C1 and M-motif were responsible for the binding to actin. However, they showed that the C3C4 domains did not bind to actin, contrary to the results shown in this study (Shaffer et al., 2009). This could mean that the co-sedimentation of the central domains seen was due to interaction between the NTFs and the C5 domain. On the other hand, the previous study utilised murine recombinant C3C4 and actin (Shaffer et al., 2009), whilst in this study human recombinant C3C5 was co-sedimented with bovine NTFs. Therefore, it is possible that these differences could be due to species sequence variations

or the interaction requires the presence of regulatory proteins present when using the NTFs as opposed to actin alone.

It should be noted that these experiments utilised fragments of cMyBP-C and therefore any interaction seen may not actually be able to occur in situ. It is possible that when the central domains are surrounded by the N- and C-terminal domains of the full-length protein and are fixed in their correct stoichiometry of the sarcomere this interaction may not be able to occur.

Previous literature, however, has hypothesised that the central domains could bind to actin, with one study stating that the C5-C10 domains are able to bind actin in a strong and specific manner (Rybakova et al., 2011), although as this is a region which also binds both myosin and titin it could be argued that this finding is unlikely. Another study suggested that the C3 domain may be involved in sensitising the thin filaments and regulating crossbridge cycling kinetics (Inchingolo et al., 2019). Therefore, this possible interaction between the central domains and actin is not completely without supporting evidence.

6.5. The central domains may have a role in the regulation of actomyosin cross-bridge cycling kinetics

Previous studies have ascertained that the N-terminal domains of cMyBP-C have an effect on the actomyosin crossbridge cycling kinetics, as it has been shown that at low concentrations of C0C2 ATP hydrolysis is activated, whilst at high concentrations it is inhibited (Belknap et al., 2014).

In order to determine whether the central domains could also be involved in the regulation of the crossbridge cycling rates, an ATPase assay was carried out using increasing concentrations of recombinantly expressed C3C5 protein. The results from this study indicated that high concentrations of C3C5 lead to a significant increase in the ATPase rate in the presence of calcium, whilst in the absence of calcium there was no change in the rate of ATP hydrolysis. This suggests that the central domains could have a role in the regulation of crossbridge cycling. Furthermore, the fact that this effect was only seen in the presence of calcium would suggest that this apparent involvement of the central domains on the regulatory state of the thin filaments does not occur during diastole.

Prior studies have also highlighted this possibility, with one study showing that the C3 domain may have a role in sensitising the thin filaments (Inchingolo et al., 2019), whilst another showed that removal of the C0-C7 domains and re-addition of just the C0C2 domains was not able to fully reverse cross-bridge cycling kinetics back to normal (Napierski et al., 2020). However, this could also be explained by the hypothesis that the central domains act as a spacer in order to maintain the N-terminal domains in their correct orientation.

The fact that the presence of the central domains increases the ATPase rate at high concentrations, whilst with the C0C2 domains high concentrations were shown to inhibit ATP hydrolysis (Belknap et al., 2014) could be due to the apparent differences in the binding affinities. As was shown in the co-sedimentation assays the C3C5 domains appeared to bind to the NTFs with a much lower affinity when compared with the C0C2 domains. This could intimate that the C3 domain may be involved in the sensitisation of the thin filaments, as suggested in previous studies (Inchingolo et al., 2019), but that the addition of the C4-C5 domains decreases or interferes with the potential interaction with actin. Alternatively, it could be that initial binding by the N-terminal domains is required in order for the C3 domain to be able to interact.

On the other hand, the fact that the C3C5 only enhanced the ATPase rate in the presence of calcium and therefore when the system has been turned on may suggest that the effect the central domains may be having on the ATP hydrolysis rate could be via an interaction with myosin and that it could be helping myosin to bind to actin in some way.

6.6. Missense mutations in the central domains lead to Hypertrophic Cardiomyopathy via several different mechanisms

Missense mutations in MYBPC3 may lead to HCM via multiple different pathogenic mechanisms. Many different theories as to the pathogenesis of the HCM-linked sarcomeric mutations have been put forward including but not limited to: an increase in the myosin motor intrinsic force (Seeböhm et al., 2009, Sommesse et al., 2013), an increase in sarcomeric Ca^{2+} sensitivity (Marston et al., 2012), inefficient or excessive ATPase activity leading to energy deficiency (Ashrafian et al., 2003) or alterations to the SRX state of myosin (McNamara et al., 2017, Spudich, 2015).

One such mechanism may be that the mutation leads to instability of the protein leading the protein to be targeted for degradation and thus leading to disease via haploinsufficiency. This loss of protein may lead to an effect similar to that seen in KO cMyBP-C animals and truncation mutations, which show signs of disruption to the thick filaments as well as an increase in crossbridge cycling kinetics (Kensler and Harris, 2008, Korte et al., 2003, Harris et al., 2002). Furthermore, it has been shown that haploinsufficiency can lead to an increase in myofilament Ca^{2+} sensitivity, with both cMyBP-C KO animals (Cazorla et al., 2006) and partial extraction of cMyBP-C from normal muscle (Kulikovskaya et al., 2003b) leading to increased sensitivity to Ca^{2+} .

Another proposed mechanism by which HCM-linked mutations could lead to disease is by disrupting the SRX state, which is when the myosin heads have a suppressed ATPase state and there are fewer myosin heads which are capable of force generation (Hooijman et al., 2011, McNamara et al., 2016, Naber et al., 2011). This decrease in the amount of sequestered heads could lead to hypercontractility which is one of the phenotypical

presentations which underpin HCM (Davies and McKenna, 1995, Hughes, 2004). A study by McNamara et al. determined the SRX state of the myosin heads in skinned muscle cells obtained from intraventricular septum samples of non-failing hearts or HCM patients either with mutations in the MYBPC3 gene or non-sarcomeric mutations. The results showed that all of the patients with mutations in cMyBP-C showed disruption of the SRX state whilst those with non-sarcomeric mutations did not. This disruption of the SRX state was seen to occur via two different mechanisms: affecting the lifetime of ATP turnover or the number of sequestered myosin heads (McNamara et al., 2017). This suggests that whilst hypercontractility due to disruption of the SRX state could be considered an overarching mechanism of pathogenesis for multiple different mutations of cMyBP-C, the precise way in which this occurs may be different depending on the mutation.

A decrease in ATP turnover lifetime has also been shown in studies utilising papillary muscle from cMyBP-C KO mice, which showed crossbridge cycles were 40% faster compared to WT. These experiments also intimated that without the presence of MyBP-C the probability of myosin being weakly bound to actin was increased and that force production during the power stroke was reduced (Lecarpentier et al., 2008). It is possible that HCM-linked missense mutations in MYBPC3 could affect the ATP turnover rate by interfering with MyBP-C's ability to prevent the inefficient weak binding of actin to myosin and thus hamper the regulatory role of MyBP-C as was seen to occur in cMyBP-C KO mice (Lecarpentier et al., 2008).

Studies looking into HCM-linked mutations of β -cardiac myosin, which along with mutations in MyBP-C make up the majority of HCM-linked mutations (Lopes et al., 2013), showed that many of these mutations occurred in positions that would be likely to weaken the IHM state

of the myosin heads making them more functionally accessible for actin interaction (Nag et al., 2017). The majority were found to be located on the so called 'myosin mesa' which is a relatively flat surface found on the myosin head domain which is highly conserved across species. This region is thought to be a binding domain, possibly for the proximal region of myosin S2 or MyBP-C, with this interaction holding the myosin head in its inactive form (Spudich, 2015, Spudich et al., 2016). Evidence supporting cMyBP-C being the binding partner for the myosin mesa come from the fact the mesa is highly conserved in striated muscle, which has MyBP-C, whilst many residues are altered in the mesa region of smooth muscle myosin and non-muscle myosin II, which do not contain MyBP-C (Spudich, 2015). Furthermore, MST studies have shown that MyBP-C binds to a shortened version of S1 myosin which lacks the RLC (which has previously been shown to bind MyBP-C (Ratti et al., 2011)). These experiments showed that the binding to S1 was 2x weaker for the C0-C2 domains of MyBP-C alone when compared with the full length protein, intimating that the central domains could be involved in the interaction with S1 and thus the stabilisation of the myosin heads in its inactive IHM form (Nag et al., 2017). It is therefore possible that some of the missense mutations found in the central domains could occur in regions that facilitate the interaction with the myosin mesa and destabilise the interaction leading to a decrease in the number of sequestered heads and thus hypercontractility.

These theories focus mainly on how mutations in cMyBP-C may affect the thick filament. However, the possibility that mutations may be affecting the ability of cMyBP-C to interact with the thin filament should also be considered. Multiple experiments have shown that MyBP-C activates the thin filament at low Ca^{2+} levels, whilst under maximal activation conditions it acts to decrease sliding velocity. It is possible that HCM-linked mutations in

cMyBP-C may occur in regions of the protein which are necessary for interaction with actin and thus interfere with this effect on the thin filaments.

These different theories as to how mutations in cMyBP-C may ultimately lead to HCM highlight the fact that it is probable that no one unifying pathogenic mechanism is responsible for causing the disease. It seems likely that the way in which a mutation leads to disease is dependent upon whether the position of the mutation is in a region which is important for the stability of the protein or for interaction with binding partners. To be able to have a clearer understanding much more needs to be elucidated about the function of the specific regions of cMyBP-C.

In this study several HCM-linked missense mutations found within the central domains were studied to elucidate potential mechanisms by which these mutations may cause HCM.

Initial studies utilising GFP-tagged mutant containing plasmids spanning the C3-C6 domains suggested that an R820Q mutation found in the C6 domain could lead to protein instability and degradation as no protein was expressed in the COS-1 cells or NRCs. The remainder of the missense mutations did express in both the COS-1 cells and the NRCs and did not show any significant difference in protein expression levels suggesting that these mutations do not lead to protein degradation.

Some of the mutants appeared to cause mislocalisation of the protein to the nucleus (G531R, G596R, A833T) whilst others showed a decrease in the co-localisation seen between the central domains and actin (E542Q, G596R) which could implicate these mutations could be causing disease via interfering with the interactions of the central domains, as the mislocalisation to the nucleus could be due to a lack of interaction with the

correct binding partners of the protein. On the other hand, mislocalisation to the nucleus could also be due to the protein being folded incorrectly.

The findings of the STED microscopy looking at the effect expression of these mutant C3C6 domains could have on the myofibrils indicated that all of the mutant proteins that were capable of being expressed in the NRCs lead to myofibrillar disarray following an extended cell culture, whilst the N755K, W792R, R817Q and A833T showed disarray even after only two days of culturing. This could indicate that these missense mutations could be causing subtle alterations to the arrangement of the thick filaments. The fact that these alterations were only seen following prolonged culture for some of the mutant proteins and for the rest this resulted in exacerbation of the disarray, could indicate that the effects of these mutations may cause alterations in the structure of the protein that can be tolerated to a certain extent. It is possible that only following prolonged exposure to mechanical stress that these mutations have an effect on the myofibrillar structure. This could suggest a reason for the fact that HCM is generally a disease which does not become apparent until later on in life.

Whilst these preliminary findings suggest multiple different mechanisms by which these missense mutations may lead to HCM, it is important to highlight the limitations of the use of the GFP-tagged mutant C3C6 domain constructs. These plasmids showed a very low transfection rate, even for the WT plasmid, therefore it is entirely possible that the cells which were expressing the mutant proteins were cells that were already not in the healthiest condition. As these experiments utilised a construct containing only the C3C6 domains, the proteins would not be expected to correctly incorporate into the sarcomere as they do not contain the C-terminal regions known to be necessary for incorporation into the

thick filament backbone (Okagaki et al., 1993, Freiburg and Gautel, 1996). Therefore, the evidence of possible mislocalisation of some of the mutant proteins and the differences seen in actin co-localisation could simply be due to the fact that the protein is not correctly incorporated into the sarcomere and in its correct stoichiometry. It is possible that when the protein is tethered to the thick filament backbone the central domains may not be physically able to show this interaction with actin. Lastly, the myofibrillar disarray in NRCs expressing the mutant proteins seen via utilisation of STED was not visualising the mutant protein but rather showing the effect that expression of the proteins had on the thick filament ultrastructure. Therefore, it was only showing the indirect effect that expression of these mutant proteins could have on the myofibrils. On the other hand, any phenotypic observations attributed to these mutations was in comparison to the WT C3C6 domain spanning protein, which would also have all of the limitations mentioned above.

The preliminary data from these experiments was then used to identify the missense mutations that showed the most interesting phenotype to carry forward into a full-length cMyBP-C adenoviral vector. This was done as adenoviruses typically show a much higher infection rate compared with plasmid transfection and as it would produce a full-length protein it would contain the C-terminal domains required for incorporation into the thick filament (Freiburg and Gautel, 1996, Okagaki et al., 1993) and so should have the correct localisation and stoichiometry in the sarcomere.

The missense mutations chosen were E542Q, G596R, N755K, W792R and R820Q. These mutations are spread across the central domains, with one each occurring in the C3, C4 and C5 domains, whilst two are found in the C6 domain. All of these mutations occur at positions that are generally well conserved between species as well as between the

different isoforms of the protein and are classified as likely pathogenic or pathogenic (except for G596R which has unknown significance) by the ClinVar database (Landrum et al., 2018). The experimentations utilising these full-length mutant cMyBP-C adenoviral vectors further added to the evidence that these missense mutations appear to cause HCM via different disease mechanisms, which are summarised in figure 64.

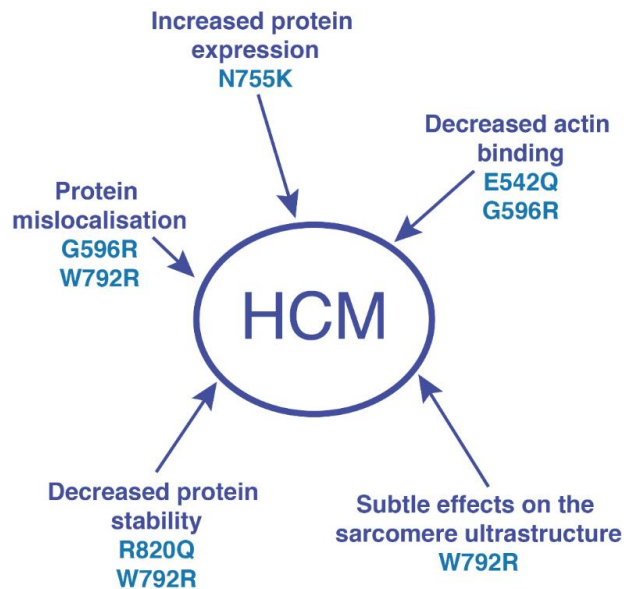


Figure 64 – Summary of potential pathogenic mechanisms by which missense mutations in the central domains may lead to Hypertrophic Cardiomyopathy

Results from this study highlighted the fact that the HCM-linked missense mutations found in the central domains of cardiac myosin binding protein C appear to lead to disease via several different pathogenic mechanisms. The N755K mutation appeared to lead to an overexpression of the protein which could lead to increased interaction with binding partners leading to aberration of protein function. Both the E542Q and G596R mutation decreased the co-localisation of the central domains with actin suggesting these mutations may interfere with binding of cMyBP-C to actin. Subtle alterations in the sarcomeric ultrastructure was seen via high resolution microscopy in cells expressing the W792R mutant protein. The G596R mutant showed an increase in nuclear localisation, whilst the W792R mutant protein showed a higher level of diffuse expression and a less distinct sarcomeric doublet pattern suggesting these mutations may lead to mislocalisation of the protein. Finally, the W792R mutant showed a decrease in protein expression levels, whilst the R820Q mutant did not express in cells, except for one anecdotal case, suggesting that these mutations may cause instability of the protein leading to degradation.

