Investigating ligands of cardiac Myosin-Binding Protein C (cMyBPC) as potential regulators of contractility and modifiers of hypertrophy.

C.C.A. Swanepoel

Dissertation presented for the Degree of Doctor of Philosophy in Medical Sciences (Molecular Biology) at the University of Stellenbosch

Promoter: Prof Johanna C. Moolman-Smook

December 2011

Faculty of Health Sciences

Department of Biomedical Sciences

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and has not previously in its entirety or in part been submitted at any university for a degree.

Signature……Date…….

This is lovingly dedicated, in memory of my uncle, Thomas Wilschutte and good friend, Rowena Keyser, who passed away before the completion of this dissertation. Your presence and spirit will remain in my heart and you'll always be remembered.

ABSTRACT

The regulation of cardiac contractility is dependent on cooperative interaction between the thick and thin filaments, as well as their accessory proteins, within the cardiac sarcomere. Alteration in cardiac contractility due to a defective sarcomere typically results in cardiomyopathies, such as hypertrophic cardiomyopathy (HCM). One of the sarcomeric genes frequently mutated and which accounts for the second most common form of HCM encodes cardiac myosin binding protein C (cMyBPC), a thick filament accessory protein whose physiological function is poorly understood. However, studies have implicated cMyBPC in thick filament structure and function as well as in the regulation of contractility. The N-terminal region of cMyBPC houses the cMyBPC-motif, which contains three phosphorylation sites, between domains C1 and C2. The hierarchical phosphorylation of this motif, by first calcium/calmodulin kinase II (CamKII) and then by cyclic AMP-activated protein kinase (PKA), is cardinal in the role of cMyBPC in the regulation of cardiac contractility in response to β-adrenergic stimulation. Moreover, phosphorylation of this motif is inversely correlated to cMyBPC proteolysis and has been shown to be cardioprotective. Thus, proteins that have an effect on cMyBPC function or turnover may also influence filament structure and hence affect contractility, which, in turn, affects the structure of the cardiac muscle.

One such protein is the Copper metabolism MURR1-domain containing protein 4 (COMMD4), which was previously identified as a novel interactor of cMyBPC during a yeast two-hybrid (Y2H) library screen in our laboratory. COMMD4 binds specifically to the cMyBPC motif in a phosphorylation-dependent manner. The exact function of COMMD4 is unknown; however, it is a member of the COMM family of proteins that has been linked to copper metabolism as well as to the ubiquitin-proteasome pathway (UPS). Intriguingly, recent studies have shown that the UPS plays a role in cMyBPC-derived HCM, while dietary copper depletion is also known to cause cardiac hypertrophy. Based on these findings, COMMD4 was considered an interesting candidate regulator of sarcomeric function and contractility, and by extension, a candidate modifier of cardiac hypertrophy.

Thus, the aim of the present study was two-fold. Firstly, COMMD4 was used as bait in a Y2H library screen to determine its distal ligands, with a view to further elucidate its function, particularly in the context of MyBPC functioning, and identified interactors were subjected to further *in vitro* and *in vivo* verification studies. Also, the phosphorylation-dependent nature of the interaction between COMMD4 and cMyBPC was further investigated using a domain/phosphorylation assay. Secondly, COMMD4 and its Y2H-identified putative interactors were assessed as possible modifiers of hypertrophy in a family-based association study, using three cohorts of South African HCM-families in which one of three founder mutations segregate.

Six putative interactors, viz. cardiac actin (ACTC1), Down syndrome critical region 3 (DSCR3), enolase 1 (ENO1), F-box and leucine rich repeat protein 10 (FBXL10), legumain (LGMN) and sorting nexin3 (SNX3) were identified and confirmed as COMMD4 interactors using Y2H analyses, followed by *in vitro* and *in vivo* co-immunoprecipitation and 3D co-localisation assays. Moreover, as some COMMD protein family members and the newly-identified interactors of COMMD4 have previously been linked to the UPS, the

functional effect of siRNA-mediated knockdown of COMMD4 on cMyBPC turnover was also investigated. Data revealed accumulation of cMyBPC in the endosomes upon COMMD4 knockdown, suggesting a functional role for COMMD4 in the turnover of cMyBPC. In addition, association analysis revealed strong evidence of association between various single nucleotide polymorphisms (SNPs) in *SNX3* and a number of hypertrophy traits, thus suggesting a role for *SNX3* as a candidate modifier of hypertrophy in HCM. No evidence of association was observed for any of the genes encoding the other COMMD4 interactors implicated in protein turnover.

The present study demonstrates that COMMD4, a little understood member of the COMM family of proteins, binds to the cMyBPC motif of cMyBPC in a phosphorylation-dependent manner. Furthermore, based on the functions of its protein interactions, we hypothesise that COMMD4 plays a role in protein trafficking and turnover. More specifically, COMMD4 seems to help to facilitate formation of protein complexes with the Skp1-Cul1-Fbxl (SCF) E3 ubiquitin ligase and probably helps to stabilise the target substrate for subsequent ubiquitin-conjugation. As COMMD4 seems to affect the protein turnover of cMyBPC and possibly other sarcomeric proteins, such as actin, these results establish a novel association between the sarcomere, HCM and the UPS. In addition, identification of *SNX3* as a hypertrophy modifier will allow for the improved understanding of HCM patho-aetiology. *SNX3* thus adds to the growing body of sarcomeric modifier genes, which, eventually, may improve risk profiling in HCM. Furthermore, as genetic modifiers appear sufficient to completely prevent disease expression in some HCM carriers, the identification of *SNX3* may point to the protein turnover pathway as a potential new target for intervention.

OPSOMMING

Die regulering van kardiale kontraktiliteit is afhanklik van die koöperatiewe interaksie tussen die dik en dun filamente, asook hul geassosieerde proteïene, in die kardiale sarkomeer. Veranderinge in kardiale kontraktiliteit as gevolg van „n defektiewe sarkomeer lei tot kardiomiopatieë soos hipertrofiese kardiomiopatie (HKM). Een van die sarkomeriese gene wat dikwels gemuteer is en wat verantwoordelik is vir die tweede algmeenste vorm van HKM,is dié van kardiale miosien-bindingsproteïen C (cMyBPC),'n proteïen geassosieer met die dik filament waarvan die fisiologiese funksie nog nie goed bekend is nie. Studies betrek cMyBPC in dik filament struktuur en funksie asook in die regulering van kontraktiliteit. Die N-terminale gebied van cMyBPC huisves die cMyBPC-motief, wat drie fosforilerings-setels tussen domeine C1 en C2 bevat. Die hiërargiese fosforilering van hierdie motief, eerstens deur kalsium/kalmodulien-gereguleerde kinase II (CamKII), gevolg deur siklies AMP-geaktiveerde proteïen kinase (PKA), is kardinaal in die rol van cMyBPC in die regulering van kardiale kontraktiliteit in reaksie op β-adrenergiese stimulasie. Verder, fosforilering van hierdie motief is omgekeerd gekorreleer aan cMyBPC proteolise en is ook bewys om kardiobeskermend te wees. Dus, proteïene wat „n uitwerking het op die funksie van cMyBPC mag ook filament struktuur en kontraktiliteit beïnvloed, wat op hul beurt die struktuur van die kardiale spier affekteer.

Die koper metabolisme MURR1-domein bevattende protein 4 (COMMD4), was voorheen geïdentifiseer as „n nuwe bindingsgenoot van cMyBPC tydens gis twee-hibried (G2H) analise in ons laboratorium. COMMD4 bind spesifiek aan die cMyBPC motief in „n fosforilasie afhanklike wyse. Die presiese funksie van COMMD4 is onbekend; maar dit is „n lid van die COMM domein familie van proteine wat geassosieerd is met koper metabolisme sowel as die "ubiquitin" proteosoom pad (UPP). Interesant genoeg, onlangse studies het bewys dat die UPP „n rol speel in cMyBPC-afgeleide HKM, terwyl koper uitputting in die dieet ook bekend is om kardiale hipertrofie te veroorsaak. Gebaseer op hierdie bevindinge was COMMD4 oorweeg as „n interessante kandidaat reguleerder van sarkomeries funksie en kontraktiliteit, asook „n kandidaat modifiseerder van kardiale hipertrofie.

Dus, die doel van die huidige studie was tweeledig. Eerstens, was COMMD4 as aas gebruik in 'n G2H biblioteek sifting om sy distale ligande te bepaal, met die oog om verdere lig te werp op sy funksie, veral in die konteks van MyBPC funksionering, en geïdentifiseerde bindingsgenote was onderwerp aan verdere „*in vitro*' en „*in vivo*' verifikasie studies. Daarbenewens was die fosforilering-afhanklike aard van die interaksie tussen COMMD4 en cMyBPC verder ondersoek met behulp van 'n domein/fosforilasie toets. Tweedens, COMMD4 en sy G2H-geïdentifiseerde vermeende bindingsgenote was geassesseer as moontlik modifiseerders van hipertrofie in 'n familie-gebaseerde assosiasie studie, met behulp van drie kohorte van Suid-Afrikaanse HKM-families waarin een van die drie stigter mutasies segregeer.

Ses vermeende interaktors, nl. kardiale aktien (ACTC1), Down-sindroom kritiese streek 3 (DSCR3), enolase 1 (ENO1), F-boks en leusien ryke herhalings proteïen 10 (FBXL10), legumain (LGMN) en sorteer nexin3 (SNX3) is geïdentifiseer en bevestig as COMMD4 bindingsgenote deur G2H analises, gevolg deur *in vitro*

en *in vivo* ko-immunopresipitasie en 3D ko-lokalisasie toetse. Die funksionele effek van siRNA-bemiddelde uitklop van COMMD4 op cMyBPC omset was ook ondersoek omdat 'n paar COMMD proteïen familielede, asook die nuut-geïdentifiseerde bindingsgenote van COMMD4, geassosieerd is met die UPP. Data toon ophoping van cMyBPC in die endosome by COMMD4 uitklop, wat dus aandui op 'n funksionele rol vir COMMD4 in die omset van cMyBPC. Daarbenewens, toon assosiasie analise sterk bewyse van assosiasie tussen die verskillende enkele nukleotied polimorfismes (SNPs) in *SNX3* en 'n aantal hipertrofiese kenmerke,wat aandui op 'n rol vir SNX3 as 'n kandidaat modifiseerder van hipertrofie in HKM. Geen bewyse van assosiasie was waargeneem vir enige van die gene wat kodeer vir die ander COMMD4 bindingsgenote wat geïmpliseer word in die proteïen omset.

Die huidige studie toon dat COMMD4, 'n min verstaande lid van die COMM familie van proteïene, aan die cMyBPC motief van cMyBPC in'n fosforilasie-afhanklike wyse bind. Verder, gebasseer op die funksies van die proteïen interaksies, hipotiseer ons dat COMMD4 'n rol speel in proteïen vervoer en omset. Meer spesifiek, COMMD4 blyk om die vorming van proteïene komplekse met die Skp1-Cul1-Fbxl (SCF) E3 „ubiquitin' ligase te fasiliteer en help waarskynlik om die teiken-substraat vir die daaropvolgende ubiquitin-konjugasie te stabiliseer. Omdat dit lyk asof COMMD4 die proteïen-omset van cMyBPC en moontlik ander sarkomeriese proteïene, soos aktien, ook beïnvloed, vestig die resultate dus 'n nuwe assosiasie tussen die sarkomeer, HKM en die UPP. Daarbenewens sal die identifisering van *SNX3* as 'n hipertrofie modifiseerder voorsiening maak vir die verbeterde begrip van HKM pato-etiologie. *SNX3* voeg dus by tot die groeiende ?getal van sarkomeriese modifiseerende gene, wat uiteindelik, die risiko-ontleding in HKM mag verbeter. Verder, omdat dit blyk dat genetiese modifiseerders voldoende is om die siekte-uitdrukking heeltemal te verhoed in sekere HKM draers, kan die identifikasie van *SNX3* na die proteïen-omset roete dui as 'n potensiële nuwe teiken vir intervensie.

INDEX	PAGE
ACKNOWLEDGEMENTS	ix
LIST OF ABBREVIATIONS	x
LIST OF FIGURES	xviii
LIST OF TABLES	xxvii
CHAPTER ONE: INTRODUCTION	1
CHAPTER TWO: MATERIALS AND METHODS	52
CHAPTER THREE: RESULTS	97
CHAPTER FOUR: DISCUSSION	143
APPENDIX I	178
APPENDIX II	189
APPENDIX III	191
APPENDIX IV	193
APPENDIX V	197
APPENDIX VI	198
APPENDIX VII	202
APPENDIX VIII	205
APPENDIX IX	208
THESIS REFERENCES	209

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following individuals who have assisted me throughout the course of this degree:

My supervisor, Professor JC Moolman-Smook, the driving force behind all of this work. Thank you so much for your scientific input, insightful suggestions and sound advice with the write-up of this thesis. Furthermore your guidance, incredible patience, and encouragement throughout the course of my studies has been most appreciated. You're an inspirational woman and brilliant scientist and I have benefited tremendously, both scientifically and personally by being your student.

Professor Valerie Corfield, for always believing in me and the encouraging words of advice. For exposing me to the various scientific outreach programs and showing me the importance of networking and communicating science to the public.

Dr. Craig Kinnear, Mrs Lundi Korkie, Dr. Amsha Ramburan and Dr. Gerrida Uys, who has been there with me every step of the way. Thank you so much for all the help in the lab, the technical support, the scientific discussions but most importantly, the friendship. I'm forever grateful for your moral support and your willingness to assist through these years.

Mrs. Jomien Mouton, thank you for all the time you had to spend in the tissue culture lab culturing and transfecting the cells needed for this study and for the help with the western blots. Thank you to Ms. Nadia Carstens for her input regarding the statistical analysis of my association data. Also thank you for taking the time to explain all the statistical data to me, so that I could fully appreciate the significance of my findings.

To all my other friends and colleagues, both old and new in the MAGIC lab and the Department of Medical Biochemistry for creating an enjoyable working environment.

None of this would ever be possible if it was not for the unconditional support and understanding of my family and friends: my parents (Cecilia and Jakobus Swanepoel), my siblings (Bianca and Elton) and my friends (Gaynor Smith, Catherine Pharo, Liesl Bloem and Gail Louw). Thank you for putting up with me during the writing of this thesis, for always believing in me and cheering me up when I doubted myself and felt overwhelmed. I am truly and deeply grateful for all your support.

Finally, and most importantly, I want to thank God for watching over me and bringing me to the end of this journey.

LIST OF ABBREVIATIONS

1β-Hsp1β	Gene encoding Heat Shock Protein
3D	3-Dimensional
α-TM	Tropomyosin
aa	Amino acid
A	Alanine
ACE	Gene encoding angiotensin converting enzyme
ACTC	α-cardiac actin gene
Ade	Adenine
ADP	Adenosine diphosphate
aIVS	Anterior intraventricular septum
AKAP	A-kinase anchoring protein
Amp	Ampicillin
ANP	Atrial natriuretic peptide
ANKRD1	Cardiac ankyrin repeat protein
ANOVA	Analysis of variance
ASREA	Allele specific restriction enzyme analysis
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
ATP7A	Gene encoding copper-transporting P-type ATPase
ARF	Alternative reading frame
AVP	Arginine-vasopressin
AW	Anterior wall
bp	Base pair
BP	Blood pressure
BIR	Baculoviral-inhibitor-of-apoptosis repeat
BLAST	Basic local alignment search tool
BLASTN	Basic local alignment search tool (nucleotide)
BLASTP	Basic local alignment search tool (protein)
BSA	Body surface area
BSA stock	Bovine serum albumin
C	Cysteine
$CaCl_2$	Calcium chloride
CALM3	Gene encoding Calmodulin III
CaMKII	Calcium-calmodulin-activated kinase II
cDNA	complementary DNA

CEU	US Utah residents with European ancestry
CHB	Han Chinese from Beijing
CHF	Congestive heart failure
CHIP	C-terminal of Hsp70-interacting protein
c-IAP2	Gene encoding cellular inhibitor of apoptosis 2 protein
CIAP	Calf intestinal alkaline phosphatase
CK	Creatine kinase
cM	Centi-Morgans
Co-IP	Co-immunoprecipitation
COMM_	Copper metabolism gene *MURR1*
COMMD4	COMM domain containing protein 4
cMyBPC	cardiac myosin binding protein C
$cMyBPC^{-/-}$	Homozygous knockout of the cMyBPC gene
$cMyBPC^{-/+}$	Heterozygous knockout of the cMyBPC gene
Comp1	Principal component score 1
CSRP3	Gene encoding muscle LIM protein
Cul1	Cullin 1
CWT	Cumulative wall thickness score
DBP	Diastolic Blood Pressure
DCM	Dilated cardiomyopathy
dATP	Deoxy-adenosine triphospate
dCTP	Deoxy-cytidine triphosphate
ddH_2O	Distilled deionised water
dGTP	Deoxy-guanosine triphosphate
DMSO	Dimethylsulfoxide
dNTPs	Deoxy-nucleotide triphosphate
dpc	Days postcoitum
dTTP	Deoxy-thymidine triphosphate
DMEM	Dulbecco's modified Eagle's medium
DMEM-HAMS-F12	DMEM containing nutrient mixture F12
DNA	Deoxyribonucleic acid
DSCR3	Down syndrome critical region 3 protein
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin ligase
ECL	Enhanced chemiluminescence
ECS^{SOCS1}	Elongins B/C/Cullin2 and SOCS box containg E3 ubiquitin ligase

EDTA	Ethylene-diamine-tetra-acetic acid
EEA1	Early endosomes antigen 1
ENaC	Epithelial sodium channel
END1	Gene encoding the protein, endothelin-1
ENO1	α-enolase
ESCRT	Endosomal sorting complex required for transport system
EYFP	Enhanced yellow fluorescent protein
FBXL10	F-Box and leucine-rich repeat protein 10
F_{max}	Maximum Ca^{2+}-activated force
FnIII	Fibronectin, type III
fsMyBPC	fast skeletal MyBPC
g	Gram
GAPDH	Gene encoding Glyceraldehyde-3-phosphate dehydrogenase
GFP	Green fluorescent protein
H3K36	Histone 3 lysine 36
H3K4me3	trimethylated histone H3 lysine 4
HA	Haemaglutinin
HCaRG	Hypertension-related, calcium regulated gene
HCM	Hypertrophic Cardiomyopathy
HECT	Homologous to E6-AP COOH terminus
HEK293	Human embryonic kidney cells
HF	Heart failure
His	Histidine
HIF-1	Hypoxia inducible factor 1
HR	Heart rate
HRP	Horseradish peroxidase
HSP90β	Heat shock protein 90
HRE	Hypoxia responsive element
Hrs	Hepatocyte growth factor-regulated tyrosine kinase substrate
HWE	Hardy-Weinberg Equilibrium
IBD	Identity-by-descent
IgG	Immunoglobulin
IκB	Inhibitory kappa B
IP	Immunoprecipitation
I/R	Ischemia-reperfusion
ITGA8	Gene encoding the protein, Integrin, alpha 8
IVST	Interventricular septum thickness
IVSTmit	Interventricular septum thickness at the mitral valve

IVSTpap	Interventricular septum thickness at the papillary level
IW	Inferior wall
JPH2	Gene encoding the protein, junctophilin-2
JPT	Japanese from Tokyo
Kan	Kanamycin
kD	Kilo Dalton
KO	Knockout
L	Litre
LAMP2	Gene encoding lysosome-associated membrane protein
LB	Luria-Bertani
LD	Linkage disequilibrium
LDU	Linkage disequilibrium unit
Leu	Leucine
LGMN	Legumain
LiAc	Lithium acetate
LMM	Light meromyosin
LRR	Leucine-rich repeats
LV	Left ventricular
LVH	Left ventricular hypertrophy
LVM	Left ventricular mass
LW	Lateral wall
LyB	Lysis Buffer
Lys	Lysine
M	MyBPC motif
M6t	Construct expressing a cMyBPC protein from domains C0 to partial deletion of domain C10
M7t	Construct expressing a cMyBPC protein from domains C0 to partial C1 only
MAF	Minor allele frequency
MDM2	Murine double minute 2
Mg^{2+}	Magnesium
MHC	Myosin heavy chain
mg	Milligram
ml	Millilitre
mISVT	Maximum interventricular septum thickness
mLVWT	Maximal left ventricular wall thickness
MMP-2	Metalloproteinases
mPWT	Maximum posterior wall thickness

mRNA	Messenger RNA
MSC	Multiple cloning site
MURR1	Mouse murine gene *U2af1-rs1* region 1
MVB	Multivesicular bodies
MW	Molecular weight
MYBPC3	cardiac myosin binding protein C gene
MyPBH	Myosin-binding protein H
MYH6	α-Myosin heavy chain gene
MYH7	β-Myosin heavy chain gene
MYL2	Myosin regulatory light chain gene
MYL3	Myosin essential light chain gene
Na^+	Sodium
NC	Non-carrier
ng	Nanogram
NES	Nuclear export signals
NFκB	Nuclear factor kappa B
NFQ	Non-fluorescent quencher
NMR	Nuclear magnetic resonance
NO	Nitric oxide
NSC	Non-silencing control
OBSCN	Gene encoding the protein, obscurin
ORF	Open reading frame
PA	Proline/alanine
PBS	Phosphate buffered saline
PCI	Phenol:Chloroform:Iso-amylalcohol
PCR	Polymerase chain reaction
PDE4DIP	Phosphodiesterase 4D interacting protein
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
PGA	2-phospho-D-glycerate
PHD	Prolyl hydroxylases
PIs	Phosphoinosititides
pIVS	Posterior intraventricular septum
PKA	Cyclic-AMP-dependent protein kinase
PKC	Calcium/phospholipid-dependent protein kinase
PLN	Gene encoding the protein, phospholamban
PMSF	Phenyl-methyl-sulphonyl-fluoride
PRKAG2	Gene encoding γ2 subunit of AMP-activated protein kinase

PTER	Gene encoding phosphotriesterase-related protein
PtdIns3P	Phosphatidylinositol 3-phosphate
PtdIns (4,5)P_2	Phosphatidylinositol 4,5-biphosphate
PTMs	Posttranslational modifications
pVHL	Von Hippel-Lindau protein
PVDF	Polyvinylidene difluoride
PW	Posterior wall
PWapx	Posterior wall thickness at the supra-apex level
PWmit	Posterior wall thickness at the mitral level
PWpap	Posterior wall thickness at the papillary level
PX	Phospholipid-binding motif
QDO	Quadruple dropout
QRT-PCR	Quantitative real-time PCR
QTDT	Quantitative transmission disequilibrium test
RAAS	Renin-angiotension-aldosterone system
RACK1	Receptor for activated protein kinase C 1
RE	Restriction enzyme
REF	Reference
RHD	Rel-homology domain
RLC	Regulatory light chain of myosin
RNA	Ribonucleic acid
rRNA	ribosomal RNA
Rpm	Revolutions per minute
RTKs	Receptor tyrosine kinase
RyR2	Gene encoding cardiac ryanodine receptor protein
SB	Sodium borate
SBP	Systolic blood pressure
SCD	Sudden cardiac death
SCF	Skp1/Cullin1/F-box protein E3 ubiquitin ligase
SD	Synthetic dropout
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Ser	Serien
SGK1	Glucocorticoid-regulated kinase 1
SHR	Spontaneously hypertensive rats
siRNA	Small interfering RNA
SNP	Single nucleotide polymorphism
SNX3	Sorting Nexin 3

ssMyBPC	slow skeletal MyBPC
T	Threonine
Ta	Annealing temperature
TBST	Tris-buffered saline
TCAP	Gene encoding the protein, telethonin
TDO	Triple dropout
TF	Transcription factor
TFR	Transferrin receptor
TG	Transgenic
TGFβ1	Gene encoding transforming growth factor, β1
TGF	Transforming growth factor
TGN	Trans-Golgi network
TNF-α	Gene encoding the protein, tumor necrosis factor α
TNI	Troponin I
TNNC1	Troponin C, type 1, cardiac gene
TNNI3	Troponin I, type3, cardiac gene
TNNT2	Troponin T, type 2, cardiac gene
TNT	Troponin T
TPM	α-Tropomyosin
TRFR	Gene encoding transferring receptor
Trp	Tryptophan
TTN	Titin gene
U	Unit
UBDs	Ub-binding domains
UPS	Ubiquitin proteasome system
Ura	Uracil
USP10	Ubiquitin-specific protease 10
UV	Ultraviolet
VCL	Gene encoding the protein, vinculin/metavinculin
VEGF	Vascular endothelial growth factor
Vps	Vacuolar protein sorting
W	Tryptophan
WB	Western blot
WT	Wild type
XIAP	X-linked inhibitor of apoptosis
Y2H	Yeast two-hybrid
YRA	Yoruba from Nigeria
YPDA	Yeast peptone dextrose adenine

ZF	Zinc Finger
µg	Microgram
µl	Microliter
^{0}C	Degree Celsius

LIST OF FIGURES

Figure 1.1. Schematic representation of the the sarcomere depicting various myofibrillar components relative to each other.

The regions of the I- and A- band, M-line and the C- zone, which make up the sarcomere, are indicated at the bottom. Thin filaments components, i.e., actin, tropomyosin, troponins C, I, and T, constitute the I-band, which overlaps with the the thick filament in the A-band. The A-band is comprised of the H-zone, which only contains the thick filament, and the C-zone, where the thick and thin filaments overlaps. (Taken from: Morimoto, 2007)

Figure 1.2. The basic domain structure of cMyBPC.

The eight immunoglobulin-like (Ig) domains are represented by the red circles, while the fibronectin type 3-like (FnIII) domains are represented by the grey squares. The green linker between domains C0 and C1 indicates the Pro-Ala-rich region, the blue rectangle indicates the cMyBPC motif between domains C1 and C2, while the purple rectangle within the cMyBPC motif indicates the LAGGGRIS insertion. Furthermore, the light blue stars indicate the three phosphorylation sites within the cMyBPC motif, while the light orange rectangle indicates the cardiac-specific insertion in C5. The arrows indicate the cMyBPC domains involved in the interactions with actin, titin and myosin.

Figure 1.3. Proposed models for cMyBPC arrangement in the sarcomere.

A) In the trimeric model, three cMyBPC molecules wrap around the thick filament via stabilised interactions between domains C5 and C8, and domains C7 and C10, while the N-terminus (C0-C4) extends into the interfilamental space. (Image taken from Moolman-Smook *et al*., 2002) (B) In the alternative model, domains C7-C10 of individual cMyBPC molecules align axially to the thick filament, parallel to titin, while the N-terminus interacts with myosin crossbridges and/or actin. (Image taken from Squire *et al*., 2003)

Figure1.4. A schematic representation of the COMMD protein family and amino acid sequence alignment of the COMM domain

A) The conserved COMMD domain is indicated with red blocks, and numbers indicate the respective amino-acid length of each protein. B) Sequence alignment of the COMM domain of the ten known human COMM proteins. (Images taken from Maine and Burstein, 2007)

Figure 1.5. Structural model of full-length COMMD1

The NMR structure of the COMMD1 N-terminal residues 1-121, identified by Sommerhalter et al (2007), are shown in cyan. This model was used as an initial constraint and matches the ab initio model as it had the same secondary structure, number of α-helices and overall fold. Residues in green represent part of the N- and C-terminal and the flexible linker (119-128) between the two domains, while the rest of the C-terminal

(residues 125-190) is shown in orange. The arrow indicates the proteolytic sensitive linker between the two-domain structures. (Image taken from Burkhead et al., 2009).

Figure 1.6. Nucleotide and amino acid sequence of human COMMD4 and schematic representation of the COMMD4 structure.

A.Nucleotide (top) and amino acid (bottom) sequence of COMMD4 (NM_017828.3). Amino acid residues 1-199 represent isoform1 of COMMD4 in black while the alternative spliced transcript (Isoform 2) correspond to amino acid residues 128-187 as indicated in red. The COMM domain spans from amino acid residue 130-199 and are highlighted in grey. **B.** Schematic representation of COMMD4 isoform1. Amino acid residues numbering at top indicate the boundaries of each exon. The calcium-binding domain spanning from amino-acid 19-198 are depicted by the blue line, while the COMMD domain which spans from amino acid 130-199 are depicted by the red line.

Figure 1.7. Schematic model of COMMD1 in the attenuation of NFκB-mediated transcription.

Various stimuli that activate the NFκB pathway result in the phosphorylation and degradation of IκBα, which enables the nuclear translocation of NFκB dimers to NFκB-dependent promoter elements to induce gene expression. Upon NFκB activation, COMMD1 is recruited to chromatin and associate with the ECSSOCS1 Ub-ligase. COMMD1 serves as an adaptor and enhances the interaction between NFκB subunits and theUb-ligase and thereby promotes the poly-ubiquitination of NFκB. The latter then undergoes proteasomal degradation, which in turn attenuates NFκB-mediated gene transcription. (Images combined from Maine and Burstein, 2007 and van de Sluis et al., 2007)

Figure 1.8. Schematic model of the mechanism whereby COMMD1 mediates HIF-1α degradation.

Under normal oxygen conditions, HIF-1α binds to HSP70 and HSP90 for maturation, but undergoes ubiquitination via the pVHL E3 Ub-complex and degradation via the 26S proteasome. In contrast, under hypoxic conditions, HIF-1α is stabilised and heterodimerises with HIF-1β to form active HIF-1, which binds to hypoxia response elements (HRE) in HIF-1-targeted genes. Upon reoxygenation, HIF-1α is ubiquitinated via pVHL, and COMMD1 facilitates its proteasomal degradation. Inhibition of HSP90 via17-AAG/Geldanamycin causes incorrect folding of HIF-1α and the formation of a tripartite complex with HSP70 and COMMD1, which promotes degradation via the 20S proteasome. (Image taken from van de Sluis et al., 2009)

Figure 1.9. A basic overview of the ubiquitin proteasome pathway.

Ub is activated by E1 in an ATP-dependent process and then transferred to the E2 enzyme. The E3 Ub-ligase confers specificity and mediates the transfer of one or more Ub molecules to the substrate. Poly-ubiquitin chains, linked via Lys48, are recognised by the 26S proteasome, which targets the protein for degradation. Deubiquitinating enzymes reverse the process of Ub conjugation and recycle Ub. (Image taken from Wilkinson et al., 2005)

Figure 1.10. A basic overview of the endosomal trafficking routes.

A). This model integrates vesicular traffic from the plasma membrane or the trans-Golgi network (TGN) to the early endosome, where cargo is sorted back to the membrane, TGN (via retromer complex) or the lysosome (via the ESCRT omplex). Schematic representation of the retromer complex (B) and the ESCRT machinery (C). Images taken from Mukhopadhyay and Rietzman, 2007 and Seaman, 2005

Figure 2.1. Schematic representation showing a cross-section of the heart transected at 3 levels.

A) Long-axis view of left ventricle, with the 3 blue dotted lines representing the mitral valve level, papillary muscle level and the apex level, where measurements were taken. B) A 2D echocardiogram of the left ventricle. Abbreviations: AV- aortic valve, LA-left atrium, LV-left ventricle, LVOT-left ventricular outflow tract, MV- mitral valve, RVOT-right ventricular outflow tract. Taken from (http://www.med.yale.edu/.../aortic_regurgitation.html) with minor modifications by JC Moolman-Smook.

Figure 3.1. The growth curves of host yeast strain AH109 transformed with an non-recombinant pGBKT7 and COMMD4-pGBKT7.

Growth rates of transformed pGBKT7-COMMD4 liquid cultures were compared to transformed non-recombinant pGBKT7 liquid cultures. The growth rate was estimated by calculating the slope of each of the curves. The slope of the blue line which represents the AH109 strain containing the COMMD4 bait construct was lower, indicating slower growth of yeast cells and thus slight toxicity on AH109.

Figure 3.2. A representative example demonstrating the *in silico* translation of the ORF of the insert of prey clone 410 and protein alignment with the predicted protein reference sequence.

A. Translation of in-frame nucleotide sequence of prey clone 410 using the DNAMAN™ software program (Lynnon Biosoft). **B**. Alignment of translated ORF sequences with protein reference sequence (REF) predicted by identity of prey insert nucleotide sequence using clustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

Figure.3.3. Linear schematic of protein and domain structure of the eight prioritised preys.

Different coloured blocks represent different domain structures in each respective prey protein. The numbers indicate the amino acid number and the red line below shows the fragment of prey protein encoded by the respective clone. Diagram not drawn to scale

Figure 3.4. Autoradiographs of S35-labeled methionine-labelled protein products from translation and transcription reactions.

Lanes 3-9 shows respectively, translated bait and prey proteins, ACTC1, LGMN, COMMD4, SNX3, ENO1 and ANKRD1. In lane 1 and 2, the 5kDa translated FBXL10 and DSCR3 are not visible and may have been obscured by the dye front. The position of sizes from non-radioactive size markers, Spectra TM Multicolor

Broad Range Protein Ladder (Fermentas) (A), and High-Range Rainbow Molecular Weight Markers (Amersham Biosciences) (B), as transferred from dried polyacrylamide gel are shown in the marker lanes.

Figure 3.5. In vitro IP and Co-IP reactions of COMMD4 and each respective prey. Autoradiograph of 20% SDS-PAGE gel indicating immunoprecipitated Met-35S-labelled in vitro transcribed/translated proteins and antibodies. COMMD4 co-immunoprecipitated with ACTC1 (lane 1), SNX3 (lane3), LGMN (lane 5) and ENO1 (lane 10). However, no interaction was detected between COMMD4 and ANKRD1 (lane 8). The protein G control lanes were clear, indicating that no non-specific occurred during pull-downs. The position of sizes from a non-radioactive size marker, Spectra TM Multicolor Broad Range Protein Ladder, (Fermentas), as transferred from dried polyacrylamide gel, is shown in the marker lane (M).

Figure 3.6. In vivo co-immunoprecipitation using H9C2 cell lysates, showing that GFP-cMyBPC specifically pulls down dsRed-tagged COMMD4.

Cells were transfected with dsRed-COMMD4 and/or GFP-cMyBPC, as indicated above each lane. Pull-downs (IP) were performed using either the anti-GFP antibody (JL8) or the anti-dsRed antibody (DsRed), and western blots (WB) subsequently stained with either of these antibodies. Anti-haemaglutinin antibody (HA) was used as a negative control antibody (lanes 4 and 9), while a protein G control (Prot G) was employed to ensure that no non-specific binding occurred during pull-downs (lanes 3 and 8). Interestingly, in the reciprocal experiments, only a small band, rather than the band of ~140kD corresponding to GFP-cMyBPC, could be detected using the anti-GFP antibody (lanes 5 and 6), although this larger band was clearly detected in cells not transfected with dsRed-COMMD4 (lane 7). This suggests that cMyBPC undergoes enhanced degradation in the presence of additional COMMD4.

Figure 3.7. In vivo co-immunoprecipitation assays of EYFP-tagged COMMD4 and its Y2H-identified putative interactors.

HEK293 or H9C2 cells were transfected with EYFP-COMMD4, or EYFP-COMMD4 and DsRed-DSCR3 (VII) as indicated above each lane. Pull-downs (IP) were performed using the anti-EYFP antibody (JL8), anti-dsRed antibody (DsRed), or prey protein-specific antibody as indicated, and Western blots (WB) subsequently stained with either of these antibodies, as indicated above each lane. Anti-haemaglutinin antibody (HA) was used as a negative control antibody (lanes 4 and 8), while a protein G control (Prot G) was employed to ensure that no non-specific binding occurred during pull-downs (lanes 3 and 7). EYFP-tagged COMMD4 could pull down endogenous SNX3 (I), ENO1 (II), LGMN (III), FBXL10 (IV), ACTC1 (V), as well as DsRed-tagged DSCR3 (VII), while these proteins also reciprocally immunoprecipitated EYFP-COMMD4 in vivo. However, no interaction was detected between EYFP-COMMD4 and endogenous ANKRD1 (VI). Both the protein G and negative control antibody lanes were clear, indicating that these precipitations were not spurious, but were the result of physical association between the relevant proteins.* denotes the extra bands seen for multiple alternatively-spliced transcripts, which are known to exist for FBXL10.

Figure 3.8. 3D Co-localisation of COMMD4 and respective preys identified in the Y2H library screen.
Representative images of fluorescence microscopy showing co-localisation of COMMD4 and its respective interactors in differentiated H9C2 cardiomyocytes. Each panel represents a single frame of the 25 images captured for the vertical Z-stack. The first three panels show a single colour channel, while the last panel shows an overlay of the three colour channels used, with the nuclei stained with DAPI for orientation purposes. Column (I) indicates the expression of EYFP-tagged COMMD4, artificially shown as green fluorescence. Column (II) indicates endogenous prey interactors stained with respective primary antibodies, followed by the relevant TxRed-conjugated secondary antibody (red fluorescence). * For DSCR3, 3D co-localisation was obtained using fluorescently-tagged proteins, as no antibodies against this protein were available. Column (III) shows co-localisation (yellow fluorescence) between EYFP-COMMD4 and each respective prey occurring in the cytoplasmic region of the cell. Scale bar: 0.02mm.

Figure 3.9. Live cell fluorescence imaging of COMMD4 and SNX3 showing different frame points from the movieclip.
Each panel represents a single frame of the 25 images captured for the vertical Z-stack Column (A) shows differentiated H9C2 cell expressing EYFP-tagged COMMD4 constructs (red fluorescence). Column (B) shows the same transfected cell expressing GFP-SNX3 constructs (green fluorescence). Column (C) shows co-localisation between COMMD4 and SNX3 (yellow fluorescence). Column (D) shows overlay of images with nuclei stained with Hoechst H-33342. Images in rows (i-iv) show images captured at different frame points from the movie clip in order to demonstrate the movement of COMMD4 and SNX3 towards the nucleus. (red block highlights point of interest). Scale bar: 0.02mm

Figure 3.10. Live cell fluorescence imaging of COMMD4, SNX3, LGMN and DSCR3 with an endosomal marker.
Representative image of live cell fluorescence microscopy showing co-localisation of EYFP-COMMD4, GFP-SNX3, GFP-LGMN and GFP-DSCR3 with the endosomal marker, RFP-Endo, in differentiated H9C2 cells. Each panel represents a single frame of the 25 images captured for the vertical Z-stack. The first three panels show a single colour channel, while the last panel shows an overlay of the three colour channels used with the nuclei stained with Hoechst H-33342. Column (I) shows representative cells expressing either EYFP-COMMD4 (artificially coloured green in the image) or one of the GFP-tagged prey proteins (green fluorescence), while column (II) shows the same cells expressing DsRed-Endo (red fluorescence). Column (III) shows co-localisation (yellow fluorescence) between COMMD4 or its interactors and the endosomal marker. This data suggests that COMMD4, SNX3, LGMN and DSCR3 all share an endosomal localisation. Scale bar: 0.02mm

Figure 3.11 Phosphorylation-dependant association between COMMD4 and the cMyBPC-motif.
Bar-graph of quantitative β-galactosidase assay results, comparing the influence of the m-domain, in various phosphorylation states, on the interaction between cMyBPC C1-C2 and COMMD4 with that of the flanking C1 and C2 domains. One unit of β-galactosidase was defined as the amount of enzyme that hydrolyzes 1μmol of OPNG to O-nitrophenol and D-galactose per minute per cell. Diploid colonies expressing mono (AAA)- or unphosphorylated (PPP) mimics of C1-C2 fragments and COMMD4 generated significantly higher β-galactosidase levels compared to the negative control (non-recomninant pGBKT7), flanking domains C1, C2 and C0-C1, as well as the tri-phosphorylated mimic (PPP), suggesting that interaction between COMMD4 and cMyBPC occurs primarily with the m-domain in a mono- or unphosphorylated state. Combined data of three independent assays, each performed with triplicate samples, were normalized to the mean β-galactosidase levels of the negative control and are presented as Mean±S.E. Different lower case letters above data bars indicate statistically significant differences, while similar letters indicate non-significant differences; Differences were calculated by ANOVA with Bonferonni post-hoc testing; $p<0.05$ was considered statistically significant (A. Ramburan, Phd Thesis, 2008).

Figure 3.12. Effect of cMyBPC phosphorylation on the interaction between COMMD4 and the cMyBPC motif
H9C2 cells transfected with YFP-COMMD4 were treated with (AI) 2.5mM CaCl2 or (AII) 2.5mM CaCl2 and 0.1uM Isoproterenol representing mono- and triphosphorylation conditions, respectively. A: Representative image of live cell fluorescence microscopy showing that co-localisation of COMMD4 and cMyBPC decreases under conditions of tri-phosphorylation. Each panel represents a single frame of the 25 images captured for the vertical Z-stack. The first three panels shows a single colour channel, while the last panel shows an overlay of the three colour channels used and the nuclei stained with DAPI. Column (I) indicates endogenous cMyBPC stained its respective primary antibody followed by the relevant TxRed-conjugated secondary antibody (red fluorescence). Column (II) indicates the expression of GFP-tagged COMMD4 as green fluorescence. Column (III) shows co-localisation (yellow fluorescence) between GFP-COMMD4 and each respective prey. Scale bar: 0.02mm (B) A higher degree of co-localisation was consistently observed in cells treated with CaCl2 only, compared to CaCl2/Isoproterenol treated cells, although the p-value did not reach statistical significance. Data are presented as ±SEM from two independent experiments. Similar letters above error bars indicate non-significance ($p>0.05$) upon unpaired t-testing.

Figure 3.13. qRT-PCR analysis of COMMD4 mRNA expression levels.
Bar graph demonstrating real-time quantification of COMMD4 cDNA, transcribed from RNA isolated from differentiated H9C2 cells treated with either 100nM COMMD4 siRNA for 24 hrs, or with a non-silencing control. COMMD4 mRNA levels were normalised according to the expression levels of rodent reference genes: Transferrin receptor (TRFR), Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and Heat Shock Protein (1β-Hsp1β). COMMD4 siRNA 2 (Commd4_Rn_LOC363068_2_HP) resulted in optimal knockdown of COMMD4 (80-05%).

Figure 3.14. siRNA-mediated knockdown of COMMD4 in differentiated H9C2 cells expressing GFP-cMyBPC and RFP-Endo.

A: Representative image of live cell fluorescence microscopy showing that co-localisation of cMyBPC and an endosomal marker in differentiated H9C2 cells increases in the (1) absence of COMMD4 (i.e. with COMMD4 knockdown) compared to the (2) presence of COMMD4 (non-silencing control, NSC). Column (I) shows H9C2 cells expressing GFP-cMyBPC (green fluorescence), while column (II) shows the same H9C2 cells expressing the endosomal marker, DsRed-Endo (red fluorescence). Column (III) shows co-localisation between cMyBPC and the endosomal marker (yellow fluorescence), while column (IV) shows an overlay of images, with nuclei stained with Hoechst H-33342 for orientation purposes. Scale bar: 0.02mm C: Quantification of the co-localisation signal shown in B shows that co-localisation of cMyBPC and the endosomal marker RFP-Endo, increased significantlyupon knockdown of COMMD4 ($p<0.05$).. Change in co-localisation was calculated using the CellR software and presented as a false colour image. Colocalisation area index is calculated as the co-localisation area indexed to the sum of the EYFP and DsRed areas for each cell layer, and is given as %. Co-localisation data from three independent experiments was pooled together and the mean values ± standard error are portrayed.

Figure 3.15. ASREA of COMMD4/rs 3809547 polymorphism (C/G).

A. Representative 2% agarose gel showing the fragment sizes of COMMD4/ rs 3809547 amplicon following digestion with restriction enzyme, MboII. M represents a 100bp molecular marker, UC represent an undigested control. The 255bp fragment represents the G allele, whereas fragments 176bp and 79bp represent the C allele. The CG heterozygote is represented by all 3 fragments. B. Schematic representation of COMMD4/ rs 3809547 amplicon outlining the position of the MboII restriction site and the fragment sizes generated following digestion.

Figure 3.16. ASREA of COMMD4/rs 1185314 polymorphism (C/T)

A. Representative 2% agarose gel showing the fragment sizes of COMMD4/rs 1185314 amplicon following digestion with restriction enzyme, NheI. M represents a 100bp molecular marker, UC represents an undigested control. The 356bp fragment represents the T allele, whereas fragments 234bp and 122bp represent the C allele. The CT heterozygote is represented by all 3 fragments. B. Schematic representation of COMMD4/rs 1185314 amplicon outlining the position of the NheI restriction site and the fragment sizes generated following digestion.

Figure 3.17. ASREA of COMMD4/ rs 11072542 polymorphism (A/G).

A. Representative 2% agarose gel showing the fragment sizes of COMMD4/rs 11072542 amplicon following digestion with restriction enzyme, DraII. M represents a 100bp molecular marker, UC represents an undigested control, The 337bp fragment represents the A allele, whereas fragments 223bp and 114bp represents the G allele. The AG heterozygote is represented by all 3 fragments. B. Schematic representation

of COMMD4/rs 11072542 amplicon outlining the position of the DraII restriction site and the fragment sizes generated following digestion.

Figure 3.18. Representative TaqMan allelic discrimination plot.
Taqman genotyping results of rs579257 in SNX3 are shown; allele X represents the G-allele and allele Y represents the T-allele at this SNP.

Figure 3.19. Linkage disequilibrium plot of SNPs within COMMD4.
D' values within blocks represent the degree of LD. A blank red or light blue block represents a D' value of 1 and indicates complete LD. Lighter shades depicting low D' values is indicative of weak LD.

Figure 3.20. Linkage disequilibrium plot block of SNPs within SNX3
D'values within blocks represents the degree of LD. A blank red or blue block represents a D' value of 1, indicating complete LD. Lighter shades depicting low D' values is indicative of weak LD.This LD plot shows a 8kb haplotype block indicative of complete LD between markers rs635051 and rs9398166 within this gene.

Figure 3.21. Linkage disequilibrium plot of SNPs within DSCR3.
D'values within blocks represents the degree of LD. A blank red or blue block represents a D' value of 1, indicating complete LD. Lighter shades depicting low D' values is indicative of weak LD.

Figure 3.22. Linkage disequilibrium plot of SNPs within LGMN.
D'values within blocks represents the degree of LD. A blank red or blue block represents a D' value of 1, indicating complete LD. Lighter shades depicting low D' values is indicative of weak LD.

Figure 3.23. Linkage disequilibrium plot and haplotype block of SNPs within FBXL10.
D'values within blocks represents the degree of LD. A blank red or blue block represents a D' value of 1, indicating complete LD. Lighter shades depicting low D' values is indicative of weak LD.This LD plot also shows 2 haplotype blocks of 9 and 10 kb each between markers, indicative of complete LD between the respective markers within this gene.

Figure 4.1. Schematic depicting the proposed role of COMMD4 in the transfer and stabilisation of cMyBPC within the SCF E3 ubiquitin ligase complex.
I.A. COMMD4 binds to the m-domain of dephosphorylated cMyBPC, which had been released from the sarcomere. **B**. COMMD4 recruits F-box protein (FBXL10), which possibly also binds to cMyBPC via its LRR domains, and which binds Skp1 via its F-box motif. Meanwhile the scaffold protein of the eventual SCF E3 ubiquitin ligase complex, Cullin, binds to the RING protein, Rbx1, which, in turn, recruits the E2 enzyme; this complex would remain inactivated due to the binding of the sequestering factor, CAND1. **C**.

Upon neddylation via Nedd8, CAND1 dissociates and the full SCF E3 ubiquitin ligase assembles and becomes activated, involving release of Rbx1 on a flexible linker, to facilitate transfer of activated ubiquitin to the cMyBPC substrate. COMMD4, by binding FBXL10, the MyBPC substrate, and Cullin, stabilises the complex and facilitates presentation of the target at the angle and distance required for efficient ubiquitin transfer by the E2 enzyme. Modified from van de Sluis *et al*., 2007

Figure 4.2. Schematic model depicting the proposed role of COMMD4 in the transfer and stabilisation of the 3% truncated cMyBPC (I) as well as lack of proper stabilisation of the 80% truncated cMyBPC (II). I. Just as with wild-type cMyBPC in figure 4.1, ubiquitin transfer can occur unimpeded with the 3% truncated cMyBPC that retains the m-domain (M6t from Sarikas *et al., 2005*.).**II. A.** The 80% truncated cMyBPC (M7t from Sarikas *et al*., 2005), which is comprised of only domain C0 and partial domain C1, is unable to bind COMMD4 due to the absence of the m-domain, although (**B.**) FBXL10 may still be able to bind the truncated protein. **C.** Without stable tethering of the cMyBPC fragment to FBXL10 and the rest of the SCF complex via COMMD4, the requisite conserved orientation of the target with regard to the rest of the SCF E3 subunits is disturbed, which affects the efficiency of ubiquitin transfer. Due to molecular movement, the severely truncated cMyBPC might randomly present in the correct orientation, allowing for mono-ubiquitination, but not for structured poly-ubiquitination. This in itself may tie up the UPS, impairing the system by effectively depleting its components from the cytoplasm through the formation of aggregates. Modified from van de Sluis *et al*., 2007

LIST OF TABLES

Table1.1. Sarcomeric genes implicated in the pathogenesis of HCM

Table 2.1. Single transfections and combinations for *in vivo* co-localisation assays

Table 2.2. Single and co-transfections for *in vivo* co-immunoprecipitations experiments

Table 2.3. Primer sequences to generate bait and prey inserts for plasmid construction

Table 2.4. Primer sequences and annealing temperatures used for the amplification of inserts from Y2H cloning vectors and generation of products used in *in vitro* transcription and translation experiments

Table 2.5 Restriction enzymes and digestion conditions used for the cloning of bait (bold font) and prey inserts into the relevant vectors.

Table 2.6. Characteristics of the relevant antibodies used in in vivo IP, Co-IP and WB analysis

Table 2.7. *Commd4* predicted siRNA sequences from the *species*, Rat

Table 2.8. Number of South African HCM-affected families analysed in present study

Table 2.9. Summary of SNP assay ID's, genotypes, chromosome position and minor allele frequency for each gene.

Table 2.10 Primer sequences, annealing temperatures and expected sizes of amplified fragment used for genotyping of polymorphisms in COMMD4

Table 2.11. Characteristics of the selected polymorphisms in *COMMD4*

Table 3.1. Phenotype verification of yeast strain AH109 after COMMD4 bait construct transformation

Table 3.2. Effect of COMMD4 bait construct on the mating efficiency of yeast strain AH109

Table 3.3. Scoring of the 28 primary and secondary clones which activated the nutritional (*HIS3* and *ADE2*) and colourimetric (*MEL1*) reporter genes by prey-COMMD4 interactions.

Table 3.4. Heterologous bait matings of the 28 putative interactors which showed specific interaction with COMMD4.

Table 3.5. Identification of the 28 COMMD4 putative interactor clones from the Y2H cardiac library cDNA library screen.

Table 3.6. Predicted and approximate observed molecular weights of bait and prey fusion proteins used in Co-IP analysis

Table 3.7. Echocardiographic characteristics and hypertrophy covariate information of affected and unaffected individuals.

Table 3.8. Comparisons of minor allele frequencies observed in study cohort compared to other populations

Table 3.9. Summary of observed LD (shown by D') between relevant markers for each gene in the study cohort.

Table 3.10 Results from population stratification analysis of COMMD4. Values in brackets indicate the number of informative individuals.

Table 3.11. Results of *COMMD 4* association analysis using the total model.
P-values for association between rs3809547, rs11853141 and rs11072542 and hypertrophy indices. Values in brackets indicate the number of informative individuals.

Table 3.12. Results from population stratification analysis of SNX3. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals. Significant p-values are highlighted and indicated in bold red

Table 3.13. Results of SNX3 association analysis using the orthogonal and total model where appropriate. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals. Significant p-values are highlighted and indicated in bold red.

Table 3.14. Effect sizes for the significant associations between markers of SNX3 and hypertrophy indices. Effect sizes are given in terms of the original measurements for all indices except for comp1.

Table 3.15. Results from population stratification analysis of *DSCR3*. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals.

Table 3.16. Results of *DSCR3* association analysis using the total model.P-values for association between *DSCR3* SNPs and hypertrophy indices**.** Values in brackets indicate the number of informative individuals. Significant p-values are highlighted and indicated in bold red.

Table 3.17. Results from population stratification analysis of *LGMN*. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals.

Table 3.18. Results of *LGMN* association analysis using the total model.P-values for association between *LGMN* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals.

Table 3.19. Results from population stratification analysis of *FBXL10*. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals. P-values for marker rs7313015 are not shown since the lack of sufficient number of informative genotyped individuals to enable QTDT analysis

Table 3.20. Results of *FBXL10* association analysis using the total model.P-values for association between *FBXL10* SNPs and hypertrophy indices.

CHAPTER 1: INTRODUCTION

INDEX **PAGE**

1.1. HYPERTROPHIC CARDIOMYOPATHY – THE DISEASE MODEL 3
1.1.1. The genetic spectrum of HCM 4
1.1.2. Clinical variability and genotype:phenotype associations in HCM 5
1.1.3. Candidate modifier genes in HCM 6
1.1.4. Possible mechanism by which mutations cause HCM 7
1.1.4.1. Altered sarcomeric function – the dominant-negative effect 7
1.1.4.2. Reduction of functional protein – haploinsufficiency 8
1.2. CARDIAC CONTRACTILITY AND SARCOMERE STRUCTURE 9
1.2.1. Localisation of MyBPC within the sarcomere 10
1.2.2. MyBPC isoforms 11
1.2.2.1. Mutations in *MYBPC3* – one of the most common cause of HCM 11
1.2.3. MyBPC structure 12
1.2.4. Biological functions of MyBPC 13
1.2.4.1. Sarcomere and thick filament assembly 13
1.2.4.1.1. MyBPC interactions with LMM, C-zone titin, LIM, Obscurin and itself 13
1.2.4.1.2. Proposed structural roles 14
1.2.4.1.3. Current proposed models of MyBPC arrangement on the thick filament 15
1.2.4.2. Regulation of cardiac contraction 17
1.2.4.2.1. Phoshorylation of MyBPC 17
1.2.4.2.2. MyBPC interaction with myosin subfragment 2 (S2) 18
1.2.4.2.3. MyBPC interactions with actin 18
1.2.4.2.4. The effect of cMyBPC phosporylation on thick filament structure and contraction 19
1.2.4.2.4.1. Structural changes 19
1.2.4.2.4.2. Altered contractile properties 20
1.2.4.2.5. Cardioprotective role of MyBPC phosphorylation 21
1.3. COMM DOMAIN CONTAINING PROTEIN 4 (COMMD4) 22
1.3.1. Identification of COMMD4 and other COMMD proteins 22
1.3.2. Structure of COMMD proteins 23
1.3.2.1. Structure of COMMD4 25
1.3.3. Expression and distribution of COMMD proteins 26
1.3.4. Biological functions of COMMD proteins 27
1.3.4.1. Involvement in the regulation of NFκB 27
1.3.4.1.1. NFκB: An overview 27

1.3.4.1.2. COMMD proteins: Inhibitors of NFκB 28
1.3.4.1.3. Ubiquitin modifications of COMMD1 in the NFκB pathway 29
1.3.4.2. Copper homeostasis 31
1.3.4.2.1. Copper dependent trafficking or ubiqitin modification of copper transporters 32
1.3.4.2.2. XIAP-COMMD1-mediated copper control 32
1.3.4.3. Regulation of sodium channels (ENAC) activity 33
1.3.4.3.1. COMMD1 interactions with ENaC subunits 33
1.3.4.3.2. Ubiquitin modification of COMMD1 in the sodium transport pathway 34
1.3.4.4. Embryogenesis 35
1.3.4.4.1. COMMD1 knockout mouse model 35
1.3.4.4.2. Regulation of HIF-1α protein levels and HIF-1 transcriptional activity 35
1.3.4.4.3. Interaction of COMMD1 with HIF-1α 36
1.3.4.5. Other reported interactions and proposed functions 37
1.3.4.5.1. Interaction of COMMD1 with Phosphatidylinositol 4.5-Biphosphate 37
1.3.4.5.2. Interaction of COMMD1 with ARF 38
1.3.4.6. Functions of other COMMD proteins 38
1.3.4.6.1. Proposed functions of COMMD5 38
1.3.4.6.2. Interaction of COMMD6 from *Amphioxus Branchiostoma belcheri* (BbCOMMD6) with creatine kinase 39
1.3.4.7. Summary of COMMD protein functions 39
1.4. SARCOMERE QUALITY CONTROL 39
1.4.1. Involvement of chaperones/co-chaperones in protein degradation 40
1.4.2. Turnover rate of myofibrillar proteins 41
1.4.3. Ubiquitin proteosome system in the heart 42
1.4.3.1. The process of ubiquitin conjugation 42
1.4.3.2. The 26S proteasome 42
1.4.3.3. E3 ubiquitin ligases – Skp1-Cullin-F-box (SCF) complex 44
1.4.3.3.1. E3 ubiquitin ligases-role in cardiac hypertrophy 44
1.4.3.4. Regulation of the UPS 45
1.4.3.5. Proteasome-independent functions of ubiquitin 46
1.5. THE ENDOSOMAL TRAFFICKING PATHWAY 46
1.5.1. Regulation of endosomal trafficking 48
1.5.1.1. The retromer complex in endosome-to-golgi trafficking 48
1.5.1.2. The ESCRT machinery in lysosomal-mediated degradation 48
1.5.2. Autophagy - lysosomal-mediated degradation 50
1.6. THE PRESENT STUDY 51

CHAPTER 1: INTRODUCTION

Cardiac hypertrophy, defined as overgrowth of the heart muscle walls, is an adaptive response to biochemical and physiological stresses and is characterised by an increase in cardiomyocyte size, protein synthesis, and increased number of sarcomeres as well as re-expression of the fetal gene program at a molecular level (Kai *et al*., 1998; Frey *et al*., 2004; Strøm *et al*., 2004). This cellular process is necessary to allow cardiomyocytes to sustain increased cardiac output. However, it can be detrimental when prolonged, because it enhances the risk for the development of cardiovascular diseases; therefore, cardiac hypertrophy, in itself is an independent predictor of morbidity and mortality (Levy *et al*., 1998; Benjamin *et al*., 1999).

Since hypertrophy is a feature of a number of common conditions such as ischaemic heart disease, heart failure, hypertension and diabetes mellitus, understanding the factors that contribute to its development is of the utmost importance. However, studying hypertrophy against the backdrop of such complex disorders is not so straightforward. Therefore, the apparently monogenic disease, hypertrophic cardiomyopathy (HCM), with left ventricular hypertrophy (LVH) as its main clinical feature, is studied as a model disease for understanding mechanisms leading to hypertrophy development (Watkins *et al*., 2005a). For the purpose of this thesis, this chapter will focus on the link between HCM, the sarcomeric protein cardiac myosin binding protein C (cMyBPC) and its novel interactor, the COMM domain containing protein 4 (COMMD4). As COMMD4 was found in this study to be implicated in the ubiquitin proteasome pathway (UPS) and in protein trafficking, a condensed and simplified overview of these pathways will also be covered to provide context for the subsequent discussion.

1.1. HYPERTROPHIC CARDIOMYOPATHY – THE DISEASE MODEL

HCM is the most common inherited cardiac muscle disorder with an estimated prevalence of 0.2% (Maron *et al*, 1995). It is characterized by LVH that occurs in the absence of other underlying causes of hypertrophy, myocyte disarray, interstitial fibrosis and the risk of sudden cardiac death (SCD) (Seidman and Seidman, 2001). The clinical presentation of this disease is diverse and ranges from individuals being asymptomatic, mildly symptomatic or developing more severe complications, such as SCD. Symptomatic HCM patients display a variety of symptoms, which include diastolic dysfunction, dyspnoea, angina, myocardial ischemia and arrythmias (Schwartz *et al*., 1995; Spirito *et al*., 1997).

The disease is also phenotypically highly variable in the degree and pattern of hypertrophy, onset of disease, penetrance and risk of SCD. With regards to the pattern of LVH, asymetrical septal hypertrophy occurs more commonly than diffuse concentric (i.e. symmetric) and apical hypertrophy patterns (Klues *et al*., 1995). Moreover, the extent of hypetrophy varies from mild (less than 11mm), or moderate (13-15mm) to severe (~ 35mm) (Spirito, 2000).

1.1.1. The genetic spectrum of HCM

HCM is transmitted as a Mendelian autosomal dominant trait and is caused by more than 630 mutations in multiple genes encoding sarcomeric- and myofilament-related proteins (reviewed in Xu *et al*., 2010). HCM has been established as a disease of the sarcomere and the 11 sarcomeric genes involved in HCM includes *MYH7* (β-myosin heavy chain), *MYH6* (α-myosin heavy chain), *MYBPC3* (cardiac myosin binding protein C), *TNNT2* (cardiac troponin T), *TNNI3* (cardiac troponin I), *TNNC1* (cardiac troponin C), *TPM1* (α-tropomyosin), *MYL3* (myosin essential light chain), *MYL2* (myosin regulatory light chain), *ACTC* (α-cardiac actin), and *TTN* (Titin) (see Table 1 [HCM mutation database: http://www.angis.org.au/Databases/Heart/]). Mutations in *MYH7* and *MYBPC3* predominate and collectively account for 75-80% of all HCM cases, but a minimum of 20-25% of cases remain for which the genetic cause remains obscure. Some of these unaccounted cases have the morphological phenotype of HCM despite the fact that they do not appear to be caused by mutations in the known HCM-associated sarcomeric protein genes (Marion *et al*., 2001; Arad *et al*., 2002). However, other diseases exist which mimic the phenotypic expression of HCM; these include the likes of Fabry's disease, Wolff-Parkinson-White syndrome, Friedreich's ataxia, Noonan and LEOPARD syndrome, which might explain a proportion of the unaccounted cases (Blair *et al*., 2001; Arad *et al*., 2002; Digilio *et al*., 2002; Sachdev *et al*., 2002). To date, the incidence of these phenocopies has been estimated to be ~10% (Marian, 2010). Causes for some of these cases include mutations in the genes that encode the γ2 subunit of AMP-activated protein kinase (*PRKAG2*) and the lysosome-associated membrane protein (*LAMP2),* which represents metabolic cardiomyopathies. While both entities present with LVH; these disorders are considered to be distinct from HCM based on the presence of prominent vacuoles and electrophysiological abnormalities and the absence of myocyte disarray and interstitial fibrosis (Arad *et al*., 2005; Yang *et al*., 2005).

Table1.1. Sarcomeric genes implicated in the pathogenesis of HCM

Protein	Gene	Chromosome	Frequency (%)	Number of mutations
β-myosin heavy chain	*MYH7*	14q1	40.3	289
α-myosin heavy chain	*MYH6*	14q1	1.3	9
Cardiac myosin binding protein C	*MYBPC 3*	11q1	37	264
Cardiac troponin T	*TNNT2*	1q3	7.0	47
Cardiac troponin I	*TNNI3*	19p1	5.2	37
Cardiac troponin C	*TNNC1*	3p	1.4	10
α-Tropomyosin	*TPM1*	15q2	2.7	19
Myosin essential light chain	*MYL3*	3p	1.5	11
Myosin regulatory light chain	*MYL2*	12q	1.8	13
α-cardiac actin	*ACTC*	11q	2.1	15
Titin	*TTN*	2q3	<1	2

Confirmed mutations as listed on HCM mutation databases at: http://www.angis.org.au/Databases/Heart, http://www.hgmd.cf.ac.uk/ac/index.php and http://cardiogenomics.med.harvard.edu/genes/gene-list.

To date the genetic spectrum of HCM has also expanded beyond the scope of sarcomeric proteins to include mutations in sarcomeric-related genes. Mutations have been found in genes that encode the Z-disk proteins muscle LIM protein (*CSRP3*), obscurin (*OBSCN*), telethonin (*TCAP*), vinculin/metavinculin (*VCL*) and in genes that encode calcium (Ca^{2+})-handling proteins, which include the cardiac ryanodine receptor (*RyR2*), junctophilin-2 (*JPH2*) and phospholamban (*PLN*) (Hayashi *et al*., 2004; Chiu *et al*., 2007; Geier *et al*., 2008; Xu *et al*., 2010). However, the evidence for their pathogenicity is less convincing than for that of sarcomeric genes.

Taken together, these different mutations in the different genes, whether sarcomeric or sarcomeric-related, illustrate the genetic heterogeneity of HCM.

1.1.2. Clinical variability and genotype:phenotype associations in HCM

Initial reports of genotype-phenotype correlations led to the labelling of mutations as being benign (viz. with low risk of SCD), intermediate, and malignant (viz. with high risk of SCD). For example, mutations in *TNNT2* were shown to be associated with a higher incidence of SCD, but with mild or no cardiac hypertrophy (Moolman *et al*., 1997). On the other hand, various studies have shown that mutations of the *MYH7* gene, in particular R403Q, R453C and R719W, are associated with an early onset of disease, a higher degree of hypertrophy and a higher incidence of SCD; *MYBPC3* mutations, on the other hand, have been associated with a milder phenotype and a later onset of the disease (Watkins *et al*., 1992; Nimura *et al*., 1998; Ackermann *et al*., 2002). However, because of significant variability both within families and between different pedigrees with the same mutation, it has been suggested that these associations are not absolute (Arad *et al*., 2002; Moolman-Smook *et al*., 2000). The latter statement is clearly illustrated with the recent identification of severe and early forms of HCM and SCD in patients with *MYBPC3* mutations (Rodiriguez-Garcia *et al*., 2010). For example, the 25-bp deletion in intron 32 of cMyBPC usually presents with a mild phenotype at the early stages of life, while older individuals present with a severe phenotype. However in the same cohort, seven probands with a mean age of just 20 years also displayed a severe phenotype (Dhandapany *et al*., 2009). Furthermore, in a follow-up study, individuals with the R92W *TNNT2* mutation, which previously showed little to mild hypertrophy, developed clinically recognised hypertrophy after the age of 35 years (Revera *et al*., 2008), again emphasising the fluidity of the phenotype associated with a given causal mutation.

Moreover, the presence of multiple mutations (gene dosage effect) has been shown to aggravate the HCM phenotype, since patients with homozygous or compound heterozygous mutations display more severe clinical phenotypes (Mohidden *et al*., 2003; Lekane *et al*., 2006). However, it has been estimated that compound mutations only account for 5% of the variability of the phenotype (Hilfiker-Kleiner and Knoll, 2008). Multiple factors, such as gender, age, blood pressure (BP) and environmental factors such as diet, lifestyle and excercise also contribute to phenotype variability in HCM (Alcalai *et al*., 2008). Together, all of these findings highlight the fact that factors other than the causal gene come into play to contribute to the disease phenotype or to modify it.

1.1.3. Candidate modifier genes in HCM

These phenotype-modifying factors can be environmental or genetic. Modifier genes, unlike disease-causing genes, are neither necessary nor adequate to cause the HCM phenotype, but affect the extent of cardiac hypertrophy and/or the risk of SCD, either directly or indirectly via gene-gene interactions (Marian, 2002).

To date, a number of studies have shown the involvement of modifier genes on the phenotypic expression of HCM, as evident from significant inter- and intrafamilial heterogeneity seen in the disease as well as in animal models using different mouse strains (Epstein *et al*., 1992; Fananapazir *et al*., 1994; Arad *et al*., 2002; Moolman-Smook *et al*., 2000; Semsarian *et al*., 2001). Over the years, several association studies have been performed in an attempt to identify the factors that modify the degree of hypertrophy that develops in a given person. The influences of a number of proposed candidate genes have been assessed, including the endothelin-1 (*END1*), tumor necrosis factor alpha (*TNFα*), transforming growth factor (*TGFβ1*), insulin-like growth factor 2 (*IGF2*) and the angiotensin converting enzyme *(ACE)* genes, all selected based on their proven ability to trigger hypertrophy in cellular systems or their involvement in the control of BP, which also influences LVH (Brugada *et al*., 1997; Patel *et al*., 2000; Marian *et al*., 1993; Marian, 2002).

More recently, the first genome-wide mapping of modifier loci for HCM was performed by Daw *et al* in an HCM cohort with the InsG791 *MYBPC3* mutation; this study further substantiated the effect of modifiers on the HCM phenotype (Daw *et al.,* 2007). This group detected linkage on loci 3q26.2, 10p13, 17q24 and suggested linkage on locus 16q12.2. Further refinement of the 10q13 locus identified candidate modifier genes namely Integrin, alpha 8 (*ITGA8*), CARP *(C10orf97*) and phosphotriesterase-related protein (*PTER)* (Daw *et al*., 2007). Both *ITGA8* and *CARP* were considered plausible candidates as they have been implicated in apoptosis and cardiac fibrosis. However, despite the identification of these candidate modifiers, the study itself was limited by various factors. Firstly, the authors chose to use left ventricular mass (LVM) as a measure of hypertrophy, as the variability in the distribution, as well as the extent, of hypertrophy between mutation-carriers made it impossible to use a single wall thickness as a measurement of ventricle-wide hypertrophy. However, echocardiographically-derived LVM is an inaccurate measure of hypertrophy, as it is calculated based on the assumption that the ventricle exhibits a uniform, geometric shape, which is not the case in the asymmetric LVH typically found in HCM. Moreover, this mutation is associated with a late onset of disease; therefore the collection of phenotypic data at a single time point at an early age may have been another limiting factor. More importantly, body weight data were not available on all the subjects, and thus LVM was not indexed to body size, which is known to be a covariate of heart size; this implies that associations found may reflect associations with covariates that were not adjusted for. All the abovementioned limiting factors increase the background „noise' of this study, making it difficult to interpret the association results.

Another genetic factor which has been extensively studied as a potential modifier of LVH in HCM is the insertion/deletion polymorphism in intron 16 of the *ACE* gene, which forms part of the renin-angiotension-aldosterone system (RAAS) (Lechin *et al*., 1995). In particular, the D/D genotype has been associated with a

higher risk of SCD and progression of hypertrophy in the context of HCM in individuals carrying the *MYH7* Arg403Gln mutation (Zhu *et al*., 2001; Doolan *et al*., 2004). In addition, proteins involved in Ca^{2+} regulation and homeostasis (e.g. phospholamban) as well as myocyte energetics (e.g. frataxin) have been studied as possible candidate modifier genes (Minamisawa *et al*., 2003; Van Driest *et al*., 2005; Medin *et al*., 2007). More recently, the EUROGENE Heart Failure project also suggested a -34T>A polymorphism in calmodulin III (*CALM3*) as a modifier of HCM in individuals carrying mutations in both the *MYH7* and *MYBPC3* genes (Friedrich *et al*., 2009).

Although association has been found between some of these genetic variants and hypertrophic indices, caution regarding the interpretation of these results is advised: many of these studies still need to be replicated and confirmed in other populations, since association studies in general are limited by confounding factors such as population stratification, phenotypic heterogeneity, statistical power, sample size and confounding variables.

However, that genetic modifiers do exist is further supported by data from animal models, as evident from the study done by Semsarian *et al* (2001). This study, which involved the breeding of transgenic (TG) *MYH7* Arg403Gln mice into different genetic backgrounds, showed that phenotypic differences are observed in terms of hypertrophy, histopathology and their exercise capacity (Semsarian *et al*., 2001). Thus, it is clear that HCM, although defined as a monogenic cardiac disorder, is in a sense genetically more complex due to the contribution of additional genetic loci and environmental and lifestyle factors.

1.1.4. Possible mechanism by which sarcomeric gene mutations cause HCM

To date, two mechanisms have been proposed by which sarcomeric gene mutations produce HCM. Most HCM missense mutations have been proposed to act as poison peptides through a dominant-negative effect, i.e. the mutant proteins are incorporated into the sarcomere where they alter normal sarcomeric functions. In contrast, nonsense and frameshift *MYBPC3* mutations have been speculated to act as „null alleles'leading to haploinsuffiency, i.e. there is not enough protein for proper sarcomeric function. Thus, for the following section, a brief overview of these two mechanisms will be given.

1.1.4.1. Altered sarcomeric function – the dominant negative effect

In an effort to understand the common mechanism whereby contractile proteins can cause hypertrophy, Redwood *et al* summarised findings which revealed that HCM-causing mutant proteins have very divergent functional effects (Redwood *et al*., 1999). Major physiological and contractile parameters that were investigated included myosin adenosine triphosphatase (Mg^{2+}-ATPase) activity, maximum Ca^{2+}-activated force (F_{max}), systolic function (contractility), diastolic function (relaxation), sliding velocity and Ca^{2+}-sensitivity of contraction (Redwood *et al.*, 1999). Interestingly, it has been shown that the majority of the mutant contractile proteins resulted in an enhanced contractile state (hypercontractility), eg., *TMP1* and *TNNT2* mutations, although others had the opposite effect with a decreased contractile state (hypocontractile), eg., some *MYH7* mutations.

Earlier functional analyses demonstrated that isolated hearts from TG mice expressing the *TNNT2* Arg92Gln mutation showed increased Ca^{2+}-sensitivity; hypercontractility and diastolic dysfunction (Tardiff *et al*., 1999). Similar results were also reported by Chandra *et al* for mice expressing the same and other *TNNT2* mutations (Chandra *et al*., 2001). Increased Ca^{2+}-sensitivity and diastolic dysfunction were also observed for both the Asp175Asn and Glu108Gly *TPM1* mutations in muscle biopsies from patients as well as TG mice expressing these mutations (Thierfelder *et al*., 1994; Bottinelli *et al*., 1998; Muthuchamy *et al*., 1999; Michele *et al*., 2002; Guinto *et al*., 2009).

In contrast, functional analyses using skinned muscle fibres and *in vitro* filament assays as well as TG mouse and rabbit model studies of the *MYH7* Arg403Gln mutation all showed a decrease in myosin motor function. Collectively, a decrease in ATPase activity, crossbridge cycling rate and reduce force has been observed in these different experimental assays (Lankford *et al.,* 1995; Geisterfer-Lowrance *et al*., 1996; Tyska *et al*., 1998; Marian *et al*., 1999). However, recent work, contradicts these earlier results regarding the *MYH7* Arg403Gln mutation, and showed that actin-dependent ATPase, *in vitro* sliding speed as well as force production was increased (Tyska *et al.,* 2000). Similar results were also observed *in vivo* by Blanchard *et al* (1999) and Keller *et al* (2004). Furthermore, in TG mice in which endogenous α-myosin heavy chain (MHC) was replaced with transgenically encoded βMHC, the *MYH7* Arg403Gln mutation showed a decrease in ATPase activity, which was opposite to the increased ATPase activity observed in the αMHC background (Lowey *et al*., 2008). Thus it is evident that the functional effect of the *MYH7* Arg403Gln mutation changed, depending on which cardiac MHC isoform was used; this might account for the differences observed in the intial studies. Nevertheless, in most of the studies, diastolic function was generally impaired, while the Ca^{2+}-sensitivity of contraction for most studies was enhanced due to the presence of both mutant and normal proteins within the sarcomere.

1.1.4.2. Reduction of functional protein - haploinsufficiency

In contrast, the mechanism by which *MYBPC3* mutations cause HCM remains unclear. Certainly, the absence of detectable truncated proteins in western blot analysis of endomyocardial biopsies obtained from various studies to date argues against a dominant-negative effect, and rather points to a haploinsufficiency model of patho-etiology. In addition, slightly lower levels of both mutant mRNAs and full length cMyBPC was observed in *MYBPC3* missense and truncation mutation-carriers than in controls (Rottbauer *et al*., 1997; Vignier *et al*., 2009; Marston *et al*., 2009; van Dijk *et al*., 2009). Consistent with this, TG mice overexpressing C-terminally truncated mutants (lacking domains C8-C10; see sections 1.2.3-1.2.4 and Figures 1.2 for a detailed discussion of the structure and function of cMyBPC) also showed lower levels of the cMyBPC mutant protein (Yang *et al*., 1999). Moreover, no truncated cMyBPC was observed in the myocardium of other mouse models of HCM and dilated cardiomyopathy (DCM) in which domains C8-C10 of this protein was truncated (McConnell *et al.*, 1999; Palmer *et al.*, 2004).The absence of the truncated protein was speculated to be the result of rapid degradation via the UPS (see section 1.4.3 for an overview of

the UPS). In 2005, Sarikas *et al* confirmed the involvement of the UPS in the degradation of cMyBPC truncated proteins, and found that certain truncated mutants, in turn, also impaired the functioning of the UPS (Sarikas *et al*., 2005). In addition, the nonsense-mediated mRNA decay pathway was also proposed to play a role in the haploinsuffiency mechanism based on the lower levels of cMyBPC mutant mRNAs observed in cardiac tissue of cMyBPC mutant carriers (Marston *et al*., 2009; Van Djik *et al*., 2009; Vignier *et al*., 2009). Therefore, these two quality control systems have been proposed to contribute to the lower levels of cMyBPC mutants seen in the various studies and most likely account for the mechanism by which *MYBPC3* mutations cause the disease phenotype.

In contrast, Flavigny *et al*, in 1999, showed that C-terminally truncated cMyBPC is incorporated into the A-band of fetal rat cardiomyocytes, which favours the poison peptide mechanism (Flavigny *et al*., 1999). They speculate that the incorporation is facilitated by the interactions between the N-terminal MyBPC motif and myosin subfragment 2 (S2) or a possible binding site for myosin in the MyBPC N-terminal domain C0 (Section 1.2.3 and Figure 1.2 for a detailed discussion of the structure and function of cMyBPC) (Witt *et al*., 2001; Flavigny *et al*., 1999). Given later evidence for extensive interaction between different domains in the N-terminal of cMyBPc and actin, the latter proposition certainly seems plausible (Kulikovskaya *et al*., 2003; Squire *et al*., 2003; Whitten *et al*., 2008; Razumova *et al*, 2008; Shaffer *et al*., 2009).

Taken together, these findings illustrate that the initial functional defect that triggers hypertrophy in HCM, is not uniform and is due to different causal mutations, in different proteins, which have different properties and roles. It also shows that mutations differentially affect the Ca^{2+}-sensitivity of contraction, via different routes. However, in the end, the underlying defect in HCM is altered contractility of a defective sarcomere. Consequently, the altered contractility of the sarcomere triggers the release of growth factors resulting in cardiac hypertrophy and fibroblast proliferation (Ahmad *et al*., 2005). Therefore, it can be hypothesised that molecules that affect contractility can act as modifiers of the hypertrophic signal, and therefore influence the development of hypertrophy.

One of the major regulators of cardiac contractility, located within the sarcomere, is MyBPC. Hence, understanding factors that influence the function of this protein is important to enhance understanding of the regulation of contractility, and the development of hypertrophy in heart muscle. The following sections will therefore focus on this integral thick filament protein and the corresponding gene, *MYBPC3*.

1.2. CARDIAC CONTRACTILITY AND SARCOMERE STRUCTURE

Cardiac contractility in the heart is tighly regulated at the levels of Ca^{2+} homeostasis, cell signalling and through the maintenance of the sarcomere, which is the smallest contractile unit of cardiac muscle. The sarcomere, which is comprised predominantly of myosin thick filaments and actin thin filaments, is further divided into the the M-line, I-band, A-band and Z-disc (Figure 1.1). The Z-discs anchors the actin thin filaments of the I-band, while the M-line anchors the myosin thick filaments of the A-band (Lang and

Nowak, 2005). Contractile force is generated as a result of actomyosin crossbridge cycling, which causes the thick and thin filaments to slide past one another (Winegrad, 1999). However, contractile force is also regulated with the help of accessory proteins, which are essential for the 3-dimensional (3D) organisation of the contractile proteins as well as the modulation of their interactions (Willis *et al*., 2009).