6.6.1 Comments on previous studies on the central domains of MyBP-C

Prior studies have been carried out looking into the missense mutations found in the central domains of MyBP-C which were studied in these experiments, many of which corroborated the findings of this study.

For the E542Q missense mutation the experiments showed no evidence of a decrease in protein expression or change to the correct incorporation of the protein into the sarcomere, corroborating previous findings (Flavigny et al., 1999). There did however appear to be a possible effect on the interaction of the central domains with actin.

Previous literature looking at the G596R mutation has only looked into the clinical effect of the mutation, with no studies undertaken in order to show the expression levels or localisation of mutant protein. The experiments undertaken in this study showed that there was no decrease in protein expression suggesting that this mutation does not lead to instability of the protein and target it for degradation. There was however some evidence in the experiments utilising the C3C6 domain containing construct of a possible mislocalisation of the mutant protein to the nucleus. This could be due to the fact that C-terminal domains required for incorporation of the protein into the sarcomere (Freiburg and Gautel, 1996, Okagaki et al., 1993) were not present, however this increase in nuclear localisation was in comparison to the WT protein which also did not contain the C-terminus. This mutation also showed a decrease in actin co-localisation which suggests that this mutation could be affecting the possible interaction of the central domains with actin. This could also account for the increased nuclear localisation of the protein seen, as a decreased interaction with the actin could lead to mislocalisation of the protein.

Both mutations that appeared to affect the possible interaction between the central domains and actin were found to be located more N-terminally in the protein, in the C3 and C4 domains, whilst the more C-terminal missense mutations did not appear to affect the interaction with actin. This implies that these two domains of the central region may be involved in actin binding. This is partially corroborated by previous work in which the C3 domain was implicated as being involved in sensitising the thin filament to calcium (Inchingolo et al., 2019).

There are no previous experiments looking at protein expression or localisation of the N755K mutation in cells, however structural studies have suggested that the mutation could affect the tight domain fold of the C5 domain and destabilise the protein. Although it is thought that the tight packing of the surrounding domains could act to stabilise the domain up to a point, with longer exposure to mechanical stress possibly leading to loss of structural integrity (Idowu et al., 2003). Due to the location of the N755K mutation, which would be at the position proposed to interact with the C8 domain in the collar model of the 3D sarcomeric localisation of MyBP-C, it is also possible that this mutation could affect the sarcomeric localisation of the protein. Experimentation has previously shown that this mutation leads to a 10x decrease in binding with the C8 domain of cMyBP-C (Moolman-Smook et al., 2002). In the current experiments, the N755K missense mutation seemed to lead to an increase in protein expression, as expression levels were seen to be slightly increased based on western blot analysis, whilst overexpression was seen in cells via confocal microscopy at a higher level than for the WT protein. If this mutation does cause overexpression of cMyBP-C then it could cause aberration of the functioning of the protein

in regulating actomyosin cross-bridge cycling as the increase in the protein levels could lead to increased interaction with MyBP-C and its binding partners.

Previous literature has shown that the W792R missense mutation led to a 70% decrease in protein expression levels in cells, a decrease in fluorescence intensity, decreased levels of bacterially expressed protein and increased levels of a proteolytic cleavage product seen by mass spectrometry (Smelter et al., 2018). The results from our experiments corroborate this finding as there was seen to be a trend to a decrease in protein expression levels shown by Western blot analysis, as well as a decrease in fluorescence intensity levels of GFP shown by confocal microscopy. However, a previous study showed that the mutant protein showed no evidence of mislocalisation (Smelter et al., 2018), whilst in this study there was a significant increase in the number of cells showing diffuse cytosolic localisation of the protein shown by confocal microscopy. Furthermore, far fewer cells were seen to show the correct doublet pattern when imaged utilising STED, with those that did show doublets being far less distinct. This would suggest that W792R mutation leads to the protein not being correctly incorporated into the sarcomere. This could imply that the C6 domain may also be involved in the incorporation of MyBP-C into the thick filament backbone along with the C7-C10 domains.

It would appear that the R820Q missense mutation could lead to HCM via protein degradation as it was not expressed in any cell type when using either the GFP-tagged C3C6 plasmid or the adenovirus containing the full-length protein. Both of these were sequenced before use in the experiments with no issues being seen and during titration of the adenovirus it was shown to be replicating by hexon staining, all of which suggests that there was nothing wrong with either of the expression vectors. There was one anecdotal cell

which was seen to be expressing the adenoviral construct in doublets suggesting the mutant protein can be expressed. The alpha-actinin staining was very weak in this image which was unexpected. On the other hand, the fact that it appeared to localise in the correct doublet pattern would suggest that this weak alpha-actinin staining was due to a technical issue during the immunostaining or the imaging process. The fact that only one cell was ever seen to express this mutant protein suggests that it is possible that the mutation causes the protein to be unstable and the cellular clearance mechanisms such as UPS may remove the protein but the clearance system may have been overwhelmed leading to expression of the protein. This missense mutation leading to an unstable protein has previously been suggested by multi-angle laser light scattering (MALLS) data which showed that the mutation led to a protein which showed non-specific aggregation (Nadvi et al., 2016). However, structural studies have suggested that as the mutation is surface exposed it would not alter the folding or stability of the protein but would be more likely to interfere with possible interactions with binding partners, with it being located in the correct region to interfere with the potential interaction with the C9 domain if the collar model of the 3D sarcomeric localisation of cMyBP-C is correct (Nadvi et al., 2016).

6.7. Limitations of the study

Whilst this study was able to elucidate many findings, particularly highlighting the heterogeneity of the pathogenic mechanisms of missense mutations of cMyBP-C, there were limitations to the experiments.

The GFP tagged C3C6 domain containing plasmids (both WT and mutants) showed a very low transfection rate. Therefore, very few cells were expressing the protein and thus any phenotypic observations that were made could have been due to the fact that the only cells which were able to express the protein were already in a bad condition. This was however remedied by the use of an adenoviral vector in the subsequent experiments, which showed much higher levels of infection rates. The use of an adenoviral vector may have also caused issues, as it is possible that the adenovirus damaged the NRCs in some way. However, as the cells expressing the mutant protein were compared with the WT adenovirus any effects occurring due to the adenovirus itself would have also been seen in cells infected with the WT adenovirus.

Another issue was the fact that the experiments overexpressed mutant proteins in NRCs with the endogenous cMyBP-C still present. It is possible that the endogenous protein may have had compensatory effects and so the true extent of the effect that the mutant protein may have was not being seen. However, as the majority of HCM cases are heterozygous this may recapitulate what is happening in the real-life scenario.

Whilst use of the central domain fragments alone allowed for the study of this region without the influence of the other domains which could obscure the interactions and functions of the central domains, this does not recapitulate what would be occurring in situ. Use of the GFP-tagged C3C6 domain protein in the NRCs would not allow for the correct

localisation of the protein, as they lack the C-terminal domains required for correct incorporation into the sarcomere (Okagaki et al., 1993, Freiburg and Gautel, 1996). Furthermore, when surrounded by the N-terminal and C-terminal domains the central domains may act in a different way than when isolated as they may be stabilised by these other domains, or these domains may affect the ability of the central domains to bind to interaction partners through steric hindrance. Additionally, the studies suggesting that the central domains could interact with actin and/or affect the actomyosin crossbridge cycling kinetics may not be relevant as it is possible that when the protein is incorporated into the sarcomere in its correct stoichiometry the central domains are not able to have these interactions or effects.

Lastly, in this study NRCs were utilised, which whilst they are a relevant model to use as they are cardiomyocytes, they are not a perfect system to utilise. NRCs exhibit an immature phenotype, which as HCM is typically a disease which occurs later in life may not be able to fully recapitulate what is occurring in the disease.

6.8. Future directions

Whilst much was learnt from the experimentations undertaken, more experiments were planned in order to gain a better understanding of the effect the chosen missense mutations of the central domains of cMyBP-C were having on the protein. However, due to restrictions imposed by the COVID-19 pandemic unfortunately these could not be carried out.

These planned experiments included recombinantly producing and purifying central domain fragments containing the chosen missense mutations in order to carry out techniques such as microscale thermophoresis (MST) or differential scanning fluorimetry (DSF) to determine whether the mutations had any impact on the interaction with known binding partners and stability of the protein, respectively. Furthermore, these mutant proteins were going to be utilised in co-sedimentation assays with the NTFs and in the ATPase assay in order to determine whether the mutations would have any effect on the levels of co-sedimentation seen or an effect on ATP hydrolysis rates. This would have been particularly interesting for the two mutations which seemed to show a decreased interaction with actin based on confocal microscopy images (E542Q and G596R) to see if this result could be further corroborated.

Whilst the ability of the central domains to interact with actin was explored in these works, whether they could also bind to myosin was not looked into. Therefore, co-sedimentation assays with the recombinant central domain protein and myosin could have been carried out in order to determine this. This would also provide more information about whether the effect of the central domains on the ATPase rate was due to interaction with the thin or thick filaments.

As the Y2H assay screen did not identify any putative binding partners for the central domains of cMyBP-C, further experiments could also be carried out. As utilising different Y2H systems has been shown to identify different interactions, the screen could be carried out utilising these different systems. Another alternative strategy would be to utilise the proximity-dependent biotin identification (BioID) tool. This system utilises an *E. coli* biotin protein ligase which is mutated to allow it to act as a promiscuous biotin ligase (Choi-Rhee et al., 2004, Cronan, 2005). This mutant protein is fused to the protein of interest allowing for proteins which are interacting with the protein of interest to be biotinylated (Roux et al., 2012). This method has the disadvantage that identification of the interacting, biotinylated proteins is via mass spectrometry which combined with the promiscuity of the biotin ligase means proteins which are not interacting with but are merely in the proximity of the protein of interest may be detected. On the other hand, it has the advantages of being able to be carried out in live mammalian cells and can also detect weak and transient interactions that may occur over a period of time (Sears et al., 2019).

The R820Q mutation led to a lack of protein expression, whilst the W792R mutation appeared to lead to a decrease in expression levels however, it is unclear whether this is due to instability and degradation at the protein or mRNA level. Therefore, mRNA levels should be determined via quantitative PCR (qPCR). The pathway by which the protein may be being degraded by could also be determined, which could be achieved by probing for components of different cellular degradation pathways by western blot. If it was seen that a specific pathway was responsible for protein degradation, experiments could then be carried out using interventions to block this pathway, such as utilisation of a proteasome inhibitor such as MG132, to see whether this could lead to expression of the mutant

protein. Previous studies using MG132 in mice with a HCM-linked MYBPC3 mutation KI have shown that this can lead to an improvement in cardiac function (Schlossarek et al., 2014), whilst another study showed that multiple different proteasome inhibitors, including MG132, were able to suppress cardiomyocyte hypertrophy (Meiners et al., 2008).

As these experiments overexpressed the mutant protein in NRCs which already have endogenous cMyBP-C it is possible that any effects that the mutation in the protein was having was being compensated for by the endogenous protein. Whilst this may be a relatively accurate portrayal of the disease, as typically HCM patients would have a heterozygous mutation, it may obscure the effect that the mutation is having on the protein. If techniques were utilised to remove or decrease the levels of endogenous protein then this may give a clearer picture of the pathogenesis of the mutation, which may not fully recapitulate what would occur in the patient, but could help to gain a better understanding of the functionality of the region in the protein where the mutation is located.

One of the limitations of this study was in the use of NRCs, which have an immature phenotype and so may not fully recapitulate what occurs in the patient, especially as HCM is generally a disease which occurs in later life. Longer culture of the NRCs was also not compatible with the adenoviral infection, as it was seen that this led to overexpression of the protein making visualisation by STED impossible. Either adult rat cardiomyocytes or engineered heart tissue (EHTs) could therefore be used in future experiments, as these would have a more mature phenotype than the NRCs. Furthermore, non-adenoviral methods of incorporating the mutant proteins could be explored, such as CRISPR or the 'cut and paste' technique recently developed by the Harris group (Napieralski et al., 2020).

Lastly, a larger number of missense mutations should be studied, as then it is possible that trends may begin to form between the region in which the mutations are located and the mechanism by which pathogenesis occurs. This could inform on the specific function of the different regions of the protein as well as classify different mutations depending on their pathogenesis, which could aid in more personalised methods of treatment avenues.

6.9. Final conclusions

The findings of this work showed that the central domains of cMyBP-C may interact with actin, particularly the C3 and C4 domains and could also have a role in regulation of actomyosin crossbridge cycling kinetics. Furthermore, it suggests that there may be instability between the C5 and C6 linker, with previous literature indicating that these domains do show a high degree of flexibility which could possibly be necessary for the function of the protein (Jeffries et al., 2011).

The major finding of this work, however, is the fact that different missense mutations of the central domains appear to lead to HCM via different pathological mechanisms. Further work would have to be carried out in order to fully elucidate these mechanisms as well as studies into a far wider range of mutations in order to see if any trends form between the position of the mutation and the mechanism of disease. This could potentially lead to being able to classify different mutations and thus help to tailor strategies to treat the disease based on the specific pathogenesis of the mutation. It may also lead to the ability to make predictions on the mechanism of disease based on the location of the mutation, as well as suggesting functional roles for the different regions of the protein.