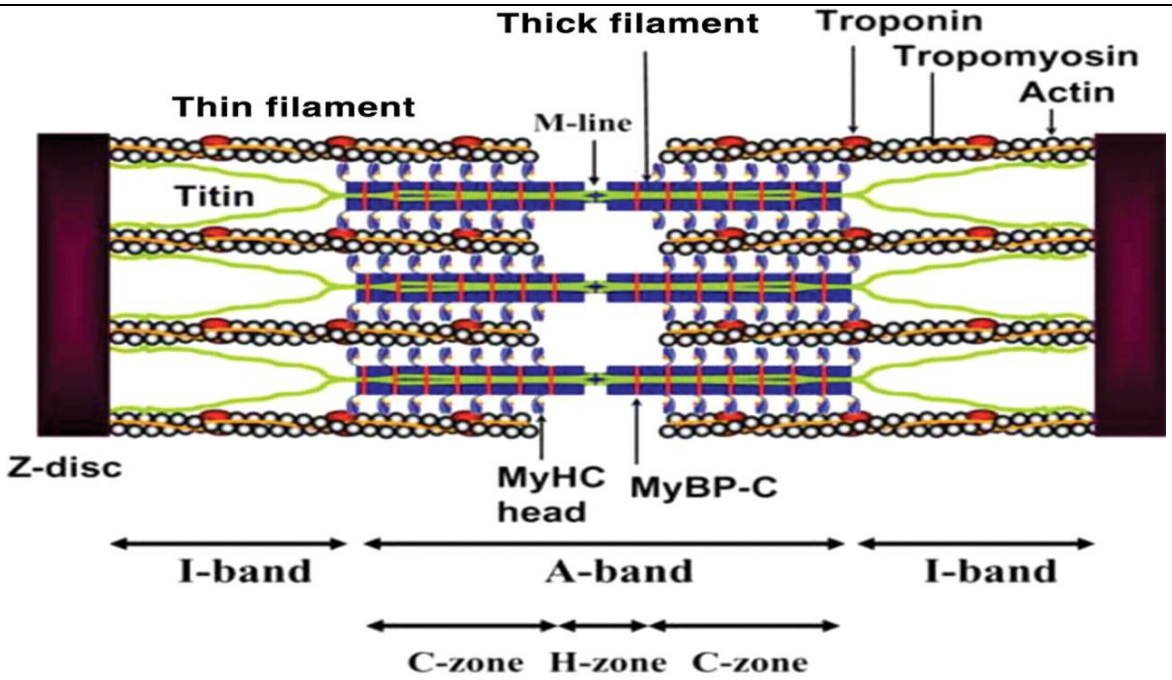


Figure 1.1. Schematic representation of the the sarcomere depicting various myofibrillar components relative to each other. The regions of the I- and A- band, M-line and the C- zone, which make up the sarcomere, are indicated at the bottom. Thin filaments components, i.e., actin, tropomyosin, troponins C, I, and T, constitute the I-band, which overlaps with the the thick filament in the A-band. The A-band is comprised of the H-zone, which only contains the thick filament, and the C-zone, where the thick and thin filaments overlaps. (Taken from: Morimoto, 2007)

1.2.1. Localisation of MyBPC within the sarcomere

Within the sarcomere, MyBPC molecules are located in the crossbridge region (C-zone) of the A-band on 7-9 transverse stripes spaced at regular intervals of 43nm (Craig and Offer, 1976). A recent study by Luther *et al* confirmed that immunolabelling of cardiac muscle revealed that the nine stripes distal to the M-line are the location of cMyBPC (Luther *et al*., 2008). Two to four MyBPC molecules are positioned on each stripe and the spacing of the transverse stripes correlate with the domain repeats of titin and with the helically arranged crossbridges of myosin; thus MyBPC binds to myosin at every third crossbridge (Rome *et al*., 1973; Hartzell and Sale, 1985).

Moreover, 3D models of vertebrate striated myosin filaments were constructed from either mice or rabbit cardiac muscle to look at the relationship of MyBPC and myosin in intact muscle. These models revealed three crown levels within each 43nm interval, indicative of the paired myosin heads, and that MyBPC is likely located between crown 1 and 2 (Al Khayat *et al*., 2008; Kensler *et al*., 2008; Luther *et al*., 2008): A mass density observed between crown 1 and 2 is more visible in contour plots at a low radius closer to the filament axis, where its not obscured by the contribution of the myosin heads (Al-Khayat *et al*., 2008). The

absence of this mass density and a weak crown 1 density from MyBPC-deficient filaments further supported that MyBPC is located between crown 1 and 2 and that it extends from the filament surface at crown 1 (Luther *et al*., 2008).

1.2.2. MyBPC isoforms

MyBPC exist in three isforms, which are species- and fibre-type specific. Two isoforms are found only in skeletal muscle: fast skeletal (fs) and slow skeletal (ss) MyBPC, and the third is found only in the heart, viz. cardiac MyBPC (cMyBPC); each isoform is encoded by a different gene (Yamamoto and Moos, 1983). The *MYBPC1* (12q23.1) and *MYBPC2* genes (19q13.33) encode the ssMyBPC and fsMyBPC proteins, respectively, while the *MYBPC3* gene, which is located on chromosome 11p11.2, encodes cMyBPC (Weber *et al*., 1993; Gautel *et al*., 1995).

1.2.2.1. Mutations in *MYBPC3* - one of the most common causes of HCM

MYBPC3 is frequently mutated and accounts for the second most common form of HCM. Ever since the identification of *MYBPC3* as an HCM-causing gene in 1995 (Bonne *et al*., 1995; Watkins *et al*., 1995), the number of mutations in this gene associated with HCM has increased. To date, at least 264 *MYBPC3* mutations have been identified, accounting for 37% of all familial HCM cases (http://www.hgmd.cf.ac.uk/ac/gene.php?gene=MYBPC3).

The prevalence of *MYBPC3* mutations has been shown to vary between different populations (Richards *et al*., 2003; Erdmann *et al*., 2003). For instance, in some European countries, *MYBPC3* mutations are more prevalent than in others (15-20% in Spain, 26% in France, 21.7% in Sweden and 24% in Finland). The high prevalence in some countries has led to the conclusion that *MYBPC3* mutations are the most common cause of HCM, rather than *MYH7* mutations, as first reported. However, as the frequency in other countries is very low (4% in South Asia) (Morner *et al*., 2003; Konno *et al*., 2003; Van Driest *et al*., 2005; Dhandapany *et al*., 2009; Garcia-Castro *et al*., 2009; Rodriguez-Garcia *et al*., 2010), this may not be the case; current consensus suggests that *MYBPC3* is one of two genes (with *MYH7*) most commonly implicated in HCM.

Unlike mutations in the other HCM-disease causing genes, which are almost exclusively missense mutations (resulting in single amino-acid substitutions), about 70% of *MYBPC3* mutations are nonsense (thus introducing a premature stop codon and truncated protein) or frameshift mutations (affecting splice donor and acceptor sites and/or insertions and deletions, resulting in the truncation of the cMYBPC protein) (Carrier *et al*., 1997; Alcalai *et al*., 2008).

Over the years, the disease profile associated with *MYBPC3* mutations has been considered to be mild, based on incomplete disease penetrance and a late onset of disease observed in many *MYBPC3* mutation carriers. Because of this profile, Richard *et al* suggested that the prevalence of *MYBPC3* mutations has been underestimated (Richard *et al*., 2003). While the majority of the *MYBPC3* mutations have been associated

with a milder phenotype, there have been exceptions to the rule, with some mutations displaying a more severe phenotype (Ehlermann *et al*., 2008; Dhandapany *et al*., 2009). Interestingly, two mutations, a 25-bp deletion in intron 32 and a frameshift at Gln791, are normally associated with a benign phenotype in the reproductive period. However, individuals with these mutations, who are over the age of 40, present with a severe phenotype (Van Driest *et al*., 2004; Dhandapany *et al*., 2009). Furthermore, Erdmann *et al* showed that a more severe phenotype has been associated with truncation mutations than that associated with missense mutations (Erdmann *et al*., 2001). A good illustration of this is the homozygous splice site mutation c.3330 + 2T>G, which produced a truncated protein lacking the C10 domain and is associated with severe neonatal HCM (Xin *et al*., 2007). More recently, cases were also identified where young children with *MYBPC3* mutations already had a severe phenotype (Alcalai *et al*., 2009).

1.2.3. MyBPC structure

MyBPC forms part of the immunoglobulin superfamily and its basic structure is comprised of a series of immunoglobulin (Ig) and fibronectin type III (FnIII) domains designated C1-C10 from the N- to the C-terminus. Of particular interest to this study is cMyBPC, which has additional cardiac-specific characteristics: an extra N-terminal IgI-like domain (C0), a proline/alanine (PA)-rich region between domains C0 and C1, a stretch of amino acids (LAGGGRRIS) inserted into the MyBPC motif (M), which lies between domains C1 and C2, three phosphorylation sites in the MyBPC motif (one within the cardiac specific LAGGGRRIS sequence) and a 28 amino-acid insertion within the C5 domain (Figure 1.2) (Gautel *et al*., 1995). Moreover, species-specific differences in the sequence of the PA region of the cMyBPC isoform have been observed (Shaffer and Harris, 2009). Small angle X-ray diffraction data revealed that the structure of the MyBPC motif is compact, is folded in solution and is similar in size and dimensions as it adjacent Ig domains (Jeffries *et al*., 2008). Thus, this modular protein is ~3nm in diameter and spans 32nm and 40-44nm in length for the skeletal and cardiac isoforms, respectively (Hartzell *et al*., 1985).

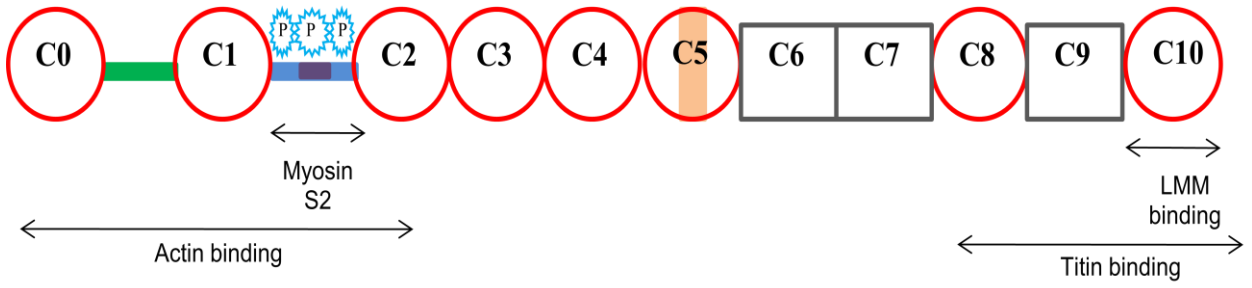


Figure 1.2. The basic domain structure of cMyBPC.
The eight immunoglobulin-like (Ig) domains are represented by the red circles, while the fibronectin type 3-like (FnIII) domains are represented by the grey squares. The green linker between domains C0 and C1 indicates the Pro-Ala-rich region, the blue rectangle indicates the cMyBPC motif between domains C1 and C2, while the purple rectangle within the cMyBPC motif indicates the LAGGGRIS insertion. Furthermore, the light blue stars indicate the three phosphorylation sites within the cMyBPC motif, while the light orange rectangle indicates the cardiac-specific insertion in C5. The arrows indicate the cMyBPC domains involved in the interactions with actin, titin and myosin.

1.2.4. Biological functions of MYBPC

Although the various aspects of the function of cMyBPC are still being unravelled, several studies have described functional roles in the structural assembly and stability of the thick filament, as well as in regulating cardiac contractility, based on the interactions of cMyBPC with various sarcomeric proteins. Therefore, the interactions and the functions of MyBPC will be discussed jointly in the following sections.

1.2.4.1. Sarcomere and thick filament assembly

1.2.4.1.1. MyBPC interactions with LMM, C-zone titin, LIM,Obscurin and itself

Over the years, a structural and stabilising role in filament and sarcomere assembly has been assigned to MyBPC, based on its interactions with the light meromyosin (LMM) region of myosin and titin. Data showed that the C-terminal part of MyBPC, in particular, domain C10, was able to bind the LMM segment of myosin (Okagaki *et al*., 1993; Flashman *et al*., 2007). Although domain C10 was shown to be vital for this interaction, the added effect of domains C7-C9 were shown to be necessary for maximum myosin binding and for the correct incorporation of MyBPC in the A-band (Welikson and Fischman, 2002).

Interestingly, the C-terminal domains (C8-C10) of MyBPC, which includes the LMM binding site of myosin, also houses the binding site for titin (Freiberg and Gautel, 1996). Despite the similar domains involved, data have shown that the binding affinity of titin for MyBPC is much weaker compared to the binding affinity between MyBPC and the LMM segment of myosin. However, the cooperative interaction of MyBPC, myosin and titin makes the thick filament a strong and stable structure (Freiberg and Gautel, 1996). Furthermore, MyBPC binding to titin *per se* is more likely to dictate its position in the C-zone of the sarcomere, as it binds to the 11 domain super repeats of A-band titin (Labeit *et al*., 1992; Trinick, 1996). In addition, mice expressing a truncated form of MyBPC, lacking the important C-terminal titin- and myosin-binding sites, displayed disorganized sarcomere patterns and ineffective MyBPC incorporation into the sarcomere, which further supported a potential role for cMyBPC in sarcomere assembly (Flavigny *et al*., 1999).

MyBPC was also shown to bind the scaffold protein, four-and-a-half LIM protein 1, which functions in myosin filament formation as well as in sarcomere assembly. Overexpression of LIM in differentiating C2C12 cells was shown to induce impaired Z-line and myosin thick filaments, a phenotype which could be rescued upon co-expression with exogenous MyBPC (McGrath *et al*., 2006).

Recently, a novel isoform of ssMyBPC (involving part of the C10 domain of ssMyBPC (amino acids 1071-1119), followed by a novel 26 amino-acid sequence) was shown to bind to the N-terminus of obscurin, a protein which is known to play a role in the organisation of contractile structures (Ackermann *et al*., 2009). Furthermore, the overexpression of the Ig2 domain of obscurin was shown to severely disrupt M-line and A-band formation, but not that of Z-disks or I-bands. Thus, the interaction between obscurin and MyBPC was

suggested to play a key role in the integrity of the M-band and thick filament assembly (Ackermann *et al*., 2009).

MyBPC was also shown to interact with itself. Moolman-Smook *et al* set out to determine interactors of the cardiac-spesific sequence in domain C5 using Y2H analysis. This study revealed that domain C5 binds domain C8 (Moolman-Smook *et al*., 2002). Further support from studies indicating that MyBPC molecules are able to multimerize (Offer *et al*., 1973; Hartzell and Sale, 1985), led to this group's proposal that three cMyBPC molecules trimerize to form a collar around the thick filament (Section 1.2.4.1.2) (Moolman-Smook *et al*., 2002).

1.2.4.1.2. Proposed structural roles

Earlier *in vitro* studies involving synthetic myosin filaments established the importance of MyBPC in thick filament formation (Moos *et al* 1975; Koretz, 1979). The presence of MyBPC was shown to alter the properties of these filaments by decreasing the diameter and increasing the length and uniformity of these filaments until it resembled native thick filaments with regard to thickness, length, bare zone and distribution of myosin heads (Koretz, 1979; Levine *et al*., 2001). Furthermore, overexpression of sarcomeric MHC in COS cells resulted in the formation of diffuse spindle-like aggregates in the cytoplasm. However, co-expression of both MyBPC and MyBPH (a myosin binding protein mainly found in fast skeletal muscle and in some regions of cardiac muscle) with MHC resulted in the formation of long peri-nuclear cable-like co-polymers containing MyBPs and MHC (Welikson and Fischman, 2002). In addition, co-expression of MHC and truncation mutants of both MyBPC and MyBPH that lacked the last C-terminal IgI domain abrogated cable formation, demonstrating the importance of domain C10 (MyBPC) and its equivalent domain in MyBPH, H4, in thick filament formation (Welikson and Fischman, 2002). These findings further substantiate the importance of myosin binding proteins, and particularly highlight the role of domain C10 of cMyBPC in the cable formation of the thick filament.

On the other hand, data obtained from a cMyBPC knockout (KO) mouse model suggest a modulatory rather than an essential role for cMyBPC in sarcomere assembly and maintenance: cMyBPC knockout mice were viable and displayed apparently well-developed sarcomeres, although the homozygous mice displayed morphological features classic for HCM, as well as impaired diastolic and systolic dysfunction (Harris *et al*., 2002). These findings suggest that, in the absence of cMyBPC, other proteins might compensate to allow formation of normal sarcomere structure.

To further clarify whether the absence of cMyBPC had an effect on the structure of the thick filament or the arrangement of the myosin heads, a new study by Kensler *et al* more closely investigated the structure of isolated cardiac myosin thick filaments from cMyBPC knockout mice (cMyBPC $^{-/-}$) (Kensler *et al*., 2008). Using electron microscopy and image analysis, this group compared thick filament structure from hearts of cMyBPC $^{-/-}$ mice with those from wild type (WT) mice (cMyBPC $^{+/+}$). Data was consistent with the results

from Harris *et al*, showing that, despite the absence of cMyBPC, thick filaments were stable in length, diameter, and appearance and retained the 43nm crossbridgeperiodicity, similar to that seen in WT mice hearts (Harris *et al*., 2002; Kensler *et al*., 2008). However, this 43nm periodicity was shown to be easily disrupted and more disordered in cMYBPC $^{-/-}$ thick filaments compared to WT mice (cMYBPC$^{+/+}$). Furthermore, the effect of cMyBPC seemed to be rather on the arrangement of the myosin heads than the structure of the thick filament *per se*. Filaments lacking cMyBPC displayed disordered myosin heads which disturbed the helical or quasihelical arrangement of crossbridges (Kensler *et al*., 2008).

Thus, it is evident from the literature to date, that MyBPC, although not essential for sarcomere assembly, is indeed needed for proper sarcomere organisation and stability.

1.2.4.1.3. Current proposed models of MyBPC arrangement on the thick filament

Taken together, all of these interactions with various sarcomeric proteins led to proposed models for the arrangement of MyBPC on the thick filament. To date, two models exist, the trimeric collar model proposed by Moolman-Smook *et al* (2002) and the alternative model proposed by Squire *et al* (2003) (Figure 1.3).

The trimeric collar model takes into account that MyBPC binds to itself and the fact that the stoichiometry of MyBPC is compatible with the presence of three MyBPC molecules. Thus, domains C5 through C10 of the three molecules would wrap around the thick filament, with successive and parallel molecules interacting with each other via domains C5-C8 from one and C7-C10 from the other MyBPC molecule, ultimately forming a threefold collar arranged perpendicular to the thick filament axis (Figure 1.3A). The N terminal domains C0-C4 would extend into the interfilamental space to interact with the myosin S2 region and possibly the thin filament (Moolman-Smook *et al*., 2002). Additionally, this group also showed in 2008 that the interaction between C5 and C8 is not dependent on the 28-residue cardiac-spesific insert in C5 and that the C5:C8 interaction occurs also in fast skeletal muscle but not in slow skeletal muscle (Flashman *et al*., 2008). Moreover, the collar model might explain the changes in thick filament structure that are observed upon phosphorylation of cMyBPC (Section 1.2.4.2.4.1) (Winegrad *et al*., 1999). It is hypothesised that upon phosphorylation of the MyBPC motif, the introduced negative charges may disrupt the postulated interactions resulting in a domain shift, which in turn could cause the expansion of the collar that allows the thick filament to adopt a loose structure with an increased diameter (Flashman *et al*., 2004).

In contrast, the alternative model of Squire *et al* proposed that the C-terminal end of individual MyBPC molecules aligned axially on the thick filament, thus MyBPC molecules would be arranged parallel to titin, and parallel to, but removed from, each other, while the N-terminus also reaches out to interact with myosin crossbridges and/or actin (Squire *et al*., 2003, Figure 1.3B). Although the model explains the C8C10 cMyBPC interactions with titin, the specific interactions between MyBPC molecules (the C5:C8 and C7:C10 domain interactions shown by Moolman-Smook *et al* [2002]), as well as MyBPC multimerisation shown by

other groups are not taken into account in this model (Craig *et al*., 1976; Winegrad *et al*., 1999; Moolman-Smook *et al*., 2002; Squire *et al*., 2003).

Evidence gained from 3D reconstructed cardiac filaments isolated from both mouse and rabbit muscle ventricles still support both models. Both studies from Zoghbi *et al* and Al-Khayat *et al* suggest that MyBPC runs parrallel to the thick filament, which gives added support to the linear model (Al-Khayat *et al*., 2008; Zoghbi *et al*., 2008). However, according to Al Khayat *et al* the trimeric collar model cannot be totally excluded since a feature of uncertain origin was observed between crowns 1 and 2 in the circumferential sections, which could correspond to a collar (Al Khayat *et al*., 2008). Unlike Zoghbi *et al*, this mass density between crown 1 and 2 was more visible in contour plots at a low radius closer to the filament axis where it was not obscured by the contribution of the myosin heads, and corresponds to the collar model suggestion of a collar running tightly against the thick filament backbone (Moolman-Smook *et al*., 2002; Al-Khayat *et al.,* 2008). Similar findings in favour of the collar model were also observed by Luther *et al* (2008). This group showed that, in the frog, the bulk of cMyBPC was at crown 1 while a mass density was observed between crown 1 and 2 in rodent heart muscles. Loss of this density in MyBPC null filaments further supports the latter observations. According to this group, this, together with the sharp transverse stripes observed, supports a perpendicular arrangement of MyBPC relative to the thick filament (Luther *et al*., 2008).

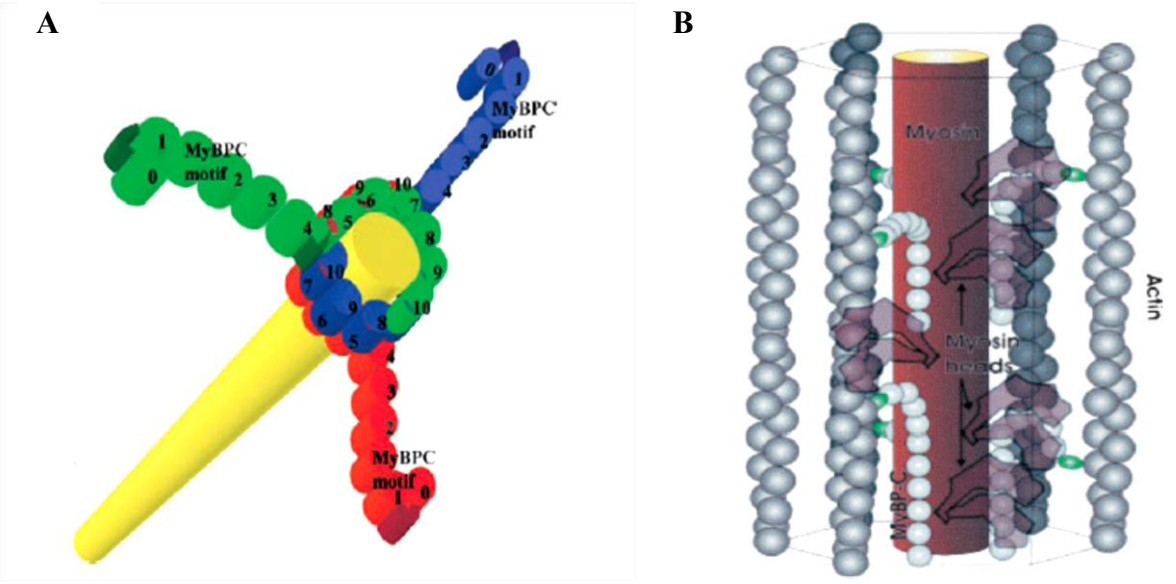


Figure 1.3. Proposed models for cMyBPC arrangement in the sarcomere.
A) In the trimeric model, three cMyBPC molecules wrap around the thick filament via stabilised interactions between domains C5 and C8, and domains C7 and C10, while the N-terminus (C0-C4) extends into the interfilamental space. (Image taken from Moolman-Smook *et al*., 2002) (B) In the alternative model, domains C7-C10 of individual cMyBPC molecules align axially to the thick filament, parallel to titin, while the N-terminus interacts with myosin crossbridges and/or actin. (Image taken from Squire *et al*., 2003)

While there is still uncertainty regarding the structural arrangement of the C-terminal domains of MyBPC with respect to myosin, both models are compatible with an arrangement where the N-terminal domains (C0-

C4) extend into the interfilamental space where interaction of these domains with both myosin crossbridges and actin is facilitated. The interaction of these N-terminal domains with thick and thin filament proteins is crucial for the role of cMyBPC in the regulation of contractility.

1.2.4.2. Regulation of cardiac contraction

A number of studies showed that the N-terminal domains of MyBPC are involved in regulation of contractility; in particular, the MyBPC motif interacts with myosin S2 in a manner that is regulated by the dephosphorylation/phosphorylation of the MyBPC motif (Gautel *et al*., 1995; Kunst *et al*., 2000). Current research also revealed the interaction of the N-terminal domains of MyBPC with actin on the thin filament (Jeffries *et al*., 2008; Shaffer *et al*., 2009); thus, it is possible that this interaction rather than, or in addition to, the MyBPC-S2 interaction contributes to MyBPC regulatory role. This actin interaction is also phosphorylation-dependent; hence it is important to briefly look at phosphorylation of MyBPC.

1.2.4.2.1. Phoshorylation of MyBPC

MyBPC, along with troponin I and T (TNI and TNT), the regulatory light chain of myosin (RLC), titin and tropomyosin (α-TM), are among the myofilament proteins that can be phosphorylated via cyclic-AMP-dependent protein kinase (PKA) and/or calcium/phospholipid-dependent protein kinase (PKC), in response to β–adrenergic and cholinergic agonists (Gruen *et al*., 1999; Yamasaki *et al*., 2002). In addition, cMyBPC can also be phosphorylated via the Ca^{2+}-calmodulin-activated kinase II (CaMKII), but only after a certain level of intracellular Ca^{2+} concentration, which is lower than the threshold for thin filament activation, is reached (McClellan *et al*., 2001).

To date, three phosphorylation sites within the cMyBPC motif have been well characterised and termed sites A (Ser273), B (Ser 282) and C (Ser 302), with site B situated within the cardiac-spesific LAGGGRRIS insertion. PKA was shown to phosphorylate all three phosphorylation sites, PKC was able to phosphorylate only site A and C, while CaMKII only phosphorylates site B. Interestingly, site B is first phosphorylated by CaMKII, before the other sites become accessible for phosphorylation by PKA upon adrenergic stimulation (Weisberg and Winegrad, 1996; McClellan *et al*., 2001). Unlike cMyBPC, the skeletal MyBPC isoforms only house one phosphorylation site, which corresponds to site A of cMyBPC; however, this site is not phosphorylated by PKA.

The interaction between myosin S2 and the MyBPC motif is common to all the MyBPC isoforms, and also involves all the different myosin isoforms. However, it was shown to be abbrogated by phosphorylation of the MyBPC motif in the cardiac isoform (Saddyappan *et al*., 2009). More recently, the MyBPC motif was also shown to bind actin in a phosphorylation-dependent manner (Shaffer *et al*., 2009). Thus, a closer look at these two interactions of N-terminal MyBPC follows.

1.2.4.2.2. MyBPC interaction with myosin subfragment 2 (S2)

The second binding site for MyBPC on myosin resides within the 126 amino-acid region of myosin (S2), while the myosin-binding site localises to the 103 residue MyBPC motif between domain C1 and C2 (Figure 1.2.3, Section 1.2.3) (Gruen and Gautel, 1999). However, the binding affinity of MyBPC for S2 was shown to be much lower compared to its affinity for LMM (Starr and Offer, 1978). Furthermore, this interaction with myosin S2 is completely abolished upon triphosphorylation of the MyBPC motif in response to β-adrenergic stimulation; thus, it was proposed that MyBPC has a role in regulating contractility in response to stress (Gruen *et al*., 1999).

Elimination of these phosphorylation sites in several animal models was also shown to alter crossbridge cycling kinetics (Yang *et al*., 2001; Sadayappan *et al*., 2005), further demonstrating the importance of phosphorylation in cMyBPC's role as a regulator of contractility. Based on these findings, it was suggested that MyBPC acts as a tether, hindering the interaction between actin and the myosin heads by acting as a brake while in its dephosphorylated state (Kunst *et al*., 2000; Harris *et al*., 2004). Consequent phosphorylation of the MyBPC motif would abolish the interaction with S2, causing myosin heads to be in a more favourible position for interaction with actin, and thus crossbridge cycling rates would increase. However, opposing data from *in vitro* studies showed that MyBPC can also have an effect on crossbridge kinetics independent of its myosin S2 interaction (Shaffer *et al*., 2007). This in turn, suggests interactions with the thin filament and/or myosin S1.

1.2.4.2.3. MyBPC interactions with actin

Earlier studies showed that MyBPC can bind actin in solution; however, the precise location and the effect of the binding on contractile function was not determined at the time (Moos *et al*., 1978). In 2003, Kulikovskaya and collegues set out to determine whether the N-terminal part of MyBPC contained the actin binding site; using cosedimentation assays, they showed that fragments of MyBPC containing the C0 domain binds to actin with an increase in Ca^{2+}-sensitivity (Kulikovskaya *et al*., 2003). At the same time, Squire *et al*, using modeling and computational analysis, identified a possible actin-binding domain in the PA-rich sequence between domains C0 and C1 (Squire *et al*., 2003). When different N-terminal fragments were exogenously added to skinned myocyte preparations, Herron *et al* showed that force production as well as crossbridge activity were affected by the PA-rich linker between domains C0 and C1, rather than the C0 domain itself as suggested by Kulikovskaya and colleagues (Herron *et al*., 2006). They also noted that the C0 fragment used by Kulikovskaya *et al* (2003) contained 54 additional residues of the PA-rich linker, which might have contributed to their results (Herron *et al*., 2006). Interestingly, studies performed by Yamasaki *et al* showed that this PA-rich region shares homology to the PEVK region of titin, which has been shown to bind actin (Yamasaki *et al*, 2001).

More recently, studies confirmed that the four N-terminal domains of cMyBPC, comprising of C0, C1, the MyBPC-motif and C2, interacts with F-actin through multiple distict binding sites in a phosphorylation-

sensitive manner, and with an effect on crossbridge cycling (Whitten *et al.*, 2008; Shaffer *et al.*, 2009). Domains C1-M-C2 binds actin when cMyBPC is dephosphosphorylated; however, upon phosphorylation, the interaction between the MyBPC motif and actin is reduced, although not abolished. In comparison, domain C1 remains bound to actin irrespective of the phosphorylation status of the cMyBPC-domain. Furthermore, the binding of the MyBPC-motif to F-actin was not affected by Ca^{2+}-activation of thin filaments and was not diminished by the competition of the MyBPC-motif with myosin S2 (Shaffer *et al.*, 2009). Similary, observation in permeabilized rat trabeculae by Razumova *et al* demonstrated that the C1 domain together with the MyBPC motif forms a functional subunit of cMyBPC, with increased effect on Ca^{2+}-sensitivity, F_{max} and activated force at low levels of Ca^{2+}. In contrast, their recombinant fragments of individual domains of C0, the PA-region and C1 showed no effect on activated force (Razumova *et al.*, 2008). However, new findings by Shaffer *et al* demonstrated that species-specific differences within the PA-region could account for the differences observed regarding force activation between these studies (Shaffer *et al.*, 2010).

Taken together, even though the data presented suggest different possible binding sites for actin in the N-terminal domains; it is clear that interaction between cMyBPC and actin might influence or contribute to crossbridge cycling.It has been proposed that upon binding of dephosphorylated cMyBPC to actin, the binding of myosin heads to actin is blocked, which limits myosin crossbridge formation with actin.

Thus, it can be concluded that cMyBPC's role in the regulation of cardiac contraction is determined by the dynamic interaction of its N-terminal with both myosin and actin in the thin filament in response to phosphorylation of its MyBPC-motif. The consequences of MyBPC phosphorylation on thick filament structure and function will now be summarised.

1.2.4.2.4. The effect of cMyBPC phosporylation on thick filament structure and contraction

1.2.4.2.4.1. Structural changes

Winegrad's laboratory has published numerous studies showing that phoshorylation of cMyBPC has a structural effect on the thick filament backbone and crossbridge arrangement, which in turn would give rise to altered contractile properties (Weisberg and Winegrad, 1996; Weisberg and Winegrad, 1998). Their studies showed that maximal PKA phosphorylation caused crossbridges to extend from the myosin backbone, placing the myosin heads in a more favourable position for attachment to the thin filament. Maximum PKA phosphorylation of cMyBPC causes the myosin filaments to become thicker, more loosly packed and also increased the order of myosin heads (Weisberg and Winegrad, 1998). Furthermore, variability in the degree of order in the crossbridges was also observed and were atrributed to the different levels of phosphorylation, this, in turn, was thought to modify force production in cardiomyocytes.

Consistent with this, a study from the same laboratory by Levine *et al* in 2001 showed that thick filaments had three different structures defined by the tightness of packing of the myosin rods within the thick filament

backbone and the degree of order of the myosin heads. In addition, they also corroborated the relation between the various thick filament structure and phosphorylation of MyBPC (Levine *et al*., 2001). The structures observed were: 1) the tight structure, defined by a narrow diameter and the myosin heads lying along the filament backbone, 2) the disordered structure, defined by myosin heads not being distributed uniformly and extended at different angles and 3) the loose structure, defined by an increase in filament diameter and looser packing of myosin heads, yet with a higher degree of order of the myosin heads (Levine *et al*., 2001). The dephosphorylated state correlated with the disordered structure, while partial phosphorylation produced the more ordered tight structure, and triphosphorylation produced the loose structure (Levine *et al*., 2001; McClellen *et al*., 2001).

It is assumed that upon PKA phosphorylation, the constraints imposed by MyBPC is alleviated, thus the myosin heads extend from the myosin backbone and are in a more favourable position for attachment to the thin filament. As the likelihood of crossbridge formation is increased, this would lead to the acceleration and enhancement of crossbridge-dependent force (Herron *et al*., 2006; Stelzer *et al*., 2006; Nagayama *et al*., 2007).

1.2.4.2.4.2. Altered contractile properties

As mentioned above, changes in phosphorylation results in changes in the thick filament structure and the order of myosin heads; this, in turn, goes hand in hand with changes in force production. The transition between the different structural forms associated with the various phosphorylation states correlated with different F_{max} values, such that thick filaments in the disordered state (dephosphorylated state) correlated with the lowest level of force generation while the loose structure (tri-phosphorylated state) correlated with maximum force production (McClellan *et al*., 2001). Evidence from several studies has shown that cMyBPC phosphorylation via PKA is associated with accelerated crossbridge kinetics (Mg^{2+}-ATPase acticity, F_{max} and Ca^{2+} sensitivity) (McCellan *et al*., 2001; Stelzer *et al*., 2007; Tong *et al*., 2008; Sadayappan *et al*., 2009).

Furthermore, ablation of cMyBPC in MyBPC KO mouse models showed accelerated crossbridge cycling and impaired contractile function, while TG mice which expressed full length cMyBPC up to the level of 40% of the normal native protein showed preserved cardiac function (Stelzer *et al*.,2006; Nagayama *et al*., 2007; Pohlmann *et al*., 2007; Stelzer *et al*., 2007). In contrast, TG mice, in which cMyBPC phosphorylation sites has been ablated, by replacement of the phosphorylatable serines with alanine in order to mimic the dephosphorylated state, showed reduced contractile performance (Sadayappan *et al*., 2005), further emphasizing the importance of cMyBPC phosphorylation in cardiac function. Likewise, van Dijk *et al*, comparing contractile performance in cardiac muscle samples from *MYBPC3* truncation mutation-carriers (c.2373dupG, c.2864_2865delCT) to those of non-failing donors, showed that reduced levels of cMyBPC (levels of full length cMyBPC was 33% lower in mutation carriers) caused a decline in the F_{max}, which they correlated to reduced expression of cMyBPC which in turn result in maladaptive alterations in Ca^{2+} homeostasis and thus altered phosphorylation (van Dijk *et al*., 2009).

Together this data shows the importance of cMyBPC phosphorylation in the maintenance of thick filament orientation and the modification of crossbridge cycling rates. Phosphorylation of cMyBPC relieves the tether-like constrains imposed by cMyBPC interaction and increases actin-myosin crossbridge formation (Colson *et al*., 2008; Lecarpentier *et al*., 2008).

1.2.4.2.5. Cardioprotective role of cMyBPC phosphorylation

From the abovementioned data, it is evident that cMyBPC phosphorylation is important in normal cardiac function. Moreover, earlier studies have showed that changes in contractile protein phosphorylation have been associated with cardiac diseases, since low levels of cTnI phosphorylation have been observed in heart failure (HF) and HCM patients (Jacques *et al*., 2008; Messer *et al*., 2007). Likewise, reduced levels of cMyBPC phosphorylation have also been observed in patients with HF, ischemia-reperfusion (I/R) injury and pathological hypertrophy (Sadayappan *et al*., 2005; Sadayappan *et al*., 2006; Jacques *et al*., 2008). This reduced phosphorylation status of cMyBPC has been associated with an increase of cMyBPC proteolysis by proteases like calpain 1 (Decker *et al*., 2005).

As will be explained in more detail later (see section 1.4.2 for an overview of myofibrillar proteins turnover), cMyBPC undergoes the most rapid turnover compared to other sarcomeric proteins, even under normal physiological conditions (halflife [$t_{1/2}$] ~9.55 hours) (Kulikovskaya *et al*., 2007; Bahrudin *et al*., 2008). However, upon pathological conditions, the rate of proteolysis of cMyBPC further increases (e.g. E334K cMyBPC truncated mutant protein, $t_{1/2}$~2.01 hours), resulting in reduced levels of cMyBPC in the sarcomere, as well as leakage of myosin from the thick filament (Decker *et al*., 2005; Bahrudin *et al*., 2008). This proteolysis of cMyBPC results in the disassembly of the thick filament from the periphery of the myofibril, which, in turn, results in the reduction of crossbridge formation as well as in contractile dysfunction (Decker *et al*., 2005; Sadayappan *et al*., 2006).

This is further underscored by a study done by Kulikovskaya *et al*, who demonstrated that, under normal physiological conditions, cMyBPC exists in two different forms *in vivo*, phosphorylated and dephosphorylated (Kulikovskaya *et al*., 2007). Phosphorylated cMyBPC is stable, and is associated with greater thick filament stability, while the dephosphorylated form is unstable and is associated with lower thick filament stability. Moreover, in intact cardiac cells, the phosphorylated from was protected against proteolysis, while the unstable, dephosphorylated form was shown to be easily proteolysed, causing the release of both cMyBPC and myosin from the thick filament. Interestingly, antibody staining of cMyBPC proteolytic peptides using different anti-cMyBPC antibodies, directed against domains C0C2, C5 and domains C8C9, revealed that proteolysis occurs in the C terminal region, involving domains C5 to C10, and that the N-terminal region remained intact during this process.

Thus, it is evident that cMyBPC undergoes normal sarcomeric turnover under normal physiological conditions. As phosphorylation of cMyBPC N-terminal is inversely correlated to cMyBPC proteolysis, and

because of the reduced phosphorylation of cMyBPC in conditions like HF and ischemia, it has been concluded that phosphorylation of cMyBPC is cardioprotective. For example, TG mice expressing mimics of the cMyBPC motif in the dephosphorylated state (cMyBPC^{AllP-}) had depressed cardiac function. However, the expression of cMyBPC^{AllP+} in a cMyBPC null background rescued cardiac function and also protected against I/R injury (Sadayappan *et al*., 2009).

1.3. COMM DOMAIN CONTAINING PROTEIN 4 (COMMD4)

In order to further understand the function of cMyBPC and to investigate its role in sarcomeric structure and function, our laboratory has over the last few years performed a series of yeast-two hybrid (Y2H) library screens, using different phospho-mimics of the cMyBPC motif, to identify ligands of the C1-C2 region under different phosphorylation conditions (Ramburan A, Phd Thesis, 2008). From one of these screens, a protein called the COMM domain containing protein 4 (COMMD4) was identified as a specific ligand of the cMyBPC motif. More interestingly, COMMD4 was found to bind to the cMyBPC motif in a manner analogous to the binding of myosin S2: i.e. when the motif is dephosphorylated or mono-phosphorylated, COMMD4 is able to bind, but when the motif is in its triphosphorylated state, the binding between COMMD4 and the cMyBPC motif is abrogated in the yeast system (Ramburan A, Phd Thesis, 2008). Very little is known about COMMD4, and elucidating its function in the context of the cMyBPC interaction was the primary objective of this present study. A summary of what is known about the COMMD protein family follows below.

1.3.1. Identification of COMMD4 and other COMMD proteins

COMMD4 forms part of a family consisting of ten COMM proteins, each with a unique conserved motif at its C-terminal, called the copper metabolism gene MURR1 (COMM) domain (Burstein *et al.*, 2005) (Figure 1.4). COMMD1, the prototype of the family, previously known as MURR1, was first identified as the orthologue of the imprinted mouse murine gene *U2af1-rs1* region 1, hence its initial name. At the time of that study, the protein showed no homology to any other protein or identifiable domains that could indicate its putative function (Nabetani *et al*., 1997). However, in 2002, Van de Sluis *et al* showed that Bedlington terriers suffering from canine copper toxicosis had a homozygous deletion of exon 2 in *MURR1* (Van de Sluis *et al.,* 2002). These dogs were characterised by inefficient excretion of copper via the bile, which resulted in copper accumulation in the liver, and ultimately lead to hepatitis and eventually liver cirrhosis (Van de Sluis *et al*., 2002). Thus, the findings of Burstein *et al* suggested a function for MURR1 involving copper homeostasis. The identification of additional proteins with structural homology to MURR1 led to the new classification of this family as the copper metabolism gene MURR1 (Mouse-U2af1-rs region 1 locus) domain (COMMD) proteins (Burstein *et al*., 2005). This nomenclature is used henceforth in the thesis.

COMMD3, -4 and -6 were initially identified as interactors of COMMD1, which led to the search for additional COMMD1 homologs. Little detail is available for most of the COMMD family members,

including COMMD4 which is of interest to this thesis; most of the available information in the literature has been deduced from COMMD1.

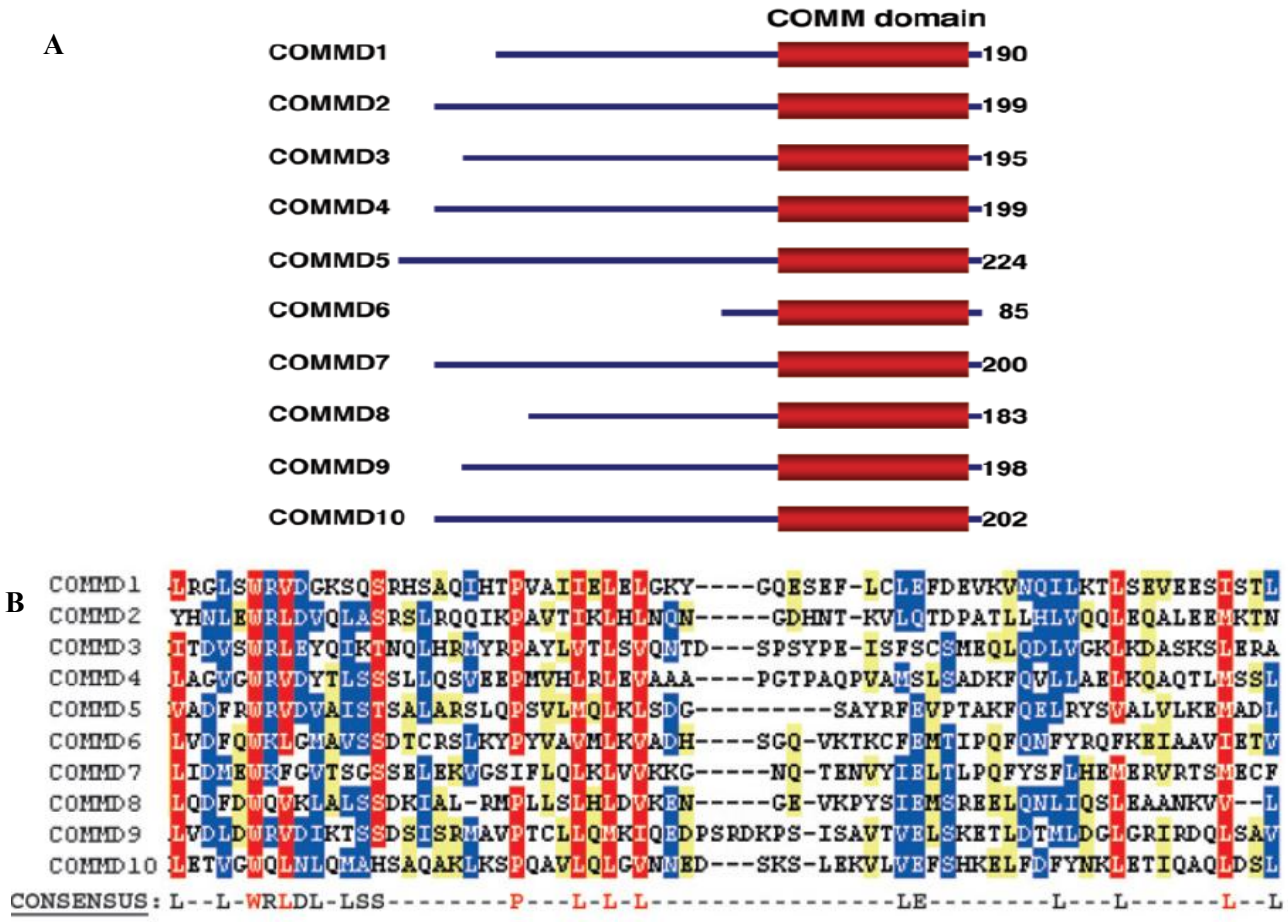


Figure1.4. A schematic representation of the COMMD protein family and amino acid sequence alignment of the COMM domain
A) The conserved COMMD domain is indicated with red blocks, and numbers indicate the respective amino-acid length of each protein. B) Sequence alignment of the COMM domain of the ten known human COMM proteins. (Images taken from Maine and Burstein, 2007)

1.3.2. Structure of COMMD proteins

Aside from COMMD1, the majority of the genes encoding other COMMD protein family members are only recognised as open reading frames and the functions of their respective proteins remain unknown. The limited data that is available about the structure and expression of these genes/proteins were derived from bioinformatics searches and tissue expression profiling at the mRNA level.

The COMMD gene family is found in all vertebrates and is also present in *Dictyostelium discoideum*. Several *COMMD* genes have also been identified in *Drosophila* and *Caenorhabditis species* and in various unicellular protozoa (Maine *et al*., 2007). The majority of the predicted proteins are ~ 200 residues long and are defined by the leucine-rich 70-85 amino-acid C-terminal COMM domain, which appears to function as an interface for protein-protein interactions (Burstein *et al*., 2005). In particular, COMMD1 appears to dimerise with itself and with all other COMMD proteins via the COMM domain. With the exception of

COMMD6, which consists mainly of a COMM domain, each COMMD protein also has an unique α-helix-rich N-terminus that is divergent across family members, but is highly conserved across various species and is thought to confer specific functions to the individual family members (de Bie *et al*., 2006).

Sequence alignments of over 90 homologs in multiple species revealed that the C-terminal of COMMD1 had conserved tryptophan and proline residues, as well as a series of conserved leucine residues, while secondary structure algorithms predicted the presence of a conserved β-sheet and an extreme C-terminal α-helix (Burstein *et al*., 2005; Sommerhalter *et al*., 2007).

Although attempts to obtain the full length COMMD1 for structural studies were complicated by protein instability and aggregation at first, the NMR solution structure of the N-terminal of COMMD1 has been resolved. This model confirmed the presence of five α-helices, with the last three adopting a distinct three-helix bundle. Comparison with other protein structures revealed a similar three-helix bundle segment in the vesicular transport protein, Sec17, in yeast and in tyrosine phosphatase SptP from *Salmonella* (Sommerhalter *et al*., 2007).

More recently, the 3D structure of full length COMMD1 was resolved (Figure 1.5). It was shown that the N-terminal of COMMD1 spans residues 1-121 and is connected via a flexible proteolytically-sensitive linker to the C-terminal COMM domain, which spans residues 125-190 (Burkhead *et al*., 2009). Consistent with previous results, the N-terminal domain in the full-length COMMD1 was also found to be entirely α-helical in this study, while the C-terminal domain was shown to contain three β-strands and two short α-helices. Interestingly, this *ab initio* modelling of COMMD1 predicted structural homology to various myosins and to Sec7 domain-containing proteins, for example, the yeast Sec 7 protein which is required for transport of proteins to the Golgi organelle (Klarlund *et al*., 1997; Burkhead *et al*., 2009). Interestingly, just as has been shown for COMMD1, these proteins also form oligomeric complexes and are involved in protein and vesicle transport (see more on COMMD1 role in protein transport in section 1.3.4.5.1). However, the copper binding properties of COMMD1 that had been shown by other studies to involve the His and Met residues within the COMM domain (Narindrasorak *et al*., 2007), was not confirmed in the Burkhead *et al* (2009) study.

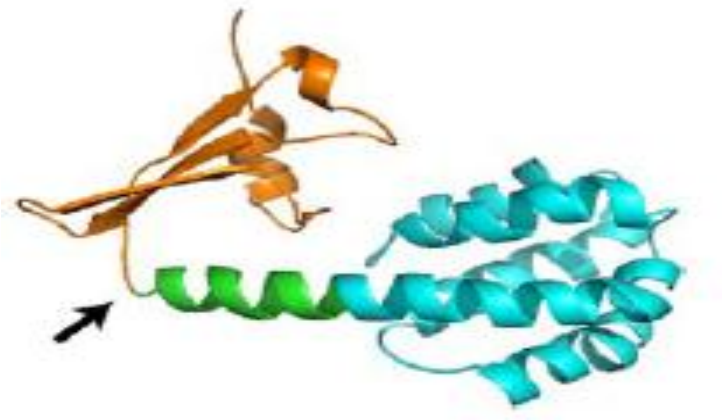


Figure 1.5. Structural model of full-length COMMD1
The NMR structure of the COMMD1 N-terminal residues 1-121, identified by Sommerhalter *et al* (2007), are shown in cyan. This model was used as an initial constraint and matches the *ab initio* model as it had the same secondary structure, number of α-helices and overall fold. Residues in green represent part of the N-and C-terminal and the flexible linker (119-128) between the two domains, while the rest of the C-terminal (residues 125-190) is shown in orange. The arrow indicates the proteolytic sensitive linker between the two-domain structures. (Image taken from Burkhead *et al*., 2009).

1.3.2.1. Structure of COMMD4

Bioinformatic searches revealed that COMMD4 has two isoforms due to alternative spicing. Isoform 1, which was identified as a specific ligand of the cMyBPC motif in our laboratory, contains 199 amino acid residues and has a molecular weight (MW) of 21.8kD. In contrast, the alternative spliced transcript, isoform 2 contains 60 amino acid residues which corresponds to amino acid residues 128-187 of isoform 1 and has a MW of 15.5kD (Figure 1.6 A) (http://www.uniprot.org/uniprot/Q9H0A8). The COMM domain in COMMD4 is 70 amino acids long (58 amino acids for isoform 2) and spans from residues 130-199 for isoform 1 of COMMD4 (Figure 1.6 A and B). In addition, COMMD4 also has an overlapping Ca^{2+}-binding domain (residues 19-198 for isoform 1 of COMMD4), originally found in the hypertension-related, calcium-regulated gene (HCaRG) (Figure 1.6 B). Interestingly, HCaRG was later identified as a COMMD family member and subsequently renamed COMMD5 (Section 1.3.4.6.2).

A

```
>COMMD4_HUMAN_TRANSCRIPT ID:NM_017828.3
ISOFORM 1:LENGTH: 199 AA, MW: 21764 Da
ISOFORM 2: POSITION: 128-187, LENGTH: 60 AA, MW: 15458 Da
COMMD DOMAIN: POSITION: 130-199, LENGTH: 70

1          ATGAGGTTCCGGTTCTGTGGTGATCTGGACTGTCCCGACTGGGTCCTGGCAGAAATCAGC
1           M  R  F  R  F  C  G  D  L  D  C  P  D  W  V  L  A  E  I  S

61         ACGCTGGCCAAGATGTCCTCTGTGAAGTTGCGGCTGCTCTGCAGCCAGGTACTAAAGGAG
21          T  L  A  K  M  S  S  V  K  L  R  L  L  C  S  Q  V  L  K  E

121        CTGCTGGGACAGGGGATTGATTATGAGAAGATCCTGAAGCTCACGGCTGACGCCAAGTTT
41          L  L  G  Q  G  I  D  Y  E  K  I  L  K  L  T  A  D  A  K  F

181        GAGTCAGGCGATGTGAAGGCCACAGTGGCAGTGCTGAGTTTCATCCTCTCCAGTGCGGCC
61          E  S  G  D  V  K  A  T  V  A  V  L  S  F  I  L  S  S  A  A

241        AAGCACAGTGTCGATGGCGAATCCTTGTCCAGTGAACTGCAGCAGCTGGGGCTGCCCAAA
81          K  H  S  V  D  G  E  S  L  S  S  E  L  Q  Q  L  G  L  P  K

301        GAGCACGCGGCCAGCCTGTGCCGCTGTTATGAGGAGAAGCAAAGCCCCTTGCAGAAGCAC
101         E  H  A  A  S  L  C  R  C  Y  E  E  K  Q  S  P  L  Q  K  H

361        TTGCGGGTCTGCAGCCTACGCATGAATAGGTTGGCAGGTGTGGGCTGGCGGGTGGACTAC
121         L  R  V  C  S  L  R  M  N  R  L  A  G  V  G  W  R  V  D  Y

421        ACCCTGAGCTCCAGCCTGCTGCAATCCGTGGAAGAGCCCATGGTGCACCTGCGGCTGGAG
141         T  L  S  S  S  L  L  Q  S  V  E  E  P  M  V  H  L  R  L  E

481        GTGGCAGCTGCCCCAGGGACCCCAGCCCAGCCTGTTGCCATGTCCCTCTCAGCAGACAAG
161         V  A  A  A  P  G  T  P  A  Q  P  V  A  M  S  L  S  A  D  K

541        TTCCAGGTCCTCCTGGCAGAACTGAAGCAGGCCCAGACCCTGATGAGCTCCCTGGGCTGA
181         F  Q  V  L  L  A  E  L  K  Q  A  Q  T  L  M  S  S  L  G  *
```

B

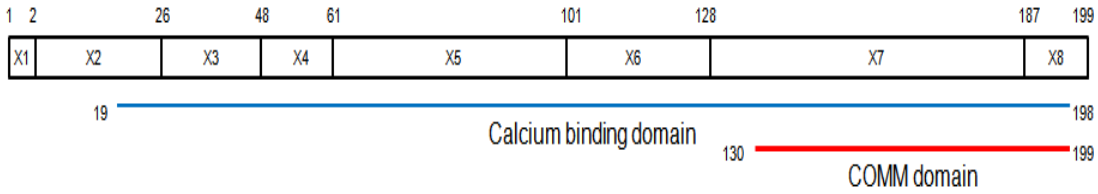


Figure 1.6. Nucleotide and amino acid sequence of human COMMD4 and schematic representation of the COMMD4 structure.
A. Nucleotide (top) and amino acid (bottom) sequence of COMMD4 (NM_017828.3). Amino acid residues 1-199 represent isoform1 of COMMD4 in black while the alternative spliced transcript (Isoform 2) correspond to amino acid residues 128-187 as indicated in red. The COMM domain spans from amino acid residue 130-199 and are highlighted in grey. **B**. Schematic representation of COMMD4 isoform1. Amino acid residues numbering at top indicate the boundaries of each exon. The calcium-binding domain spanning from amino-acid 19-198 are depicted by the blue line, while the COMMD domain which spans from amino acid 130-199 are depicted by the red line.

1.3.3. Expression level and distribution of COMMD proteins

The COMMD proteins are widely expressed in all tissues and the protein family is thought to have multiple functions, depending on cell type or the environment of the cell (Klomp *et al*., 2003). Although initial reports classified COMMD1 as a cytosolic protein, later studies have shown distribution in the nucleus, perinuclear organelles, lysosomes and endosomes (Klomp *et al*., 2003; Burkhead *et al*., 2009). The presence of COMMD1 in punctate vesicles throughout the cytosol was confirmed by Burkhead *et al* (2009). Interestingly, a high degree of co-localisation of COMMD1 was observed with the markers for early endosomes antigen 1 (EEA1) and the late endosomes/ multivesicular bodies (MVB) (ESCRT-III subunit CHMP2B), which localises COMMD1 in a post-Golgi vesicular compartment, and thus suggests a role for COMMD1 in the endosomal trafficking/recycling/degradation pathway. Moreover, the targeting of

COMMD1 to various compartments was independent of copper. Furthermore, COMMD1 was shown to form stable multimeric complexes at the membrane and in solution (Burkhead *et al*., 2009).

The identification of two highly conserved nuclear export signals (NESs), designated NES1 and NES2, in the COMMD domain of COMMD1 further underscores a role within the nucleus, which was later shown to involve nuclear factor κB (NFκB) (Section 1.3.4.1.2.) (Muller *et al*., 2009). Furthermore, mutations in NESs resulted in the accumulation of COMMD1 in the nucleus, with effects on the activity of NFκB. Following nuclear export, COMMD1 apparently undergoes poly-ubiquitination and degradation, as its cellular levels are tightly regulated within a physiological range through post-translational mechanisms. In fact, this regulation is so tight that every attempt to generate cell lines that stably overexpress COMMD1, whether through plasmid transfection or lentiviral infection, has resulted in very modest or no overexpression of COMMD1 (Muller *et al*., 2009; Maine *et al.*, 2009).

Recent findings have shown that the levels of COMMD1 and -4 are regulated in part through their interaction with X-linked inhibitor of apoptosis (XIAP), the expression of which is, in turn, regulated by the levels of copper in the cell (Section 1.3.4.2.2). Interestingly, mutations in the NESs resulted in decreased interaction of COMMD1 with XIAP, which resulted in a decrease in the ubiqitination of COMMD1 (Maine *et al.*, 2009). In addition, the half-life of COMMD1, when expressed alone, is 1hour, but upon binding to the tumor suppressor protein, alternative reading frame (ARF), the half-life of COMMD1 is extended to four hours (Section 1.3.4.5.2) (Huang *et al*., 2008). Furthermore, HSCARG, a NADPH redox sensor protein located in the cytoplasm, has also been shown to interact with COMMD1 and to promote COMMD1 poly-ubiquitination and proteasomal degradation, thereby reducing COMMD1 protein levels (Lian and Zheng, 2009).

1.3.4. Biological functions of COMMD proteins

Over the past few years, a number of disparate functions have been assigned to the COMMD family, based on the interactions of COMMD1 with various molecular partners. It has been shown to function in NFκB regulation, copper homeostasis, sodium channel regulation and embryogenesis and more recently, ubiquitin-modification. Each function will be discussed briefly below, in order to acquire a basic understanding of this protein family.

1.3.4.1. Involvement in the regulation of NFκB

1.3.4.1.1. NFκB: An overview

The transcription factor, NFκB, has been implicated in numerous biological processes via its transcriptional regulation of various gene products, which have been associated with apoptosis, cell survival, cell division, cell growth, innate immunity, cellular differentiation, cellular responses to stress and hypoxia (Jones *et al*., 2003). In the heart, it has been implicated in (I/R) injury, congestive heart failure (CHF), the development of cardiac hypertrophy and DCM (Brown *et al*., 2005).

The highly conserved mammalian NFκB family of proteins consist of p65 (RelA), Rel B, c-Rel, p50/p105 (NFκB1) and p52/p100 (NFκB2), each sharing the conserved Rel-homology domain (RHD) (Hayden and Ghosh, 2004). Via the RHD, NFκB homo- or heterodimerises and interacts with inhibitory κB (IκB), maintaining it in an inactive form in the cytoplasm, as the latter interaction masks the nuclear localisation signal present in the RHD of NFκB. Subsequent phosphorylation of serine or tyrosine residues within IκB, in response to various stimuli, are key steps in NFκB activation (Huxford *et al*., 1998). Signalling pathways that lead to NFκB activation include the likes of PKC, phosphoinositol 3-kinase (PI3K/Akt/mTOR), tumor necrosis factor-α (TNF), calcium/nuclear factor of activated Tcells (NFAT), calcium-calmodulin kinase IV, Ras and the Janus kinase pathways (Zhang *et al*., 2003; Li *et al*., 2006; Ha *et al*., 2005).

Serine phosphorylation of IκB leads to its ubiquitination and degradation, while tyrosine phosphorylation causes IκB to dissociate from NFκB without degradation (Brown *et al*., 2005). However, both mechanisms free NFκB to translocate to the nucleus, where it activates the transcription of its target genes (Zhang *et al*., 2003; Brown *et al*., 2005). The termination of NFκB activity requires the ubiquitination of chromatin-bound NFκB, which is mediated by the re-synthesis of IκB; this eventually leads to the nuclear export of NFκB (Brown *et al*., 2005).

1.3.4.1.2. COMMD proteins: Inhibitors of NFκB

The role of COMMD1 in regulating NFκB activity had first been demonstrated by Ganesh *et al* who studied the effects of potential NFκB regulators on viral replication. It was demonstrated that COMMD1 had an inhibitory effect on HIV-1 growth in $CD4^+$ T cells, which was mediated in part through its inhibitory effect on NFκB activity. Subsequent siRNA knockdown of COMMD1 also showed increased NFκB activity, decreased IκBα (an isoform of IκB) concentration, as well as increased HIV-1 replication, which substantiated their initial results and revealed a role for COMMD1 in protection against HIV infection (Ganesh *et al*., 2003).

In addition, the site of action for COMMD1 in the NFκB signalling pathway was narrowed down to the NFκB-IκBα complex (Ganesh *et al*., 2003). In agreement with this, Burstein *et al* mapped the COMMD1-NFκB interactions to the RHD of all five NFκB subunits (Burstein *et al*., 2005).

In turn, the effect of COMMD1 on the NFκB-IκBα complex was further investigated by treating human embryonic kidney (HEK293) cells, in which COMMD1 was either present or knocked down via siRNA interference, with TNFα, in order to induce phosphorylation of IκBα and subsequent NFκB activation. However, expression of COMMD1 inhibited the ubiquitination or degradation of IκBα. This, in turn, was suggested to attenuate NFκB activation, since degradation of phospho-IκBα is necessary for NFκB translocation to the nucleus and subsequent NFκB activation. Given the fact that the E3 ubiquitin-ligase, the Skp1/Cullin1/F-box protein (SCF) ubiquitin-ligase complex, was required for the degradation of phospho-IκBα (Tan *et al*., 1999), a possible interaction between COMMD1 and the SCF ubiquitin-ligase complex was

also investigated as a likely mechanism for this inhibitory effect of COMMD1 on IκBα. Data from immunoprecipitation experiments revealed an interaction between COMMD1 and Cul1, but not Skp1. However, the interaction between COMMD1 and the F-box protein was not investigated (Ganesh *et al*., 2003; Greene, 2004). Taken together, this data showed that COMMD1 binds to both the NFκB-IκBα complex and the SCF ubiquitin-ligase complex and inhibits NFκB activation by preventing the degradation of IκBα.

Subsequently, all known COMMD proteinswere shown to interact with NFκB and to strongly inhibit TNFα-mediated NFκB activition in HEK293 cells, similar to COMMD1. Of these proteins, COMMD1, -2, -4, -5, -6, -7, -9, and 10 showed stronger inhibitory effects on NFκB than did COMMD3 and -8 (Burstein *et al*., 2005; de Bie *et al*., 2006). Furthermore, mutations of two amino-acid residues within the COMM domain, Trp24 and Pro41 that are conserved in all human COMMD proteins, completely abolished the inhibitory effect of COMMD6 on TNFα-induced NFκB activation (de Bie *et al*., 2006).

In addition, siRNA suppression of COMMD1, -4, and -6 expression lead to increased mRNA levels of *c-IAP2*, a gene that is transcribed in response to NFκB upregulation upon TNFα stimulation (Burstein *et al.,* 2005). This further substantiates the effects of COMMD proteins on the regulation of NFκB transcriptional activity.

1.3.4.1.3. Ubiquitin modifications of COMMD1 in the NFκB pathway

Whereas earlier studies suggested that COMMD1 exerts its regulatory effect on NFκB by affecting the nuclear translocation of NFκB via its ability to regulate IκBα turnover (Ganesh *et al*., 2003), an alternative mechanism by which COMMD1 inhibits NFκB activity has also been established. According to this proposal, COMMD1 does not inhibit the nuclear translocation of NFκB, despite its effect on IκBα; instead, COMMD1 excerts its effect by controlling the duration of NFκB on chromatin (Burstein *et al*., 2005; Maine *et al*., 2007). Since most of the COMMD proteins favour association with the RelA NFκB subunit, the effect of COMMD1 levels on the recruitment of RelA to the NFκB-responsive *BIRC3*/*c-IAP2* promoter (a gene that is transcribed in response to NFκB) was studied. Endogenous depletion of COMMD1 prolonged the duration of RelA on the *BIRC3*/*c-IAP2* promoter, whereas expression of COMMD1 had the opposite effect (Burstein *et al*., 2005). Thus, it was proposed that, upon NFκB activation, COMMD1 is recruited to chromatin of the NFκB-responsive gene and seems to destabilise the association of RelA with the chromatin, thereby affecting the half-life of this NFκB-chromatin complex.

Observations that COMMD1 influences RelA-chromatin interactions together with evidence that ubiquitination and protein degradation of nuclear RelA contribute to the termination of NFκB-chromatin interactions, led Maine and collegues to investigate a possible role for COMMD1 in RelA ubiquitination (Maine *et al*., 2007). It appeared that COMMD1 had an effect on the protein stability of RelA, since, in COMMD1-deficient cells, RelA protein levels were elevated and had a prolonged half-life, but RelA mRNA

expression remained unchanged. Thus, it was suggested that COMMD1 destabilises RelA by promoting its ubiquitination and targeting it for proteasomal degradation (Maine *et al*., 2007). This effect of COMMD1 is mediated by its interaction with the multimeric E3 ubiquitin-ligase, consisting of Elongins B/C, Cullin 2 and the SOCS box containing protein 1 (ECS^{SOCS1}) (Figure 1.7), which is similar in structure to the Cul1–containing SCF ubiquitin-ligase complex that had previously been reported to interact with COMMD1 (Section 1.3.4.1.2) (Ganesh *et al*., 2003; Ryo *et al*., 2003; Maine *et al*., 2007). COMMD1 was shown to bind independently to the Cul2 and SOCS subunits of the complex, but not to the adaptor Elongin B/C. Furthermore, COMMD1 associated with both SOCS1 and RelA, and enhanced the interaction of SOCS1 with RelA (Maine *et al*., 2007). These findings demonstrate that COMMD1 serves as an adaptor protein, facilitating the binding of SOCS1 to RelA, and thus resembling the role of the Elongin B/C complex in ECS^{SOCS1}. Through this complex, COMMD1 enhances the binding of NFκB to the E3 ubiquitin-ligase by mediating the protein stability between ECS^{SOCS1} and NFκB subunits, leading to subsequent ubiquitination and degradation of NFκB, thereby influencing NFκB-mediated transcription.

Most recently, it has been shown that HSCARG, which interacts with COMMD1 (Section.1.3.3), translocates to the nucleus in response to decreased NADPH levels or increased nitric oxide (NO), where it inhibits NFκB activity by promoting the ubiquitination and degradation of RelA or COMMD1 (Lian and Zheng, 2009).

Thus, to summarise, it is evident that COMMD proteins inhibit NFκB activity by regulating turnover of IκBα in the cytoplasm, while it also regulates the stability of NFκB in the nucleus; these two functions are made possible by the nuclear import and export capabilities of COMMD1.

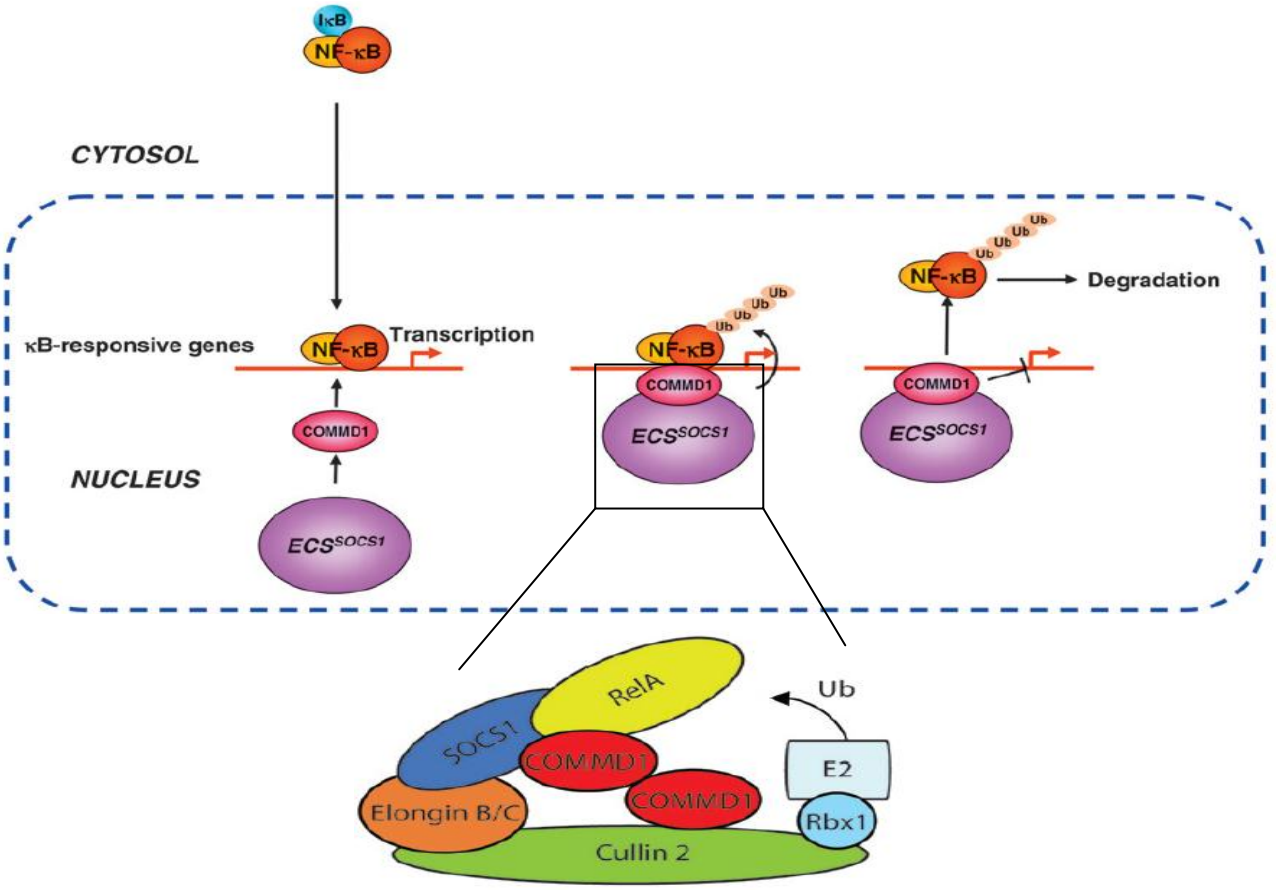


Figure 1.7. Schematic model of COMMD1 in the attenuation of NFκB-mediated transcription.
Various stimuli that activate the NFκB pathway result in the phosphorylation and degradation of IκBα, which enables the nuclear translocation of NFκB dimers to NFκB-dependent promoter elements to induce gene expression. Upon NFκB activation, COMMD1 is recruited to chromatin and associate with the ECS^{SOCS1} ubiquitin-ligase. COMMD1 serves as an adaptor and enhances the interaction between NFκB subunits and the ubiquitin-ligase and thereby promotes the poly-ubiquitination of NFκB. The latter then undergoes proteasomal degradation, which in turn attenuates NFκB-mediated gene transcription. (Images combined from Maine and Burstein, 2007 and van de Sluis *et al*., 2007)

1.3.4.2. Copper homeostasis

Lately, copper-binding properties have also been assigned to COMMD1. Narindrasorasak *et al* showed that COMMD1 binds copper in the Cu(II) form and mapped the Cu(II) binding site to exon 2 within the COMM domain (Narindrasorasak *et al*., 2007). It is therefore relevant to briefly discuss copper homeostasis.

Copper is an important trace metal that is essential as a cofactor in copper-dependent enzymes; therefore, tight regulation of the metal is crucial for cellular copper homeostasis. Human diseases associated with copper metabolism include Menkes syndrome and Wilson disease. Menkes is a lethal X-linked disorder caused by mutations in a gene encoding a copper-transporting P-type ATPase, *ATP7A,* and is characterised by a copper-deficiency due to inadequate copper-uptake from the placenta and diet (Kaler, 1998). On the other hand, Wilson disease is autosomal recessive and is caused by mutations in a gene encoding the copper transporter, *ATP7B,* which is homologous to *ATP7A.* In contrast to Menkes syndrome, Wilson disease is characterised by a defect in biliary excretion of copper which results in copper accumulation in the liver and subsequently in the brain and other organs (Scheinberg and Sternlieb, 1996). Menkes syndrome manifests with neurological degeneration and growth defects due to dysfunctional copper-dependent enzymes, while

Wilson disease results in liver disease, neurological defects and a mild form of cardiac involvement characterised by left ventricular (LV) thickening, increased concentric LV remodelling and high frequencies of benign tachycardias (Hlubocka *et al*., 2002).

Copper toxicosis in Bedlington terriers resembles the human form of Wilson disease and was therefore originally though to be a model for human autosomal recessive copper diseases (van de Sluis *et al*., 1999). Studies aimed at the identification of susceptibility genes in these terriers led to identification of *COMMD1* as the defective gene that underlies copper toxicosis, and thus implicated COMMD1 as a novel regulator in copper homeostasis (van de Sluis *et al*., 2002). That COMMD1 plays a role in regulating intracellular copper levels was further demonstrated by Burstein *et al* and Spee *et al*, who observed increased accumulation ofcopper in HEK293 and dog hepatic cells (bile-duct epithelial cells, BDE) after siRNA-mediated depletion of COMMD1 (Burstein *et al*., 2004; Spee *et al.,* 2007).

1.3.4.2.1. Copper dependent trafficking or ubiqitin modification of copper transporters

How COMMD1 exerts its function on copper homeostasis was partially explained with the demonstration of individual interactions between COMMD1, 2, 8 and 10 with ATP7B, the protein defective in Wilson disease (Tao *et al*., 2003; de Bie *et al*., 2007). Similarly, individual interactions were also observed between COMMD1, 2, 3, 4, 5, 8 as well as 10 and ATP7A, the protein involved in Menkes syndrome (de Bie *et al*., 2007). These interactions were shown to be mediated via the copper-binding N-termini of ATP7A and ATP7B, and implicated a role for COMMD proteins in the regulation of biliary copper excretion and uptake (Tao *et al*., 2003). Under basal conditions, these copper transporters are located in the trans-Golgi network (TGN), but upon elevation of copper levels, these transporters are translocated to vesicular compartments (Petris *et al*., 1996; Hung *et al*., 1997). Interestingly, COMMD1 was shown not to be involved in the copper-dependent trafficking of either ATP7A or ATP7B (de Bie *et al*., 2007; Burkhead *et al*., 2009) and, since COMMD1 is localised in a post-Golgi vesicular compartment, the interactions of these copper transporters with COMMD1 at the TGN is excluded. However, expresssion of COMMD1 increased the protein degradation of ATP7B; this, together with the fact that an ATP7B mutant had a higher affinity for COMMD1 than WT-ATP7B, suggested a role for COMMD1 in promoting protein degradation of ATP7B (de Bie *et al*., 2007).

1.3.4.2.2. XIAP-COMMD1-mediated copper control

Further studies also demonstrated a role for COMMD1 in copper homeostasis via its interactions with XIAP (Burstein *et al*., 2004).

XIAP, which inhibits caspase-mediated cell death, is comprised of three N-terminal baculoviral-inhibitor-of-apoptosis repeat (BIR) domains and a C-terminal RING domain; the latter domain confers its E3 Ub-ligase acitivity. XIAP binds to COMMD1 via its third BIR domain (BIR3), in a manner that does not affect its anti-apoptotic properties (Mufti *et al*., 2007; Maine *et al*., 2009). XIAP was shown to negatively regulate

COMMD1; overexpression of XIAP results in a decrease in COMMD1 levels, while the suppression of endogenous XIAP causes COMMD1 levels to increase. This inverse effect is mediated by XIAP's ability to promote degradation of COMMD1. The suppression of COMMD1, in cultured HEK293 cells, resulted in increased copper levels, while fibroblast and liver tissue derived from Xiap-deficient mice showed decreased copper levels (Burstein *et al*., 2004). Thus, XIAP and COMMD1 have inverse effects on intracellular copper, and may explain why XIAP overexpression promotes copper retention: through its interaction with COMMD1 (Mufti *et al.*, 2006; Mufti *et al*., 2007).

However, XIAP itself also binds to copper, which results in a conformational change and blocks XIAP in effecting its anti-apoptotic functions. In particular, copper destabilises XIAP and promotes its degradation, thereby lowering its steady-state levels. Thus, from these observations, it was suggested that XIAP acts as a copper sensor: upon elevation of copper levels, XIAP levels decrease, which in turn increases COMMD1 levels in order to aid in the efflux of copper from the cell (Mufti *et al*., 2007).

More recently, the region for the interaction between XIAP and COMMD1 was mapped to leucine repeats within the COMMD domain. COMMD proteins 2, 4 and 5 showed similar binding properties to those of COMMD1 (Maine *et al*, 2009). Thus, these studies implicate a role for XIAP in the regulation of intracellular copper levels via its ability to tighly control the expression levels of several COMMD proteins.

1.3.4.3. Regulation of sodium channels (ENAC) activity

The amiloride-sensitive epithelial sodium channel (ENaC), consisting of α-, β-, γ-and δ-subunits, is important in controlling salt homeostasis, and thus BP. The latter is tightly regulated by hormones such as aldosterone, insulin and arginine-vasopressin (AVP), which, in turn, is regulated via glucocorticoid-regulated kinase 1 (SGK1), a family member of the AGC kinases. This kinase family also includes Akt1/PKBα .While SGK1 mediates aldosterone- and AVP-induced sodium transport, both SGK1 and Akt1 mediate insulin-induced Na^+ transport via the PI3K pathway (Lang *et al*., 2006). Under normal conditions, the E3 Ub-ligase Nedd4-2 binds to the PY motif in the C-terminal domains of the ENaCs and subsequently promotes its degradation; this, in turn, regulates ENaC expression on the cell surface (Debonneville *et al*., 2001; Snyder, 2005). Both SGK1 and Akt1 bind and phopshorylate Nedd4-2 and so reduce its interactions with ENaC, which results in more cell surface-expression of these channels and thus increased sodium absorption (Lee *et al*., 2007).

1.3.4.3.1. COMMD1 interactions with ENaC subunits

Interestingly, COMMD1 has been shown to interact with the C-terminus of the δ ENaC subunit (Basio *et al*., 2004). Similarly, COMMD1 also binds to the β- and γENaC subunits, but not to the αENaC subunit.

To determine the functional effect of these interactions, COMMD1 was coexpressed with β-, γ- and δ-ENaC subunits in COS7 cells and the results indicated that COMMD1 is able to inhibit the δβγENaC current in

Xenopus oocytes in a dose-dependent manner. Although, the mechanism of how COMMD1 inhibits these sodium channels is still unclear; findings that COMMD1 also interacts with ATP7B, a copper transporter (Section 1.3.4.2.1), suggested at first that COMMD1 is involved in the intracellular trafficking of these transporters, and thus could possibly be involved in the trafficking of sodium channels as well (Tao *et al*., 2003). Together with the ubiquitious nature of COMMD1, a role for it in protein trafficking seems plausible, since the activity of these ENaCs is regulated by relocalisation from the TGN to the cell membrane (Klomp *et al*., 2003). Interestingly, a link between sodium and copper transport has been shown in fish gills, in which the uptake of copper was inhibited by elevated sodium channel expression (Pyle *et al*., 2003).