- ABABOU, A., GAUTEL, M. & PFUHL, M. 2007. Dissecting the N-terminal myosin binding site of human cardiac myosin-binding protein C. Structure and myosin binding of domain C2. *J Biol Chem*, 282, 9204-15.
- ABABOU, A., ROSTKOVA, E., MISTRY, S., LE MASURIER, C., GAUTEL, M. & PFUHL, M. 2008. Myosin binding protein C positioned to play a key role in regulation of muscle contraction: structure and interactions of domain C1. *J Mol Biol*, 384, 615-30.
- ABBE, E. 1873. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9, 413-468.
- ACKERMANN, M. A. & KONTROGIANNI-KONSTANTOPOULOS, A. 2011. Myosin binding protein-C: a regulator of actomyosin interaction in striated muscle. *J Biomed Biotechnol*, 2011, 636403.
- AL-KHAYAT, H. A., KENSLER, R. W., SQUIRE, J. M., MARSTON, S. B. & MORRIS, E. P. 2013. Atomic model of the human cardiac muscle myosin filament. *Proc Natl Acad Sci U S A*, 110, 318-23.
- ALAMO, L., WRIGGERS, W., PINTO, A., BÁRTOLI, F., SALAZAR, L., ZHAO, F. Q., CRAIG, R. & PADRÓN, R. 2008. Three-dimensional reconstruction of tarantula myosin filaments suggests how phosphorylation may regulate myosin activity. *J Mol Biol*, 384, 780-97.
- ALYONYCHEVA, T., COHEN-GOULD, L., SEWERT, C., FISCHMAN, D. A. & MIKAWA, T. 1997. Skeletal muscle-specific myosin binding protein-H is expressed in Purkinje fibers of the cardiac conduction system. *Circ Res*, 80, 665-72.
- ANDERSON, R. L., TRIVEDI, D. V., SARKAR, S. S., HENZE, M., MA, W., GONG, H., ROGERS, C. S., GORHAM, J. M., WONG, F. L., MORCK, M. M., SEIDMAN, J. G., RUPPEL, K. M., IRVING, T. C., COOKE, R., GREEN, E. M. & SPUDICH, J. A. 2018. Deciphering the super relaxed state of human β -cardiac myosin and the mode of action of mavacamten from myosin molecules to muscle fibers. *Proc Natl Acad Sci U S A*, 115, E8143-e8152.
- ANTUNES, M. O. & SCUDELER, T. L. 2020. Hypertrophic cardiomyopathy. *Int J Cardiol Heart Vasc*, 27, 100503.
- ARAD, M., MONSERRAT, L., HARON-KHUN, S., SEIDMAN, J. G., SEIDMAN, C. E., ARBUSTINI, E., GLIKSON, M. & FREIMARK, D. 2014. Merits and pitfalls of genetic testing in a hypertrophic cardiomyopathy clinic. *Isr Med Assoc J*, 16, 707-13.
- ARYAL, B., JEONG, J. & RAO, V. A. 2014. Doxorubicin-induced carbonylation and degradation of cardiac myosin binding protein C promote cardiotoxicity. *Proc Natl Acad Sci U S A*, 111, 2011-6.
- ASHRAFIAN, H., REDWOOD, C., BLAIR, E. & WATKINS, H. 2003. Hypertrophic cardiomyopathy: a paradigm for myocardial energy depletion. *Trends Genet*, 19, 263-8.
- AVKIRAN, M., ROWLAND, A. J., CUELLO, F. & HAWORTH, R. S. 2008. Protein kinase d in the cardiovascular system: emerging roles in health and disease. *Circ Res*, 102, 157-63.
- AZIMZADEH, J., HERGERT, P., DELOUVÉE, A., EUTENEUER, U., FORMSTECHE, E., KHODJAKOV, A. & BORNENS, M. 2009. hPOC5 is a centrin-binding protein required for assembly of full-length centrioles. *J Cell Biol*, 185, 101-14.
- BARDSWELL, S. C., CUELLO, F., ROWLAND, A. J., SADAYAPPAN, S., ROBBINS, J., GAUTEL, M., WALKER, J. W., KENTISH, J. C. & AVKIRAN, M. 2010. Distinct sarcomeric substrates are responsible for protein kinase D-mediated regulation of cardiac myofilament Ca^{2+} sensitivity and cross-bridge cycling. *J Biol Chem*, 285, 5674-82.
- BAREFIELD, D. & SADAYAPPAN, S. 2010. Phosphorylation and function of cardiac myosin binding protein-C in health and disease. *J Mol Cell Cardiol*, 48, 866-75.
- BELKNAP, B., HARRIS, S. P. & WHITE, H. D. 2014. Modulation of thin filament activation of myosin ATP hydrolysis by N-terminal domains of cardiac myosin binding protein-C. *Biochemistry*, 53, 6717-24.
- BENNETT, P., CRAIG, R., STARR, R. & OFFER, G. 1986. The ultrastructural location of C-protein, X-protein and H-protein in rabbit muscle. *J Muscle Res Cell Motil*, 7, 550-67.
- BENNETT, P., REES, M. & GAUTEL, M. 2020. The Axial Alignment of Titin on the Muscle Thick Filament Supports Its Role as a Molecular Ruler. *J Mol Biol*, 432, 4815-4829.

- BENNETT, P., STARR, R., ELLIOTT, A. & OFFER, G. 1985. The structure of C-protein and X-protein molecules and a polymer of X-protein. *J Mol Biol*, 184, 297-309.
- BERS, D. M. 2002. Cardiac excitation–contraction coupling. *Nature*, 415, 198-205.
- BERSHITSKY, S. Y., KOUBASSOVA, N. A., FERENCZI, M. A., KOPYLOVA, G. V., NARAYANAN, T. & TSATURYAN, A. K. 2017. The Closed State of the Thin Filament Is Not Occupied in Fully Activated Skeletal Muscle. *Biophys J*, 112, 1455-1461.
- BETTS, M. J. & RUSSELL, R. B. 2003. Amino Acid Properties and Consequences of Substitutions. *Bioinformatics for Geneticists*.
- BEZOLD, K. L., SHAFFER, J. F., KHOSA, J. K., HOYE, E. R. & HARRIS, S. P. 2013. A gain-of-function mutation in the M-domain of cardiac myosin-binding protein-C increases binding to actin. *J Biol Chem*, 288, 21496-505.
- BHUIYAN, M. S., GULICK, J., OSINSKA, H., GUPTA, M. & ROBBINS, J. 2012. Determination of the critical residues responsible for cardiac myosin binding protein C's interactions. *J Mol Cell Cardiol*, 53, 838-47.
- BOUSSOUF, S. E. & GEEVES, M. A. Tropomyosin and Troponin Cooperativity on the Thin Filament. 2007 Tokyo. Springer Japan, 99-109.
- BRENNER, B. M., BALLERMANN, B. J., GUNNING, M. E. & ZEIDEL, M. L. 1990. Diverse biological actions of atrial natriuretic peptide. *Physiol Rev*, 70, 665-99.
- BRICKSON, S., FITZSIMONS, D. P., PEREIRA, L., HACKER, T., VALDIVIA, H. & MOSS, R. L. 2007. In vivo left ventricular functional capacity is compromised in cMyBP-C null mice. *Am J Physiol Heart Circ Physiol*, 292, H1747-54.
- BRONDYK, W. H. 2009. Chapter 11 Selecting an Appropriate Method for Expressing a Recombinant Protein. In: BURGESS, R. R. & DEUTSCHER, M. P. (eds.) *Methods in Enzymology*. Academic Press.
- BROOKS, C. L. & GU, W. 2003. Ubiquitination, phosphorylation and acetylation: the molecular basis for p53 regulation. *Curr Opin Cell Biol*, 15, 164-71.
- BRÜCKNER, A., POLGE, C., LENTZE, N., AUERBACH, D. & SCHLATTNER, U. 2009. Yeast two-hybrid, a powerful tool for systems biology. *International journal of molecular sciences*, 10, 2763-2788.
- CAI, Z. & YAN, L. J. 2013. Protein Oxidative Modifications: Beneficial Roles in Disease and Health. *J Biochem Pharmacol Res*, 1, 15-26.
- CALAGHAN, S. C., TRINICK, J., KNIGHT, P. J. & WHITE, E. 2000. A role for C-protein in the regulation of contraction and intracellular Ca²⁺ in intact rat ventricular myocytes. *J Physiol*, 528 Pt 1, 151-6.
- CARRIER, L., BONNE, G., BAHREND, E., YU, B., RICHARD, P., NIEL, F., HAINQUE, B., CRUAUD, C., GARY, F., LABEIT, S., BOUHOUE, J. B., DUBOURG, O., DESNOS, M., HAGEGE, A. A., TRENT, R. J., KOMAJDA, M., FISZMAN, M. & SCHWARTZ, K. 1997. Organization and sequence of human cardiac myosin binding protein C gene (MYBPC3) and identification of mutations predicted to produce truncated proteins in familial hypertrophic cardiomyopathy. *Circ Res*, 80, 427-34.
- CARRIER, L., KNOLL, R., VIGNIER, N., KELLER, D. I., BAUSERO, P., PRUDHON, B., ISNARD, R., AMBROISINE, M. L., FISZMAN, M., ROSS, J., JR., SCHWARTZ, K. & CHIEN, K. R. 2004. Asymmetric septal hypertrophy in heterozygous cMyBP-C null mice. *Cardiovasc Res*, 63, 293-304.
- CARRIER, L., MEARINI, G., STATHOPOULOU, K. & CUELLO, F. 2015. Cardiac myosin-binding protein C (MYBPC3) in cardiac pathophysiology. *Gene*, 573, 188-97.
- CAZORLA, O., SZILAGYI, S., VIGNIER, N., SALAZAR, G., KRÄMER, E., VASSORT, G., CARRIER, L. & LACAMPAGNE, A. 2006. Length and protein kinase A modulations of myocytes in cardiac myosin binding protein C-deficient mice. *Cardiovasc Res*, 69, 370-80.
- CHALOVICH, J. M. & EISENBERG, E. 1982. Inhibition of actomyosin ATPase activity by troponin-tropomyosin without blocking the binding of myosin to actin. *J Biol Chem*, 257, 2432-7.

- CHARRON, P., DUBOURG, O., DESNOS, M., BENNACEUR, M., CARRIER, L., CAMPROUX, A. C., ISNARD, R., HAGEGE, A., LANGLARD, J. M., BONNE, G., RICHARD, P., HAINQUE, B., BOUHOUR, J. B., SCHWARTZ, K. & KOMAJDA, M. 1998. Clinical features and prognostic implications of familial hypertrophic cardiomyopathy related to the cardiac myosin-binding protein C gene. *Circulation*, 97, 2230-6.
- CHAUVEAU, C., ROWELL, J. & FERREIRO, A. 2014. A rising titan: TTN review and mutation update. *Hum Mutat*, 35, 1046-59.
- CHEN, Z., ZHAO, T. J., LI, J., GAO, Y. S., MENG, F. G., YAN, Y. B. & ZHOU, H. M. 2011. Slow skeletal muscle myosin-binding protein-C (MyBPC1) mediates recruitment of muscle-type creatine kinase (CK) to myosin. *Biochem J*, 436, 437-45.
- CHOI-RHEE, E., SCHULMAN, H. & CRONAN, J. E. 2004. Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Sci*, 13, 3043-50.
- COLSON, B. A., BEKYAROVA, T., FITZSIMONS, D. P., IRVING, T. C. & MOSS, R. L. 2007. Radial displacement of myosin cross-bridges in mouse myocardium due to ablation of myosin binding protein-C. *J Mol Biol*, 367, 36-41.
- COLSON, B. A., BEKYAROVA, T., LOCHER, M. R., FITZSIMONS, D. P., IRVING, T. C. & MOSS, R. L. 2008. Protein kinase A-mediated phosphorylation of cMyBP-C increases proximity of myosin heads to actin in resting myocardium. *Circ Res*, 103, 244-51.
- COLSON, B. A., PATEL, J. R., CHEN, P. P., BEKYAROVA, T., ABDALLA, M. I., TONG, C. W., FITZSIMONS, D. P., IRVING, T. C. & MOSS, R. L. 2012. Myosin binding protein-C phosphorylation is the principal mediator of protein kinase A effects on thick filament structure in myocardium. *J Mol Cell Cardiol*, 53, 609-16.
- CONSORTIUM, I. H. G. S. 2004. Finishing the euchromatic sequence of the human genome. *Nature*, 431, 931-45.
- COPELAND, O., SADAYAPPAN, S., MESSER, A. E., STEINEN, G. J., VAN DER VELDEN, J. & MARSTON, S. B. 2010. Analysis of cardiac myosin binding protein-C phosphorylation in human heart muscle. *J Mol Cell Cardiol*, 49, 1003-11.
- CRAIG, R. & LEHMAN, W. 2001. Crossbridge and tropomyosin positions observed in native, interacting thick and thin filaments. *J Mol Biol*, 311, 1027-36.
- CRAIG, R. & OFFER, G. 1976. Axial arrangement of crossbridges in thick filaments of vertebrate skeletal muscle. *J Mol Biol*, 102, 325-32.
- CRONAN, J. E. 2005. Targeted and proximity-dependent promiscuous protein biotinylation by a mutant *Escherichia coli* biotin protein ligase. *J Nutr Biochem*, 16, 416-8.
- CUELLO, F., BARDSWELL, S. C., HAWORTH, R. S., EHLE, E., SADAYAPPAN, S., KENTISH, J. C. & AVKIRAN, M. 2011. Novel role for p90 ribosomal S6 kinase in the regulation of cardiac myofilament phosphorylation. *J Biol Chem*, 286, 5300-10.
- DAVIES, M. J. & MCKENNA, W. J. 1995. Hypertrophic cardiomyopathy--pathology and pathogenesis. *Histopathology*, 26, 493-500.
- DAVIS, J. S. 1988. Interaction of C-protein with pH 8.0 synthetic thick filaments prepared from the myosin of vertebrate skeletal muscle. *J Muscle Res Cell Motil*, 9, 174-83.
- DESAI, R., GEEVES, M. A. & KAD, N. M. 2015. Using fluorescent myosin to directly visualize cooperative activation of thin filaments. *J Biol Chem*, 290, 1915-25.
- DUAN, G. & WALTHER, D. 2015. The roles of post-translational modifications in the context of protein interaction networks. *PLoS Comput Biol*, 11, e1004049.
- EBASHI, S. 1980. The Croonian lecture, 1979: Regulation of muscle contraction. *Proc R Soc Lond B Biol Sci*, 207, 259-86.
- EDWARDS, A. M., KUS, B., JANSEN, R., GREENBAUM, D., GREENBLATT, J. & GERSTEIN, M. 2002. Bridging structural biology and genomics: assessing protein interaction data with known complexes. *Trends Genet*, 18, 529-36.
- EHLE, E. 2016. Cardiac cytoarchitecture - why the "hardware" is important for heart function! *Biochim Biophys Acta*, 1863, 1857-63.

- EHLER, E. 2018. Actin-associated proteins and cardiomyopathy-the 'unknown' beyond troponin and tropomyosin. *Biophys Rev*, 10, 1121-1128.
- EISNER, D. A., ISENBERG, G. & SIPIDO, K. R. 2003. Normal and pathological excitation-contraction coupling in the heart -- an overview. *J Physiol*, 546, 3-4.
- ERDMANN, J., RAIBLE, J., MAKI-ABADI, J., HUMMEL, M., HAMMANN, J., WOLLNIK, B., FRANTZ, E., FLECK, E., HETZER, R. & REGITZ-ZAGROSEK, V. 2001. Spectrum of clinical phenotypes and gene variants in cardiac myosin-binding protein C mutation carriers with hypertrophic cardiomyopathy. *J Am Coll Cardiol*, 38, 322-30.
- ESCHENHAGEN, T., MUMMERY, C. & KNOLLMANN, B. C. 2015. Modelling sarcomeric cardiomyopathies in the dish: from human heart samples to iPSC cardiomyocytes. *Cardiovasc Res*, 105, 424-38.
- ESPOSTI, M. D., DE VRIES, S., CRIMI, M., GHELLI, A., PATARNELLO, T. & MEYER, A. 1993. Mitochondrial cytochrome b: evolution and structure of the protein. *Biochim Biophys Acta*, 1143, 243-71.
- FAGERBERG, L., HALLSTRÖM, B. M., OKSVOLD, P., KAMPF, C., DJUREINOVIC, D., ODEBERG, J., HABUKA, M., TAHMASEBPOOR, S., DANIELSSON, A., EDLUND, K., ASPLUND, A., SJÖSTEDT, E., LUNDBERG, E., SZIGYARTO, C. A., SKOGS, M., TAKANEN, J. O., BERLING, H., TEGEL, H., MULDER, J., NILSSON, P., SCHWENK, J. M., LINDSKOG, C., DANIELSSON, F., MARDINOGLU, A., SIVERTSSON, A., VON FEILITZEN, K., FORSBERG, M., ZWAHLEN, M., OLSSON, I., NAVANI, S., HUSS, M., NIELSEN, J., PONTEN, F. & UHLÉN, M. 2014. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*, 13, 397-406.
- FANANAPAZIR, L. & EPSTEIN, N. D. 1994. Genotype-phenotype correlations in hypertrophic cardiomyopathy. Insights provided by comparisons of kindreds with distinct and identical beta-myosin heavy chain gene mutations. *Circulation*, 89, 22-32.
- FERT-BOBER, J. & SOKOLOVE, J. 2014. Proteomics of citrullination in cardiovascular disease. *Proteomics Clin Appl*, 8, 522-33.
- FIELDS, S. & SONG, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature*, 340, 245-6.
- FIGUEIREDO-FREITAS, C., DULCE, R. A., FOSTER, M. W., LIANG, J., YAMASHITA, A. M., LIMA-ROSA, F. L., THOMPSON, J. W., MOSELEY, M. A., HARE, J. M., NOGUEIRA, L., SORENSON, M. M. & PINTO, J. R. 2015. S-Nitrosylation of Sarcomeric Proteins Depresses Myofilament Ca²⁺Sensitivity in Intact Cardiomyocytes. *Antioxid Redox Signal*, 23, 1017-34.
- FLAVIGNY, J., SOUCHET, M., SÉBILLON, P., BERREBI-BERTRAND, I., HAINQUE, B., MALLET, A., BRIL, A., SCHWARTZ, K. & CARRIER, L. 1999. COOH-terminal truncated cardiac myosin-binding protein C mutants resulting from familial hypertrophic cardiomyopathy mutations exhibit altered expression and/or incorporation in fetal rat cardiomyocytes. *J Mol Biol*, 294, 443-56.
- FLOTHO, A. & MELCHIOR, F. 2013. Sumoylation: a regulatory protein modification in health and disease. *Annu Rev Biochem*, 82, 357-85.
- FOTH, B. J., GOEDECKE, M. C. & SOLDATI, D. 2006. New insights into myosin evolution and classification. *Proc Natl Acad Sci U S A*, 103, 3681-6.
- FRANKLIN, A. J., BAXLEY, T., KOBAYASHI, T. & CHALOVICH, J. M. 2012. The C-terminus of troponin T is essential for maintaining the inactive state of regulated actin. *Biophys J*, 102, 2536-44.
- FREIBURG, A. & GAUTEL, M. 1996. A molecular map of the interactions between titin and myosin-binding protein C. Implications for sarcomeric assembly in familial hypertrophic cardiomyopathy. *Eur J Biochem*, 235, 317-23.
- FREY, N., LUEDDE, M. & KATUS, H. A. 2011. Mechanisms of disease: hypertrophic cardiomyopathy. *Nat Rev Cardiol*, 9, 91-100.
- FRIES, R., HEANEY, A. M. & MEURS, K. M. 2008. Prevalence of the myosin-binding protein C mutation in Maine Coon cats. *J Vet Intern Med*, 22, 893-6.