1.3.4.3.2. Ubiquitin modification of COMMD1 in the sodium transport pathway

On the other hand, similar to the eventual findings regarding the copper transporters, it is also possible that COMMD1 might be involved in the ubiquination and degradation of the sodium channels rather than in their trafficking, since it is well known that the UPS regulates the stability of the ENaCs via Nedd 4-2 (Rotin *et al*., 2001; Malik *et al*., 2006). More recently, COMMD1 was shown to be co-expressed with the ENaC subunits in the mammalian renal medulla, and, consistent with previous results, to inhibit sodium transport. Moreover, siRNA knockdown of COMMD1 caused an increase in the sodium transport current (Ke *et al*., 2010). It was shown that the inhibitory effect of COMMD1 on sodium transport was due to an increase in the ubiquitination of ENaCs and subsequent decrease in the cell surface-expression of these channels. COMMD1 was also shown to interact with SGK1 and Akt 1 (kinases which induce sodium transport) via the COMM domain, but not Nedd4-2, and COMMD1 reduced both insulin and SGK1-induced sodium transport (Ke *et al*., 2010). Subsequent knockdown of COMMD1 enhanced the stimulatory effect of both SGK1 and Akt1 on sodium transport (Ke *et al*., 2010).

Both COMMD1 and Nedd4-2 are known to interact with ENaC; further findings suggested that SGK1, Akt1, a COMMD1 dimer and Nedd4-2 form a complex within the cell, which subsequently facilitates the ubiquitination and endocytosis of ENaCs in order to control the cell surface-expression of these channels (Ke *et al*., 2010). Thus, both COMMD1 and Nedd4-2 would influence the same pathway, which was verified when subsequent evidence showed that COMMD1 had no effect on the sodium transport pathway when the Nedd4-2 pathway was blocked. Thus, a potential mechanism whereby COMMD1 could have its inhibitory effect could be that COMMD1 promotes degradation of SGK1/Akt1, which in turn prevents the association of these kinases with Nedd4-2, and thus inhibits the phosphorylation of Nedd4-2 (Ke *et al*., 2010). These data indicate that COMMD1 negatively regulates the activity of sodium channels through Nedd4-2, even though it does not physically interact with the latter.

Taken together, it is evident from the findings above that COMMD1 plays a key role in regulating the activity of sodium channels. The effect of this pathway on sodium homeostasis is, however, apparently modest, since SGK and Nedd4-2 knockout mice appear normal. However, it is significant in the appropriate

environment, as these mice showed changes in sodium homeostasis when their dietary salt conditions were altered (Shi *et al*., 2008).

1.3.4.4. Embryogenesis

1.3.4.4.1. COMMD1 knockout mouse model

In an attemp to investigate the *in vivo* function of COMMD1, van de Sluis and collegues generated a homozygous *Commd1* knockout (*Commd1*$^{-/-}$) mouse which mimicked the in-frame exon 2 deletion that was identified in homozygous Bedlington terriers affected with copper toxicosis (Section 1.3.1) (van de Sluis *et al*., 2007). This in-frame deletion leads to a *Commd1* null allele, with no detectable Commd1 protein expression, as opposed to the WT and heterozygous embryos. Loss of *Commd1* ultimately led to a recessive embryonic lethal phenotype, since no homozygous null mice were detected among 259 offspring. Further analysis of embryos at different stages of gestation also revealed that *Commd1*$^{-/-}$ embryos died between 9.5 and 10.5 days postcoitum (dpc) and that development was retarded compared to that of WT and heterozygous mice (van de Sluis *et al*., 2007). That this homozygous deletion could lead to such a severe phenotype in rodents, but not in dogs, further demonstrates that the functional significance of COMMD1 is species-dependent.

This group also investigated whether embryonic copper-overload could explain the lethal phenotype of these *Commd1* null mice, since COMMD1 has been shown to be essential in copper homeostasis, while findings elsewhere also demonstrated that copper metabolism is important for normal placenta and fetus development (van de Sluis *et al*., 2007; Donadio *et al*., 2007). Interestingly, COMMD1 was shown to be expressed in human placenta throughout pregnancy (Donadio *et al*., 2007). However, investigation of the embryonic *Commd1* null mice showed no elevated copper levels. Previous *in vitro* studies had demonstrated that severe copper -deficiency leads to early embryonic death (Mercer, 1998; Kuo *et al*., 2001; Lee *et al*., 2001), which led to the suggestion that ineffecient copper transport across the placenta, due to the loss of *Commd1*, and thus a copper-deficiency, might explain the lethal phenotype in the *Commd1* null mice. However, even feeding pregnant heterozygous mice a high copper diet in order to circumvent this possible copper-deficiency, could not rescue the lethal phenotype of *Commd1*$^{-/-}$ embryos. This result therefore suggests that copper-deficiency is not the main cause of the phenotype. Further characterization of these *Commd1*-deficient mice ruled out defects in the NFκB signaling, sodium transport and the XIAP pathway as a cause for the embryonic lethality and rather established a vital role for Commd1 in the regulation of Hypoxia-inducible factor 1 (HIF-1) activity during early mouse embryogenesis as the probable cause.

1.3.4.4.2. Regulation of HIF-1α protein levels and HIF-1 transcriptional activity

Hypoxia inducible factor 1 is a heterodimeric transcription factor (TF), comprised of HIF-1α and HIF-1β, and is important in embryogenesis, cardiovascular development, glycolysis, apoptosis and cell proliferation (Wang *et al*., 1995). The regulation of HIF-1activity is analogous to the regulation of NFκB activity. In

Commd1$^{-/-}$ embryos, mRNA expression revealed increased HIF-1 activity, which also correlated with increased expression of HIF-1α (van de Sluis B *et al*., 2007).

The protein stability of the α-subunit is tightly regulated by oxygen concentrations via the UPS. Thus, under normal oxygen conditions (pO_2-21%), HIF-1α is constituvely hydroxylated by prolyl hydroxylases (PHD), which recruit the Von Hippel-Lindau protein (pVHL), Elongin B/C, Cullin2 E3 Ub-ligase complex (similar to ECS^{SOCS1} complex, see section 1.3.4.1.3) for ubiquitination and rapid degradation of HIF-1α. In contrast, under hypoxic conditions, (pO_2<5%), the activity of the PHD is decreased, which, in turn, increases the half-life of HIF-1α (Hirota and Semenza, 2005). Upon stabilisation, HIF-1α is translocated to the nucleus and forms a dimer with HIF-1β to form active HIF-1, which binds to hypoxia responsive element (HRE) in the promoters of HIF-1-targeted genes (Figure 1.8) and initiates transcription of factors such as vascular endothelial growth factors (VEGF) and transforming growth factor (TGF) (Wang *et al*., 1995; Willis *et al*., 2010).

1.3.4.4.3. Interaction of COMMD1 with HIF-1α

In vitro and *in vivo* data showed that COMMD1 modulates the protein levels of HIF-1α and thus HIF-1 activity, possibly via a COMMD1-HIF-1α protein interaction (van de Sluis *et al*., 2007). Moreover, siRNA-mediated knockdown of COMMD1 in HEK293 cells caused an increase in the protein stability of HIF-1α and HIF-1β reporter activity. In addition, an increase in the mRNA expression of a HIF-1 target gene, namely *BNIP3* is also observed under these conditions (van de Sluis *et al*., 2007). Conversely, expression of COMMD1 resulted in the downregulation of HIF-1 activity in cells under normal and hypoxic conditions (van de Sluis *et al*., 2007). More recently, van de Sluis *et al* was abled to show that COMMD1 binds to the N-terminal region of HIF-1α and competes with HIF-1β for binding to the former (van de Sluis *et al*., 2009).

Interestingly, COMMD1 also interacts with pVHL and both these proteins can shuttle between the nucleus and the cytoplasm. Based on these observations, a role was suggested for COMMD1 in pVHL-mediated HIF-1α degradation, which would lead to abrogation of HIF-1-mediated ranscription in a manner similar to the role of COMMD1 in the attenuation of NFκB transcription. Furthermore, the nuclear export capability of COMMD1 also provides a mechanism to terminate the inhibitory role of COMMD1 in HIF-1 pathway when necessary.

While HIF1 activity is mainly controlled via its ubiquitination and degradration by pVHL in an oxygen-dependent manner, other mechanisms, which involve the heat shock protein 90 (HSP90β) and deacetylase inhibitors, have been identified that regulate HIF-1α in a manner independent of oxygen and pVHL (van de Sluis *et al*., 2009). For example, binding of HSP90β to HIF-1α was shown to be vital for proper maturation and stabilisation of HIF-1α under both normal and hypoxic conditions (Liu and Semenza, 2007). Furthermore, it was shown that COMMD1 competes with HSP90β for binding to HIF-1α and, in complex

with HSP70, facilitates the degradation of HIF-1α. This tripartite complex thus regulates the stability of HIF-1α in a manner independent of pVHL, oxygen and Ub (Figure 1.8) (van de Sluis *et al*., 2009).

Interaction between COMMD1 and the 20S proteasome has also been observed, which suggests a novel role for COMMD1 in escorting target proteins to the proteasome (van de Sluis *et al*., 2009). Likewise, others have also reported ubiquitin-independent degradation of HIF-1α directly via the 20S proteasome (Kong *et al*., 2007). Taken together, these findings demonstrate the importance of COMMD1 in the regulation of HIF-1α, whether in an ubiquitin-dependent or -independent manner.

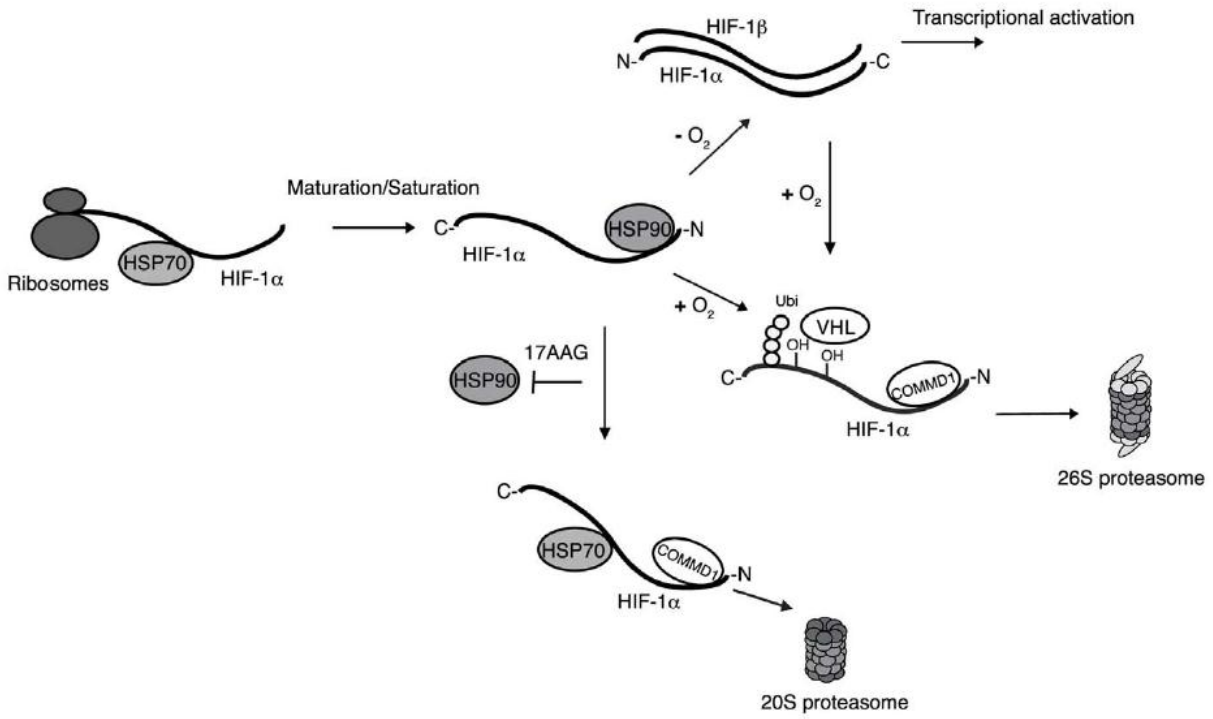


Figure 1.8. Schematic model of the mechanism whereby COMMD1 mediates HIF-1α degradation.
Under normal oxygen conditions, HIF-1α binds to HSP70 and HSP90 for maturation, but undergoes ubiquitination via the pVHL E3 ubiquitin-complex and degradation via the 26S proteasome. In contrast, under hypoxic conditions, HIF-1α is stabilised and heterodimerises with HIF-1β to form active HIF-1, which binds to hypoxia response elements (HRE) in HIF-1-targeted genes. Upon reoxygenation, HIF-1α is ubiquitinated via pVHL, and COMMD1 facilitates its proteasomal degradation. Inhibition of HSP90 via17-AAG/Geldanamycin causes incorrect folding of HIF-1α and the formation of a tripartite complex with HSP70 and COMMD1, which promotes degradation via the 20S proteasome. (Image taken from van de Sluis *et al*., 2009)

1.3.4.5. Other reported interactions and proposed functions

1.3.4.5.1. Interaction of COMMD1 with Phosphatidylinositol 4.5-Biphosphate

More recently, a role in the trafficking of membrane proteins has been assigned to COMMD1, since an interaction between COMMD1 and a signalling and regulatory lipid, phosphatidylinositol 4,5-biphosphate (PtdIns (4,5)P_2), which is mediated via the COMMD domain, has been demonstrated (Burkhead *et al*., 2009). These PtdIns or their phosphorylated deratives, phosphoinosititides (PIs), which are regulated by PI3-

kinases, are important in membrane trafficking, cytoskeletal dynamics, cell signalling, and the regulation of nuclear processes (Lindmo and Stenmark, 2006).

Based on the observation that COMMD1 can form stable multimeric complexes at the membrane and in solution (Burkhead *et al*., 2009) (Section 1.3.3) it was suggested that COMMD1 could act as a scaffold protein, and might resemble the hexameric hepatocyte growth factor-regulated tyrosine kinase substrate (Hrs) protein (described in section 1.5.1.2), another phosphatidylinositol-binding molecule, which functions in vesicular sorting by offering multiple sites for protein-protein interactions (Pullan *et al*., 2006).

1.3.4.5.2. Interaction of COMMD1 with ARF

Interaction between COMMD1 and the tumour suppressor protein, ARF, has also been described. ARF inhibits murine double minute 2 (MDM2), an E3 ubiquitin-ligase for p53, and protects p53 from degradation, but the interaction of ARF with COMMD1 is independent of its role in the p53 pathway (Huang *et al*., 2008). Furthermore, it was shown that ARF stabilises COMMD1 by promoting Lys63-linked poly-ubiquitination of COMMD1, which signals either for protein kinase activation, DNA repair and/or vesicle trafficking, but not for degradation (Huang *et al*., 2008). The authors further speculated that the Lys63-linked poly-ubiquitination may regulate and stabilise the structure of COMMD1, protect it from being recognised by an unknown protease, or regulate protein-protein interactions by affecting its binding affinity for other proteins (Huang *et al*., 2008).

1.3.4.6. Functions of other COMMD proteins

1.3.4.6.1. Proposed functions of COMMD5

As mentioned earlier, COMMD5 was previously known as HCaRG, first identified in spontaneously hypertensive rats (SHR). COMMD5 is associated with hypertension and also plays a role in $Ca2^+$ homeostasis, since it is negatively regulated by extracellular calcium concentrations, i.e. lower extracelllar calcium concentrations increase the mRNA expression of COMMD5 (Solban *et al*., 2002). COMMD5 is predominantly a nuclear protein, consisting of four overlapping putative leucine-zipper motifs, an EF-hand $Ca2^+$ binding-like motif and a nuclear receptor motif.

It is chiefly expressed in the parathyroid gland, kidney and the heart (Solban *et al*., 2002). Furthermore, COMMD5 mRNA levels are negatively correlated to the proliferative status of the cell, suggesting a negative effect of COMMD5 on growth. Consistent with this, overexpression of COMMD5 in HEK293 cells resulted in growth suppression and cell cycle arrest at the G_2M phase. Interestingly, this COMMD5-induced growth suppression is also associated with cellular hypertrophy and increased levels of atrial natriuretic peptide (ANP) (Devlin *et al*., 2003).

COMMD5 mRNA levels are also decreased in tumor and cancerous cell lines and rapidly declines after renal I/R with a concurrent increase in *c-myc* mRNA. These observations, together with the location of the

COMMD5 gene on chromosome 8q21-24, a region also associated with cancers, suggest that disruption of the *COMMD5* gene due to chromosome abnormalities could lead to uncontrolled proliferation (Solban *et al*., 2000; Solban *et al.,* 2002). Thus, a role for COMMD5 in the regulation of cell proliferation has been implicated, although the precise downstream targets are not known (Devlin *et al*., 2003).

1.3.4.6.2. Interaction of COMMD6 from *Amphioxus Branchiostoma belcheri* (BbCOMMD6) with creatine kinase

More recently, Li *et al* identified COMMD6 from *Amphioxus Branchiostoma belcheri* (a fish-like marine organism) and designated it BbCOMMD6 (Li *et al*., 2009). Moreover, this group showed that BbCOMMD6 is expressed in the cytoplasm and interacts with creatine kinase (CK), with an inhibitory effect on CK activities. Since BbCOMMD6 does not result in the degradation of muscle CK; it was proposed that the interaction causes a conformational change (Li *et al*., 2009). CK functions to catalyse a reversible phosphotransferase reaction between creatine and adenosine triphosphate (ATP) that generates creatine phosphate and adenosine diphosphate (ADP) (Au, 2004). Therefore, these results implicate a role for BbCOMMD6 in energy transduction. Whether this interaction of COMMD6 is true for other species needs to be clarified.

1.3.4.7. Summary of COMMD proteins functions

To summarise, although the exact function of COMMD1 and its family members in the pathways discussed are still not completely understood, it is evident from these findings that the basic mechanism that underlies the majority of these interactions/functions implicate a role for these COMMD proteins in the UPS and protein trafficking pathways. These COMMD proteins seem to facilitate the assembly of protein complexes with the E3 ubiquitin ligase and help to stabilise target substrates for subsequent ubiquitin conjugation.

Our protein of interest, COMMD4, has been shown to interact with cMyBPC when it is either in a dephosphorylated or monophosphorylated state. Since these forms of cMyBPC have been shown to be more prone to proteolysis, the question arises as to whether COMMD4 may play a role in cMyBPC degradation. Thus, we further need to consider sarcomeric turnover; however, these pathways are very complex, and for the purpose of this thesis, a very condensed and simplified overview of these pathways will be discussed in the following sections.

1.4. SARCOMERE QUALITY CONTROL

Muscle cells, whether skeletal or cardiac, are continuously being remodelled in order to adapt to strain such as contraction-induced heat stress and/or oxidative stress due to high energy consumption at the sarcomere and due to mitochondrial activity (Arndt *et al*., 2007). This is even more true in the context of a cardiac muscle disease such as HCM. As mentioned earlier (Section 1.1.4), in HCM, the underlying defect is altered contractility caused by a defective sarcomere, which in turn leads to increased hypertrophic signalling (Redwood *et al.*, 1999; Ahmad *et al*., 2005). However, increased hypertrophic signalling goes hand in hand

with increased protein synthesis as well as increased protein degradation (Meiners *et al*., 2008); processes which are often prone to errors (Coleman *et al*., 1989). Since cardiac contractility is partly regulated through the maintenance of the cardiac sarcomere (Willis *et al*., 2009), the assembly and turnover of this contractile unit is of the utmost importance for proper contractile functioning. While little is known about the precise mechanisms involved; the process of sarcomeric protein turnover forms part of the quality control system that is specifically and tightly regulated through the actions of chaperones and the UPS (Willis *et al*., 2009). If the functions of the chaperones and UPS are impaired, misfolded proteins tend to form aggregates with detrimental effects (Galvez *et al*., 2007).

Interestingly, protein aggregates are a common feature of amyloid-related neurodegenerative diseases such as Alzheimer's, Parkinson's and Huntington's; illustrating that UPS impairment is an important pathogenic mechanism (Schlossarek and Carrier, 2011). Similar observations of aberrant protein aggregation, accumulated ubiquitinated proteins, as well as amyloid oligomers in cardiomyopathies such as HCM, DCM and desmin-related cardiomyopathy (DRM), resembles the pathology of neurodegenerative diseases and highlights the possible role of the UPS in their pathogenesis (Sanbe *et al*., 2004; Chen *et al*., 2005; Schlossarek and Carrier, 2011). Moreover, UPS impairment has been detected prior to cardiac hypertrophy, thus suggesting that defective protein turnover likely contributes to cardiac remodelling (Chen *et al*., 2005).

Furthermore upon impairment of the UPS, lysosomal-mediated degradation via autophagy is activated (see section 1.5.2), which also contribute to the quality control system. Although the involvement of chaperones will be briefly discussed in the context of chaperone-assisted degradation, this section of the thesis will focus more on the involvement of the UPS, since it has been shown to be intimately involved in regulating cMyBPC levels, *MYBPC3*-derived HCM and because it seems to underlie the function of the COMMD proteins to a great extent.

1.4.1. The involvement of chaperones/co-chaperones in protein degradation

Chaperones are factors that protect proteins from misfolding and prevent aggregation by directing whether proteins should be re-folded to their native state or targeted to the UPS through the cooperation of distinct co-chaperones (Arndt *et al*., 2007). In addition, chaperones also aid in the transport of proteins across cellular membranes, in the assembly of protein complexes and in the conformational regulation of signalling proteins and regulators of apoptosis (Caplan *et al*., 2007). However, for the purpose of this section of the thesis, our interests in these factors will only extend to their involvement in the chaperone-assisted degradation of sarcomeric components.

Throughout the UPS, distinct co-chaperones have been shown to aid in each step of the pathway; from recognising the target protein through to its subsequent delivery and docking at the proteasome complex. They aid in the prevention of aggregate formations, stimulate sorting to the proteasome and also protect

against the cleavage of poly-ubiquination chains via de-ubiquitinating enzymes (DUBs) (Westhoff *et al*., 2005).

An example of chaperone/co-chaperone complexes involved in sarcomeric protein assembly and degradation are Hsp 90 and 70, along with their co-chaperone UNC-45 (Kachur and Pilgram, 2008). UNC-45 has been implicated in myosin assembly, as it binds directly to myosin and prevents its aggregation. However, the regulation of UNC-45 is mediated by the UPS, in particular through the actions of C-terminal of Hsp70-interacting protein (CHIP). Thus CHIP, a protein with both chaperone and E3 Ub-ligase activity, is very important to both the folding and the degradation of myosin (Nyamsuren *et al*., 2007; Boateng and Goldspink, 2008; Willis *et al*., 2009).

1.4.2. Turnover rate of myofibrillar proteins

The $t_{1/2}$ of proteins in general vary from a few minutes to years, but myofibrillar proteins have $t_{1/2}$ that vary from hours to days. The $t_{1/2}$ of myosin is ~5-8 days; that of troponin subunits (T/I/C) is ~3-5 days, while actin and tropomyosin have $t_{1/2}$ of ~7-10 days (Martin *et al*., 1977; Willis *et al*., 2009). In contrast, WT cMyBPC has been shown to be a short-lived protein with a $t_{1/2}$ of ~10 hours under normal conditions, while a truncated cMyBPC mutant (e.g. E334) had a $t_{1/2}$ of ~2 hours; WT ssMyBPC also has a $t_{1/2}$ of 3 hours (Bosch-Comas *et al*., 2006; Kulikovskaya *et al*., 2007; Bahrudin *et al.,* 2008). Evidence from skeletal muscle cells shows that all myofibrillar proteins are degraded by the UPS (Purintrapiban *et al*., 2003; Kedar *et al*., 2004).

Myofibrillar proteins can only be degraded via the UPS after they have been removed from the surface of the myofibril without disrupting the continuously functioning muscle (Solomon and Goldberg, 1996; Goll *et al*., 2008). Several groups have shown that calpains, which are Ca^{2+}-dependent, non-lysosomal, cysteine-proteases, initiate sarcomeric protein turnover by disassembling proteins from the myofibril and releasing them as individual filament systems (Tidball and Spencer, 2002; Du *et al*., 2004; Goll, 2008). Cardiac calpain-1 was shown to cleave titin, nebulin as well as desmin and filamin at the Z-disk, with subsequent release of α-actinin, which would, in turn, release thin filaments from the surface of the myofibril (Goll *et al*., 1991; Goll et al., 2003). Furthermore, calpain cleavages of the tropomyosin dimer and troponin complex, would in turn facilitate the disassembly of the actin in the thin filament to G-actin monomers (Goll *et al*., 2008). Likewise, calpain cleavage of M proteins and titin would sever the thick filaments from the myofibril (Goll *et al*., 2008). Furthermore, calpain cleavage of MyBPC and selective loss of the myosin light chains would promote the disassembly of the thick filament to myosin monomers, as these proteins, in their phosphorylated state, are important in the stabilisation of the thick filament (Goll *et al*., 1992; Goll *et al*., 1999; Russel *et al*., 2000; Decker *et al*., 2005). In this way, both the sarcomeric proteins actin and myosin are made susceptible to degradation by the UPS.

However, in the process of atrophy of skeletal muscle, the myosin light chains and MyBPC can be ubiquitinated via the E3 ubiquitin-ligase, MuRF1 (see section 1.4.3.3.1), even while they were still

associated with the myofibril, unlike myosin heavy chains and actin (Cohen *et al*., 2009). Following ubiquitination and degradation of myosin light chain and MyBPC the thick filament becomes unstable. Consequently, this in turn prompts the ubiquitination, disassembly and degradation of myosin heavy chains. On the other hand, the thin filaments components are degradaded via a distinct mechanism independent of MuRF1 and most likely involve other ubiquitin-ligases (Cohen *et al*., 2009). Thus, the ubiquitination of certain sarcomeric proteins via E3 ubiquitin-ligases (e.g. MuRF1) under atrophying conditions provides an additional mechanism whereby thick filaments can be released from the myofibril and subsequently be degraded, resulting in muscle loss. However, it is still unclear whether the mechanisms or pathways that stimulate muscle wasting are similar to those involved in myofibrillar turnover during muscle growth (Goll *et al*., 2008).

Nevertheless, the importance of the UPS in sarcomeric protein turnover is quite clear from the abovementioned facts.

1.4.3. Ubiquitin proteosome system in the heart

The UPS, present in the nucleus and the cytosol, are known as the intracellular ‚garbage disposal system' since it degrades 80-90% of all intracellular proteinsthat are misfolded, mutated or damaged.

For the process of protein degradation, an appropriate substrate, the ubiquitin-conjugation machinery and the 26S proteasome are required. However, for optimal degradation, additional supplementary proteins such as chaperone/co-chaperone complexes and ubiquitin-binding proteins are also essential (Hartmann-Petersen *et al*., 2003). Furthermore, evidence to date has also shown that this pathway is an integral part of key cellular processes such as cell signalling, apoptosis, DNA repair, endocytosis and transcriptional control.

1.4.3.1. The process of ubiquitin conjugation

The process of ubiquitination is ATP-dependent, and involves the concerted action of three enzymes. E1, the ubiquitin-activating enzyme, uses ATP to activate ubiquitin. The C-terminal glycine (G76) of ubiquitin, which is linked to a sulphydryl group of E1 via a thiolester bond, becomes adenylated with the release of pyrophosphate and subsequent transfer of activated ubiquitin from E1 to E2, the ubiquitin-conjugating enzyme (Powell, 2006). E2s can also directly ubiquitinate proteins that have ubiquitin-binding domains (UBDs), but in most cases an E3 ubiquitin-ligase is required to do this (Hoeller *et al*., 2007). Different E3 ubiquitin-ligases recognise specific substrates and mediate the transfer of the activated ubiquitin from the E2 to a lysine residue of the substrate protein (Figure 1.9) (Glickman and Cienchanover, 2002; Zolk *et al*., 2006).

Since ubiquitin can be conjugated at one of its seven lysine residues, the type of lysine-ubiquitin linkages, together with the manner in which ubiquitin is attached to the substrate, either as a single ubiquitin moiety

(mono-ubiquitination) or the covalent linkage of multiple ubiquitin moieties (poly-ubiquitination), determines the fate and activity of the target protein.

Poly-ubiquitination, which occurs via the successive addition of more than four activated ubiquitins to Lysine48 or Lysine29 of the first ubiquitin, generally specify degradation of the substrate by the 26S proteasome (Hartmann-Petersen *et al*., 2003). In some cases, the elongation of these ubiquitin chains is also facilitated by E4 ubiquitin-ligases, which have been proposed by some to be a subset of E3 ubiquitin-ligases because of their ubiquitin-ligase properties (Koegl *et al*., 1999).

1.4.3.2. The 26S proteasome

Substrates marked to be degraded are either directly recognised by the 26S proteasome or are bound to chaperone/co-chaperone complexes or an ubiquitin-binding domain within a shuttling protein. Both these proteasome-binding proteins and chaperone complexes are accessory proteins that are required for optimum degradation of target proteins, as they aid in the transport of the substrate to the proteosome and also act as adaptors to facilitate proteasomal binding (Hartmann-Petersen *et al*., 2003; Westhoff *et al*., 2005).

The 26S proteasome, which degrades poly-ubiquitinated substrates in an ATP-dependent manner, consist of a proteolytically active 20S cylinder core and a 19S regulatory particle, which is attached to either one or both ends of the 20S cylinder. Furthermore, access to the active sites of the 26S proteasome is tighly regulated via the the 19S particle, which recognises, deubiquitinates and unfolds ubiquitinated substrates (Hartmann-Petersen *et al*., 2003).

Ubiquitin molecules are first released for reuse via the actions of DUBs, which are associated with the 19S regulatory particle. Aside from a role in recycling, these enzymes are also needed for proof-reading in case substrates are improperly ubiquitinated (Glickman and Cienchanover, 2002).

Furthermore, partial unfolding of the substrate is a prerequisite for the targeted substrate to be fed into the proteolytic core of the proteasome (Pickart and Cohen, 2004). Simultaneously, while the substrate is processed and inserted, the folded state of chaperones and associated co-chaperones prevents their own insertion into the proteolytic core (Arndt *et al*., 2007).

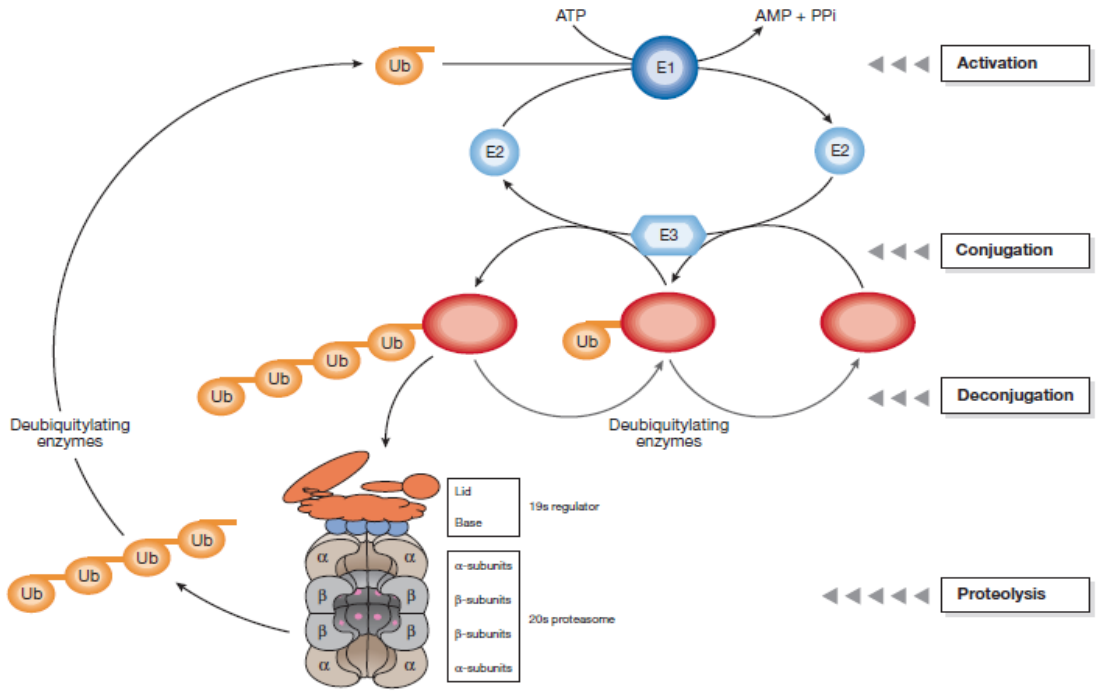


Figure 1.9. A basic overview of the ubiquitin proteasome pathway.
Ubiquitin is activated by E1 in an ATP-dependent process and then transferred to the E2 enzyme. The E3 ubiquitin-ligase confers specificity and mediates the transfer of one or more ubiquitin molecules to the substrate. Poly-ubiquitin chains, linked via Lysine48, are recognised by the 26S proteasome, which targets the protein for degradation. Deubiquitinating enzymes reverse the process of ubiquitin conjugation and recycle ubiquitin. (Image taken from Wilkinson *et al*., 2005)

1.4.3.3. The E3 ubiquitin ligase - Skp1-Cullin-F-box (SCF) complex

Since E3 ubiquitin-ligases confer substrate specifity, the amount and type of E3 enzyme complexes reflects the diversity of proteins regulated by the UPS. Compared to the single E1 enzyme and the dozen E2 enzymes, there are approximately 1000 E3 enzymes grouped into three different classes. There are the "homologous to E6-AP COOH terminus" (HECT) domain E3 ubiquitin-ligases, the RING finger-containing E3 ubiquitin-ligases and the SCF complex ubiquitin-ligases, which are considered to be a subset of the RING-finger domain E3s (Powell, 2006). This cullin-based RING finger-type E3 ubiquitin-ligase is the largest family of ubiquitin ligases and is involved in the ubiquitination of various regulatory and signalling proteins. Since it was shown that COMMD proteins *per se* interact with several Cullin subunits, a closer look at the multisubunit RING SCF complex follows.

Cul1 provides a scaffold for the components of the complex and binds both a RING finger protein, RBX1, which recruits the E2 enzyme, and Skp1, which serves as an adaptor for bridging to different F-box proteins that recognises the particular substrate (Zheng *et al*., 2002). However, for the complex to become activated, Cul1 must first be neddylated via Nedd8; neddylation is a posttranslational modification similar to ubiquitination. This is necessary, since unmodified Cul1 is bound to a sequestration factor called CAND1, which in turn blocks the binding of the Skp1-F-box complex to the N-terminus of Cul1. However, Nedd8 conjugation to Cul1 causes CAND1 to dissociate and allows the formation of the active SCF complex (Skaar

and Pagano, 2009). More importantly, neddylation has a conformational effect as it causes closure of the large gap of 50Å that at first exists between the F-box protein and the E2 binding site; this closure, in turn, is crucial for the intial ubiquitin transfer as well as for ubsequent ubiquitin-chain extension (Duda *et al*., 2008; Saha and Deshales, 2008).

1.4.3.3.1. E3 ubiquitin ligases - role in cardiac hypertrophy

The role of the UPS in cardiac disease, in particular cardiac hypertrophy, has only recently been elucidated with the identification of certain E3Ub-ligases within the sarcomere of striated muscle. These include the Muscle Ring Finger family of proteins (MuRF1 [Section 1.4.2], MuRF2, and MuRF3) and MAFbx/atrogin-1 (muscle atrophy F-box). But, exactly how these enzymes are involved in the context of hypertrophy and other cardiomyopathies is still under investigation.

Interestingly, both MuRF1 and atrogin-1 have anti-hypertrophic properties and were shown to be upregulated in an *in vivo* rat model of chronic heart failure as well as in a cardiac model of hypertrophy (pressure overload generated by constricting the ascending aorta) (Razeghi *et al*., 2006; Adams *et al*., 2007). Atrogin-1, a SCF E3 complex subunit, was shown in *in vitro* and *in vivo* experiments to inhibit pathologically-induced cardiac hypertrophy, by interacting with and targeting calcineurin A for degradation (Li *et al*., 2004). Calcineurin A is a Ca^{2+}-activated serine/threonine phosphatase that is an important intermediary in the hypertrophic signalling cascade (Molkentin *et al*., 1998).

Unlike atrogin 1, the MuRF family of proteins has been shown to interact with sarcomeric proteins such as myosin heavy chain, titin, troponin I and T, regulatory myosin light chain and telethonin, at least under *in vitro* conditions, and thus might be a key player in sarcomeric protein turnover (Kedar *et al*., 2004; Witt *et al*., 2005). MuRF1, in particular was shown to interact and degrade troponin I and myosin heavy chain, as well as inhibiting protein kinase C signalling through its interaction with receptor for activated protein kinase C 1 (RACK1), an anchoring protein that mediates the translocation and activation of PKC (Kedar *et al*., 2004; Clarke *et al*., 2007). Moreover, mice deficient in MuRF1 developed cardiac hypertrophy during development (Willis *et al*., 2007).More recently, using adenoviral gene transfer in neonatal rat cardiac myocytes and MuRF1 mouse models, Mearini *et al* (2009) demonstrated that both atrogin-1 and MuRF1 interacts with WT and truncated cardiac MyBPC. Moreover, atrogin-1 specifically targets truncated cMyBPC and not WT-cMyBPC for degradation, while MuRF1 indirectly reduces both WT and truncated cMyBPC levels by regulating the transcription of MHC (Mearini *et al*., 2009). The authors speculated that MuRF1 might degrade a MHC-specific transcription factor. Furthermore, as the stoichiometry of the sarcomere is tighly regulated by co-translational and co-assembly mechanisms, it was assumed that lower expression levels of MHC is most likely associated with a lower levels of cMyBPC and vice versa (Mearini *et al*., 2009).

1.4.3.4. Regulation of the UPS

As protein degradation is higly regulated, the following question arises: how and when does the UPS recognise and determine that a specific protein, which is not required at the time, needs to be ubiquitinated for degradation?

Earlier studies by Bachmair and Varshavsky (1989) have shown that the UPS recognises short-lived proteins as proteolytic substrates via their amino terminal residues. This is called the N-end rule pathway, which states that the nature of the N-terminal amino acids is an important factor that determines the proteins half-life (Bachmair and Varshavsky, 1989). Stabilising N-terminal residues such as methionine, glycine, alanine, serine and threonine was shown to confer a long half-life of ~20 hours, while de-stabilising N-terminal residues such as leucine, phenylalanine, aspartate, lysine and in particular arginine, confers half-lives from 2-3 minutes (Bachmair and Varshavsky, 1989). These features of a protein that confer metabolic instability are called its degradation signal or its N-degron. In eukaryotes, the N-degron comprises of both a destabilising N-terminal residue and an internal lysine, which functions as an acceptor for the poly-ubiquitin chain. In addition, a functional N-degron can also be created via N-terminal modifications (deamidation, oxidation and/or arginylation) of a pre-N-degron (Tasaki *et al*., 2007). Interestingly, 60% of muscle proteins can be degraded via this pathway; however, long-lived sarcomeric proteins such as myosin and actin are not N-end rule substrates (Solomon *et al*., 1998; Reid, 2005).

Following recognition of proteolytic substrates, degradation is regulated at the ubiquitination step where specificity exists. Thus, ubiquitination of a specific substrate is controlled by the exposure or maturation of its ubiquitination signal (N-degron), the availability and activity of its specific E3 Ub-ligase and the interaction between the substrate and the E3 ubiquitin-ligase (Chew and Hagen, 2007; Wang *et al*., 2008). Post-translational modifications (PTMs) of a substrate within or nearby its N-degron can regulate its ubiquitination by changing the conformation of the substrate to either expose or mask its ubiquitination signal (Wang *et al*., 2008).

Furthermore, some E2/E3 complexes are not constitutively active and also require PTMs, such as phosphorylation or neddylation, as is the case for cullin-based E3 ubiquitin ligases (Skaar and Pagano, 2009). Binding of substrate and the substrate recognition subunit increases neddylation and thereby activation of the E3 complex, as described in section 1.4.3.3. Degradation signals on substrates may also be masked by binding to other proteins, which protect the bound protein from ubiquitin-conjugation. Furthermore, muscle fibres contain DUBs, which reverses the ubiquitin signal and protect against degradation of the target protein (Reid, 2005).

1.4.3.5. Proteasome-independent functions of ubiquitin

Unlike Lysine48 chains, Lysine63 poly-ubiquitin chains function in endocytosis, DNA repair, transcriptional regulation and cellular signalling (Powell, 2006). Moreover, Lysine63 poly-ubiquitinations have been implicated in the regulation of physiological cardiac hypertrophy (Li *et al*., 2007).

Similar to Lysine63 poly-ubiquitination, mono-ubiquitination is associated with non-proteolytic functioning and is important in the regulation of protein activity and function, protein trafficking (such as receptor internalization for endocytosis, and nuclear export), histone modification, tagging and sorting of proteins as well as transcription (Woelk *et al*., 2007). In addition, several lysine residues of a substrate can simultaneously be mono-ubiquitinated (multi-mono-ubiquitination) for receptor internalization and endocytosis (Haglund *et al*., 2003), while ubiquitination of the α-NH_2 group on the N-terminal residue of a substrate is also possible, which has important negative effects on protein stability and activity (Powell, 2006).

1.5. ENDOSOMAL PROTEIN TRAFFICKING

Within the eukaryotic cell, proteins carry out their specific cellular processes within membrane-defined intracellular compartments. Vesicle transport of cargo proteins, whether inwards (endocytosis) or outwards (exocytosis), allows proteins from various compartments to communicate with the extracellular and intracellular environment (Jovic *et al*., 2010). While exocytosis is more well-defined, the intracellular trafficking is of more interest to us in the context of this thesis, and thus this section will briefly focus on endosomal protein trafficking. In addition, the role of autophagy-mediated protein degradation via the lysosome will also be briefly discussed as it forms part of the sarcomeric quality control system and has been implicated in hypertrophy development.

The endosomal compartment functions in one of two ways: endocytosed material are either recycled to be returned to the plasma membrane or transported in a series of steps to digestive lysosomes (Jovic *et al*., 2010). The early endosome serves as a point of divergence for degradative and recycling pathways, as well as the point of delivery of internalised proteins to the biosynthetic pathway (TGN). For lysosomal-protein degradation, materials are transported from the early/sorting endosome to the late endosome. Subsequently, mature late endosomes, also known as MVB (vesicles within vesicles), either become lysomes or fuse with other lysosomes for degradation of the target proteins (Piper and Katzmann, 2007). With regards to myofibrillar turnover, it has been established that lysosomal proteases are unlikely to be involved in this process since myofibrils are too large (Goll *et al*., 2008). Furthermore, muscle cells contain few lysosomes; however under pathological conditions this lysosomal proteolytic system is significantly enhanced (Bechet *et al*., 2005).

Endosomes use different adaptor proteins and coats to accomplish their multiple functions. The retromer complex and ESCRT machinery (endosomal sorting complex required for transport system) are coat/adapter

combinations that regulate endosomal trafficking. While the retromer mediates recycling of sorting receptors back to the Golgi apparatus, the ESCRT system is needed for sorting of transmembrane cargo into the MVB through to the lysosome (Hurley, 2008) (Figure 1.10).

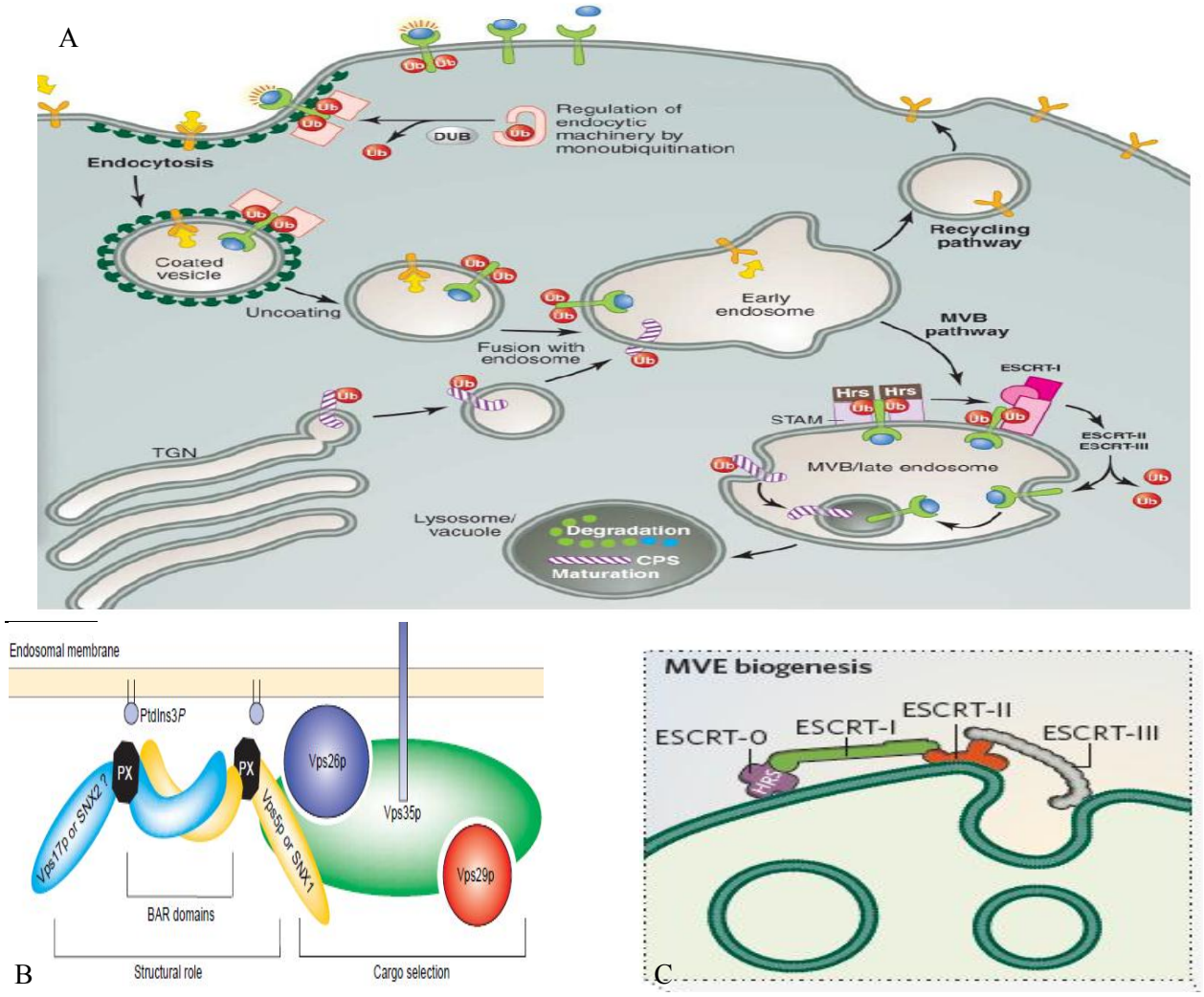


Figure 1.10. A basic overview of the endosomal trafficking routes.
A). This model integrates vesicular traffic from the plasma membrane or the trans-Golgi network (TGN) to the early endosome, where cargo is sorted back to the membrane, TGN (via retromer complex) or the lysosome (via the ESCRT omplex). Schematic representation of the retromer complex (B) and the ESCRT machinery (C). Images taken from Mukhopadhyay and Rietzman, 2007 and Seaman, 2005

1.5.1 Regulation of endosomal trafficking

1.5.1.1 The retromer complex in endosome-to-golgi trafficking

The retromer is a pentameric complex at the early endosome which mediates the endosome-to-Golgi trafficking of various proteins (Seaman, 2005; Bujny *et al*., 2007). The retromer consists of a cargo-recognition subcomplex, containing proteins vacuolar protein sorting (Vps) 26, Vps29 and Vps35, and a sorting nexin dimer (eg. SNX1 and 2) (Figure 1.10 B). Interestingly, these sorting nexins (SNX) proteins contain a PX domain, which specifically binds phosphatidylinositol 3-biphosphate (PtdIns(3)P). Binding to

PtdIns(3)P recruits the retromer to the endosomal membranes and mediates vesicle budding (Lindmo and Stenmark, 2006).

Vps35 is proposed to be the core component of the retromer, since it binds directly to Vps26, Vps29 and SNX1 (Haft *et al*., 2000). The N-terminal of Vps35 interacts with Vps26, while the C-terminal binds to Vps29. Vps26 exists as two isoforms, encoded by seperate genes, *Vps26A* and *Vps26B*. There are thus two retromer cargo-recognition complexes, since both form part of the retromer subcomplex (Collins *et al*., 2008).

Interestingly, in yeast, adaptor proteins are needed to act as regulators of the retromer. Grd19, a yeast sorting nexin, which is an orthologue of human SNX3, physically associates with the retromer and provides it with more cargo-recognition specificity (Strochlic *et al*., 2007).

1.5.1.2. The ESCRT machinery in lysosomal-mediated degradation

This lysosomal pathway is important in the degradation of internalised proteins, especially in receptor downregulation.

As mentioned in section 1.4.3.5, ubiquitination of signalling receptor proteins *en route* to the lysosome is a key signal for entry and sorting. As ubiquitin-conjugation works in concert with the 26S proteasome, inhibition of the UPS would in turn also affect internalisation of proteins such as the growth factor-activated receptor tyrosine kinase (RTKs) and growth hormone receptors.

The initial recognition of ubiquitinated internalised proteins in the early endosome is carried out by MVB sorting receptors, eg. Vps27 (in yeast)/Hrs (in mammals), which, together with the ESCRT multiprotein complex (ESCRT -I, -II, and III), target ubiqitinated proteins into the MVB sorting pathway (Slagsvold *et al*., 2005, Saksena *et al*., 2007; Hurley, 2008). Interestingly, ubiquitin itself binds only weakly to Hrs, thus it requires the presence of additional ubiquitin-binding proteins to stabilise its association with the ubiquitinated cargo (Bache *et al*., 2003). The highly conserved Hrs connects to the endosome via its FYVE domain, which binds PtdIns(3)P (Mukhopadhyay and Riezman, 2007).

Vsp27/Hrs, now renamed ESCRT-0, facilitates the assembly of the ESCRT proteins; these complexes consist of soluble cytosolic Vps proteins and delivers the ubiquitinated cargo to ESCRT-I, from where they are then passed on to ESCRT-II and then to ESCRT-III. ESCRT-III, in turn, promotes invagination of the endosomal membrane (budding inwards) to form luminal vesicles and thus the formation of MVBs (Figure 1.10C) (Luzio *et al*., 2009). Subsequent fusion with lysosomes results in the degradation of these sorted, internalised proteins.

The model for the interplay of theses ESCRT proteins in yeast have been well-described, however, their role in mammalian MVB formation and sorting is more complex and and not well understood.

1.5.2. Autophagy - lysosomal mediated degradation

In addition to its role in the disposal of internalised proteins, the lysosome also digests cargo derived from autophagy, which, along with the UPS, is responsible for intracellular protein degradation. While the UPS predominantly degrades short-lived nuclear and cytosolic proteins, autophagy, on the other hand, degrades long-lived proteins, aggregated proteins and organelles (Korolchuk *et al*., 2009). Either UPS malfunction and/or aberrant protein aggregation may activate autophagy (Wang *et al*., 2008).

Induction of autophagy starts with the formation of a small isolation membrane, which elongates upon recruitment of Atg proteins. The membrane encloses protein or organelles to be sequestered and forms the autophagosome, which in turn fuses with lysosomes for degradation (Gustafsson and Gottlieb, 2008).

More recently, a connection between the ESCRT machinery and autophagy has also come to light (Raiborg and Stenmark, 2009). Upon inactivation of the ESCRT machinery, an increase in the number of autophagosomes is observed, which suggests a direct or indirect involvement of the ESCRT machinery in the regulation of autophagy. The accumulation was suggested to be due to the inhibition of fusion between the autophagosome and the lysosome, since these fusions are rare in ESCRT-depleted cells (Raiborg and Stenmark, 2006).

Furthermore, both the PI3K and mTOR pathways have been implicated in the regulation of autophagy-induction (Wang *et al*., 2008).

Under normal conditions, the process of autophagy occurs at low levels, but it is upregulated under cellular stress, such as hypoxia, nutrient deprivation, mitochondrial dysfunction and conditions such as I/R, cardiac hypertrophy, DCM and HF (Korolchuk *et al*., 2009). However, the role of increased autophagy in these heart conditions is still unclear due to the conflicting results from various studies. For example, deletion of *atg5* (involved in vesicle elongation) leads to sarcomeric disarray impairs autophagy and results in the accumulation of abnormal organelles and subsequent cardiac hypertrophy and contractile dysfunction in adulthood (Nakai *et al*., 2007). Similar results are also observed in Danon's cardiomyopathy, which is due to a deficiency in the lysosomal protein, Lamp-2 (Tanaka *et al*., 2000). In contrast, increased numbers of autophagosomes are observed in DCM, but the functional significance of this is unclear. Whether these increased numbers of autophagosomes observed in dying cells contributes directly to cell death, or whether they represent an effort to prevent cell death, remains controversial (Gustafsson and Gottlieb, 2008). However, evidence has suggested a dual role for autophagy in the heart; to protect cells or contribute to cell death, depending on the stimuli.

At first the UPS and autophagy were considered two separate degradation pathways, however, recent evidence suggests cross-talk between these two pathways, especially in the heart. Both *in vitro* and *in vivo* studies show that UPS impairment induces autophagy. Alternatively, knockout of Atg7 or Atg5 in mouse models resulted in increased poly-ubiquitinated proteins, which led to formation of Ub-positive protein aggregates (Wang *et al*., 2008). The reduced autophagy in these mice caused both enhanced protein synthesis as well as degradation via the UPS; however, the clearance of protein aggregates and organelles depended greatly on autophagy rather than the UPS, as size constraints regulate efficient proteosomal degradation (Nakai *et al*., 2007; Wang *et al*., 2008; Verhoef *et al*., 2009). Furthermore, protein aggregates also lead to the depletion of free ubiquitin and inactivation of UPS components and results in inefficient functioning of the UPS (Dantuma *et al*., 2006). Thus, these findings delineate a link between the UPS and autophagy.

1.6. THE PRESENT STUDY

From the literature, it is evident that the N-terminal region of cardiac MyBPC plays an important role in regulating cardiac contractility in a manner dependent on the phosphorylation statusof its cMyBPC motif. It was, therefore, previously hypothesised in our laboratory that the cMyBPC-motif under various phosphorylated states might interact with other ligands that could also influence cardiac contractility. In a search for other interactors of the N-terminal of cMyBPC, COMMD4 was found to bind to the N-terminal of cMyBPC in a phosphorylation dependent manner. This protein of unknown function belongs to the COMMD family of proteins that have been linked to cardiac hypertrophy, Cu metabolism and the UPS, as described in this chapter. Moreover, studies have implicated the UPS in at least *MYBPC3*-derived HCM, while dietary Cu-depletion has also been shown to cause cardiac hypertrophy. Based on these findings, COMMD4 was viewed as an intriguing candidate modifier of cardiac hypertrophy in HCM, but its role remained unclear.

Since the project forms part of a greater study focused on identifying sarcomeric-bound regulators of contractility and modifiers of hypertrophy, the aim of this present study was two-fold. First, COMMD4 was used as bait in a Y2H library screen to determine its distal ligands with a view to elucidating the function of COMMD4, particularly in the context of MyBPC functioning. Putative COMMD4 interactors were subjected to *in vitro* and *in vivo* co-immunoprecipitation (Co-IP) analysis and 3D live-cell fluorescent colocalisation, with Z-stacking, to verify the putative interactions. In addition, a domain/phosphorylation interaction assay was performed to determine whether the interaction between COMMD4 and the cMyBPC-motif was phosphorylation-dependent as had been suggested in a previous study, while the effect of RNAi-mediated COMMD4 knockdown in H9C2 cardiomyocytes on MyBPC turnover was monitored using live cell 3D-fluorescence imaging.

Secondly, variants in genes encoding COMMD4 and certain of its identified interactors were assessed as possible modifiers of cardiac hypertrophy in three cohorts of South African HCM-families, in which one of

three-founder mutations segregate. Appropriate family-based statistical methods were applied to assess association of these gene variants with hypertrophy development.

CHAPTER 2: MATERIALS AND METHODS

INDEX PAGE

A. YEAST TWO HYBRID ASSAY 55

2.1. BACTERIAL, YEAST STRAINS AND CELL LINES 55

2.1.1. Bacterial strain 55

2.1.2. Yeast strain 55

2.1.3. Cell lines 55

2.2. PRODUCTION OF *E.COLI* DH5α COMPETENT CELLS FOR BACTERIAL TRANSFORMATION 56

2.3. CULTURING OF CELL LINES 56

2.3.1. Culture of cell lines from frozen stocks 56

2.3.1.1. Thawing of cells 56

2.3.1.2. Removing DMSO from frozen stocks and culturing of cells 56

2.3.2. Subculturing of H9C2/ HEK293 cell cultures 57

2.4. TRANSFORMATION AND TRANSFECTION OF PLASMIDS INTO *E.COLI*, *S.CEREVISIAE*, AND H9C2/HEK293 CELL LINES 57

2.4.1. Bacterial plasmid Transformation (*E.coli)* 57

2.4.2. Yeast plasmid transformation (*S. cerevisae*) 57

2.4.3. Mammalian cells transfection (H9C2/HEK293) 58

2.4.3.1. Single and co-transfections of plasmid constructs 58

2.4.3.2. Differentiation of H9C2 cells 61

2.5. DNA ISOLATION AND PURIFICATION 61

2.5.1. Bacterial plasmid purification 61

2.5.2. Plasmid isolation from yeast 62

2.5.3. Purification of PCR and plasmid products 62

2.5.4. Gel purification of PCR-amplified product from agarose gels 62

2.6. POLYMERASE CHAIN REACTION (PCR) 63

2.6.1. Primer design 63

2.6.1.1. Primers design for generating plasmid constructs 63

2.6.1.2. Primers for Y2H insert screening and *in vitro* transcription and translation 65

2.6.2. High fidelity PCR amplification of bait/prey inserts 65

2.6.3. PCR amplification for *in vitro* transcription and translation 66

2.6.4. Bacterial colony PCR 66

2.7. GEL ELECTROPHORESIS 66

2.7.1. Agarose gel electrophoresis for the visualisation of PCR amplified products

and isolated plasmids 66
2.7.2. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) 67
2.8. AUTORADIOGRAPHY 67
2.9. AUTOMATED SEQUENCING 68
2.10. DNA SEQUENCING ANALYSIS 68
2.11. Y2H PREY IDENTIFICATION 68
2.12. RESTRICTION ENZYME DIGESTION 69
2.12.1. Restriction enzyme digestion for cloning 69
2.12.1.1. Digest for cloning bait/prey inserts into relevant vectors 69
2.13. GENERATION OF BAIT AND PREY CONSTRUCTS 70
2.13.1. Generation of plasmid constructs 70
2.13.2. Calf intestinal alkaline phosphatase (CIP) treatment of double digested vectors 71
2.13.3. DNA Ligation 71
2.14. QUALITY CONTROL TESTING OF THE Y2H CONSTRUCTS 71
2.14.1. Phenotypic verification of yeast strains AH109 and Y187 72
2.14.2. Toxicity test of the transformed cells 72
2.14.3. Establishment of mating efficiency 72
2.15. YEAST-TWO-HYBRID ANALYSIS 73
2.15.1. Cardiac cDNA library 73
2.15.2. Establishment of bait culture 73
2.15.3. Library mating 74
2.15.4. Establishment of the library titre 74
2.15.5. Detection and measurement of reporter gene activation 74
2.15.5.1. Selection of diploid yeast colonies containing putative interactor peptides 75
2.15.5.2. Detection of colourimetric reporter gene activation 75
2.15.6. Rescuing prey plasmids from diploid cells 75
2.15.7. Interaction specificity tests-Heterologous bait matings 76
2.16. PROTEIN-PROTEIN INTERACTIONS OF COMMD4 76
2.16.1. *IN VITRO* CO-IMUNOPRECIPITATION 76
2.16.1.1. Transcription and translation of baits and preys 76
2.16.1.2. *In vitro* co-immunoprecipitation of translated PCR products 77
2.16.2. *IN VIVO* CO-IMUNOPRECIPITATION 78
2.16.2.1. Cell lysis and sample preparation of monolayer cells 78
2.16.2.2. Bradford Assay for determination of protein concentration 78
2.16.2.3. Optimisation of optimal antibody dilution concentrations 79
2.16.2.4. Pre-clearing with protein G agarose 79
2.16.2.5. *In vivo* immunoprecipitation / co-immunoprecipitation reactions 81
2.16.2.6. Western blots / Immunoblotting 81

2.16.3. *IN VIVO* CO-LOCALISATION 82

2.16.3.1. Immunocytochemistry 82

2.16.3.2. Co-localisation assay 83

B. FUNCTIONAL STUDIES 84

2.17 DOMAIN/PHOSPHORYLATION ASSAY 84

2.18 siRNA-MEDIATED COMMD4 KNOCKDOWN 84

2.18.1. siRNA preparation and transfection 85

2.18.2. Total RNA extraction 86

2.18.3. cDNA synthesis and two-step Q-RT-PCR 86

2.18.4. Live cell fluorescence 86

C. EVALUATION OF *COMMD4* AND ITS LIGANDS AS HYPERTROPHY MODIFIERS 87

2.19. STUDY SUBJECTS, CLINICAL EVALUATION AND BLOOD COLLECTION 87

2.19.1. Study subjects 87

2.19.2. Clinical Evaluation 87

2.19.3. Blood collection 89

2.20. BIOINFORMATICS 90

2.20.1. SNP selection 90

2.21. ALLELE SPECIFIC RESTRICTION ENZYME ANALYSIS (ASREA) OF COMMD4 92

2.21.1. Primers for *COMMD4* SNP genotyping using ASREA 92

2.21.2. PCR amplification conditions for genotyping 92

2.21.3. SNP genotyping by means of ASREA 92

2.22. TAQMAN®SNP GENOTYPING ASSAYS OF PUTATIVE COMMD4 INTERACTOR GENES 93

2.22.1. PCR genotyping by Taqman® SNP Genotyping assays 93

2.23. STATISTICS 94

2.23.1. Descriptive statistics 95

2.23.2. Linkage disequilibrium test 95

2.23.3. Quantitative transmission disequilibrium test 95

CHAPTER 2: MATERIAL AND METHODS

A. YEAST TWO HYBRID ASSAY

2.1. BACTERIAL, YEAST STRAINS AND CELL LINES

2.1.1. Bacterial strains

Ligation reactions were transformed into the *E.coli* DH5α bacterial strain, to aid in the selection, amplification and purification of plasmid constructs to be used in the various assays. Transformed colonies were selected on the basis of their ability to grow on Luria-Bertani (LB)-agar plates supplemented with the appropriate selection antibiotics (Appendix I). Kanamycin (Kan) (50mg/ml) was used as a selection antibiotic for the pGBKT7 and DsRed-Monomer-C1vectors, Ampicillin (Amp) (50mg/ml) for the selection of the pACT2 and EYFP vectors, and Zeocin^{TM} (100mg/ml) was used for the selection of the pGFP^2 vector. Subsequently bacterial colony PCR (Section 2.6.4) was performed to identify colonies containing the recombined plasmids.

2.1.2. Yeast strains

The pGBKT7 bait construct was transformed into the MATa strain, AH109 and the cardiac protein clones present in the pre-transformed CLONTECH cDNA library had been transformed by the manufacturer into the MATα strain, Y187, the mating partner for AH109. Both yeast strains AH109 and Y187 are modified strains that are deficient for the selection genes:*ADE2*, *HIS3*, *LEU2* and *TRP1*, therefore, they are unable to grow on synthetic dropout (SD)-agar plates that lack the nutrients adenine (Ade$^-$), histidine (His$^-$), leucine (Leu$^-$) and tryptophan (Trp$^-$). On the contrary, both AH109 and Y187 are able to grow on agar plates that lack the nutrient, Uracil (Ura$^+$) because of the presence of the relevant *URA3* gene that facilitates production of Ura in the genetic make-up of both yeast stains.

2.1.3. Cell lines

Rattus norvegicus (rat) cardiac (atrial) myocyte, H9C2 and human embryonic kidney (HEK 293) cell lines were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA) for use in subsequent experiments. Cell lines were grown and maintained, respectively, in Dulbecco's modified Eagle's medium (DMEM) or DMEM containing nutrient mixture F12 at a ratio of 1:1 [(DMEM-HAMS-F12), Lonza, Walkersville, MD, USA], both supplemented with 10% fetal bovine serum (Lonza, Walkersville, MD, USA) and 1% penicillin-streptomycin (Lonza, Walkersville, MD, USA).The latter is an antibiotic mixture used as a preventative measure against Gram negative and positive bacterial organisms.For subcellular *in vivo* co-localisation (Section 2.16.3), the appropriate combination of expression constructs were co-transfected into the cardiomyocyte cell line, H9C2, whereas both cell lines were used to prepare cell lysates for use in *in vivo* co-immunoprecipitation (Co-IP) experiments (Section 2.16.2).

2.2. PRODUCTION OF *E.COLI* DH5α COMPETENT CELLS FOR BACTERIAL TRANSFORMATION

Fifty microlitres of an *E.coli* DH5α (-70°C) frozen glycerol stock was inoculated into 10ml LB-media without any selection antibiotics. The inoculated culture was subjected to an overnight incubation step at 37°C in a YIH Der model LM-530 shaking incubator (SCILAB instrument Company Ltd, Taipei, Taiwan) at 200rpm.

Following this incubation period, 1ml of this culture was inoculated into a 2L Erlenmeyer flask containing 200ml LB media (Appendix I), which was then incubated at room temperature for 24 hours, shaking at 70rpm to mid-log phase (OD_{600nm}=0.6) on a Labcon orbital shaker (Labcon Pty, Ltd, Maraisburg, RSA).

When spectrophotometer optical density readings of between 0.6-0.8nm were reached, the culture was split into 4x 50ml polypropylene tubes, which were centrifuged at 3000rpm for 15 minutes at 4°C in a Multitex centrifuge (MSE instruments, England). The supernatant was decanted and the pellet was resuspended in 8ml of ice-cold CAP buffer (Appendix I). Following this, the cells were re-pelleted by centrifugation at 3000rpm for 15 minutes at 4°C in a Multitex centrifuge (MSE instruments, England). After decanting the supernatant, the pellet was resuspended in 4ml of ice-cold CAP buffer (Appendix I) and transferred in 200µl aliquots into 2ml microfuge tubes. The cell samples were stored overnight on ice in a walk-in freezer at 4°C before it was stored at –70°C until they were required.

2.3. CULTURING OF HEK293 / H9C2 CELL LINES

2.3.1. Culture of H9C2 / HEK 293 cells from frozen stock

2.3.1.1. Thawing of cells

After long-term storage of frozen aliquots in a liquid nitrogen container at -196°C, frozen H9C2 or HEK 293 cells were rapidly thawed. The vial containing the frozen stock was immersed for 10 minutes in a 37°C waterbath (Memmert ®, Schwabach, Germany). Thereafter, the outside of the vial was immediately sterilized with 70% ethanol.

2.3.1.2. Removing DMSO from frozen stocks and culturing of cells

Frozen stocks contain dimethylsulfoxide (DMSO), a cryoprotective agent, which had to be removed in order to achieve maximum viability of the cells upon plating. One millilitre of prewarmed (37°C) growth media (Appendix I), was added to the thawed stock and mixed by gentle pipetting. The mixture was transferred to a 12ml Greiner tube (Greiner Bio-one, Frickenhausen, Germany), whereupon another 5ml growth media was added. The cells were centrifuged at 10000rpm for 1 minute using a Sorval® GLC-4 General Laboratory centrifuge (Separations Scientific, Johannesburg, South Africa), followed by the removal of the supernatant. The pellet was resuspended in 5ml growth media and the cells were re-centrifuged at 10000rpm for 1 minute. Subsequently the cells were resuspended in 10ml growth media and transferred into a T25 culture flask.

Lastly, the flask was gently swirled in order to distribute the cells evenly over the growth surface of the flask, whereupon it was incubated at 37°C in a Farma termosteri-cycle 5% carbon dioxide, humidified incubator (Farma International, Miami, Florida, USA).

2.3.2. Subculturing of H9C2 / HEK 293 cell cultures

H9C2 and HEK293 cells were subcultured every 2-4 days when 80%-90% confluency was reached. Firstly, the growth media was removed from the flask whereupon the adhering cells were washed with sterile phosphate buffered saline (PBS) lacking calcium and magnesium (Appendix I). The cells were treated with 2ml trypsin (0.5g/L Trypsin) (Lonza, Walkersville, MD, USA) to facilitate the detachment of cells from the growth surface of the flask. After 3 minutes, the cells were gently resuspended in 5ml of growth media, after which the cells were transferred into four flasks, each containing 10ml of growth medium.

2.4. TRANSFORMATION AND TRANSFECTION OF PLASMIDS INTO *E.COLI*, *S.CEREVISIAE*, AND MAMMALIAN CELLS

2.4.1. Bacterial plasmid Transformation (*E.coli*)

In preparation for the bacterial transformation, competent *E. coli* DH5α cells stored at -70^{o}C were thawed on ice for 20-30 minutes. To this, either 20ul of purified plasmid isolated from yeast or 3-5µl of ligation reaction was added and the mixture incubated on ice for a further 20-30 minutes. Following this step, the mixture was placed in a Lasec 102 circulating waterbath (Lasec laboratory and Scientific Company Pty Ltd, Cape Town, RSA) at 42^{o}C for 45 seconds and thereafter placed at room temperature for 2 minutes. Subsequently, 1ml of LB media was added to the mixture and incubated for one hour at 37^{o}C, shaking at 200rpm, in a YIH DER model LM-530 shaking incubator (SCILAB instrument Company Ltd, Taipei, Taiwan). After this incubation period, 200µl of the mixture was plated on LB agar plates supplemented with the appropriate selection antibiotic (Appendix I). The remaining mixture (800ul) was centrifuged at 14000rpm for 2 minutes in a Beckman Microfuge Lite (Beckman Instruments Inc, CA, USA); the supernatant was discarded and the pellet was resuspended in 200ul of LB media. This concentrated mixture was also plated on the appropriate LB agar plates, which were incubated, inverted, for 16 hours at 37^{o}C in a model 329 stationary CO2 incubator (Former Scientific, Marieta, Ohio U.S.A). The following morning, colonies were either picked for a midi preparation (Section 2.5.3) or a bacterial colony PCR (Section 2.6.4).

2.4.2. Yeast plasmid transformation (*S. cerevisae*)

Yeast was streaked from frozen stocks onto YPDA agar plates (Appendix I), which were then incubated upside down at 30^{o}C for 2-3 days in a Sanyo MIR262 stationary ventilated incubator (Sanyo, Electrical Company Ltd, Ora-Gun, Japan). Following the incubation period, a large colony or a number of smaller colonies of yeast were picked using an inoculation loop and resuspended in 1ml of sterile millipore H_2O in a sterile 2ml microfuge tube by vortexing. The cells were then re-pelleted in a benchtop Beckman Microfuge Lite centrifuge (Beckman Instruments Inc, CA, USA) at 13000rpm for 30 seconds. The supernatant was removed and resuspended in 1ml of 100mM lithium acetate (LiAc) (Appendix 1), after which it was

incubated for 5 minutes at 30°C in a Sanyo MIR262 stationary ventilated incubator (Sanyo, Electrical Company Ltd, Ora-Gun, Japan). The cells were re-pelleted by centrifugation at 13000rpm for 20 seconds in a benchtop Beckman Microfuge Lite centrifuge (Beckman Instruments Inc, CA, USA) after which all the LiAC was removed by pippeting. Next, 240µl of 50% Polyethylene glycol (PEG) (Appendix 1), 36µl 1M LiAc (Appendix 1), 25µl of 2mg/ml heat denatured and snap-cooled sonicated herring sperm DNA (Promega, Madison, WI, USA) and 10-20µl plasmid preparation and ddH_2O to a final volume of 350µl were added sequentially. The sample was then mixed by vortexing for at least 1 minute on a Snijders model 34524 press-to-mix (Snijders, Tilburg, Holland), thereafter the mixture was incubated at 42°C for 20-30 minutes in a Lasec 102 circulating water-bath (Lasec Laboratory and Scientific Company Pty Ltd, Cape Town, RSA). Following incubation, the cells were centrifuged in a benchtop Beckman Microfuge Lite centrifuge (Beckman Instruments Inc, CA, USA) at 13000rpm, after which all supernatant was removed and cells resuspended in 200µl sterile millipore ddH_2O. One hundred and fifty to two hundred microlitres of this sample was plated onto the appropriate selection plates, which contained SD^{-Trp} for the selection of the pGBKT7 clones, SD^{-Leu} for the pACT2 clones, and $SD^{-LeuTrp}$ for selection of diploid clones containing both plasmids. The plates were incubated upside down at 30°C for 2-5 days in a Sanyo MIR262 stationary ventilated incubator (Sanyo, Electrical Company Ltd, Ora-Gun, Japan).

2.4.3. Mammalian cells transfection (H9C2 and/or HEK 293)

2.4.3.1 Single and co-transfection of plasmid constructs

H9C2 cells were grown and maintained (Section 2.3.1.2) in $DMEM^{10\%FBS+1\%PS}$ (Lonza, Walkersville, MD, USA) for transfection of the appropriate combination of plasmid constructs for *in vivo* co-localisation assays (Section 2.16.3), and transfection of COMMD4 small interfering RNA (siRNA) for subsequent knockdown studies (Section 2.18). Both HEK 293 and H9C2 cells lines were maintained in $DMEM\text{-}HAMF12^{10\%FBS+1\%PS}$ and $DMEM^{10\%FBS+1\%PS}$ (Lonza, Walkersville, MD, USA), respectively, in order to prepare cell lysates for subsequent use in *in vivo* Co-IP experiments (Section 2.16.2).

Forty-eight hours prior to the transfection of cells, approximately 1-3 x 10^4 cells per well were seeded in 6-well tissue culture plates (Costar®, Corning Incorporated, Corning, NY, USA) containing the relevant growth media. For immunocytohistochemistry, cells were cultured on autoclaved coverslips placed inside the 6-well tissue culture plates. Plates were incubated at 37°C in a Farma-thermosteri-cycle 5% carbon dioxide humidified incubator (Farma, international, Miami, Florida, U.S.A). Two days later, the level of confluence was determined by visualizing the cells under a Nikon TMS light microscope (Nikon, Tokyo, Japan) and cells were transfected once approximately 80% confluence was reached. For each transfection, 100µl of serum-free medium was aliquoted into a sterile 1.5ml eppendorf tube and three microlitres GeneJuice® (EMD Biosciences, Darmstadt, Germany) was subsequently added. The mixture was vortexed and incubated for 5 minutes at room temperature.

For *in vivo* co-localisation using fluorescently-tagged proteins, a total of 1µg of the two plasmids (see Table 2.1 for transfection combinations and single transfections) was added to the Genejuice/serum-free medium mixture and mixed by gentle pipetting. Single transfection for the EYFP construct (COMMD4), Endo-DsRed-monomer-C1 construct (DsRed-Endo; kind gift of Dr. Craig Kinnear) and each of the pGFP^2-C3 constructs (SNX3, LGMN, ENO1, DSCR3, ANKRD1, ACTC1 and cMyBPC [last two plasmid constructs are kind gifts of Dr. Amsha Ramburan] were also performed and 1µg of each plasmid was used for these single transfections. For immunocytochemistry, 1µg of the EYFP-COMMD4 plasmid construct was transfected into H9C2 cells, since antibodies against COMMD4 are not commercially available. The GeneJuice/DNA/medium was incubated at room temperature for 15 minutes and the entire volume of the mixture was then added drop-wise to the cells in the growth media. The 6-well tissue culture plates were gently rocked back and forth in order to evenly distribute the drops across the surface of the plate. Next, the cells were incubated at 37°C, in a 5% carbon dioxide humidified incubator (Farma, International, Miami, Florida) for 48 hours before use in co-localisation assay (Section 2.16.3). In the case of H9C2 cells, the cells were allowed to differentiate into cardiac myotubes (Section 2.4.3.2), prior image acquisition.

Similarly, for *in vivo* Co-IP experiments, 1ug of plasmids (Table 2.2 for single and co-transfections) which includes EYFP-COMMD4 or DsRed-COMMD4, GFP-cMyBPC and DsRed-DSCR3 and were transfected into both the H9C2 and HEK293 cell lines.

Table 2.1. Single transfections and combinations for *in vivo* co-localisation assays

Transfection (1:1[#])	**EYFP- construct**	**GFP[2]- construct**	**DsRed construct**
Test 1	COMMD4	ENO1	-
Test 2	COMMD4	ANKRD1	-
Test 3	COMMD4	SNX3	-
Test 4	COMMD4	DSCR3	-
Test 5	COMMD4	ACTC1	-
Test 6	COMMD4	LGMN	-
Test 7	COMMD4	-	Endo-DsRed
Test 8	-	ENO1	Endo-DsRed
Test 9	-	ANKRD1	Endo-DsRed
Test 10	-	SNX3	Endo-DsRed
Test 11		LGMN	Endo-DsRed
Test 12	-	DSCR3	Endo-DsRed
Test 13	-	cMyBPC	Endo-DsRed
Control 1	COMMD4	-	-
Control 2	-	ENO1	-
Control 3	-	ANKRD1	-
Control 4	-	SNX3	-
Control 5	-	DSCR3	-
Control 6	-	ACTC1	-
Control 7	-	LGMN	-
Control 8	-	cMyBPC	-
Control 9	-	-	Endo-DsRed
Gene juice control	-	-	-
Untransfected control	-	-	-

Abbreviations: ACTC1 – cardiac actin, ANKRD – ankyrin repeat domain 1, cMyBPC –cardiac myosin binding protein C, COMMD4 – COMM domain containing protein 4, DSCR3 - Down Syndrome critical region 3, ENO1 - alpha-enolase, LGMN - Legumain, SNX3 – Sorting nexin 3, [#] indicate the ratio to a total of 1µg

Table 2.2. Single and co-transfections for *in vivo* co-immunoprecipitations experiments

Transfection (1:1[#])	EYFP- construct	DsRed-construct	GFP-construct
Test 1	COMMD4	-	-
Test 2	-	DSCR3	-
Test 3	COMMD4	DSCR3	-
Test 4	-	-	cMyBPC
Test 5	-	COMMD4	-
Test 6	-	COMMD4	cMyBPC
Gene juice control	-	-	-
Untransfected control	-	-	-

Abbreviations: COMMD4 – COMM domain containing protein 4, DSCR3 - Down Syndrome critical region 3, [#] indicate the ratio to a total of 1µg

2.4.3.2. Differentiation of H9C2 cells

In order to differentiate cardiac myocytes into myotubes, the growth medium was changed 48 hours after transfection. In brief, the growth media was removed, the cells were washed with sterile PBS (containing no calcium or magnesium) and 3ml of differentiating growth medium (Appendix I) was added. Subsequently, the cells were incubated at 37^0C in a Farma-thermosteri-cycle 5% carbon dioxide humidified incubator for seven days. On day four, the cells were washed again and fresh differentiating media was added. The cells were visualised from day four through to day seven, using an Olympus IX 81 motorised inverted microscope (Olympus, Hamburg, Germany) to monitor that the cells were differentiated. Afterwards the cells were imaged for co-localisation (Section 2.16.3).

2.5. DNA ISOLATION AND PURIFICATION

2.5.1. Bacterial plasmid purification

In order to purify plasmid DNA, a colony containing the plasmid of interest was added to 10ml of LB media (Appendix I), supplemented with the appropriate selection antibiotic (Kan [50mg/ml] or Amp [50mg/ml]), in a 50ml polypropylene tube. The culture was incubated overnight at 37^oC, shaking at 250rpm in a YIH DER model LM-530 shaking incubator (SCILAB instrument Co LTD., Taipei, Taiwan).