- FÜRST, D. O., OSBORN, M., NAVE, R. & WEBER, K. 1988. The organization of titin filaments in the half-sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: a map of ten nonrepetitive epitopes starting at the Z line extends close to the M line. *J Cell Biol*, 106, 1563-72.
- FURST, D. O., VINKEMEIER, U. & WEBER, K. 1992. Mammalian skeletal muscle C-protein: purification from bovine muscle, binding to titin and the characterization of a full-length human cDNA. *J Cell Sci*, 102 (Pt 4), 769-78.
- FUSI, L., BRUNELLO, E., YAN, Z. & IRVING, M. 2016. Thick filament mechano-sensing is a calcium-independent regulatory mechanism in skeletal muscle. *Nat Commun*, 7, 13281.
- GARNEAU, N. L., WILUSZ, J. & WILUSZ, C. J. 2007. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol*, 8, 113-26.
- GAUTEL, M., ZUFFARDI, O., FREIBURG, A. & LBEIT, S. 1995. Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? *Embo j*, 14, 1952-60.
- GE, Y., RYBAKOVA, I. N., XU, Q. & MOSS, R. L. 2009. Top-down high-resolution mass spectrometry of cardiac myosin binding protein C revealed that truncation alters protein phosphorylation state. *Proc Natl Acad Sci U S A*, 106, 12658-63.
- GEDICKE, C., BEHRENS-GAWLIK, V., DREYFUS, P., ESCHENHAGEN, T. & CARRIER, L. 2010. Specific Skipping of Exons Using Antisense Oligoribonucleotides Results in a Novel Molecule in cMyBP-C Knock-in Mouse Myocytes. *Circ*, 122.
- GEEVES, M. A., LEHRER, S. S. & LEHMAN, W. 2019. The mechanism of thin filament regulation: Models in conflict? *J Gen Physiol*, 151, 1265-1271.
- GILBERT, R., KELLY, M. G., MIKAWA, T. & FISCHMAN, D. A. 1996. The carboxyl terminus of myosin binding protein C (MyBP-C, C-protein) specifies incorporation into the A-band of striated muscle. *J Cell Sci*, 109 (Pt 1), 101-11.
- GLUZMAN, Y. 1981. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell*, 23, 175-82.
- GOVADA, L., CARPENTER, L., DA FONSECA, P. C., HELLIWELL, J. R., RIZKALLAH, P., FLASHMAN, E., CHAYEN, N. E., REDWOOD, C. & SQUIRE, J. M. 2008. Crystal structure of the C1 domain of cardiac myosin binding protein-C: implications for hypertrophic cardiomyopathy. *J Mol Biol*, 378, 387-97.
- GOVINDAN, S., SARKEY, J., JI, X., SUNDARESAN, N. R., GUPTA, M. P., DE TOMBE, P. P. & SADAYAPPAN, S. 2012. Pathogenic properties of the N-terminal region of cardiac myosin binding protein-C in vitro. *J Muscle Res Cell Motil*, 33, 17-30.
- GREEN, E. M., WAKIMOTO, H., ANDERSON, R. L., EVANCHIK, M. J., GORHAM, J. M., HARRISON, B. C., HENZE, M., KAWAS, R., OSLOB, J. D., RODRIGUEZ, H. M., SONG, Y., WAN, W., LEINWAND, L. A., SPUDICH, J. A., MCDOWELL, R. S., SEIDMAN, J. G. & SEIDMAN, C. E. 2016. A small-molecule inhibitor of sarcomere contractility suppresses hypertrophic cardiomyopathy in mice. *Science*, 351, 617-21.
- GRESHAM, K. S. & STELZER, J. E. 2016. The contributions of cardiac myosin binding protein C and troponin I phosphorylation to β -adrenergic enhancement of in vivo cardiac function. *J Physiol*, 594, 669-86.
- GROVE, B. K., KURER, V., LEHNER, C., DOETSCHMAN, T. C., PERRIARD, J. C. & EPPENBERGER, H. M. 1984. A new 185,000-dalton skeletal muscle protein detected by monoclonal antibodies. *J Cell Biol*, 98, 518-24.
- GRUEN, M. & GAUTEL, M. 1999. Mutations in beta-myosin S2 that cause familial hypertrophic cardiomyopathy (FHC) abolish the interaction with the regulatory domain of myosin-binding protein-C. *J Mol Biol*, 286, 933-49.
- GRUEN, M., PRINZ, H. & GAUTEL, M. 1999. cAPK-phosphorylation controls the interaction of the regulatory domain of cardiac myosin binding protein C with myosin-S2 in an on-off fashion. *FEBS Lett*, 453, 254-9.

- GRUNER, C., CARE, M., SIMINOVITCH, K., MORAVSKY, G., WIGLE, E. D., WOO, A. & RAKOWSKI, H. 2011. Sarcomere protein gene mutations in patients with apical hypertrophic cardiomyopathy. *Circ Cardiovasc Genet*, 4, 288-95.
- GUPTA, M. K. & ROBBINS, J. 2014. Post-translational control of cardiac hemodynamics through myosin binding protein C. *Pflugers Arch*, 466, 231-6.
- GYÖRGY, B., TÓTH, E., TARCSA, E., FALUS, A. & BUZÁS, E. I. 2006. Citrullination: a posttranslational modification in health and disease. *Int J Biochem Cell Biol*, 38, 1662-77.
- HALPAIN, S. & DEHMELT, L. 2006. The MAP1 family of microtubule-associated proteins. *Genome Biol*, 7, 224.
- HANSON, J. & LOWY, J. 1963. The structure of F-actin and of actin filaments isolated from muscle. *Journal of Molecular Biology*, 6, 46-IN5.
- HARRIS, S. P., BARTLEY, C. R., HACKER, T. A., MCDONALD, K. S., DOUGLAS, P. S., GREASER, M. L., POWERS, P. A. & MOSS, R. L. 2002. Hypertrophic cardiomyopathy in cardiac myosin binding protein-C knockout mice. *Circ Res*, 90, 594-601.
- HARRIS, S. P., LYONS, R. G. & BEZOLD, K. L. 2011. In the thick of it: HCM-causing mutations in myosin binding proteins of the thick filament. *Circ Res*, 108, 751-64.
- HARRIS, S. P., ROSTKOVA, E., GAUTEL, M. & MOSS, R. L. 2004. Binding of myosin binding protein-C to myosin subfragment S2 affects contractility independent of a tether mechanism. *Circ Res*, 95, 930-6.
- HARRISON, P. M. & AROSIO, P. 1996. The ferritins: molecular properties, iron storage function and cellular regulation. *Biochim Biophys Acta*, 1275, 161-203.
- HARTZELL, H. C. & SALE, W. S. 1985. Structure of C protein purified from cardiac muscle. *J Cell Biol*, 100, 208-15.
- HEDHLI, N., LIZANO, P., HONG, C., FRITZKY, L. F., DHAR, S. K., LIU, H., TIAN, Y., GAO, S., MADURA, K., VATNER, S. F. & DEPRE, C. 2008. Proteasome inhibition decreases cardiac remodeling after initiation of pressure overload. *Am J Physiol Heart Circ Physiol*, 295, H1385-93.
- HEELEY, D. H., WHITE, H. D. & TAYLOR, E. W. 2019. Investigation into the mechanism of thin filament regulation by transient kinetics and equilibrium binding: Is there a conflict? *J Gen Physiol*, 151, 628-634.
- HEITNER, S. B., JACOBY, D., LESTER, S. J., OWENS, A., WANG, A., ZHANG, D., LAMBING, J., LEE, J., SEMIGRAN, M. & SEHNERT, A. J. 2019. Mavacamten Treatment for Obstructive Hypertrophic Cardiomyopathy: A Clinical Trial. *Ann Intern Med*, 170, 741-748.
- HELL, S. W. & WICHMANN, J. 1994. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19, 780-782.
- HELMS, A. S., TANG, V. T., O'LEARY, T. S., FRIEDLINE, S., WAUCHOPE, M., ARORA, A., WASSERMAN, A. H., SMITH, E. D., LEE, L. M., WEN, X. W., SHAVIT, J. A., LIU, A. P., PREVIS, M. J. & DAY, S. M. 2020a. Effects of MYBPC3 loss-of-function mutations preceding hypertrophic cardiomyopathy. *JCI Insight*, 5.
- HELMS, A. S., THOMPSON, A. D., GLAZIER, A. A., HAFEEZ, N., KABANI, S., RODRIGUEZ, J., YOB, J. M., WOOLCOCK, H., MAZZAROTTO, F., LAKDAWALA, N. K., WITTEKIND, S. G., PEREIRA, A. C., JACOBY, D. L., COLAN, S. D., ASHLEY, E. A., SABERI, S., WARE, J. S., INGLES, J., SEMSARIAN, C., MICHELS, M., OLIVOTTO, I., HO, C. Y. & DAY, S. M. 2020b. Spatial and Functional Distribution of MYBPC3 Pathogenic Variants and Clinical Outcomes in Patients With Hypertrophic Cardiomyopathy. *Circ Genom Precis Med*, 13, 396-405.
- HENDERSON, C. A., GOMEZ, C. G., NOVAK, S. M., MI-MI, L. & GREGORIO, C. C. 2017. Overview of the Muscle Cytoskeleton. *Compr Physiol*, 7, 891-944.
- HERRON, T. J., ROSTKOVA, E., KUNST, G., CHATURVEDI, R., GAUTEL, M. & KENTISH, J. C. 2006. Activation of myocardial contraction by the N-terminal domains of myosin binding protein-C. *Circ Res*, 98, 1290-8.
- HILL, J. A. & OLSON, E. N. 2012. *An introduction to muscle*, Academic Press.

- HILL, T. L., EISENBERG, E. & GREENE, L. 1980. Theoretical model for the cooperative equilibrium binding of myosin subfragment 1 to the actin-troponin-tropomyosin complex. *Proc Natl Acad Sci U S A*, 77, 3186-90.
- HO, C. Y., DAY, S. M., ASHLEY, E. A., MICHELS, M., PEREIRA, A. C., JACOBY, D., CIRINO, A. L., FOX, J. C., LAKDAWALA, N. K., WARE, J. S., CALESHU, C. A., HELMS, A. S., COLAN, S. D., GIROLAMI, F., CECCHI, F., SEIDMAN, C. E., SAJEEV, G., SIGNOROVITCH, J., GREEN, E. M. & OLIVOTTO, I. 2018. Genotype and Lifetime Burden of Disease in Hypertrophic Cardiomyopathy: Insights from the Sarcomeric Human Cardiomyopathy Registry (SHaRe). *Circulation*, 138, 1387-1398.
- HOFMANN, P. A., GREASER, M. L. & MOSS, R. L. 1991. C-protein limits shortening velocity of rabbit skeletal muscle fibres at low levels of Ca²⁺ activation. *J Physiol*, 439, 701-15.
- HOOIJMAN, P., STEWART, M. A. & COOKE, R. 2011. A new state of cardiac myosin with very slow ATP turnover: a potential cardioprotective mechanism in the heart. *Biophys J*, 100, 1969-76.
- HOUMEIDA, A., HOLT, J., TSKHOVREBOVA, L. & TRINICK, J. 1995. Studies of the interaction between titin and myosin. *J Cell Biol*, 131, 1471-81.
- HOWARTH, J. W., RAMISETTI, S., NOLAN, K., SADAYAPPAN, S. & ROSEVEAR, P. R. 2012. Structural insight into unique cardiac myosin-binding protein-C motif: a partially folded domain. *J Biol Chem*, 287, 8254-62.
- HUANG, F. & NAU, W. M. 2003. A conformational flexibility scale for amino acids in peptides. *Angew Chem Int Ed Engl*, 42, 2269-72.
- HUGHES, S. E. 2004. The pathology of hypertrophic cardiomyopathy. *Histopathology*, 44, 412-27.
- HUMPHREY, S. J., JAMES, D. E. & MANN, M. 2015. Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. *Trends Endocrinol Metab*, 26, 676-687.
- HUNTER, T. 2007. The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol Cell*, 28, 730-8.
- HUTAGALUNG, A. H., LANDSVERK, M. L., PRICE, M. G. & EPSTEIN, H. F. 2002. The UCS family of myosin chaperones. *J Cell Sci*, 115, 3983-90.
- HUXLEY, A. F. 1957. Muscle structure and theories of contraction. *Prog Biophys Biophys Chem*, 7, 255-318.
- HUXLEY, A. F. & NIEDERGERKE, R. 1954. Structural changes in muscle during contraction; interference microscopy of living muscle fibres. *Nature*, 173, 971-3.
- HUXLEY, H. & HANSON, J. 1954. Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. *Nature*, 173, 973-6.
- HUXLEY, H. E. 1963. ELECTRON MICROSCOPE STUDIES ON THE STRUCTURE OF NATURAL AND SYNTHETIC PROTEIN FILAMENTS FROM STRIATED MUSCLE. *J Mol Biol*, 7, 281-308.
- IDOWU, S. M., GAUTEL, M., PERKINS, S. J. & PFUHL, M. 2003. Structure, stability and dynamics of the central domain of cardiac myosin binding protein C (MyBP-C): implications for multidomain assembly and causes for cardiomyopathy. *J Mol Biol*, 329, 745-61.
- INCHINGOLO, A. V., PREVIS, S. B., PREVIS, M. J., WARSHAW, D. M. & KAD, N. M. 2019. Revealing the mechanism of how cardiac myosin-binding protein C N-terminal fragments sensitize thin filaments for myosin binding. *Proc Natl Acad Sci U S A*, 116, 6828-6835.
- IQBAL, S., PÉREZ-PALMA, E., JESPERSEN, J. B., MAY, P., HOKSZA, D., HEYNE, H. O., AHMED, S. S., RIFAT, Z. T., RAHMAN, M. S., LAGE, K., PALOTIE, A., COTTRELL, J. R., WAGNER, F. F., DALY, M. J., CAMPBELL, A. J. & LAL, D. 2020. Comprehensive characterization of amino acid positions in protein structures reveals molecular effect of missense variants. *Proceedings of the National Academy of Sciences*, 117, 28201-28211.
- IRVING, M. 2017. Regulation of Contraction by the Thick Filaments in Skeletal Muscle. *Biophys J*, 113, 2579-2594.
- ISKRATSCH, T., LANGE, S., DWYER, J., KHO, A. L., DOS REMEDIOS, C. & EHLER, E. 2010. Formin follows function: a muscle-specific isoform of FHOD3 is regulated by CK2 phosphorylation and promotes myofibril maintenance. *J Cell Biol*, 191, 1159-72.
- JAMES, J. & ROBBINS, J. 2011. Signaling and myosin-binding protein C. *J Biol Chem*, 286, 9913-9.