Following the overnight incubation step, the culture was centrifuged for 10 minutes at 3000rpm in a Beckman model TJ-6 centrifuge (Beckman Coulter, Scotland, UK). Afterwards the supernatant was discarded and the plasmid DNA was extracted using the GeneJET ^{TM} Plasmid Miniprep Kit (Fermentas, Burlington, Canada) according to the manufacturer's protocol. The purified plasmid DNA was stored at –20^oC prior to use in automated sequencing (Section 2.9) or bacterial or yeast transformation (Section 2.4.1 and 2.4.2). Conversely, the Wizard® Purefection Plasmid DNA purification kit (Promega Corp. Madison Wisconsin, USA) was used to isolate plasmid DNA, free of endotoxins, for subsequent use in transfections of H9C2 and HEK 293 cells (Section 2.4.3).

2.5.2. Plasmid isolation from yeast

Prey plasmids were isolated from yeast cells for subsequent *E.coli* transformation.Yeast cells containing the plasmid of interest were grown in 1ml of SD media containing the appropriate dropout supplement (BD Bioscience, Clontech, Paulo Alto, CA, USA). The yeast culture was incubated overnight at 30°C in a shaking incubator at 250rpm. The following morning, 4ml of YPDA (Appendix I) was added to the overnight culture, which was incubated for a further 4 hours at 30°C. The culture was then centrifuged at 13000rpm for 5 minutes in a Beckman model TJ-6 centrifuge (Beckman Coultier, Scotland, UK).

Afterwards, the supernatant was decanted and the pellet was resuspended in the residual medium and transferred to a sterile 2ml Eppendorf microfuge tube. To this, 200µl of Smash-and-Grab buffer (Appendix I), 200µl Phenol:Chloroform:Iso-amylalcohol (PCI) and 0.3g sterile 450µm glass beads were added. The contents were vortexed for 2.5 minutes using a Snijders model 34524 press-to-mix vortex (Snijders, Tilburg, Holland). In order to allow for phase separation, the tube containing the mixture was centrifuged (Beckman microfuge Lite, Beckman Instruments Inc., CA, USA) at 14000rpm for 5 minutes at room temperature. Lastly, the aqueous phase (upper clear plasmid-containing layer) was transferred into a new sterile 1.5ml eppendorf microfuge tube and purified as described in section 2.5.3.

2.5.3. Purification of PCR and plasmid products

PCR-amplified products as well as plasmid products were purified using the GenElute ™ PCR Clean-up kit (Sigma, Saint Louis, Missouri, USA) and the Zyppy™ Plasmid Miniprep kit (Zymo Research Corp., USA), respectively, according to the manufacturers' directions, in order to obtain purified products for subsequent sequencing, cloning or transformations reactions.

The purified PCR product or plasmid DNA was electrophoresed on a 2% or 1% agarose gel, respectively, for verification and stored at –20°C prior to use in the abovementioned experiments.

2.5.4. Gel purification of PCR-amplified product from agarose gels

PCR amplified products that were excised from agarose gels were cleaned using the Promega Wizard® SV Gel and PCR Clean-Up System according to manufacturer's protocols, prior to use in cloning procedures. The microcentrifuge tube containing the eluted DNA was either stored at 4°C or used in subsequent cloning steps.

2.6. POLYMERASE CHAIN REACTION (PCR)

2.6.1. Primer design

In order to facilitate PCR amplification of cDNA for construction of plasmids, gene-specific primers were designed using published sequence data (complementary DNA) obtained from the NCBI database (http://www.ncbi.nlm.nih.gov/Entrez).

Each primer sequence was subjected to self-complimentarity and primer-primer complimentarity tests as well as melting temperature compatibility tests, using the DNAman™ version 4 computer programme (Lynnion Biosoft Corp©). Synthesis of primers was carried out at the Synthetic DNA Laboratory, Department of Molecular and Cell Biology, University of Cape Town (UCT), Cape Town, South Africa.

2.6.1.1 Primers for generating bait and prey contructs

In order to generate plasmid constructs for both Y2H and co-localisation studies, gene-specific primers were designed from sequence data obtained from the NCBI database for COMMD4 and each of its putative interactors. Engineered restriction enzyme (RE) sites were incorporated into the primer design to allow for convenient subcloning into the respective vectors (Appendix VI). Both the forward and reverse primer were designed to contain a universal "seater" (tag) at their 5' end, to ensure successful restriction enzyme digestion at the adjacent restriction enzyme recognition site, followed by the gene-specific sequence for binding to the DNA template. The reverse primers were designed to additionally contain the relevant restriction site and a "stop" codon.

These RE sites were chosen based on their presence in the multiple cloning site (MCS) of the vector and their absence in the proposed insert. Table 2.3 shows the primer designs of the forward and reverse primers for COMMD4 and each prey as well as the respective restriction enzymes sites used for each insert.

Table 2.3. Primer sequences to generate bait and prey inserts for plasmid construction

Primer Name	Sequence 5'-3'	Vector	Product size (bp)	Ta°C
COMMD4	F-ATCGTATCGCATATGATGAGGTTCCGGTTCTGTGG	pGBKT-7	676	55
	R-ATCGTATCGGAATTCTTATACGGGCAGGGCGGCT			
ENO1	F – ATCGTATCGGATATCATGTCTATTCTCAAGATCCATGCC	pGFP-C3	1335	48
	R – ATCGTATCGCCGCGGTTGCTTGGCCAAGGGGTTTCTG			
LGMN	F – ATCGTATCGGAGCTCATGGTTTGGAAAGTAGCTGTATTC	pGFP-C3	1332	62
	R – ATCGTATCGGATATCTCAGTAGTGACCAAGGCACACG			
DSCR3	F – ATCGTATCGAGATCTATGGGGACCGCCCTGGAC	pGFP-C1	924	57
	R – ATCGTATCGGGTACCCTATATCCTGCAGAGCTTCAGC			
ANKRD 1	F - ATCGTATCG CTCGAGATGATGGTACTGAAAGTAGAGGAACTGG	pGFP-C3	2024	62
	R - ATCGTATCG GGATCC TCAGAATGTAGCTATGCGAG AGGTC			
SNX3	F – ATCGTATCGAGATCTTACAGCGAAATGGCGGAGACC	pGFP-C3	1526	70
	R - ATCGTATCG GGATCCTCAGGCATGTCTTATTTTAGATGG			
DSCR3	F – ATCGTATCGCTCGAGATGGGGACCGCCCTGGAC	pDsRed-C1	924	65
	R – ATCGTATCGGGTACCCTATATCCTGCAGAGCTTCAGC			

The black font represents the gene-specific sequences that anneals to the DNA template in the PCR reaction. The colored-coded sequences indicate the seater tag, restriction enzyme sites and stop codon. **Blue – universal seater, red - *Nde I* restriction enzyme site, green – *EcoRI* restriction enzyme site, *grey - EcoRV* restriction enzyme site, light blue – *Sac II* restriction enzyme site, orange – *Sac I* restriction enzyme site, turquoise – *BgI II* restriction enzyme site, purple – *Kpn I* restriction enzyme site, dark brown – *BamH I,* light green – *XhoI*, pink and underlined – stop codon**

Abbreviations: ANKRD – ankyrin repeat domain 1, COMMD4 – COMM domain containing protein 4, DSCR3 - Down Syndrome critical region 3, ENO1 - alpha-enolase, LGMN - Legumain, SNX3 – Sorting nexin 3, SNP - single nucleotide polymorphism, bp - base pair, Ta - Annealing temperature, °C - degrees Celcius.

2.6.1.2. Primers for Y2H insert screening and *in vitro* transcription and translation

Primers for the generation of products used in *in vitro* transcription and translation experiments and to amplify inserts cloned into Y2H cloning vectors, were designed using specific vector sequences flanking the multiple cloning site (MCS) of the pGBKT7 and pACT2 vectors (BD Bioscience, Clontech, Paulo Alto, CA, U.S.A) (Appendix VI), obtained from the Clontech™ Matchmaker™ vector handbook (www.clontech.com). The primer sequences and annealing temperatures are shown in Table 2.4.

Table 2.4. Primer sequences and annealing temperatures used for the amplification of inserts from Y2H cloning vectors and generation of products used in *in vitro* transcription and translation experiments

	Primer Name	Sequence 5'-3'	Ta (°C)
Y2H insert primers	pGBKT7-F	5`-TCATCGGAAGAGAGTAG-3	45
	pGBKT7	5`-TCACTTTAAAATTTGTATACA-3`	
	pACT2-F	5`-CTATTCGATGATGAAGATACCCCACCAAACC-3`	57
	pACT2-R	5`-GTGAACTTGCGGGGTTTTTCAGTATCTACGA-3`	
In vitro transcription and translation primers	BK-Myc-F	5`AAATAAAATTGTAATACGACTCACTATAGGGCGAGCC GCCACCATGGAGGAGCAGAAGCTGATGTCA-3`	49
	BK-R	5`-TCACTTTAAAATTTGTATACAC-3`	
	ADHA-F	5`AAATAAAATTGTAATACGACTCACTATAGGGCGAGCC GCCACCATGTACCCATACGACGTTCCAGAT-3`	42
	ADHA-R	5`-GGGGTTTTTCAGTATCTACGAT-3	

Abbreviations: ^{0}C, degrees Celsius; Ta, annealing temperature

2.6.2. High fidelity PCR of bait/prey inserts

High fidelity PCR was used to amplify the bait or various prey inserts from a heart cDNA library. The resulting PCR products for each bait and prey amplified were subsequently cloned into the relevant vector as set out in Table 2.3.

For the amplification of each bait or prey fragment, 50ng of a heart cDNA library (BD Bioscience, Clontech, Palo Alto, CA, U.S.A) was used as template in a 50ul reaction mix, which contained the following reagents: 150ng of each primer (Tables 2.2) 1.5µl of an equimolar dNTP (2.5mM of each, dATP, dCTP, dGTP and dTTP) solution (TaKaRa Shuzo Co.Ltd, Shiga, Japan), 5µl Ex Taq™ Mg^{2+}-containing 10x reaction buffer (TaKaRa Shuzo Co.Ltd, Shiga, Japan), 2U Ex Taq™ (TaKaRa Shuzo Co.Ltd, Shiga, Japan) and ddH_2O to a final volume of 50µl. Amplification was performed in a thermal cycler (Perkin–Elmer, Applied Biosystems Inc, Foster city CA, USA) using the following thermal cycle profile: a denaturing step of 94^{0}C for 2 minutes followed by 30 cycles of 94^{0}C for 30 seconds, annealing temperature (Table 2.3) for 30 seconds and 72^{0}C for 1 minute. A final extension step was performed for 5 minutes at 72^{0}C. PCR–amplified products were subsequently electrophoresed on a 1% agarose gel for verification (Section 2.7.1).

2.6.3. PCR amplification for *in vitro* transcription and translation

The pGBKT7-COMMD4 construct and putative prey clones in pACT2 vector were used as template for the PCR amplification of inserts for *in vitro* transcription and translation reactions, to generate radio-labeled proteins for *in vitro* Co-IP experiments. Appropriate primer pairs (BK-Myc-F/R and ADHA-F/R) (Table 2.4) were used; the amplification reaction mix and amplification conditions were the same as described in section 2.6.2. Amplified products were visualised by gel electrophoresis on a 1% agarose gel for verification (Section 2.7.1).

2.6.4. Bacterial colony PCR

In order to identify *E.coli* colonies that harbored the desired recombinant plasmid, and as the Y2H plasmids did not allow blue-white selection of recombinant plasmids, the presence of the insert was confirmed by performing a bacterial colony PCR.

A small quantity of a bacterial colony was used as template in the PCR reaction. Y2H vector-specific primers (Table 2.4) or insert primers (Table 2.3) were used in the PCR reaction mixtures using the same cycling conditions as described in Section 2.6.2. PCR–amplified products were subsequently verified on a 1% agarose gel.

2.7. GEL ELECTROPHORESIS

In this study, agarose gel electrophoresis was used to visualize PCR-amplified fragments, to verify plasmid preparations, in size separation of restriction enzyme digested PCR products and/or to excise DNA fragments for subsequent purification.

2.7.1 Agarose gel electrophoresis for the visualization of PCR amplified products and plasmid isolated from *E.coli*

In order to verify that PCR amplification was achieved, a 1-2% agarose gel (depending on size of amplified product) containing 5µl of ethidium bromide (10mg/ml) (Whitehead Scientific, RSA) was prepared for gel electrophoresis. Next, 8µl of each PCR amplified product was mixed with 1µl bromophenol blue loading dye (Appendix I), which was subsequently loaded into the wells of an agarose gel submerged in 1X SB (Appendix I) running buffer in an electrophoresis chamber. A 100bp DNA ladder (Promega, USA) was used as a size marker and was co-electrophoresed with the PCR-amplified products. Following the gel electrophoresis, the samples were visualized and photographed using the SYNGENE UV gel documentation system (SYNGENE, Synoptics Ltd Beacon House, Cambridge, UK).

Similarly, 1-5µl of plasmid DNA isolated from *E.coli* was electrophoresed in the same manner as mentioned above. Bacteriophage λ DNA (Promega, USA), digested with *Pst*I restriction enzyme (Promega, USA), was used as an approximate size marker and was co-electrophoresed with the plasmid DNA.

A 2.5% agarose gel was used for the size separation of restriction enzyme digested PCR products. In the case of some high fidelity PCR amplifications, many non-specific DNA bands were also visualized when gels were viewed. To overcome this problem, the specific amplified products were pooled together and electrophoresed on a 2% agarose gel with big, deep wells for excision of the correct amplified band.

2.7.2. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate (SDS) PAGE was used to separate *in vitro* transcribed/translated protein products (Section 2.16.1.1), as well as immunoprecipitated (IP) and co-immunoprecipitated (Co-IP) (Section 2.16.1 and 2.16.2) proteins on the basis of size.

Protein products of *in vitro* Co-IP experiments were resolved on a 20% SDS-polyacryalamide gel, after 15µl of SDS loading buffer (Appendix I) was mixed with 2µl of translated protein or 15µl of Co-IP reaction products. These samples were resolved at 230V for 4-6 hours in 1X SDS running buffer (Appendix I) and protein bands were detected via autoradiography (Section 2.8).

For *in vivo* Co-IP protein products, 40µl of 2X SDS sample application buffer (Appendix I) were mixed with ~60µl IP/Co-IP reaction products and denatured for 5 minutes at 95^{o}C. The protein G-agarose beads were pelleted by centrifugation at 2500rpm for 20 seconds and the supernatant then loaded on a 10-12% polyacrylamide gel containing 1% SDS along with a size marker (Spectra TM Multicolor Broad Range Protein Ladder, Fermentas). Electrophoresis was carried out at 100V in 1x SDS running buffer (Appendix I) for 1 hour. Afterwards, the separated proteins were transferred from the electrophoresis gel to a polyvinylidene difluoride (PVDF) membrane via electrophoretic transfer (Section 2.16.2.6).

2.8. AUTORADIOGRAPHY

Following SDS-PAGE (Section 2.7.2), gels on which *in vitro* Co-IP products were separated were transferred to Whatman 3M paper (Whatman International Lt, Maidstone, England) and subsequently heat and vacuum-dried on a Drygel Sr^{TM} slab gel drier (Hoeffer Scientific Instruments, San Francisco, CA, USA) for one hour. Afterwards the dried gel was exposed to autoradiography film Kodak (Eastman Kodak Company, Rochester, New York, USA) inside an autoradiography cassette, for 1-7 days (depending on the strength of the radioactive signal and concentration of proteins).

For Western blot analysis, the PVDF membrane containing the transferred proteins was exposed to the respective primary and secondary antibodies, after which the blots were detected with the enhanced chemiluminescence (ECL) kit (SuperSignal® West Pico Chemiluminescent Substrate, Thermo Scientific, Rockford, USA) according to manufacturer's recommendations; the subsequent light emitted from the bands on the membrane were captured on a autoradiography film [Kodak (Eastman Kodak Company, Rochester, New York, U.S.A)] inside an autoradiography cassette.

Afterwards, films were developed in a Hyperprocessor ™ automatic autoradiography film processor (Amersham pharmacia biotech U.K Ltd., Little Chalfont, Bucks, U.K).

2.9. AUTOMATED SEQUENCING

The purified DNA samples of PCR-amplified products as well as cloned inserts were bidirectionally sequenced, with PCR primers or vector-specific primers, respectively, on an ABI Prism ™ 3100 automated sequencer at the Central DNA Sequencer, University of Stellenbosch, Stellenbosch, South Africa.

2.10. DNA SEQUENCE ANALYSIS

The sequences obtained from the ABI automated sequencing were analysed using the ChromasPro computer program (Techelysium Pty limited, Helensvale, Queensland, Australia). This was done in order to ascertain that the reading frame of the bait and prey vectors, as well as the nucleotide sequences of the bait and prey inserts, were maintained during cloning.

2.11. Y2H PREY IDENTIFICATION

The sequences obtained from the prey constructs were analysed by comparing them to publicly available databases such as NCBI (www.ncbi.nlm.nih.gov/Entrez) and Ensembl (www.Ensembl.org) using the BLAST-search programs to assign identities to these putative interactors. The nucleotide sequence of the in-frame prey insert was also translated in DNAMAN followed by a BLASTP search of the translated sequence with reading frame dictated by the upstream GAL4 ORF, in order to determine whether the prey insert encoded for a known in-frame protein. Lastly, publicly available databases such as the ExPASy proteomics server (http://au.expasy.org/), GENATLAS database (http://www.dsi.univ-paris5.fr/genatlas/) or the Bioinformatic Harvester database (http://harvester.fzk.de/harvester/) were used to determined the subcellular localisation, as well as function of the prey proteins.

2.12. RESTRICTION ENZYME DIGESTION

2.12.1. Restriction enzyme digestion for cloning

2.12.1.1 Digest for cloning bait and prey inserts into relevant vectors

To generate fusion protein expression constructs encoding the insert of interest, DNA fragments with in-frame restriction enzyme sites at both ends were generated using PCR amplification (Table 2.3, Section 2.6.1.1). PCR-generated fragments along with the appropriate vectors were subjected to double digestions with the relevant restriction enzymes (Table 2.5). This included the COMMD4 bait insert into pGBKT7, prey inserts into pGFP-C3 or pGFP-C1 for in *vivo* co-localisation and both COMMD4 and the DSCR3 prey insert into pDsRed-C1 for *in vivo* Co-IP experiments (Section 2.13.1).

As an example, double digestions for the COMMD4 bait insert and pGBKT7 vector is described below: The first digest was prepared in a 50µl cocktail mix containing: 30µl COMMD4 PCR product or 10µl of vector DNA, 2-3U of the *Nde I* restriction enzyme in the appropriate buffer supplied by the manufacturer (Fermentas, Burlington, Canada) and the appropriate volume water. The mixture was incubated for 3 hours at 37°C and then heat-inactivated for 15 minutes at 65°C. After the first digestion, the mixture was purified (Section 2.5.3) and eluted in 30µl sterile water prior to digestion with the second enzyme.

For the second digestion using the restriction enzyme, *EcoRI*, a 50µl cocktail mix was prepared containing 30µl purified *NdeI*-digested PCR product or 10µl of *NdeI*-digested vector DNA, 2-3U of the *EcoRI* restriction enzyme in the appropriate buffer (Fermentas, Burlington, Canada) and the appropriate volume of water. This mixture was also incubated for 3 hours at 37°C, then heat-inactivated for 15 minutes at 65°C, after which the DNA was re-purified (Section 2.5.3). Both the insert and vector were subsequently electrophoresed on a 1% agarose gel in order to ascertain that the DNA was recovered after cleanup. The double-digested vector was also treated with *Calf Intestinal Alkaline Phophatase (CIAP)* (Section 2.13.2), prior to use in the ligation reaction to minimize vector self ligation, whereas double digested insert was directly used in ligation reactions.

Double-digestions was also carried out in the same manner as described above for each prey insert and the relevant vector used; the restriction enzymes used and conditions are set out in Table 2.5.

Table 2.5 Restriction enzymes and digestion conditions used for the cloning of bait (bold font) and prey inserts into the relevant vectors.

Preys	Vector used	5' RE	3' RE	Incubation time
ANKRD1	pGFP^2 -C1	*XhoI*	*BamHI*	37^{o}C
SNX3	pGFP^2 -C1	*BgIII*	*BamHI*	37^{o}C
ENO1	pGFP^{2-}C3	*EcoRV*	*SacII*	37^{o}C
LGMN	pGFP^{2-}C3	*SacI*	*EcoRV*	37^{o}C
DSCR3	pGFP^2 -C1	*BgIII*	*KpnI*	37^{o}C
DSCR3	pDsRed monomer–C1	*XhoI*	*KpnI*	37^{o}C
COMMD4	pDsRed monomer–C1	*HindIII*	*BamHI*	37^{o}C
COMMD4	pGBKT7	*NdeI*	*EcoRI*	37^{o}C

Abbreviations: ANKRD – ankyrin repeat domain 1, COMMD4 – COMM domain containing protein 4, DSCR3 - Down Syndrome critical region 3, ENO1 - alpha-enolase, LGMN - Legumain, SNX3 – Sorting nexin 3, RE – restriction enzyme

2.13. GENERATION OF BAIT AND PREY CONSTRUCTS

2.13.1. Generation of plasmid constructs

Plasmid constructs were generated for use in Y2H analysis, *in vivo* co-localisation analysis and for *in vivo* Co-IP experiments.

In order to generate an expression vector encoding COMMD4, the Y2H bait-insert was cloned into the pGBKT7 shuttle-vector (Appendix VI). The COMMD4 sequence integrity, cloning sites integrity and conservation of the GAL4-DNA-BD reading frame were verified by automated sequencing and the bait-construct was subsequently transformed into the yeast strain AH109. This transformed yeast strain was then used to screen a CLONTECH MATCHMAKER pre-transformed human cardiac cDNA library, comprising of cardiac cDNAs fused into the pACT2 prey-vector and transformed into the yeast strain Y187.

For subcelullar co-localisation, each putative prey interactor chosen for further analysis was cloned into a pGFP^2-C3 or pGFP^2-C1 vector, to construct a green fluorescent protein (GFP) fusion. After the verification of sequence integrity, cloning site integrity and conservation of the pGFP^2-C3 or pGFP^2-C1 reading frame by automated sequencing, the prey constructs were co-transfected with a construct encoding enhanced yellow fluorescent protein (EYFP)-tagged COMMD4 (purchased from the German Resource Center for Genome Research (RZPD, Berlin, Germany) into cardiomyocytes cells H9C2 for subsequent fluorescence microscopy.

The putative prey interactor, DSCR3 and COMMD4, was cloned into the shuttle vector, DsRed-Monomer-C1 (Appendix VI), in order to generate a plasmid constructs (DsRed-DSCR3 and DsRed-COMMD4) for use in subsequent *in vivo* Co-IP experiments. After verification of sequence integrity and conservation of the

reading frame by automated sequencing, the construct was transfected into HEK 293 and/or H9C2 cells for subsequent pull-down assays.

2.13.2. Calf intestinal alkaline phosphatase (CIP) treatment of double digested vectors

Following the final restriction enzyme digestion step, the ends of the linearized plasmid were subjected to a CIP-treatment to remove the phosphate groups. This was done to prevent vectors from re-circularising by self–ligation. The CIP-treatment was carried out by mixing 50µl of the digested vector with 1µl of CIP (Promega, Madison, WI, USA), 10µl of CIP buffer (Promega, Madison, WI, USA) and 38µl ddH_2O. The sample was then incubated for 30 minutes at 37^{o}C, after which another 2µl of CIP was sequentially added and the mixture was incubated for an additional 30 minutes. Following this, 2µl 0.5M EDTA (Appendix I) was added to the mixture, after which the mixture was incubated for 20 minutes at 65^{o}C to inactivate the enzyme. The vector was subsequently purified as described in section 2.5.3.

2.13.3. DNA ligation

After restriction enzyme digestions of bait and prey inserts and the CIP treatment of vectors, DNA ligations were carried out in order to generate Y2H bait constructs for Y2H analysis and prey constructs for use in subsequent co-localisation analysis. For the ligation, 2µl of the double-digested purified inserts was added to 1µl CIP-treated vector. Five microlitres 2X T4 ligase buffer (Promega, Madison, WI, USA), 5U T4 DNA ligase and ddH_2O to a final volume of 10µl was sequentially added, after which the mixture was incubated for 16 hours at 4^{o}C. After the incubation period, 5µl of the sample was transformed into the bacterial strain DH5α (Section 2.4.1).

2.14. QUALITY CONTROL TESTING OF THE Y2H CONSTRUCTS

2.14.1. Phenotypic verification of yeast strains AH109 and Y187

The phenotypes of each yeast strain (AH109 and Y187) used for Y2H analysis were first verified before use. This was done in order to control for contamination, the optimal viability of each strain and their nutritional requirements according to their genotypes.

To test for the nutritional requirements, 3-4 colonies grown from frozen stocks on YPDA plates were streaked onto all appropriate SD agar plates and incubated at 30^{o}C for 4-6 days. Only non-transformed yeast cells that were unable to grow on SD^{-Leu}, SD^{-Ade}, SD^{-Trp}, SD^{-His}, but which grew on $SD^{-Ura,}$ were chosen for further use in the transformation with the appropriate vectors. After the transformation of the bait construct (pGBKT7-COMMD4) into yeast strain AH109, the transformed yeast were again plated onto these respective SD agar plates in order to assess whether the transformed bait construct was able to autonomously activate the transcription of the reporter genes and that transformation had not changed the yeast phenotype inappropriately.

2.14.2. Toxicity test of the transformed cells

To ensure that the yeast strain AH109 was compatible with the bait construct before proceeding with the library screen, the toxicity of the bait construct on AH109 was first assessed. Growth curves of cells transformed with a non-recombinant pGBKT7 (control–AH109-pGBKT7) and cells transformed with the bait construct (AH109- pGBKT7-COMMD4) were generated in liquid cultures. These growth curves were simultaneously set up under the same experimental conditions.

Each of the transformed yeast strains was grown to a stationary phase in a 1ml tryptophan deficient (SD^{-Trp}) liquid culture prepared in a 2ml microfuge tube. The liquid cultures were incubated at 30^{o}C in a YIH DER model LM-530 shaking incubator (SCILAB instrument Co LTD., Taipei, Taiwan) for 24 hours shaking at 150-200rpm. Following the incubation period, the 1ml overnight liquid cultures for each of the transformed yeast strain were transferred to separate 250ml Erlenmeyer flasks containing 50ml of SD^{-Trp} medium and incubated for an additional 24 hours at 30^{o}C, shaking at 150-200 rpm in a YIH DER model LM-530 shaking incubator (SCILAB instrument Co LTD., Taipei, Taiwan). During this incubation period, a 1ml aliquot of each culture was taken every two hours over a period of eight hours, as well as an overnight (24hr) aliquot, and the OD_{600nm} of each aliquot was measured in order to generate the growth curves. The measurement of these aliquots at an absorbance of 600nm was plotted against time and represented as log-linear curves using the GraphPad Prism version 4 computer software programme. The slopes of the graphs generated for the recombinant and non-recombinant transformants were then compared to assess whether there were significant growth differences. If so, this would indicate that the bait construct might have an effect on the growth of the yeast strain, which could possibly hinder further Y2H steps.

2.14.3. Establishment of mating efficiency

In order to establish whether the bait construct affected the mating efficiency of the yeast strain AH109, small scale yeast matings were performed. The yeast strain AH109 transformed with the pGBKT7-COMMD4 bait construct was mated with the prey host strain, Y187 transformed with the control prey vector, pTD1.1, supplied by the manufacturer (BD Bioscience, Clontech, Paulo Alto, CA, U.S.A). Control matings were performed concurrently, in which yeast strain AH109 transformed with the control bait vector, pGBKT7-53 (encoding a DNA-BD/murine p53 fusion protein) was mated with pTD1.1-Y187 (both control vectors supplied by manufacturer, BD Bioscience, Clontech, Paulo Alto, CA, U.S.A).

First, each of the yeast strains was plated onto the relevant SD agar plates (AH109-pGBKT7-COMMD4, AH109-pGBKT7-53) on SD^{-Trp} plates; (pTD1.1-Y187) on SD^{-Leu} plates), which were incubated for 2-5 days in a Sanyo MIR262 stationary gravity-ventilated incubator (Sanyo, Electronic Company Ltd, Ora-Gun, Japan). A single colony from each type was picked and used in the small-scale mating experiments. The respective bait strain colonies (AH109-pGBKT7-COMMD4 or AH109-pGBKT7-53) and control prey strain colonies (pTD1.1-Y187) were inoculated together in 1ml YPDA media (Appendix I) in a 2ml microfuge tube. These mating cultures were incubated overnight at 30^{o}C, shaking at 200rpm in a YIH DER model LM-

530 shaking incubator (SCILAB instrument Co LTD., Taipei, Taiwan). Following the overnight incubation, serial dilutions (1:10; 1:100; 1:1000 and 1:10000) of the mating cultures were made and plated onto SD^{-Trp}, SD^{-Leu} and $SD^{-Leu\ -Trp}$ plates and incubated for 4-5 days at 30°C in a Sanyo MIR262 stationary gravity-ventilated incubator. The number of colonies on each plate was then scored and used to calculate the mating efficiency (Appendix II).

2.15 YEAST-TWO-HYBRID ANALYSIS

2.15.1. Cardiac cDNA library

A pre-transformed human MATCHMAKER cardiac cDNA library (BD Bioscience, Clontech, Palo Alto, CA, U.S.A) which consists of *S. cerevisiae* Y187 transformed with a cardiac cDNA library constructed in Y2H prey-vector pACT2, was used in the Y2H library assay. The mRNA source for the library was pooled from the normal whole hearts of three male caucasians; ages 28-47. The library contained approximately 3.5 x 10^6 independent clones inserted in pACT2 through the *EcoR1* and *Xho1* restriction enzyme sites. Xho 1-$(dT)_{15}$ priming had been used to generate cDNAs from mRNA. The average insert size was 2kb, with a range between 0.4 and 4.0kb (pre-transformed human heart MATCHMAKER cDNA library product analysis certificate, BD Bioscience, Clontech, Palo Alto, CA, U.S.A).

2.15.2. Establishment of bait culture

The yeast strain AH109 transformed with the bait construct was streaked onto SD^{-Trp} plates and incubated for 4 days at 30°C. Following this, an overnight culture was prepared by inoculating a colony in 1ml of SD^{-Trp} media (Appendix I), after which it was incubated at 30°C while shaking at 200rpm in a YIH DER model LM-530 shaking incubator (SCILAB instrument CO. Ltd, Taipei, Taiwan). The following morning, 250µl of the 1ml overnight culture was transferred to four separate 250ml Erlenmeyer flasks, each containing 50ml of SD^{-Trp} media and subsequently incubated at 30°C while shaking at 200rpm. Four separate bait cultures were produced, which, when pooled together, allowed for the generation of a final bait culture with a titre of $1x10^{10}$ (Section 2.15.4), in order to achieve the excess of bait strain required for a high mating efficiency with the particular library. After overnight incubation, the cultures were transferred to separate 50ml polypropylene tubes, centrifuged at 3000rpm for 10 minutes at room temperature in a Beckman model TJ-6 centrifuge (Beckman Coulter, Scotland, UK). The supernatants were discarded and the pellets were resuspended and pooled together in a new 50ml polypropylene tube. The titer of the bait culture was estimated by measuring the OD_{600nm} of a 1ml aliquot of the bait culture and this estimation was further confirmed by means of a haemocytometric cell count (Appendix II).

2.15.3. Library mating

Prior to use, a 1ml frozen aliquot of the pre-transformed cardiac cDNA library (BD Bioscience, Clontech, Paulo Alto, CA, U.S.A) was removed from the -70°C freezer and thawed at room temperature. Afterwards, the library aliquot was gently vortexed using a Snijders model 34524 press-to-mix vortex (Snijders

Scientific, Tilburg, Holland) and a 10µl of the library mix was aliquoted in a sterile 1.5ml microfuge tube for subsequent library titering. The bait culture (Section 2.15.2) was resuspended in 45ml 2X YPDA media (Appendix I) supplemented with 10µg/µl Kan in a 2L Erlenmeyer flask and the remaining library culture (990µl) was added to this mixture. This mating culture was incubated at 30°C, gently swirling at 50rpm in a YIR DER model LM-530 shaking incubator (SCILAB instrument CO. Ltd, Taipei, Taiwan). After the overnight mating, the mating mix was transferred into a sterile 50ml polypropylene tube and the cells pelleted by centrifugating at 3000rpm for 10 minutes in a Beckman model TJ-6 centrifuge (Beckman Coulter, Scotland, UK), after which the supernatant was removed. The Erlenmeyer flask used for library mating was rinsed twice with 40ml 2X YPDA supplemented with 10ng/µl Kan. For each rinse, 2X YPDA media was used to resuspend the bait yeast pellet, which was then re-pelleted by centrifuging at 3000rpm for 10 minutes at room temperature in a YIR DER model LM-530 shaking incubator. After the final centrifugating step, the supernatant was discarded and the pellet was resuspended in 10 ml 0.5X YPDA supplemented with 10ng/µ Kan and the total volume of cells and media was measured and noted for subsequent calculations (Appendix II)

Hundred microlitre aliquots of a serial dilution of the mating mix (1:10; 1:100; 1:1000 and 1:10000) were plated onto 90mm SD^{-Trp}, SD^{-Leu} and $SD^{-Leu\ -Trp}$ agar plates in order to determine the bait:library mating efficiency. The remaining mating mix was plated onto 140mm TDO (SD media lacking, trp, leu and his [Appendix I]) plates in 200µl aliquots. The TDO plates were incubated inverted at 30°C for two weeks in a Sanyo MIR stationary ventilated incubator.

2.15.4. Establishing the library titre

The serial dilutions of the library culture (Section 2.15.3) was plated onto SD^{-Leu} agar plates (Appendix I) and incubated for 4 days at 30°C in a Sanyo MIR262 stationary ventilated incubator (Sanyo, Electronic Company Ltd, Ora-Gun, Japan). After the incubation period, colony counts were performed, in order to calculate the number of library plasmids screened (Appendix II).

2.15.5. Detection and measurement of reporter gene activation

2.15.5.1. Selection of diploid yeast colonies containing putative interactor peptides

In order to select for diploid yeast colonies expressing interacting proteins, colonies were plated onto TDO and Quadruple dropout (QDO) (media lacking leu, his, trp and ade) plates (Appendix I). Growth of yeast cells on TDO plates indicated the transcriptional activation of the *HIS3* nutritional reporter gene, whereas growth on QDO plates indicated transcriptional activation of the *HIS3* and *ADE2* nutritional reporter genes. The activation of these nutritional genes in the diploid yeast cells is indicative of an interaction between the COMMD4 bait and its respective preys.

After plating library mating mix on TDO plates as described in section 2.15.3, diploid cells were allowed to form visible colonies. TDO plates were observed every 4 days and full-grown colonies were picked and re-

streaked onto TDO and QDO plates. These plates were incubated for 3-6 days at 30°C and colonies growing on QDO plates were subsequently picked and re-streaked on QDO plates containing X-α-galactose in order to detect the activation of the colourimetric reporter gene, *MEL1* (Section 2.15.5.2).

2.15.5.2. Detection of colourimetric reporter gene activation

The *MEL1* gene, which encodes the enzyme α-galactosidase, is a colourimetric reporter gene that can be assayed directly on X-α-Gal indicator plates, which employs blue/white screening. Thus, selected diploid colonies were assessed for their ability to activate the *MEL1* reporter gene by assessing the intensity of the blue coloring of the yeast colonies, which is proportional to the strength of interaction between the bait and prey.

In brief, yeast colonies in which the *HIS3* and *ADE2* reporter genes had been activated, as determined by their growth on QDO agar plates (Section 2.15.5.1), were replicated from QDO plates onto Hybond N^+ nylon membranes (Amersham pharmacia biotech Ltd, England). Subsequently, the membranes were placed colony-side up onto a QDO plate which had been spread with 200µl of X-α-Gal solution (20mg/ml) (BD Bioscience, Clontech, Palo Alto, CA, USA). The plates were incubated at 30°C for 16-48hr in a Sanyo MIR262 stationary ventilated incubator (Sanyo, Electronic Company Ltd, Ora-Gun, Japan). Following the incubation period, the activation of the *MEL1* reporter gene was assessed and scored semi-quantitatively based on the intensity of the blue coloring of the yeast colonies.

2.15.6. Rescuing prey plasmids from diploid clones

In order to identify the interacting protein, each of the putative prey plasmid clones which had shown transcription of the *HIS3*, *ADE2* and *MEL1* reporter genes needed to be isolated from the diploid yeast colonies. The plasmid DNA for each of the diploid cells was isolated using the protocol described in section 2.5.2 and subsequently transformed into *E.coli* strain DH5α (Section 2.4.1). The transformants were inoculated into LB^{Amp} liquid media (Appendix I) allowing for the growth of transformants which only contain prey plasmids. Afterwards, these prey plasmids were transformed into yeast strain Y187 for subsequent heterologous bait matings (Section 2.15.7).

2.15.7. Interaction specificity test i.e. heterologous bait mating

Following the nutritional and colourimetric selection screens, Y187 colonies expressing the putative prey interactors were subjected to matings with AH109 transformed with heterologous baits to establish whether interaction with COMMD4 was specific, and to exclude non-specific bait and prey interactions. Briefly, each prey plasmid was re-transformed into yeast strain Y187 and these prey strains were individually mated with yeast strain AH109, which had been transformed with either of four heterologous baits (pGBKT7-COMMD4, pGBKT7-Reeler (encoding the reeler domain of the brain-expressed reelin protein, a kind gift of

Dr Craig Kinnear), non-recombinant pGBKT7, and pGBKT7-53 control bait [both supplied by manufacturer, BD Bioscience, Clontech, Palo Alto, CA, USA]).

Following the selection of diploid clones from $SD^{-Leu-Trp}$ plates, these clones were streaked onto TDO and QDO selection plates and incubated at 30°C in a Sanyo MIR262 stationary ventilated incubator (Sanyo, Electronic Company Ltd, Ora-Gun, Japan) in order to assess the activation of nutritional reporter genes, thereby testing whether interactions occurred non-specifically with these heterologous baits or only, specifically, with the COMMD4-bait construct. Those prey clones which specifically interacted with the COMMD4-bait construct were considered putative true interactors and the inserts of the relevant prey plasmids were subsequently nucleotide sequenced to determine their identities.

2.16. PROTEIN-PROTEIN INTERACTIONS OF COMMD4

After the identification of putative interactors by means of the Y2H experiments, the interactions were verified using additional, independent approaches which included *in vitro* Co-IP analysis, *in vivo* Co-IP followed by western blot (WB) analysis and live-cell 3D co-localisation with Z-stacking.

2.16.1. *IN VITRO* CO-IMMUNOPRECIPITATIONS

2.16.1.1. Transcription and translation of baits and preys

To reduce the possibility of RNase contamination, pipette tips and microcentrifuge tubes certified RNase free by the manufacturer (Porex, Fairburn, Georgia, USA) were used and all surfaces and instrumentation were wiped thoroughly with RNase Zap wipes (Ambion Inc, Austin, Texas, USA) prior to use in transcription/translation and Co-IP experiments.

Inserts of the bait and putative interactor plasmids were PCR-amplified using specific primers (Table 2.4, Section 2.6.1.2) in order to generate PCR fragments comprising of the relevant insert linked to an HA-(for prey inserts) and/or Myc– (for bait inserts) antibody epitope-encoding sequence, as well as a T7 promoter sequence, to facilitate *in vitro* transcription. The bait and prey PCR fragments were subsequently transcribed and translated in the presence of S^{35}-labeled methionine using TNT® Quick Coupled Transcription/Translation system kit (Promega Corp.), according to the manufacturer's instructions.

After translation, 2µl of the reaction mixture was added to 15µl SDS loading dye (Appendix I), resolved by SDS-PAGE (Section 2.7.2) and *in vitro* translated proteins were subsequently visualised by autoradiography (Section 2.8).