- JANSSEN-HEININGER, Y. M., NOLIN, J. D., HOFFMAN, S. M., VAN DER VELDEN, J. L., TULLY, J. E., LAHUE, K. G., ABDALLA, S. T., CHAPMAN, D. G., REYNAERT, N. L., VAN DER VLIET, A. & ANATHY, V. 2013. Emerging mechanisms of glutathione-dependent chemistry in biology and disease. *J Cell Biochem*, 114, 1962-8.
- JANSSEN, P. M. L., CANAN, B. D., KILIC, A., WHITSON, B. A. & BAKER, A. J. 2018. Human Myocardium Has a Robust α 1A-Subtype Adrenergic Receptor Inotropic Response. *J Cardiovasc Pharmacol*, 72, 136-142.
- JEFFRIES, C. M., LU, Y., HYNSON, R. M., TAYLOR, J. E., BALLESTEROS, M., KWAN, A. H. & TREWHELLA, J. 2011. Human cardiac myosin binding protein C: structural flexibility within an extended modular architecture. *J Mol Biol*, 414, 735-48.
- JEFFRIES, C. M., WHITTEN, A. E., HARRIS, S. P. & TREWHELLA, J. 2008. Small-angle X-ray scattering reveals the N-terminal domain organization of cardiac myosin binding protein C. *J Mol Biol*, 377, 1186-99.
- JENSEN, B. C., O'CONNELL, T. D. & SIMPSON, P. C. 2011. Alpha-1-adrenergic receptors: targets for agonist drugs to treat heart failure. *J Mol Cell Cardiol*, 51, 518-28.
- JENSEN, O. N. 2004. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol*, 8, 33-41.
- JUNG, H. S., BURGESS, S. A., BILLINGTON, N., COLEGRAVE, M., PATEL, H., CHALOVICH, J. M., CHANTLER, P. D. & KNIGHT, P. J. 2008a. Conservation of the regulated structure of folded myosin 2 in species separated by at least 600 million years of independent evolution. *Proc Natl Acad Sci U S A*, 105, 6022-6.
- JUNG, H. S., KOMATSU, S., IKEBE, M. & CRAIG, R. 2008b. Head-head and head-tail interaction: a general mechanism for switching off myosin II activity in cells. *Mol Biol Cell*, 19, 3234-42.
- KAMPOURAKIS, T., SUN, Y. B. & IRVING, M. 2016. Myosin light chain phosphorylation enhances contraction of heart muscle via structural changes in both thick and thin filaments. *Proc Natl Acad Sci U S A*, 113, E3039-47.
- KAMPOURAKIS, T., YAN, Z., GAUTEL, M., SUN, Y. B. & IRVING, M. 2014. Myosin binding protein-C activates thin filaments and inhibits thick filaments in heart muscle cells. *Proc Natl Acad Sci U S A*, 111, 18763-8.
- KARCZEWSKI, K. J., FRANCIOLI, L. C., TIAO, G., CUMMINGS, B. B., ALFÖLDI, J., WANG, Q., COLLINS, R. L., LARICCHIA, K. M., GANNA, A., BIRNBAUM, D. P., GAUTHIER, L. D., BRAND, H., SOLOMONSON, M., WATTS, N. A., RHODES, D., SINGER-BERK, M., ENGLAND, E. M., SEABY, E. G., KOSMICKI, J. A., WALTERS, R. K., TASHMAN, K., FARJOUN, Y., BANKS, E., POTERBA, T., WANG, A., SEED, C., WHIFFIN, N., CHONG, J. X., SAMOCHA, K. E., PIERCE-HOFFMAN, E., ZAPPALA, Z., O'DONNELL-LURIA, A. H., MINIKEL, E. V., WEISBURD, B., LEK, M., WARE, J. S., VITTAL, C., ARMEAN, I. M., BERGELSON, L., CIBULSKIS, K., CONNOLLY, K. M., COVARRUBIAS, M., DONNELLY, S., FERRIERA, S., GABRIEL, S., GENTRY, J., GUPTA, N., JEANDET, T., KAPLAN, D., LLANWARNE, C., MUNSHI, R., NOVOD, S., PETRILLO, N., ROAZEN, D., RUANO-RUBIO, V., SALTZMAN, A., SCHLEICHER, M., SOTO, J., TIBBETTS, K., TOLONEN, C., WADE, G., TALKOWSKI, M. E., AGUILAR SALINAS, C. A., AHMAD, T., ALBERT, C. M., ARDISSINO, D., ATZMON, G., BARNARD, J., BEAUGERIE, L., BENJAMIN, E. J., BOEHNKE, M., BONNYCASTLE, L. L., BOTTINGER, E. P., BOWDEN, D. W., BOWN, M. J., CHAMBERS, J. C., CHAN, J. C., CHASMAN, D., CHO, J., CHUNG, M. K., COHEN, B., CORREA, A., DABELEA, D., DALY, M. J., DARBAR, D., DUGGIRALA, R., DUPUIS, J., ELLINOR, P. T., ELOSUA, R., ERDMANN, J., ESKO, T., FÄRKKILÄ, M., FLOREZ, J., FRANKE, A., GETZ, G., GLASER, B., GLATT, S. J., GOLDSTEIN, D., GONZALEZ, C., GROOP, L., et al. 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581, 434-443.
- KAWAS, R. F., ANDERSON, R. L., INGLE, S. R. B., SONG, Y., SRAN, A. S. & RODRIGUEZ, H. M. 2017. A small-molecule modulator of cardiac myosin acts on multiple stages of the myosin chemomechanical cycle. *J Biol Chem*, 292, 16571-16577.

- KENSLER, R. W., CRAIG, R. & MOSS, R. L. 2017. Phosphorylation of cardiac myosin binding protein C releases myosin heads from the surface of cardiac thick filaments. *Proc Natl Acad Sci U S A*, 114, E1355-e1364.
- KENSLER, R. W. & HARRIS, S. P. 2008. The structure of isolated cardiac Myosin thick filaments from cardiac Myosin binding protein-C knockout mice. *Biophys J*, 94, 1707-18.
- KITTLESON, M. D., MEURS, K. M., MUNRO, M. J., KITTLESON, J. A., LIU, S. K., PION, P. D. & TOWBIN, J. A. 1999. Familial hypertrophic cardiomyopathy in maine coon cats: an animal model of human disease. *Circulation*, 99, 3172-80.
- KOHR, M. J., APONTE, A. M., SUN, J., WANG, G., MURPHY, E., GUCEK, M. & STEENBERGEN, C. 2011. Characterization of potential S-nitrosylation sites in the myocardium. *Am J Physiol Heart Circ Physiol*, 300, H1327-35.
- KONNO, T., SHIMIZU, M., INO, H., MATSUYAMA, T., YAMAGUCHI, M., TERA, H., HAYASHI, K., MABUCHI, T., KIYAMA, M., SAKATA, K., HAYASHI, T., INOUE, M., KANEDA, T. & MABUCHI, H. 2003. A novel missense mutation in the myosin binding protein-C gene is responsible for hypertrophic cardiomyopathy with left ventricular dysfunction and dilation in elderly patients. *J Am Coll Cardiol*, 41, 781-6.
- KOOIJ, V., HOLEWINSKI, R. J., MURPHY, A. M. & VAN EYK, J. E. 2013. Characterization of the cardiac myosin binding protein-C phosphoproteome in healthy and failing human hearts. *J Mol Cell Cardiol*, 60, 116-20.
- KORETZ, J. F. 1979. Effects of C-protein on synthetic myosin filament structure. *Biophys J*, 27, 433-46.
- KORTE, F. S., MCDONALD, K. S., HARRIS, S. P. & MOSS, R. L. 2003. Loaded shortening, power output, and rate of force redevelopment are increased with knockout of cardiac myosin binding protein-C. *Circ Res*, 93, 752-8.
- KULIKOVSKAYA, I., MCCLELLAN, G., FLAVIGNY, J., CARRIER, L. & WINEGRAD, S. 2003a. Effect of MyBP-C binding to actin on contractility in heart muscle. *J Gen Physiol*, 122, 761-74.
- KULIKOVSKAYA, I., MCCLELLAN, G., LEVINE, R. & WINEGRAD, S. 2003b. Effect of extraction of myosin binding protein C on contractility of rat heart. *Am J Physiol Heart Circ Physiol*, 285, H857-65.
- KUNST, G., KRESS, K. R., GRUEN, M., UTENWEILER, D., GAUTEL, M. & FINK, R. H. 2000. Myosin binding protein C, a phosphorylation-dependent force regulator in muscle that controls the attachment of myosin heads by its interaction with myosin S2. *Circ Res*, 86, 51-8.
- KUSTER, D. W., SEQUEIRA, V., NAJAFI, A., BOONTJE, N. M., WIJNKER, P. J., WITJAS-PAALBERENDS, E. R., MARSTON, S. B., DOS REMEDIOS, C. G., CARRIER, L., DEMMERS, J. A., REDWOOD, C., SADAYAPPAN, S. & VAN DER VELDEN, J. 2013. GSK3beta phosphorylates newly identified site in the proline-alanine-rich region of cardiac myosin-binding protein C and alters cross-bridge cycling kinetics in human: short communication. *Circ Res*, 112, 633-9.
- LABEIT, S. & KOLMERER, B. 1995. Titins: giant proteins in charge of muscle ultrastructure and elasticity. *Science*, 270, 293-6.
- LANDRUM, M. J., LEE, J. M., BENSON, M., BROWN, G. R., CHAO, C., CHITIPIRALLA, S., GU, B., HART, J., HOFFMAN, D., JANG, W., KARAPETYAN, K., KATZ, K., LIU, C., MADDIPATLA, Z., MALHEIRO, A., MCDANIEL, K., OVETSKY, M., RILEY, G., ZHOU, G., HOLMES, J. B., KATTMAN, B. L. & MAGLOTT, D. R. 2018. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*, 46, D1062-d1067.
- LANGE, S., XIANG, F., YAKOVENKO, A., VIHOLA, A., HACKMAN, P., ROSTKOVA, E., KRISTENSEN, J., BRANDMEIER, B., FRANZEN, G., HEDBERG, B., GUNNARSSON, L. G., HUGHES, S. M., MARCHAND, S., SEJERSEN, T., RICHARD, I., EDSTRÖM, L., EHLE, E., UDD, B. & GAUTEL, M. 2005. The kinase domain of titin controls muscle gene expression and protein turnover. *Science*, 308, 1599-603.
- LAWRENZ, T., BORCHERT, B., LEUNER, C., BARTELSMEIER, M., REINHARDT, J., STRUNK-MUELLER, C., MEYER ZU VILSENDORF, D., SCHLOESSER, M., BEER, G., LIEDER, F., STELLBRINK, C. & KUHN, H. 2011. Endocardial radiofrequency ablation for hypertrophic obstructive cardiomyopathy: acute results and 6 months' follow-up in 19 patients. *J Am Coll Cardiol*, 57, 572-6.

- LEAVIS, P. C. & GERGELY, J. 1984. Thin filament proteins and thin filament-linked regulation of vertebrate muscle contraction. *CRC Crit Rev Biochem*, 16, 235-305.
- LECARPENTIER, Y., VIGNIER, N., OLIVIERO, P., GUELICH, A., CARRIER, L. & COIRAULT, C. 2008. Cardiac Myosin-binding protein C modulates the tuning of the molecular motor in the heart. *Biophys J*, 95, 720-8.
- LEE, K., HARRIS, S. P., SADAYAPPAN, S. & CRAIG, R. 2015. Orientation of myosin binding protein C in the cardiac muscle sarcomere determined by domain-specific immuno-EM. *J Mol Biol*, 427, 274-86.
- LIN, Z., LU, M. H., SCHULTHEISS, T., CHOI, J., HOLTZER, S., DILULLO, C., FISCHMAN, D. A. & HOLTZER, H. 1994. Sequential appearance of muscle-specific proteins in myoblasts as a function of time after cell division: evidence for a conserved myoblast differentiation program in skeletal muscle. *Cell Motil Cytoskeleton*, 29, 1-19.
- LINARI, M., BRUNELLO, E., RECONDITI, M., FUSI, L., CAREMANI, M., NARAYANAN, T., PIAZZESI, G., LOMBARDI, V. & IRVING, M. 2015. Force generation by skeletal muscle is controlled by mechanosensing in myosin filaments. *Nature*, 528, 276-9.
- LINKE, W. A., IVEMEYER, M., LABEIT, S., HINSEN, H., RUEGG, J. C. & GAUTEL, M. 1997. Actin-titin interaction in cardiac myofibrils: probing a physiological role. *Biophys J*, 73, 905-19.
- LINKE, W. A., RUDY, D. E., CENTNER, T., GAUTEL, M., WITT, C., LABEIT, S. & GREGORIO, C. C. 1999. I-band titin in cardiac muscle is a three-element molecular spring and is critical for maintaining thin filament structure. *The Journal of cell biology*, 146, 631-644.
- LIVERSAGE, A. D., HOLMES, D., KNIGHT, P. J., TSKHOVREBOVA, L. & TRINICK, J. 2001. Titin and the sarcomere symmetry paradox. *J Mol Biol*, 305, 401-9.
- LODISH, H., BERK, A., ZIPURSKY, S., MATSUDAIRA, P., BALTIMORE, D. & DARNELL, J. 2000. *Molecular Cell Biology*, New York, W.H. Freeman.
- LOPES, L. R. & ELLIOTT, P. M. 2014. A straightforward guide to the sarcomeric basis of cardiomyopathies. *Heart*, 100, 1916-23.
- LOPES, L. R., RAHMAN, M. S. & ELLIOTT, P. M. 2013. A systematic review and meta-analysis of genotype-phenotype associations in patients with hypertrophic cardiomyopathy caused by sarcomeric protein mutations. *Heart*, 99, 1800-11.
- LORENZ, M., POOLE, K. J., POPP, D., ROSENBAUM, G. & HOLMES, K. C. 1995. An atomic model of the unregulated thin filament obtained by X-ray fiber diffraction on oriented actin-tropomyosin gels. *J Mol Biol*, 246, 108-19.
- LU, Y., KWAN, A. H., JEFFRIES, C. M., GUSS, J. M. & TREWHELLA, J. 2012. The motif of human cardiac myosin-binding protein C is required for its Ca²⁺-dependent interaction with calmodulin. *J Biol Chem*, 287, 31596-607.
- LU, Y., KWAN, A. H., TREWHELLA, J. & JEFFRIES, C. M. 2011. The C0C1 fragment of human cardiac myosin binding protein C has common binding determinants for both actin and myosin. *J Mol Biol*, 413, 908-13.
- LUTHER, P. K., BENNETT, P. M., KNUPP, C., CRAIG, R., PADRÓN, R., HARRIS, S. P., PATEL, J. & MOSS, R. L. 2008. Understanding the organisation and role of myosin binding protein C in normal striated muscle by comparison with MyBP-C knockout cardiac muscle. *J Mol Biol*, 384, 60-72.
- LUTHER, P. K., WINKLER, H., TAYLOR, K., ZOGHBI, M. E., CRAIG, R., PADRÓN, R., SQUIRE, J. M. & LIU, J. 2011. Direct visualization of myosin-binding protein C bridging myosin and actin filaments in intact muscle. *Proc Natl Acad Sci U S A*, 108, 11423-8.
- LYMN, R. W. & TAYLOR, E. W. 1971. Mechanism of adenosine triphosphate hydrolysis by actomyosin. *Biochemistry*, 10, 4617-24.
- MA, H., MARTI-GUTIERREZ, N., PARK, S. W., WU, J., LEE, Y., SUZUKI, K., KOSKI, A., JI, D., HAYAMA, T., AHMED, R., DARBY, H., VAN DYKEN, C., LI, Y., KANG, E., PARK, A. R., KIM, D., KIM, S. T., GONG, J., GU, Y., XU, X., BATTAGLIA, D., KRIEG, S. A., LEE, D. M., WU, D. H., WOLF, D. P., HEITNER, S. B., BELMONTE, J. C. I., AMATO, P., KIM, J. S., KAUL, S. & MITALPOV, S. 2017. Correction of a pathogenic gene mutation in human embryos. *Nature*, 548, 413-419.

- MARIAN, A. J. & BRAUNWALD, E. 2017. Hypertrophic Cardiomyopathy: Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. *Circ Res*, 121, 749-770.
- MARON, B. J. 2002. Hypertrophic cardiomyopathy: a systematic review. *Jama*, 287, 1308-20.
- MARSTON, S., COPELAND, O., GEHMLICH, K., SCHLOSSAREK, S. & CARRIER, L. 2012. How do MYBPC3 mutations cause hypertrophic cardiomyopathy? *J Muscle Res Cell Motil*, 33, 75-80.
- MARSTON, S., COPELAND, O., JACQUES, A., LIVESEY, K., TSANG, V., MCKENNA, W. J., JALILZADEH, S., CARBALLO, S., REDWOOD, C. & WATKINS, H. 2009. Evidence from human myectomy samples that MYBPC3 mutations cause hypertrophic cardiomyopathy through haploinsufficiency. *Circ Res*, 105, 219-22.
- MARUYAMA, K. & EBASHI, S. 1965. Alpha-actinin, a new structural protein from striated muscle. II. Action on actin. *J Biochem*, 58, 13-9.
- MARUYAMA, K., NATORI, R. & NONOMURA, Y. 1976. New elastic protein from muscle. *Nature*, 262, 58-60.
- MASON, P., BAYOL, S. & LOUGHNA, P. T. 1999. The novel sarcomeric protein telethonin exhibits developmental and functional regulation. *Biochem Biophys Res Commun*, 257, 699-703.
- MATSUYAMA, S., KAGE, Y., FUJIMOTO, N., USHIJIMA, T., TSURUDA, T., KITAMURA, K., SHIOSE, A., ASADA, Y., SUMIMOTO, H. & TAKEYA, R. 2018. Interaction between cardiac myosin-binding protein C and formin Fhod3. *Proc Natl Acad Sci U S A*, 115, E4386-e4395.
- MCCLELLAN, G., KULIKOVSKAYA, I. & WINEGRAD, S. 2001. Changes in cardiac contractility related to calcium-mediated changes in phosphorylation of myosin-binding protein C. *Biophys J*, 81, 1083-92.
- MCGRATH, M. J., COTTLE, D. L., NGUYEN, M. A., DYSON, J. M., COGHILL, I. D., ROBINSON, P. A., HOLDSWORTH, M., COWLING, B. S., HARDEMAN, E. C., MITCHELL, C. A. & BROWN, S. 2006. Four and a half LIM protein 1 binds myosin-binding protein C and regulates myosin filament formation and sarcomere assembly. *J Biol Chem*, 281, 7666-83.
- MCKAY, R. T., TRIPET, B. P., HODGES, R. S. & SYKES, B. D. 1997. Interaction of the second binding region of troponin I with the regulatory domain of skeletal muscle troponin C as determined by NMR spectroscopy. *J Biol Chem*, 272, 28494-500.
- MCKILLOP, D. F. & GEEVES, M. A. 1993. Regulation of the interaction between actin and myosin subfragment 1: evidence for three states of the thin filament. *Biophys J*, 65, 693-701.
- MCNAMARA, J. W., LI, A., LAL, S., BOS, J. M., HARRIS, S. P., VAN DER VELDEN, J., ACKERMAN, M. J., COOKE, R. & DOS REMEDIOS, C. G. 2017. MYBPC3 mutations are associated with a reduced super-relaxed state in patients with hypertrophic cardiomyopathy. *PLoS One*, 12, e0180064.
- MCNAMARA, J. W., LI, A., SMITH, N. J., LAL, S., GRAHAM, R. M., KOOIKER, K. B., VAN DIJK, S. J., REMEDIOS, C. G. D., HARRIS, S. P. & COOKE, R. 2016. Ablation of cardiac myosin binding protein-C disrupts the super-relaxed state of myosin in murine cardiomyocytes. *J Mol Cell Cardiol*, 94, 65-71.
- MEINERS, S., DREGER, H., FECHNER, M., BIELER, S., ROTHER, W., GUNTHER, C., BAUMANN, G., STANGL, V. & STANGL, K. 2008. Suppression of cardiomyocyte hypertrophy by inhibition of the ubiquitin-proteasome system. *Hypertension*, 51, 302-8.
- MEURS, K. M., NORGARD, M. M., EDERER, M. M., HENDRIX, K. P. & KITTLESON, M. D. 2007. A substitution mutation in the myosin binding protein C gene in ragdoll hypertrophic cardiomyopathy. *Genomics*, 90, 261-4.
- MEURS, K. M., SANCHEZ, X., DAVID, R. M., BOWLES, N. E., TOWBIN, J. A., REISER, P. J., KITTLESON, J. A., MUNRO, M. J., DRYBURGH, K., MACDONALD, K. A. & KITTLESON, M. D. 2005. A cardiac myosin binding protein C mutation in the Maine Coon cat with familial hypertrophic cardiomyopathy. *Hum Mol Genet*, 14, 3587-93.
- MOHAMED, A. S., DIGNAM, J. D. & SCHLENDER, K. K. 1998. Cardiac myosin-binding protein C (MyBP-C): identification of protein kinase A and protein kinase C phosphorylation sites. *Arch Biochem Biophys*, 358, 313-9.

- MOOLMAN-SMOOK, J., FLASHMAN, E., DE LANGE, W., LI, Z., CORFIELD, V., REDWOOD, C. & WATKINS, H. 2002. Identification of novel interactions between domains of Myosin binding protein-C that are modulated by hypertrophic cardiomyopathy missense mutations. *Circ Res*, 91, 704-11.
- MOOLMAN, J. A., REITH, S., UHL, K., BAILEY, S., GAUTEL, M., JESCHKE, B., FISCHER, C., OCHS, J., MCKENNA, W. J., KLUES, H. & VOSBERG, H. P. 2000. A newly created splice donor site in exon 25 of the MyBP-C gene is responsible for inherited hypertrophic cardiomyopathy with incomplete disease penetrance. *Circulation*, 101, 1396-402.
- MOOS, C. 1981. Fluorescence microscope study of the binding of added C protein to skeletal muscle myofibrils. *J Cell Biol*, 90, 25-31.
- MOOS, C., MASON, C. M., BESTERMAN, J. M., FENG, I. N. & DUBIN, J. H. 1978. The binding of skeletal muscle C-protein to F-actin, and its relation to the interaction of actin with myosin subfragment-1. *J Mol Biol*, 124, 571-86.
- MROWKA, R., PATZAK, A. & HERZEL, H. 2001. Is there a bias in proteome research? *Genome Res*, 11, 1971-3.
- MUELLER, H. & PERRY, S. V. 1962. The degradation of heavy meromyosin by trypsin. *Biochem J*, 85, 431-9.
- MUN, J. Y., GULICK, J., ROBBINS, J., WOODHEAD, J., LEHMAN, W. & CRAIG, R. 2011. Electron microscopy and 3D reconstruction of F-actin decorated with cardiac myosin-binding protein C (cMyBP-C). *J Mol Biol*, 410, 214-25.
- MUN, J. Y., PREVIS, M. J., YU, H. Y., GULICK, J., TOBACMAN, L. S., BECK PREVIS, S., ROBBINS, J., WARSHAW, D. M. & CRAIG, R. 2014. Myosin-binding protein C displaces tropomyosin to activate cardiac thin filaments and governs their speed by an independent mechanism. *Proc Natl Acad Sci U S A*, 111, 2170-5.
- MYBURGH, K. H., FRANKS-SKIBA, K. & COOKE, R. 1995. Nucleotide turnover rate measured in fully relaxed rabbit skeletal muscle myofibrils. *J Gen Physiol*, 106, 957-73.
- NABER, N., COOKE, R. & PATE, E. 2011. Slow myosin ATP turnover in the super-relaxed state in tarantula muscle. *J Mol Biol*, 411, 943-50.
- NADVI, N. A., MICHIE, K. A., KWAN, A. H., GUSS, J. M. & TREWHELLA, J. 2016. Clinically Linked Mutations in the Central Domains of Cardiac Myosin-Binding Protein C with Distinct Phenotypes Show Differential Structural Effects. *Structure*, 24, 105-115.
- NAG, S., TRIVEDI, D. V., SARKAR, S. S., ADHIKARI, A. S., SUNITHA, M. S., SUTTON, S., RUPPEL, K. M. & SPUDICH, J. A. 2017. The myosin mesa and the basis of hypercontractility caused by hypertrophic cardiomyopathy mutations. *Nature structural & molecular biology*, 24, 525-533.
- NAPIERSKI, N. C., GRANGER, K., LANGLAIS, P. R., MORAN, H. R., STROM, J., TOUMA, K. & HARRIS, S. P. 2020. A Novel 'Cut and Paste' Method for In Situ Replacement of cMyBP-C Reveals a New Role for cMyBP-C in the Regulation of Contractile Oscillations. *Circulation Research*, 126, 737-749.
- NAZER, B., GERSTENFELD, E. P., HATA, A., CRUM, L. A. & MATULA, T. J. 2014. Cardiovascular applications of therapeutic ultrasound. *J Interv Card Electrophysiol*, 39, 287-94.
- NIRAULA, T. N., TOCHIO, N., KOSHIBA, S., INOUE, M., KIGAWA, T. & YOKOYAMA, S. 2008. Solution Structure of 4th Immunoglobulin Domain of Slow Type Myosin-Binding Protein C. *TO BE PUBLISHED*.
- NISHIMURA, R. A., OMMEN, S. R. & TAJIK, A. J. 2003. Cardiology patient page. Hypertrophic cardiomyopathy: a patient perspective. *Circulation*, 108, e133-5.
- OAKLEY, C. E., HAMBLY, B. D., CURMI, P. M. & BROWN, L. J. 2004. Myosin binding protein C: structural abnormalities in familial hypertrophic cardiomyopathy. *Cell Res*, 14, 95-110.
- OBERMANN, W. M., PLESSMANN, U., WEBER, K. & FURST, D. O. 1995. Purification and biochemical characterization of myomesin, a myosin-binding and titin-binding protein, from bovine skeletal muscle. *Eur J Biochem*, 233, 110-5.

- OFFER, G., MOOS, C. & STARR, R. 1973. A new protein of the thick filaments of vertebrate skeletal myofibrils. Extractions, purification and characterization. *J Mol Biol*, 74, 653-76.
- OJIMA, K., ONO, Y., DOI, N., YOSHIOKA, K., KAWABATA, Y., LABEIT, S. & SORIMACHI, H. 2007. Myogenic stage, sarcomere length, and protease activity modulate localization of muscle-specific calpain. *J Biol Chem*, 282, 14493-504.
- OKA, T., SAYANO, T., TAMAI, S., YOKOTA, S., KATO, H., FUJII, G. & MIHARA, K. 2008. Identification of a novel protein MICS1 that is involved in maintenance of mitochondrial morphology and apoptotic release of cytochrome c. *Mol Biol Cell*, 19, 2597-608.
- OKAGAKI, T., WEBER, F. E., FISCHMAN, D. A., VAUGHAN, K. T., MIKAWA, T. & REINACH, F. C. 1993. The major myosin-binding domain of skeletal muscle MyBP-C (C protein) resides in the COOH-terminal, immunoglobulin C2 motif. *J Cell Biol*, 123, 619-26.
- ORLOVA, A., GALKIN, V. E., JEFFRIES, C. M., EGELMAN, E. H. & TREWHELLA, J. 2011. The N-terminal domains of myosin binding protein C can bind polymorphically to F-actin. *J Mol Biol*, 412, 379-86.
- PARRY, D. A. & SQUIRE, J. M. 1973. Structural role of tropomyosin in muscle regulation: analysis of the x-ray diffraction patterns from relaxed and contracting muscles. *J Mol Biol*, 75, 33-55.
- PATEL, B. G., WILDER, T. & SOLARO, R. J. 2013. Novel control of cardiac myofilament response to calcium by S-glutathionylation at specific sites of myosin binding protein C. *Front Physiol*, 4, 336.
- PFUHL, M. & GAUTEL, M. 2012. Structure, interactions and function of the N-terminus of cardiac myosin binding protein C (MyBP-C): who does what, with what, and to whom? *J Muscle Res Cell Motil*, 33, 83-94.
- PINOTSIS, N., CHATZIEFTHIMIOU, S. D., BERKEMEIER, F., BEURON, F., MAVRIDIS, I. M., KONAREV, P. V., SVERGUN, D. I., MORRIS, E., RIEF, M. & WILMANN, M. 2012. Superhelical architecture of the myosin filament-linking protein myomesin with unusual elastic properties. *PLoS Biol*, 10, e1001261.
- POHLMANN, L., KRÖGER, I., VIGNIER, N., SCHLOSSAREK, S., KRÄMER, E., COIRAULT, C., SULTAN, K. R., EL-ARMOUCHE, A., WINEGRAD, S., ESCHENHAGEN, T. & CARRIER, L. 2007. Cardiac myosin-binding protein C is required for complete relaxation in intact myocytes. *Circ Res*, 101, 928-38.
- PONNAM, S., SEVRIEVA, I., SUN, Y.-B., IRVING, M. & KAMPOURAKIS, T. 2019. Site-specific phosphorylation of myosin binding protein-C coordinates thin and thick filament activation in cardiac muscle. *Proceedings of the National Academy of Sciences*, 116, 15485-15494.
- POOLE, K. J., LORENZ, M., EVANS, G., ROSENBAUM, G., PIRANI, A., CRAIG, R., TOBACMAN, L. S., LEHMAN, W. & HOLMES, K. C. 2006. A comparison of muscle thin filament models obtained from electron microscopy reconstructions and low-angle X-ray fibre diagrams from non-overlap muscle. *J Struct Biol*, 155, 273-84.
- POTTER, J. D., SHENG, Z., PAN, B. S. & ZHAO, J. 1995. A direct regulatory role for troponin T and a dual role for troponin C in the Ca²⁺ regulation of muscle contraction. *J Biol Chem*, 270, 2557-62.
- PREVIS, M. J., MUN, J. Y., MICHALEK, A. J., PREVIS, S. B., GULICK, J., ROBBINS, J., WARSHAW, D. M. & CRAIG, R. 2016. Phosphorylation and calcium antagonistically tune myosin-binding protein C's structure and function. *Proc Natl Acad Sci U S A*, 113, 3239-44.
- PRONDZYNSKI, M., KRAMER, E., LAUFER, S. D., SHIBAMIYA, A., PLESS, O., FLENNER, F., MULLER, O. J., MUNCH, J., REDWOOD, C., HANSEN, A., PATTEN, M., ESCHENHAGEN, T., MEARINI, G. & CARRIER, L. 2017. Evaluation of MYBPC3 trans-Splicing and Gene Replacement as Therapeutic Options in Human iPSC-Derived Cardiomyocytes. *Mol Ther Nucleic Acids*, 7, 475-486.
- RAHMANSERESHT, S., LEE, K. H., O'LEARY, T. S., MCNAMARA, J. W., SADAYAPPAN, S., ROBBINS, J., WARSHAW, D. M., CRAIG, R. & PREVIS, M. J. 2021. The N terminus of myosin-binding protein C extends toward actin filaments in intact cardiac muscle. *J Gen Physiol*, 153.

- RAJAGOPALA, S. V., HUGHES, K. T. & UETZ, P. 2009. Benchmarking yeast two-hybrid systems using the interactions of bacterial motility proteins. *Proteomics*, 9, 5296-302.
- RATTI, J., ROSTKOVA, E., GAUTEL, M. & PFUHL, M. 2011. Structure and interactions of myosin-binding protein C domain CO: cardiac-specific regulation of myosin at its neck? *J Biol Chem*, 286, 12650-8.
- RAYNAUD, F., ASTIER, C. & BENYAMIN, Y. 2004. Evidence for a direct but sequential binding of titin to tropomyosin and actin filaments. *Biochim Biophys Acta*, 1700, 171-8.
- RAYNAUD, F., FERNANDEZ, E., COULIS, G., AUBRY, L., VIGNON, X., BLEIMLING, N., GAUTEL, M., BENYAMIN, Y. & OUALI, A. 2005. Calpain 1-titin interactions concentrate calpain 1 in the Z-band edges and in the N2-line region within the skeletal myofibril. *Febs j*, 272, 2578-90.
- RAZUMOVA, M. V., BEZOLD, K. L., TU, A.-Y., REGNIER, M. & HARRIS, S. P. 2008. Contribution of the Myosin Binding Protein C Motif to Functional Effects in Permeabilized Rat Trabeculae. *Journal of General Physiology*, 132, 575-585.
- RAZUMOVA, M. V., SHAFFER, J. F., TU, A.-Y., FLINT, G. V., REGNIER, M. & HARRIS, S. P. 2006a. Effects of the N-terminal Domains of Myosin Binding Protein-C in an in Vitro Motility Assay: EVIDENCE FOR LONG-LIVED CROSS-BRIDGES. *Journal of Biological Chemistry*, 281, 35846-35854.
- RAZUMOVA, M. V., SHAFFER, J. F., TU, A. Y., FLINT, G. V., REGNIER, M. & HARRIS, S. P. 2006b. Effects of the N-terminal domains of myosin binding protein-C in an in vitro motility assay: Evidence for long-lived cross-bridges. *J Biol Chem*, 281, 35846-54.
- RHEE, D., SANGER, J. M. & SANGER, J. W. 1994. The premyofibril: evidence for its role in myofibrillogenesis. *Cell Motil Cytoskeleton*, 28, 1-24.
- RIBEIRO EDE, A., JR., PINOTSIS, N., GHISLENI, A., SALMAZO, A., KONAREV, P. V., KOSTAN, J., SJOBLUM, B., SCHREINER, C., POLYANSKY, A. A., GKOUKOULIA, E. A., HOLT, M. R., AACHMANN, F. L., ZAGROVIC, B., BORDIGNON, E., PIRKER, K. F., SVERGUN, D. I., GAUTEL, M. & DJINOVIC-CARUGO, K. 2014. The structure and regulation of human muscle alpha-actinin. *Cell*, 159, 1447-60.
- RICHARD, P., CHARRON, P., CARRIER, L., LEDEUIL, C., CHEAV, T., PICHEREAU, C., BENAICHE, A., ISNARD, R., DUBOURG, O., BURBAN, M., GUEFFET, J. P., MILLAIRE, A., DESNOS, M., SCHWARTZ, K., HAINQUE, B. & KOMAJDA, M. 2003. Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. *Circulation*, 107, 2227-32.
- RODRÍGUEZ-GARCÍA, M. I., MONSERRAT, L., ORTIZ, M., FERNÁNDEZ, X., CAZÓN, L., NÚÑEZ, L., BARRIALES-VILLA, R., MANEIRO, E., VEIRA, E., CASTRO-BEIRAS, A. & HERMIDA-PRIETO, M. 2010. Screening mutations in myosin binding protein C3 gene in a cohort of patients with Hypertrophic Cardiomyopathy. *BMC Med Genet*, 11, 67.
- ROTTBAUER, W., GAUTEL, M., ZEHELEIN, J., LABEIT, S., FRANZ, W. M., FISCHER, C., VOLLRATH, B., MALL, G., DIETZ, R., KUBLER, W. & KATUS, H. A. 1997. Novel splice donor site mutation in the cardiac myosin-binding protein-C gene in familial hypertrophic cardiomyopathy. Characterization Of cardiac transcript and protein. *J Clin Invest*, 100, 475-82.
- ROUX, K. J., KIM, D. I., RAID, M. & BURKE, B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol*, 196, 801-10.
- RUPPEL, K. M. & SPUDICH, J. A. 1996. Structure-function analysis of the motor domain of myosin. *Annu Rev Cell Dev Biol*, 12, 543-73.
- RYBAKOVA, I. N., GREASER, M. L. & MOSS, R. L. 2011. Myosin binding protein C interaction with actin: characterization and mapping of the binding site. *J Biol Chem*, 286, 2008-16.
- SABATER-MOLINA, M., PEREZ-SANCHEZ, I., HERNANDEZ DEL RINCON, J. P. & GIMENO, J. R. 2018. Genetics of hypertrophic cardiomyopathy: A review of current state. *Clin Genet*, 93, 3-14.
- SADAYAPPAN, S. & DE TOMBE, P. P. 2012. Cardiac myosin binding protein-C: redefining its structure and function. *Biophys Rev*, 4, 93-106.

- SADAYAPPAN, S., GULICK, J., OSINSKA, H., BAREFIELD, D., CUELLO, F., AVKIRAN, M., LASKO, V. M., LORENZ, J. N., MAILLET, M., MARTIN, J. L., BROWN, J. H., BERS, D. M., MOLKENTIN, J. D., JAMES, J. & ROBBINS, J. 2011. A critical function for Ser-282 in cardiac Myosin binding protein-C phosphorylation and cardiac function. *Circ Res*, 109, 141-50.
- SADAYAPPAN, S., GULICK, J., OSINSKA, H., MARTIN, L. A., HAHN, H. S., DORN, G. W., 2ND, KLEVITSKY, R., SEIDMAN, C. E., SEIDMAN, J. G. & ROBBINS, J. 2005. Cardiac myosin-binding protein-C phosphorylation and cardiac function. *Circ Res*, 97, 1156-63.
- SADAYAPPAN, S., OSINSKA, H., KLEVITSKY, R., LORENZ, J. N., SARGENT, M., MOLKENTIN, J. D., SEIDMAN, C. E., SEIDMAN, J. G. & ROBBINS, J. 2006. Cardiac myosin binding protein C phosphorylation is cardioprotective. *Proc Natl Acad Sci U S A*, 103, 16918-23.
- SAHDEV, S., KHATTAR, S. K. & SAINI, K. S. 2008. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem*, 307, 249-64.
- SANCHEZ, O., CAMPUZANO, O., FERNÁNDEZ-FALGUERAS, A., SARQUELLA-BRUGADA, G., CESAR, S., MADEMONT, I., MATES, J., PÉREZ-SERRA, A., COLL, M., PICO, F., IGLESIAS, A., TIRÓN, C., ALLEGUE, C., CARRO, E., GALLEGU, M., FERRER-COSTA, C., HOSPITAL, A., BARDALET, N., BORONDO, J. C., VINGUT, A., ARBELO, E., BRUGADA, J., CASTELLÀ, J., MEDALLO, J. & BRUGADA, R. 2016. Natural and Undetermined Sudden Death: Value of Post-Mortem Genetic Investigation. *PLoS One*, 11, e0167358.
- SCHLENDER, K. K. & BEAN, L. J. 1991. Phosphorylation of chicken cardiac C-protein by calcium/calmodulin-dependent protein kinase II. *J Biol Chem*, 266, 2811-7.
- SCHLENDER, K. K., HEGAZY, M. G. & THYSERIL, T. J. 1987. Dephosphorylation of cardiac myofibril C-protein by protein phosphatase 1 and protein phosphatase 2A. *Biochim Biophys Acta*, 928, 312-9.
- SCHLOSSAREK, S., MEARINI, G. & CARRIER, L. 2011. Cardiac myosin-binding protein C in hypertrophic cardiomyopathy: mechanisms and therapeutic opportunities. *J Mol Cell Cardiol*, 50, 613-20.
- SCHLOSSAREK, S., SINGH, S. R., GEERTZ, B., SCHULZ, H., REISCHMANN, S., HÜBNER, N. & CARRIER, L. 2014. Proteasome inhibition slightly improves cardiac function in mice with hypertrophic cardiomyopathy. *Front Physiol*, 5, 484.
- SCHULTHEISS, T., LIN, Z. X., LU, M. H., MURRAY, J., FISCHMAN, D. A., WEBER, K., MASAKI, T., IMAMURA, M. & HOLTZER, H. 1990. Differential distribution of subsets of myofibrillar proteins in cardiac nonstriated and striated myofibrils. *J Cell Biol*, 110, 1159-72.
- SEARS, R. M., MAY, D. G. & ROUX, K. J. 2019. BioID as a Tool for Protein-Proximity Labeling in Living Cells. *Methods Mol Biol*, 2012, 299-313.
- SEEBOHM, B., MATINMEHR, F., KÖHLER, J., FRANCINO, A., NAVARRO-LOPÉZ, F., PERROT, A., OZCELIK, C., MCKENNA, W. J., BRENNER, B. & KRAFT, T. 2009. Cardiomyopathy mutations reveal variable region of myosin converter as major element of cross-bridge compliance. *Biophys J*, 97, 806-24.
- SEIBEL, N. M., ELJOUNI, J., NALASKOWSKI, M. M. & HAMPE, W. 2007. Nuclear localization of enhanced green fluorescent protein homomultimers. *Anal Biochem*, 368, 95-9.
- SEIDMAN, J. G. & SEIDMAN, C. 2001. The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. *Cell*, 104, 557-67.
- SELLERS, J. R. 2000. Myosins: a diverse superfamily. *Biochim Biophys Acta*, 1496, 3-22.
- SEMSARIAN, C., INGLES, J., MARON, M. S. & MARON, B. J. 2015. New perspectives on the prevalence of hypertrophic cardiomyopathy. *J Am Coll Cardiol*, 65, 1249-1254.
- SEQUEIRA, V., NIJENKAMP, L. L., REGAN, J. A. & VAN DER VELDEN, J. 2014. The physiological role of cardiac cytoskeleton and its alterations in heart failure. *Biochim Biophys Acta*, 1838, 700-22.
- SHAFFER, J. F., KENSLER, R. W. & HARRIS, S. P. 2009. The myosin-binding protein C motif binds to F-actin in a phosphorylation-sensitive manner. *J Biol Chem*, 284, 12318-27.
- SHAFFER, J. F. J., W LEARY, J.A. HARRIS S.P. 2009. PKA Phosphorylates Serine 307 of Murine Cardiac Myosin Binding Protein-C In Vitro. *biophys. J.*, 96.

- SHEIKH, F., RASKIN, A., CHU, P. H., LANGE, S., DOMENIGHETTI, A. A., ZHENG, M., LIANG, X., ZHANG, T., YAJIMA, T., GU, Y., DALTON, N. D., MAHATA, S. K., DORN, G. W., 2ND, BROWN, J. H., PETERSON, K. L., OMENS, J. H., MCCULLOCH, A. D. & CHEN, J. 2008. An FHL1-containing complex within the cardiomyocyte sarcomere mediates hypertrophic biomechanical stress responses in mice. *J Clin Invest*, 118, 3870-80.
- SMELTER, D. F., DE LANGE, W. J., CAI, W., GE, Y. & RALPHE, J. C. 2018. The HCM-linked W792R mutation in cardiac myosin-binding protein C reduces C6 FnIII domain stability. *Am J Physiol Heart Circ Physiol*, 314, H1179-h1191.
- SOMMESE, R. F., SUNG, J., NAG, S., SUTTON, S., DEACON, J. C., CHOE, E., LEINWAND, L. A., RUPPEL, K. & SPUDICH, J. A. 2013. Molecular consequences of the R453C hypertrophic cardiomyopathy mutation on human β -cardiac myosin motor function. *Proc Natl Acad Sci U S A*, 110, 12607-12.
- SOTERIOU, A., GAMAGE, M. & TRINICK, J. 1993. A survey of interactions made by the giant protein titin. *J Cell Sci*, 104 (Pt 1), 119-23.
- SPOLADORE, R., FRAGASSO, G., PANNONE, L., SLAVICH, M. & MARGONATO, A. 2020. Pharmacotherapy for the treatment of obstructive hypertrophic cardiomyopathy. *Expert Opin Pharmacother*, 21, 233-242.
- SPUDICH, J. A. 2015. The myosin mesa and a possible unifying hypothesis for the molecular basis of human hypertrophic cardiomyopathy. *Biochem Soc Trans*, 43, 64-72.
- SPUDICH, J. A., AKSEL, T., BARTHOLOMEW, S. R., NAG, S., KAWANA, M., YU, E. C., SARKAR, S. S., SUNG, J., SOMMESE, R. F., SUTTON, S., CHO, C., ADHIKARI, A. S., TAYLOR, R., LIU, C., TRIVEDI, D. & RUPPEL, K. M. 2016. Effects of hypertrophic and dilated cardiomyopathy mutations on power output by human β -cardiac myosin. *J Exp Biol*, 219, 161-7.
- SQUIRE, J. M., LUTHER, P. K. & KNUPP, C. 2003. Structural evidence for the interaction of C-protein (MyBP-C) with actin and sequence identification of a possible actin-binding domain. *J Mol Biol*, 331, 713-24.
- SQUIRE, J. M., ROESSLE, M. & KNUPP, C. 2004. New X-ray diffraction observations on vertebrate muscle: organisation of C-protein (MyBP-C) and troponin and evidence for unknown structures in the vertebrate A-band. *J Mol Biol*, 343, 1345-63.
- STANSFIELD, W. E., TANG, R. H., MOSS, N. C., BALDWIN, A. S., WILLIS, M. S. & SELZMAN, C. H. 2008. Proteasome inhibition promotes regression of left ventricular hypertrophy. *Am J Physiol Heart Circ Physiol*, 294, H645-50.
- STARR, R. & OFFER, G. 1978. The interaction of C-protein with heavy meromyosin and subfragment-2. *Biochem J*, 171, 813-6.
- STATHOPOULOU, K., WITTIG, I., HEIDLER, J., PIASECKI, A., RICHTER, F., DIERING, S., VAN DER VELDEN, J., BUCK, F., DONZELLI, S., SCHRÖDER, E., WIJNKER, P. J., VOIGT, N., DOBREV, D., SADAYAPPAN, S., ESCHENHAGEN, T., CARRIER, L., EATON, P. & CUELLO, F. 2016. S-glutathiolation impairs phosphoregulation and function of cardiac myosin-binding protein C in human heart failure. *Faseb j*, 30, 1849-64.
- STELLBERGER, T., HÄUSER, R., BAIKER, A., POTHINENI, V. R., HAAS, J. & UETZ, P. 2010. Improving the yeast two-hybrid system with permuted fusions proteins: the Varicella Zoster Virus interactome. *Proteome Sci*, 8, 8.
- STELZER, J. E., FITZSIMONS, D. P. & MOSS, R. L. 2006. Ablation of myosin-binding protein-C accelerates force development in mouse myocardium. *Biophys J*, 90, 4119-27.
- STEWART, M. A., FRANKS-SKIBA, K., CHEN, S. & COOKE, R. 2010. Myosin ATP turnover rate is a mechanism involved in thermogenesis in resting skeletal muscle fibers. *Proc Natl Acad Sci U S A*, 107, 430-5.
- STÖHR, A., FRIEDRICH, F. W., FLENNER, F., GEERTZ, B., EDER, A., SCHAAF, S., HIRT, M. N., UEBELER, J., SCHLOSSAREK, S., CARRIER, L., HANSEN, A. & ESCHENHAGEN, T. 2013. Contractile abnormalities and altered drug response in engineered heart tissue from Mybpc3-targeted knock-in mice. *J Mol Cell Cardiol*, 63, 189-98.

- SULIMAN, B. A., XU, D. & WILLIAMS, B. R. 2012. The promyelocytic leukemia zinc finger protein: two decades of molecular oncology. *Front Oncol*, 2, 74.
- SZENT-GYORGYI, A. G. 1953. Meromyosins, the subunits of myosin. *Arch Biochem Biophys*, 42, 305-20.
- TABUSE, Y., IZUMI, Y., PIANO, F., KEMPHUES, K. J., MIWA, J. & OHNO, S. 1998. Atypical protein kinase C cooperates with PAR-3 to establish embryonic polarity in *Caenorhabditis elegans*. *Development*, 125, 3607-14.
- TANOKURA, M., TAWADA, Y., ONO, A. & OHTSUKI, I. 1983. Chymotryptic subfragments of troponin T from rabbit skeletal muscle. Interaction with tropomyosin, troponin I and troponin C. *J Biochem*, 93, 331-7.
- TARDIFF, J. C. 2011. Thin filament mutations: developing an integrative approach to a complex disorder. *Circ Res*, 108, 765-82.
- TEARE, D. 1958. Asymmetrical hypertrophy of the heart in young adults. *Br Heart J*, 20, 1-8.
- THEIS, J. L., BOS, J. M., THEIS, J. D., MILLER, D. V., DEARANI, J. A., SCHAFF, H. V., GERSH, B. J., OMMEN, S. R., MOSS, R. L. & ACKERMAN, M. J. 2009. Expression patterns of cardiac myofilament proteins: genomic and protein analysis of surgical myectomy tissue from patients with obstructive hypertrophic cardiomyopathy. *Circ Heart Fail*, 2, 325-33.
- TIBERTI, M., LECHNER, B. D. & FORNILI, A. 2019. Binding Pockets in Proteins Induced by Mechanical Stress. *J Chem Theory Comput*, 15, 1-6.
- TROMBITAS, K., REDKAR, A., CENTNER, T., WU, Y., LABEIT, S. & GRANZIER, H. 2000. Extensibility of isoforms of cardiac titin: variation in contour length of molecular subsegments provides a basis for cellular passive stiffness diversity. *Biophys J*, 79, 3226-34.
- TRYBUS, K. M. & TAYLOR, E. W. 1980. Kinetic studies of the cooperative binding of subfragment 1 to regulated actin. *Proc Natl Acad Sci U S A*, 77, 7209-13.
- UHLÉN, M., FAGERBERG, L., HALLSTRÖM, B. M., LINDSKOG, C., OKSVOLD, P., MARDINOGLU, A., SIVERTSSON, Å., KAMPF, C., SJÖSTEDT, E., ASPLUND, A., OLSSON, I., EDLUND, K., LUNDBERG, E., NAVANI, S., SZIGYARTO, C. A., ODEBERG, J., DJUREINOVIC, D., TAKANEN, J. O., HOBER, S., ALM, T., EDQVIST, P. H., BERLING, H., TEGEL, H., MULDER, J., ROCKBERG, J., NILSSON, P., SCHWENK, J. M., HAMSTEN, M., VON FEILITZEN, K., FORSBERG, M., PERSSON, L., JOHANSSON, F., ZWAHLEN, M., VON HEIJNE, G., NIELSEN, J. & PONTÉN, F. 2015. Proteomics. Tissue-based map of the human proteome. *Science*, 347, 1260419.
- USHIJIMA, T., FUJIMOTO, N., MATSUYAMA, S., KAN, O. M., KIYONARI, H., SHIOI, G., KAGE, Y., YAMASAKI, S., TAKEYA, R. & SUMIMOTO, H. 2018. The actin-organizing formin protein Fhod3 is required for postnatal development and functional maintenance of the adult heart in mice. *J Biol Chem*, 293, 148-162.
- VAN CRIEKINGE, W. & BEYAERT, R. 1999. Yeast Two-Hybrid: State of the Art. *Biological procedures online*, 2, 1-38.
- VAN DIJK, S. J., BEZOLD KOOIKER, K., MAZZALUPO, S., YANG, Y., KOSTYUKOVA, A. S., MUSTACICH, D. J., HOYE, E. R., STERN, J. A., KITTLESON, M. D. & HARRIS, S. P. 2016. The A31P missense mutation in cardiac myosin binding protein C alters protein structure but does not cause haploinsufficiency. *Arch Biochem Biophys*, 601, 133-40.
- VAN DIJK, S. J., DOOIJES, D., DOS REMEDIOS, C., MICHELS, M., LAMERS, J. M., WINEGRAD, S., SCHLOSSAREK, S., CARRIER, L., TEN CATE, F. J., STIENEN, G. J. & VAN DER VELDEN, J. 2009. Cardiac myosin-binding protein C mutations and hypertrophic cardiomyopathy: haploinsufficiency, deranged phosphorylation, and cardiomyocyte dysfunction. *Circulation*, 119, 1473-83.
- VAN DRIEST, S. L., VASILE, V. C., OMMEN, S. R., WILL, M. L., TAJIK, A. J., GERSH, B. J. & ACKERMAN, M. J. 2004. Myosin binding protein C mutations and compound heterozygosity in hypertrophic cardiomyopathy. *J Am Coll Cardiol*, 44, 1903-10.
- VIBERT, P., CRAIG, R. & LEHMAN, W. 1997. Steric-model for activation of muscle thin filaments. *J Mol Biol*, 266, 8-14.

- VIGNIER, N., SCHLOSSAREK, S., FRAYSSE, B., MEARINI, G., KRÄMER, E., POINTU, H., MOUGENOT, N., GUIARD, J., REIMER, R., HOHENBERG, H., SCHWARTZ, K., VERNET, M., ESCHENHAGEN, T. & CARRIER, L. 2009. Nonsense-mediated mRNA decay and ubiquitin-proteasome system regulate cardiac myosin-binding protein C mutant levels in cardiomyopathic mice. *Circ Res*, 105, 239-48.
- VON DER ECKEN, J., MULLER, M., LEHMAN, W., MANSTEIN, D. J., PENCZEK, P. A. & RAUNSER, S. 2015. Structure of the F-actin-tropomyosin complex. *Nature*, 519, 114-7.
- VYDYANATH, A., GURNETT, C. A., MARSTON, S. & LUTHER, P. K. 2012. Axial distribution of myosin binding protein-C is unaffected by mutations in human cardiac and skeletal muscle. *J Muscle Res Cell Motil*, 33, 61-74.
- WALLIMANN, T., WYSS, M., BRDICZKA, D., NICOLAY, K. & EPPENBERGER, H. M. 1992. Intracellular compartmentation, structure and function of creatine kinase isoenzymes in tissues with high and fluctuating energy demands: the 'phosphocreatine circuit' for cellular energy homeostasis. *Biochemical Journal*, 281, 21-40.
- WANG, J., WAN, K., SUN, J., LI, W., LIU, H., HAN, Y. & CHEN, Y. 2018a. Phenotypic diversity identified by cardiac magnetic resonance in a large hypertrophic cardiomyopathy family with a single MYH7 mutation. *Scientific Reports*, 8, 973.
- WANG, K., MCCLURE, J. & TU, A. 1979. Titin: major myofibrillar components of striated muscle. *Proc Natl Acad Sci U S A*, 76, 3698-702.
- WANG, L., GEIST, J., GROGAN, A., HU, L. R. & KONTROGIANNI-KONSTANTOPOULOS, A. 2018b. Thick Filament Protein Network, Functions, and Disease Association. *Compr Physiol*, 8, 631-709.
- WEBER, F. E., VAUGHAN, K. T., REINACH, F. C. & FISCHMAN, D. A. 1993. Complete sequence of human fast-type and slow-type muscle myosin-binding-protein C (MyBP-C). Differential expression, conserved domain structure and chromosome assignment. *Eur J Biochem*, 216, 661-9.
- WEEDS, A. G. & LOWEY, S. 1971. Substructure of the myosin molecule. II. The light chains of myosin. *J Mol Biol*, 61, 701-25.
- WEISBERG, A. & WINEGRAD, S. 1996. Alteration of myosin cross bridges by phosphorylation of myosin-binding protein C in cardiac muscle. *Proc Natl Acad Sci U S A*, 93, 8999-9003.
- WHITTEN, A. E., JEFFRIES, C. M., HARRIS, S. P. & TREWHELLA, J. 2008. Cardiac myosin-binding protein C decorates F-actin: implications for cardiac function. *Proc Natl Acad Sci U S A*, 105, 18360-5.
- WILLOTT, R. H., GOMES, A. V., CHANG, A. N., PARVATIYAR, M. S., PINTO, J. R. & POTTER, J. D. 2010. Mutations in Troponin that cause HCM, DCM AND RCM: what can we learn about thin filament function? *J Mol Cell Cardiol*, 48, 882-92.
- WITT, C. C., GERULL, B., DAVIES, M. J., CENTNER, T., LINKE, W. A. & THIERFELDER, L. 2001. Hypercontractile properties of cardiac muscle fibers in a knock-in mouse model of cardiac myosin-binding protein-C. *J Biol Chem*, 276, 5353-9.
- WOLFE, J. E., ISHIWATA, S. I., BRAET, F., WHAN, R., SU, Y., LAL, S. & DOS REMEDIOS, C. G. 2011. SPontaneous Oscillatory Contraction (SPOC): auto-oscillations observed in striated muscle at partial activation. *Biophysical Reviews*, 3, 53-62.
- WOODHEAD, J. L., ZHAO, F. Q., CRAIG, R., EGELMAN, E. H., ALAMO, L. & PADRÓN, R. 2005. Atomic model of a myosin filament in the relaxed state. *Nature*, 436, 1195-9.
- XIONG, Y., UYS, J. D., TEW, K. D. & TOWNSEND, D. M. 2011. S-glutathionylation: from molecular mechanisms to health outcomes. *Antioxid Redox Signal*, 15, 233-70.
- YAMADA, Y., NAMBA, K. & FUJII, T. 2020. Cardiac muscle thin filament structures reveal calcium regulatory mechanism. *Nature Communications*, 11, 153.
- YAMAMOTO, K. 1986. The binding of skeletal muscle C-protein to regulated actin. *FEBS Lett*, 208, 123-7.
- YANG, F., AIELLO, D. L. & PYLE, W. G. 2008. Cardiac myofilament regulation by protein phosphatase type 1alpha and CapZ. *Biochem Cell Biol*, 86, 70-8.

- YANG, Q., OSINSKA, H., KLEVITSKY, R. & ROBBINS, J. 2001. Phenotypic deficits in mice expressing a myosin binding protein C lacking the titin and myosin binding domains. *J Mol Cell Cardiol*, 33, 1649-58.
- YANG, Q., SANBE, A., OSINSKA, H., HEWETT, T. E., KLEVITSKY, R. & ROBBINS, J. 1998. A mouse model of myosin binding protein C human familial hypertrophic cardiomyopathy. *J Clin Invest*, 102, 1292-300.
- YANG, Q., SANBE, A., OSINSKA, H., HEWETT, T. E., KLEVITSKY, R. & ROBBINS, J. 1999. In vivo modeling of myosin binding protein C familial hypertrophic cardiomyopathy. *Circ Res*, 85, 841-7.
- YANG, X., PABON, L. & MURRY, C. E. 2014. Engineering adolescence: maturation of human pluripotent stem cell-derived cardiomyocytes. *Circ Res*, 114, 511-23.
- YASUDA, M., KOSHIDA, S., SATO, N. & OBINATA, T. 1995. Complete primary structure of chicken cardiac C-protein (MyBP-C) and its expression in developing striated muscles. *J Mol Cell Cardiol*, 27, 2275-86.
- YOKOTA, T., OTSUKA, T., MOSMANN, T., BANCHEREAU, J., DEFRANCE, T., BLANCHARD, D., DE VRIES, J. E., LEE, F. & ARAI, K. 1986. Isolation and characterization of a human interleukin cDNA clone, homologous to mouse B-cell stimulatory factor 1, that expresses B-cell- and T-cell-stimulating activities. *Proc Natl Acad Sci U S A*, 83, 5894-8.
- YOUNG, P., EHRLER, E. & GAUTEL, M. 2001. Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor protein involved in sarcomere assembly. *J Cell Biol*, 154, 123-36.
- YOUNG, P., FERGUSON, C., BANUELOS, S. & GAUTEL, M. 1998. Molecular structure of the sarcomeric Z-disk: two types of titin interactions lead to an asymmetrical sorting of alpha-actinin. *Embo j*, 17, 1614-24.
- YOUNG, V. R. & AJAMI, A. M. 2000. Glutamate: an amino acid of particular distinction. *J Nutr*, 130, 892s-900s.
- YU, B., FRENCH, J. A., CARRIER, L., JEREMY, R. W., MCTAGGART, D. R., NICHOLSON, M. R., HAMBLY, B., SEMSARIAN, C., RICHMOND, D. R., SCHWARTZ, K. & TRENT, R. J. 1998a. Molecular pathology of familial hypertrophic cardiomyopathy caused by mutations in the cardiac myosin binding protein C gene. *J Med Genet*, 35, 205-10.
- YU, B., FRENCH, J. A., JEREMY, R. W., FRENCH, P., MCTAGGART, D. R., NICHOLSON, M. R., SEMSARIAN, C., RICHMOND, D. R. & TRENT, R. J. 1998b. Counselling issues in familial hypertrophic cardiomyopathy. *J Med Genet*, 35, 183-8.
- YUAN, C., SHENG, Q., TANG, H., LI, Y., ZENG, R. & SOLARO, R. J. 2008. Quantitative comparison of sarcomeric phosphoproteomes of neonatal and adult rat hearts. *Am J Physiol Heart Circ Physiol*, 295, H647-56.
- ZAREMBA, R., MERKUS, D., HAMDANI, N., LAMERS, J. M., PAULUS, W. J., DOS REMEDIOS, C., DUNCKER, D. J., STIENEN, G. J. & VAN DER VELDEN, J. 2007. Quantitative analysis of myofilament protein phosphorylation in small cardiac biopsies. *Proteomics Clin Appl*, 1, 1285-90.
- ZHANG, X. L., DE, S., MCINTOSH, L. P. & PAETZEL, M. 2014. Structural characterization of the C3 domain of cardiac myosin binding protein C and its hypertrophic cardiomyopathy-related R502W mutant. *Biochemistry*, 53, 5332-42.
- ZHOU, Z. & FU, X. D. 2013. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma*, 122, 191-207.
- ZOGHBI, M. E., WOODHEAD, J. L., MOSS, R. L. & CRAIG, R. 2008. Three-dimensional structure of vertebrate cardiac muscle myosin filaments. *Proc Natl Acad Sci U S A*, 105, 2386-90.