2.16.1.2. *In vitro* Co-IP of translated PCR products

Translated bait and prey fusion peptides were co-immunoprecipitated to assess the protein interactions identified by Y2H analysis. In brief, 5µl bait and 5µl prey were mixed in a 1.5ml microcentrifuge tube and incubated at room temperature for 1 hour. Next, 1µl Myc antibody (5µg/ml) [Roche Biosciences, Palo Alta,

CA, USA] was added to the mixture and samples were further incubated for 1 hour at room temperature. Following incubation, 10µl pre-washed protein G agarose (Appendix I) (Kirkegaard and Perry Laboratories, Gaithersburg, ML, USA) and 135µl Co-IP buffer (Appendix I) were added to each mixture. Samples were subsequently rotated on a Labnet rotor (Labnet Inc, NJ, USA) at 10rpm at 4^0C for 1 hour and after which they were washed five times with 0.1% Tween Tris-buffered saline (TBST) (Appendix I).

Single IP experiments for the bait, using the Myc antibody (Roche Biosciences), and for each of the putative preys, using the HA antibody (Roche Biosciences), were carried out along with the Co-IP experiments, to serve as controls. Briefly, 5µl of the bait was mixed with 1µl Myc antibody in a 1.5ml microcentrifuge tube and incubated at room temperature for 1 hour. Conversely, 5µl of each putative prey interactor was incubated with 5µl HA antibody (Roche Biosciences) in separate 1.5ml microcentrifuge tubes. Additionally, a protein G agarose control was included to assess whether the protein G itself was interacting non-specifically with the prey protein; for this reaction, 5µl of each prey protein was incubated in the absence of an antibody and incubated at room temperature for 1 hour. Next, 10µl pre-washed protein G agarose and 135µl Co-IP was added to each tube and the samples were rotated on a Labnet rotor (Labnet Inc, NJ, USA) at 10rpm at 4^0C for 1 hour and subsequently washed five times with TBST. Prior to SDS-PAGE analysis, 15µl SDS loading dye was added to each sample; samples were then incubated at 95^0C for 5 minutes after which they resolved on a 15-20% SDS polyacrylamide gel (Section 2.7.2) and visualised by autoradiography (Section 2.8).

2.16.2. *IN VIVO* CO-IMMUNOPRECIPITATIONS

For the *in vivo* Co-IP's and IP experiments, endogenous proteins were precipitated using the relevant antibody specific to the protein of interest. In the case of COMMD4 and DSCR3, specific antibodies were not commercially available for each of these proteins, so the plasmid constructs encoding the YFP- and DsRed-tagged versions of these proteins, respectively, were transfected into appropriate cell lines. Thus, these proteins could be precipitated via their tags, for which antibodies are commercially available.

2.16.2.1. Cell lysis and sample preparation of monolayer cells

Cells used in *in vivo* Co-IP, IP experiments and WB analysis were cultured and transfected as described in section 2.4.3. After growing cells to confluency in either a 100mm x 35mm 6-well tissue culture plate (Costar®, Corning Incorporated, Corning, NY, USA) or a 145mm petri dish; the monolayer cells were dislodge with a cell scraper and the resulting cell mixture was transferred to a 50ml polypropylene tube. Cells were subsequently pelleted at a 1000rpm for 1 minute in a Multitex centrifuge (MSE instruments, England) followed by the removal of supernatant. While the cells were being centrifuged, the 6-well tissue culture plate or the 145mm petri dish was washed with 1X PBS (Appendix I), which was also used for the resuspension of the pellet, after removal of the supernatant. The resuspension mix was transferred to 1.5ml microcentrifuge tube and centrifuged at a 1000rpm for 1 minute in a benchtop Beckman Microfuge Lite centrifuge (Beckman Instruments Inc, CA, USA). After discarding the supernatant, the resultant pellet was lysed on ice for 30 minutes with the addition of 350µl of ice cold 1X Lysis Buffer (LyB) supplemented with complete protease inhibitor cocktail tablets (Roche) and 50mM phenyl-methyl-sulphonyl-fluoride (PMSF) (Appendix I). Afterwards, the supernatant, representing the whole cell lysate was transferred to a new 1.5ml microcentrifuge tube for further usag-e. Prior to use in subsequent *in vivo* Co-IP, IP and WB experiments, the protein concentration of the cell lysates were determined via Bradford assay (Section 2.16.2.2).

2.16.2.2. Bradford Assay for determination of protein concentrations

The protein concentration of cell lysates was first determined via the Bradford assay in order to facilitate use of lysates with equal protein concentrations in WB experiments. BSA standards ranging from 0-1000ug/ml was prepared from 10mg/ml BSA stock (New England, Biolabs) (Appendix I) and 10µl of each BSA standard was pipetted in duplicate in a 96-well plate. Volumes of 1-, 2- and 5µl lysate samples were also pipetted in duplicate per well along with a duplicated lysis buffer control. Next, 200µl Bradford reagent (Appendix I) were added to BSA standards and lysate samples and the 96-well plate was read at 595nm absorbance on a Synergy HT luminometer (BioTek Instruments Inc., Vermont, USA) using KC4 software. The mean concentration was used to determine the amount of lysate required to obtain 50-100µg protein per well on an SDS-PAGE gel.

2.16.2.3. Optimisation of optimal antibody dilution

Antibodies against SNX3 (goat, polyclonal), HA (mouse, monoclonal), ANKRD1 (goat, monoclonal), ACTC1 (mouse, monoclonal), cMyBPC (K16-mouse, monoclonal) and HRP- and TxRed-conjugated secondary antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA), while antibodies against ENO1 (mouse, polyclonal), LGMN (mouse, polyclonal) and FBXL10 (mouse, polyclonal) were purchased from ABNOVA (Taiwan). Antibodies against the YFP/GFP (JL8-mouse, monoclonal) and DsRed (mouse, monoclonal) tags were purchased from CLONTECH. Prior to the *in vivo* IP and Co-IP reactions, the optimal dilution of each of the antibodies was first experimentally determined ranging from 1/100-1/500,000 dilutions (Table 2.6), while at the same time ascertaining that the antibody both pulled down and detected the particular protein.

2.16.2.4. Pre-clearing with protein G agarose

Prior to IP and Co-IP reactions, a pre-clearing step was performed in order to reduce background noise caused by non-specific adsorption of cellular proteins to protein G-agarose. Briefly, 20µl of protein G-agarose slurry (Kirkegaard and Perry Laboratories, Gaithersburg, ML, USA) was added to 200µg of whole cell lysate made up to 200µl with LyB supplemented with protease inhibitors cocktails and PMSF. The mixture was rotated on a Labnet rotor (Labnet Inc, NJ, USA) at 10rpm for 2 hours at 4^{o}C. Afterwards, the agarose beads were pelleted by centrifugation at 7000rpm for 30 minute at 4^{o}C, and the supernatant was subsequently transferred to a fresh 1.5ml microcentrifuge tube.

Table 2.6. Characteristics of the relevant antibodies used in in vivo IP, Co-IP and WB analysis

Target Protein	Antibody	Primary Ab dilution range	Secondary Ab dilution range	Expected size (kDa)	Company
SNX3	goat/poly	1:1000	1:5000	21	Santa-Cruz (sc-10619)
ENO1	mouse/poly	1:1000	1:5000	47	Abnova (H00002023-A01)
LGMN	mouse/poly	1:500	1:5000	49	Abnova (H00005641-B01)
ANKRD1(CARP)	goat/mono	1:1000	1:5000	36	Santa-Cruz (sc-23251)
FBXL10	mouse/poly	1:1000	1:5000	37	Abnova (H00084678-A01)
ACTC1	mouse/mono	1:1000	1:5000	42	Santa-Cruz (sc-32251)
EYFP# -COMMD4	mouse/mono	1:1000	1:5000	21 + 30~51	Clontech (JL8 -632381)
RFP#-COMMD4	mouse/mono	1:1000	1:5000	21 + 29~50	Clontech (632392)
RFP#-DSCR3-	mouse/mono	1:1000	1:5000	33 + 29~62	Clontech (632392)
GFP#-cMyBPC	mouse/mono	1:1000	1:5000	140 +30~170	Clontech (JL8 -632381)
HA	mouse/mono	1:1000	1:5000	-	Santa-Cruz (sc7392-)
β-Actin*	rabbit/mono	1:2500	1:5000	45	Cell Signalling Technologies (#4970)

Abbreviations: Ab: antibody, ACTC1 – cardiac actin, ANKRD – ankyrin repeat domain 1, COMMD4 – COMM domain containing protein 4, DSCR3 - Down Syndrome critical region 3, ENO1 - alpha-enolase, FBXL10 – F-box and leucine-rich repeat protein 10, GNB2L1, kDa –kilodalton, LGMN - Legumain, mono- monoclonal, cMyPBC – cardiac myosin binding protein C, poly-polyclonal, SNX3 – Sorting nexin 3*represent β-actin which were used as a loading control, # represents the tags for which Ab were available when antibodies against the primary protein were not available.

2.16.2.5. *In vivo* Immunoprecipitation / Co-immunoprecipitation reactions

Co-IP reactions were performed in order to verify the interactions identified by Y2H analysis as well the interaction between COMMD4 and cMyBPC. Single IP experiments were performed along with Co-IP experiments to serve as controls using relevant antibodies directed against the bait and each of the putative preys, respectively (Table 2.6, Section 2.16.2.3). In addition, a protein G agarose control was also included to assess whether the protein G itself was interacting non-specifically with the prey proteins while anti-Haemaglutinin antibody (HA) was used as a negative control antibody

For each IP reaction, pre-cleared cell lysates corresponding to 200µg of protein made up to 200µl with LyB (Section 2.16.2.4) was incubated overnight along with 1-10ul (1-2µg) of the relevant primary antibody at 4°C. The following morning, 60µl of prewashed (3X with LyB supplemented with protease inhibitors cocktails and PMSF) protein G agarose beads (Kirkegaard and Perry Laboratories, Gaithersburg, ML, USA) was added to the mixture and incubated at 4°C shaking gently on a Labcon orbital shaker (Labcon Pty, Ltd, Maraisburg, RSA) for 2 hours. Pellets were collected by centrifugation at 7000rpm for 30 seconds at 4°C. Subsequently, the samples containing the antibody-protein complex were washed 2-4 times with LyB supplemented with protease inhibitors cocktails and PMSF (Appendix I), each time repeating the centrifugation step. After the final wash step, the pellet was resuspendend in 40µl of 2X electrophoresis sample buffer (Appendix I) and boiled for 5 minutes at 95°C. The protein G agarose beads was pelleted at 13000rpm for 3 minute in a benchtop Beckman Microfuge Lite centrifuge (Beckman Instruments Inc, CA, USA), and the supernatant was loaded onto a 10-12% SDS-polyacrylamide gel (Section 2.7.2.) for subsequent immunoblotting (Section 2.16.2.6).

2.16.2.6. Western blots / Immunoblotting

Following the separation of proteins by electrophoresis, the proteins on the gel were transferred to a PVDF membrane, which was then probed for proteins of interest using the relevant primary antibody (Table 2.6, Section 2.16.2.3).

Prior to electrophoretic transfer, the gel was soaked in ddH_2O for 5 minutes, and then in transfer buffer (Appendix I) for 20 minutes, shaking at 4^0C on a Labnet rotor (Labnet Inc, NJ, USA) at 10rpm. The PVDF membrane was first soaked in methanol for 15 seconds, followed by ddH_2O for 5 minutes and lastly in transfer buffer (Appendix I) for 10 minutes, along with WB sponge pads and pieces of Whatman 3M paper (Whatman International Ltd, Maidstone, England) cut to size. The blotting apparatus was stacked in the following order: 1X WB sponge pad, 2-3X Whatman 3M paper, gel, PVDF membrane, 2-3X Whatman 3M paper and 1X WB sponge pad. The sandwich was subsequently inserted into the WB apparatus submerged in ice cold transfer buffer (Appendix I), along with an icepack and a magnetic stirrer. Transfer was carried out at 100V for 1 hour.

Following the transfer, the PVDF membrane was soaked in 0.1% TBST (Appendix I) and subsequently blocked at room temperature for 1 hour shaking on Labcon orbital shaker (Labcon Pty, Ltd, Maraisburg, RSA). Five percent fat free milk dissolved in TBST (Appendix I) was used as blocking agent, in order to prevent non-specific binding of antibodies to the surface of membrane. The membrane was rinsed 3 times with TBST and subsequently probed with the relevant primary antibody (Table 2.6, Section 2.16.2.3) diluted in blocking agent shaking overnight at 4^{o}C.

Following overnight incubation, the membrane was briefly rinsed with TBST and subsequently washed with TBST for 15 minutes shaking at room temperature on a rotating shaker (Labcon Pty, Ltd, Maraisburg, RSA). This was followed again by, 5 minutes washes for 3 times, replenishing the TBST between each wash step. Next, the membrane was incubated for 1 hour with the relevant horseradish peroxidase (HRP)-conjugated secondary antibody (Table 2.6, Section 2.16.2.3) which was directed against the primary antibody. Thereafter, the membrane was again washed in TBST for 5 minutes, and then subjected to detection using the ECL kit (Section 2.8).

2.16.3. *IN VIVO* CO-LOCALISATION

Verification of interactions by means of 3D-live-cell fluorescent co-localisation with Z-stacking was performed in order to assess the proposed protein:protein interactions in a cellular environment. For this, either immunocytochemistry were performed or fluorescently-tagged prey proteins were co-expressed with EYFP-COMMD4 in H9C2 cardiomyocytes and were viewed using a fluorescence microscope. The images acquired were subsequently processed for co-localisation. In the present study, the GFP, EYFP and DsRed tags were utilised.

2.16.3.1. Immunocytochemistry

Following transfection with EYFP-COMMD4 and subsequent differentiation, H9C2 cells cultured on cover glasses were washed with sterile PBS, fixed with a 1:1 methanol/acetone solution and incubated for 10minutes on ice at 4˚C. Ten minutes prior to slide preparations, cells to be used in the domain/phosphorylation interaction assay (section 2.17), were respectively treated with 65mM $CaCl_2$, or with 65mM $CaCl_2$ followed by 0.1µM isoproterenol. Following the methanol/acetone incubation, samples were then incubated at room temperature in either 5% donkey or goat serum (depending on antibody) for 20min. Cells were then incubated for 90 minutes at room temperature with appropriate primary antibody (1:50-1:200; table 2.6, section 2.16.2.3), followed by a 30 minutes RT incubation with a TxRed-conjugated secondary antibody (1:200). Cells were rinsed 3X with PBS. Thereafter, Hoechst 33342 (10mg/ml; Sigma, St Louis, MO, USA), which stains the nuclear material blue and is required for orientation purposes during the acquisition of images for the Z-stack, was added in a 1:200 dilution for a further 10min at room temperature incubation. Cover glasses were then washed three times with sterile PBS and transferred to glass microscope slides, mounted with fluorescent mounting medium (Dako Cytomation, Glostrup, Denmark) and sealed with nail varnish.

2.16.3.2. Co-localisation assay

Images were acquired following the co-transfection of the appropriate combination of plasmids (Section 2.4.3), and subsequent differentiation of transfected cells for 5-7 days (Section 2.4.3.1). Prior to image acquisition, culture media containing a 1:200 dilution of the nucleic acid stain, Hoechst H-33342 (Sigma) was added to the cells.

Cells or slides were then photographed immediately using an Olympus IX 81 motorised inverted fluorescence microscope equipped with a Z-motor, a cooled CCD camera (F-view-II Soft Imaging Systems) and automated excitation and emission filter wheel controlled by MT20 Cell^{R} (Olympus Biosystems GMBH, Tokyo, Japan). Image acquisition and analysis was performed using the Cell^{R} software (Olympus). Fluorescence was excited through an excitation filter for DAPI (Hoechst H-33342; 360nm), yellow fluorescent protein (YFP; 500nm), green fluorescent protein (GFP; 395nm) and for red fluorescent protein (RFP; 557nm). Emission was collected using Ultraviolet, blue, green (UBG) bandpass and YFP emission filter cubes, using an oil immersion 60x objective.

To facilitate the co-localisation of tagged proteins in three dimensions, Z-stacks were performed. Double-labelled images, using the co-transfected samples, were acquired at different focal planes, which were processed by the CellR software to determine co-localisation. The order of image acquisition in multi-colour were as follows: a GFP image was first taken at the lowest stack level and then a RFP image was taken at the same level. The camera then moved to the second lowest image, repeating acquisition of a GFP image followed by an RFP image. This order was repeated for each layer until the last image was acquired. The highest and lowest stack positions were determined by first focusing on the nucleus of the cell to be imaged, then taking the image out of focus by moving the objective up (top of the cell) and then down (bottom of the cell). Concurrently, co-ordinates of the top and bottom of the cell were recorded by the software. 3D co-localised images were produced from an average of 25 Z-stacked images collected at 0.26 µm intervals in the plane of the optical axis. Thus, all images represented maximum intensity projections from a Z-stack.

The image acquisition and analysis for the co-localisation of the GFP:YFP and GFP:DsRed protein pairs were performed differently. Since the spectra of GFP and YFP overlap, the excitation and visualization of the YFP-labelled structures occur with the GFP filter set and vice versa. Spectral imaging and linear unmixing were performed to separate and resolve the respective contribution of GFP and YFP to the total signal in each colour channel. To „unmix' the spectral information, it was necessary to determine the spectral properties of the individual fluorochromes from single transfection samples (reference samples, Table 2.1, Section 2.4.3) under the same imaging conditions used for the multi-labelled samples (co-transfections; Table 2.1, Section 2.4.3): i.e. the system had to be calibrated for each fluorochrome. Calibration was performed by first taking reference GFP and YFP images of single-labelled samples with the same filter set (excitation and emission filter) as used for the images of double-labelled specimen. Spectral unmixing was

performed using the CellR software (Olympus) and involved initial calibration (using the reference images) and subsequent unmixing. The processed double-labelled GFP:YFP images were then co-localised. Furthermore, the YFP-expression of the EYFP-COMMD4 was artificially coloured green or red instead of yellow in order to distinguish and improve contrast of the overlayed GFP:YFP images.

For the GFP:RFP co-localisations, images were acquired using the co-transfected samples and were processed for co-localisation in the same way as the GFP:YFP processed double-labelled images, except that unmixing was not required.

B. FUNCTIONAL STUDIES

2.17. DOMAIN/PHOSPHORYLATION ASSAY

Previous studies performed in our lab, implied that the interaction between COMMD4 and cMyBPC is mediated by the MyBPC motif in a phosphorylation-dependant manner (A Ramburan, Phd Thesis, 2008). Thus, in order to confirm whether the interaction between COMMD4 and cMyBPC was phosphorylation-dependent, a domain/phosphorylation assay involving immunocytochemistry and *in vivo* co-localisation was performed as described in section 2.16.3.

For this, cells were treated either with 65mM $CaCl_2$, or a combination of 65mM $CaCl_2$ and 0.1µM isoproterenol (1:1) in order to induce mono- and triphosphorylation conditions (McClellan *et al*., 2001). Co-localisation data from three independent experiments were pooled together and the mean values were used for quantification. A p value < 0.05 following unpaired t-testing was considered significant.

2.18. siRNA MEDIATED COMMD4 KNOCKDOWN

To bring about the specific knockdown of *COMMD4*, RNA interference was performed using small interfering RNA (siRNA). This was done to determine whether *COMMD4* suppression in H9C2 cardiomyocytes had an effect on cMyBPC turnover, which were, in turn, monitored using live cell fluorescence studies (Section 2.18.4). In addition, the co-localisation of GFP-tagged cMyBPC with a DsRed-tagged endosomal marker was also monitored in live cells with and without knockdown of COMMD4, in order to examine whether COMMD4 played a role in the accumulation of ubiquitin conjugates and also to view the location of accumulation within the cell.

2.18.1. siRNA preparation and transfection

Knockdown of *COMMD4* was achieved using "Genomewide designed" non-validated siRNAs from Qiagen (Hilden, Germany). Four different siRNA from the species *Rattus norvegius* were optimised in order to assess which of the four siRNA sequences achieved optimal knockdown of the targeted gene; these were subsequently used for downstream knockdown experiments. The siRNA sequences are summarised in table 2.7.

Table 2.7. *Commd4* predicted siRNA sequences from the *species*, Rat

Product name	si RNA sequences
Commd4 _Rn_LOC363068_1_HP siRNA	Sense: r(GAA AGU ACU UGU CAA CCA U)dTdT
	Antisense:r(AUG GUU GAC AAG UAC UUU C)dTdT
Commd4 _Rn_LOC363068_2_HP siRNA	Sense: r(AUA AGA AUA AAC AAG UGA A)dTdt
	Antisense:r(UAC ACU UGU UUA UUC UUA U)dTdG
Commd4 _Rn_LOC363068_3_HP siRNA	Sense: r(CCU UCG GGC CUG UAG CCU A)dTdT
	Antisense:r(UAG GCU ACA GGC CCG AAG G)dTdG
Commd4 _Rn_LOC363068_4_HP siRNA	Sense: r(CGG UUG CAG UGC UGA GUU U)dTdT
	Antisense:r(AAA CUC AGC ACU GCA ACC G)dTdG
NCS	Sense:UUC UCC GAA CGU GUC ACG UdT
	Antisense: ACG UGA CAC GUU AGA AdTdT

Abbreviations: NCS-non-silencing control and siRNA-small interfering RNA

H9C2 cells were cultured and subcultured as described in section 2.3.2; for subsequent transfection of siRNAs, $2x10^5$ cells were seeded in 6-well plates (Costar®, Corning Incorporated, Corning, NY, USA) at 80% confluency. Prior to transfection, stock siRNAs of 20µM was prepared by the addition of 250µl siRNA suspension buffer (Qiagen, Hilden, Germany) to a 5nmol synthesis. The tubes were heated to 90^{o}C for 1 minute, and incubated for an hour at 37^{o}C, after which the stock was either used or frozen at -20^{o}C.

For siRNA transfection, 12µl of the 20µM siRNA stock was added to 88µl serum-free medium, after which 7µl of HiPerfect transfection agent (Qiagen, Hilden, Germany) was added to each tube to achieve a final siRNA concentration of 100nM. In order to allow for complex formation, the samples were incubated at room temperature for 5 minutes. The culture medium in the 6-well plates was replaced with 2.3ml of fresh DMEM$^{10\%FBS+1\%PS}$ (Lonza, Walkersville, MD, USA) and the complexes were added drop-wise to cells. Plates were gently shaken to ensure uniform complex distribution and subsequently incubated at 37°for 24-48hrs in a Farma-thermosteri-cycle 5% carbon dioxide humidified incubator (Farma, international, Miami, Florida, U.S.A). Transfection of a non-silencing (NSC) siRNA served as a negative control, whereas an untransfected sample, with only transfection agent but no siRNA added to the cells, served as a mock transfection control. Suppresion of COMMD4 via siRNAs was subsequently assessed at the mRNA level, using two-step quantitative real-time RT-PCR (2.18.3), and at the protein level, via WB analysis (Section 2.16.2).

2.18.2. Total RNA extraction

Following transfection of siRNAs, H9C2 cells were washed with PBS and subsequently disaggregated by the addition of trypsin (Lonza, Walkersville, MD, USA). Cells were centrifuged for 1 minute in a benchtop Beckman Microfuge Lite centrifuge (Beckman Instruments Inc, CA, USA) at 2000rpm. Total RNA was extracted from H9C2 cells transfected with or without NSC or COMMD4 siRNAs, according to the manufacturer's instructions of the RNeasy® Plus Mini Kit (Qiagen, Hilden, Germany).

2.18.3. cDNA synthesis and two-step Q-RT-PCR

Subsequent to RNA extraction, 100ng of RNA was reverse transcribed into cDNA using the Quantitect® Reverse Transcription Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Prior to real-time PCR, a 100X dilution of this cDNA stock was prepared for future experiments; the resultant cDNA was used along with the Quantifast™ SYBR® green PCR kit (Qiagen, Hilden, Germany) according to manufacturer's instructions to perform real-time quantification of the cDNA.

Two-step Q-RT-PCR was performed with the Rotor-Gene 6000 real-time rotary analyzer (Corbett Life Sciences, Australia) using Quantitect® primer assays specific for *COMMD4* and rodent house-keeping genes, which included Transferrin receptor (*TRFR*), Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and Heat Shock Protein (*1β-Hsp1β*), purchased from Qiagen (Hilden, Germany), along with a non-template control. PCR-amplifications were performed in a 10µl volume containing 1µl 100X cDNA stock, 1µl Quantitect ® primer assay mix,5µl 2X Quantifast™ SYBR® green PCR master mix and 3µl nuclease-free water. Cycling conditions included a PCR initial activation step at 95^{o}C for 5 minutes, followed by two-step cycling; denaturing at 95^{o}C for 10 seconds and 40 cycles of combined annealing and extension step at 60^{o}C for 30 seconds.

For each sample, reactions were carried out in triplicate and *COMMD4* mRNA levels were normalised according to the expression of the three housekeeping genes using the normalisation factor calculation, based on the geometric mean of multiple house-keeping genes, as described by Vandesompele *et al* (2002).

2.18.4. Live cell fluorescent studies

In order to monitor co-localisation between GFP-tagged cMyBPC and Endo-DsRed-monomer-C1 in the presence and absence of COMMD4 siRNA, the co-transfection of the plasmid contructs were first carried out as described in section 2.4.3. After two days, the transfection of COMMD4 siRNA was performed as described in section 2.18.1. Afterwards cells were allowed to differentiate (Section 2.4.3.1), prior to live cell monitoring and image acquisition using the Olympus IX 81 motorised inverted microscope (Olympus). Co-localisation data from three independent experiments was pooled together and the mean values were used for quantification. A p value < 0.05 was considered significant following unpaired t-testing.

C. EVALUATION OF COMMD4 AND ITS LIGANDS AS HYPERTROPHY MODIFIERS

2.19. STUDY SUBJECTS, CLINICAL EVALUATION AND BLOOD COLLECTION

2.19.1. Study subjects

The present study has been approved by the Research Ethics Committee of the Faculty of Health Sciences, University of Stellenbosch (Project registration number: N04/03/062).

Informed consent was obtained from all participating subjects, or parents in case of minors, and blood samples were subsequently collected from each subject for molecular genetic testing. During routine screening for HCM-causing mutations, 22 referred index cases bearing one of the South African HCM-founder mutations ($R92W_{TNNT2}$, $R403W_{MYH7}$ and $A797T_{MYH7}$ [previously described within a South African population by Moolman-Smook et al., 1999]) were identified; pedigree tracing was performed and family members were asked to participate in the study.

A panel of 358 genetically and clinically affected and unaffected family members representing 22 pedigrees which were either of South African Caucasian or of Mixed Ancestry descent (Table 2.8), were used in the family-based association study.

2.19.2. Clinical Evaluation

Consenting subjects of 18 years and older were comprehensively clinically assessed by echocardiography performed by a single cardiologist (Dr. Miriam Revera, Pavia University, Italy), who was blinded to the mutation status. M-mode and 2D echocardiographic investigations as well as Doppler blood-flow imaging were made using a 2,5 Hz transducer in standard parasternal long-axis and short-axis, apical four- and two-chamber views with a GE Healthcare Vivid7 cardiovascular ultrasound system. In order to determine maximal left ventricular wall thickness (MLVWT), the heart muscle was divided in three levels, i.e. mitral valve, papillary muscle and supra-apex level (Figure 2.1). Six segments each at mitral valve level and papillary muscle level of the left ventricular wall, viz. anterior (aIVS) and posterior intraventricular septum (pIVS), anterior wall (AW), lateral wall (LW), inferior wall (IW) and posterior wall (PW) were measured by 2D echocardiography. Four segments at the immediate supra-apex level, viz. IVS, AW, LW and PW as per four-chamber view were also evaluated. All these dimensions (16 segments) were measured according to recommendation from the American Society of Echocardiography (http://www.asecho.org/guidelines.php). The atrial volumes were measured using the biplane area-length method (Schiller et al., 1989).

Table 2.8. Number of South African HCM-affected families analysed in present study

Gene	Location	Mutation	Pedigree	Ethnic group	n
MYH7	Exon 21	A797T	101	Caucasian	22
			104	Mixed Ancestry	14
			123	Mixed Ancestry	16
			124	Caucasian	4
			131	Caucasian	25
			138	Caucasian	32
			145	Mixed Ancestry	4
			147	Mixed Ancestry	10
			158	Caucasian	5
			159	Mixed Ancestry	11
			163	Caucasian	9
			172	Caucasian	8
MYH7	Exon 13	R403W	106	Mixed Ancestry	69
			134	Mixed Ancestry	9
			157	Mixed Ancestry	4
TNNT2	Exon 9	R92W	100	Mixed Ancestry	43
			103	Mixed Ancestry	5
			109	Mixed Ancestry	8
			137	Mixed Ancestry	7
			139	Mixed Ancestry	41
			149	Mixed Ancestry	10
			173	Mixed Ancestry	2

Abbreviations: n = number of individuals screened for SNPs in the present study which include mutation and non-mutation carriers, *MYH 7* – myosin heavy chain gene 7; *TNNT2* – troponin T gene 2

Current and past cardiovascular-related treatment, athletic activity and covariates for LVH, which included systolic blood pressure (SBP), diastolic BP (DBP) age, sex, body surface area (BSA), presence of hypertension and heart rate (HR) were recorded for each subject to allow adjustment for these traits during statistical analysis of association. Subjects were identified as hypertensive if they were on anti-hypertensive medication or had systolic BP $\geq$ 140 mmHg or diastolic BP $\geq$ 90 mmHg.

Additional hypertrophy indices, namely LVM and Spirito-Maron score (Spirito and Maron, 1990) were also calculated to describe the extent of hypertrophy. LVM was calculated using the American Society of Echocardiography's recommended formula for estimation of LVM from 2D-LV linear dimensions:

$$LVM = 0.8 \times (1.04[(LVIDd + PWTd + SWTd)^3 - (LVIDd)^3]) + 06g.$$

Furthermore, a new cumulative wall thickness score (CWT) was also determined by the addition of the wall thickness in all three levels, thus the 16 segments in total. Additionally, principle component analysis was performed to statistically define composite scores that best reflect the variability in hypertrophy throughout the left ventricle. One of these components, namely Comp1, was used in the association analysis since it explained 76% of the variance in the 16 wall thickness measurements and also showed

significant heritability. Each wall thickness measurement was transformed to approximate normality using quantile normalization prior to analyses (Pilia *et al*., 2006).

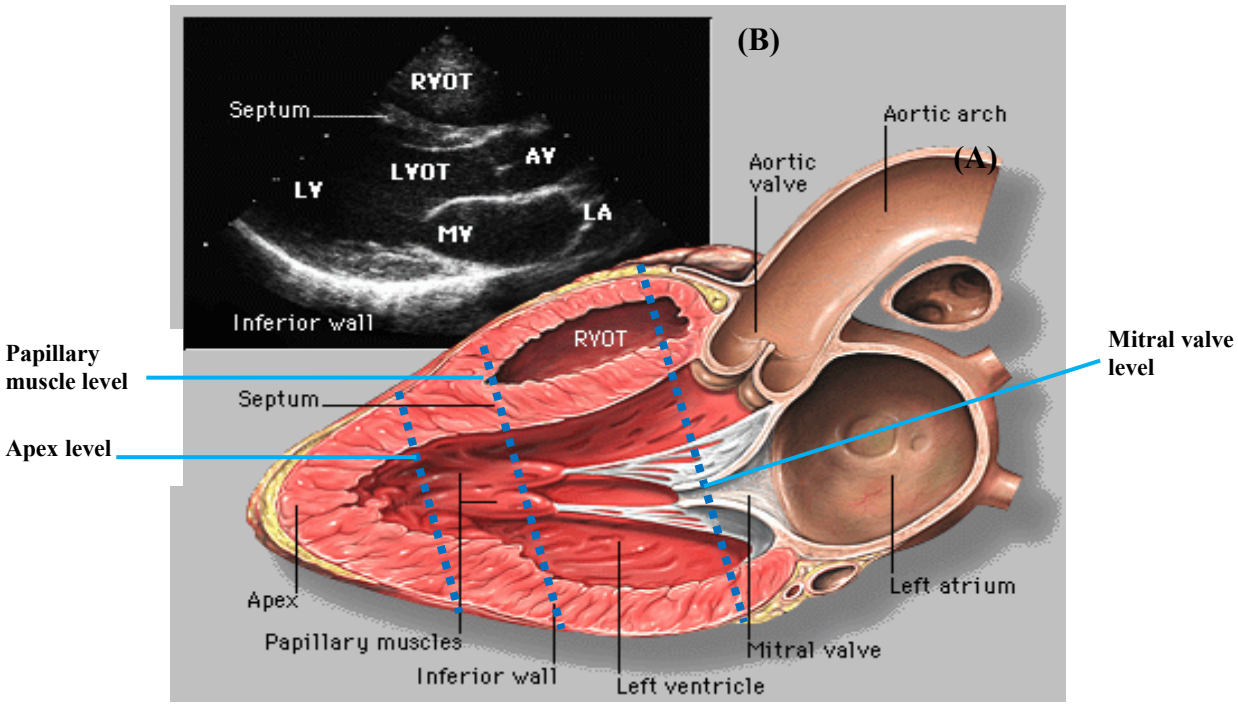


Figure 2.1: Schematic representation showing a cross-section of the heart transected at 3 levels.
A) Long-axis view of left ventricle, with the 3 blue dotted lines representing the mitral valve level, papillary muscle level and the apex level, where measurements were taken. B) A 2D echocardiogram of the left ventricle.
Abbreviations: AV- aortic valve, LA-left atrium, LV-left ventricle, LVOT-left ventricular outflow tract, MV- mitral valve, RVOT-right ventricular outflow tract. Taken from (http://www.med.yale.edu/.../aortic_regurgitation.html) with minor modifications by JC Moolman-Smook.

2.19.3. Blood collection and DNA extraction

Blood samples were collected from each patient in 2x 5ml ethylene-diamine-tetra-acetic acid (EDTA) tubes (Vacutainer, RSA) for DNA extraction and in 1x 10ml heparin tubes (Vacutainer, RSA) in order to establish permanent lymphoblast cell lines using the technique as described by Neitzel et al., (1986). These were used to ensure long-term provision of DNA as far as possible.

DNA was extracted from nucleated blood cells or transformed cells using a phenol-chloroform method, as described by Corfield *et al*., (1993). Mrs Ina Le Roux performed the DNA extractions as well as lymphoblastoid cell EBV transformations and maintenance, as described in Appendix III.

2.20. BIOINFORMATICS

2.20.1. SNP selection

For *COMMD4* and its verified interactors, intragenic polymorphisms spread throughout the genes were selected for use in the association studies, in order to best capture the spectrum of genetic variation within these candidate genes. The most informative single nucleotide polymorphisms (SNPs) were chosen from

public domain databases, (NCBI dbSNP, the Genome Database, Genecards and HAPMAP). For *COMMD4*, the SNPs were prioritized not only according to high heterozygosity values and validation status, but also according to whether they affected a restriction enzyme site, for subsequent allele-specific restriction enzyme analysis (ASREA) (Section 2.21).

Subsequently, the SNPbrowser™ Software 3.0 developed by Applied Biosystems (Foster City, CA) was used to select validated SNPs for which Taqman Genotyping assays were available in the genes encoding putative COMMD4-interactors. Validated SNPs by Applied Biosystems as well as data from the International Hapmap project from four populations (US Utah residents with European ancestry [CEU], Japanese from Tokyo [JPT], Han Chinese from Beijing [CHB] and Yoruba from Nigeria [YRI]) are incorporated in this software package. This program facilitates easy visualization of the spread of SNPs across the respective gene as well as physical and linkage disequilibrium (LD) maps and putative haplotype blocks for the specified populations (De La Vega et al. 2006).

Maniatis et al. (2002) constructed a metric LD map, expressed in LD Units (LDUs) which were calculated by the LDMAP software. This map places SNPs on a coordinate system where distances between SNPs are additive and proportional to the degree of LD between them. These LD maps are analogous to the genetic map which are expressed in centi-Morgans (cM) and thus can be used to position markers for population-based disease association studies (Collins et al., 2004).

For the present study, SNPs at an even spacing of 0.5 LDUs were selected for the CEU and YRI populations using the default parameters of the software (De la Vega, 2007). Table 2.9 summarises all the chosen SNPs of the candidate genes evaluated in this study, their genotypes, Taqman assay identification numbers and minor allele frequencies.

Table 2.9. Summary of SNP assay ID's, genotypes, chromosome position and minor allele frequency for each gene.

Gene	Chromosome	SNP	Taqman Assay	Genotype	MAF / HT
COMMD4	15	rs 3809547	-	C/G	0.2 / 0.40
		rs1185314	-	C/T	0.5 / 0.48
		rs1107254	-	A/G	0.2 / 0.47
SNX3	6	rs579257	C___2248075_10	G/T	0.27
		rs628424	C__26107227_10	A/C	0.25
		rs588409	C___8945545_10	C/T	0.21
		rs635051	C___2235563_10	C/T	0.30
		rs9398166	C__29986797_10	A/C	0.28
FBXL10	12	rs12427382	C__30926953_10	A/G	0.41
		rs3751131	C__25759495_10	A/G	0.36
		rs6489811	C__30926984_10	A/G	0.43
		rs11065585	C__30709465_10	C/G	0.43
		rs4980986	C__31117848_10	A/T	0.33
		rs7307400	C__31117814_10	A/G	0.40
		rs10849893	C_____86516_10	C/G	0.24
		rs7313015	C__43488864_10	G/T	0.33
LGMN	14	rs2402189	C___1686518_10	A/C	0.30
		rs3783933	C___1686512_10	C/T	0.42
		rs4904980	C__11472525_10	C/T	0.33
		rs1242102	C___7585598_20	C/T	0.41
		rs1242095	C___3071183_10	A/C	0.39
DSCR3	21	rs3165	C___9510305_1_	C/T	0.26
		rs2835680	C__16076326_10	C/T	0.41
		rs9305614	C___2438787_20	A/G	0.47
		rs2835684	C___2438799_10	A/G	0.37
		rs3787802	C__27482280_10	A/T	0.25

Abbreviations: bp- base pair, COMMD4 – COMM domain-containing protein 4, DSCR3 – Down syndrome critical region 3, FBXL10- F-box and Leucine rich repeat protein 10, HT – heterozygosity values, LGMN-Legumain, MAF – minor allele frequency, SNP- single nucleotide polymorphism, SNX3-Sorting nexin 3. The values in red indicate heterozygosity values.

2.21. ALLELE SPECIFIC RESTRICTION ENZYME ANALYSIS (ASREA) OF COMMD4

2.21.1. Primers for COMMD4 SNP genotyping using ASREA

Primers were designed to amplify the SNP-containing regions of *COMMD4;* primer sequences used for each of the targeted *COMMD4* polymorphisms are summarized in table 2.10. PCR-amplification and cycles of the SNP genotypes are described in section 2.21.2.

Table 2.10. Primer sequences, annealing temperatures and expected sizes of amplified fragment used for genotyping of polymorphisms in COMMD4

Polymorphism	Primer sequences 5'- 3'	Size (bp)	Tm ^{o}C
rs 11853141	F - GTAAGGCCAACTGAGGACC	356	60
	R - CTGAGGCAGAAGAATTGCTTG		
rs 11072542	F - CCAGCCTGGATAACAGAGC	337	57
	R - GAGACAGGTTTGCCATGTTG		
rs 3809547	F – CAT CCCCAAGAGGACCTC	255	58
	R - CCTTGGCAGCTCAAGTCC		

Abbreviations: bp-base pairs and Tm- melting temperature

2.21.2. PCR-amplification conditions for genotyping

The conventional PCR reaction-mix prepared for ASREA genotyping contained 100ng genomic DNA as template, 2.5mM of each of dATP, dCTP, dGTP, dTTP (Bioline UK Ltd, London, UK), 10X Taq DNA polymerase buffer (Bioline UK Ltd, London, UK), 150ng of each nucleotide primers (Tables 2.10), 0.5 units (U) Taq DNA polymerase (Bioline UK Ltd, London, UK), 1.5mM $MgCl_2$, 5% formamide or Dimethylsulfoxide (DMSO) [Sigma Chemical Company, St. Louis, Missouri, USA] and water to a final volume of 50µl. Amplification was carried out in a thermal cycler (Perkin–Elmer, Applied Biosystems Inc, Foster city CA, USA) according to standard protocols and PCR conditions for each set of oligonucleotide primers (Table 2.10). Annealing temperatures were optimized for each genotype and the cycling profile for 30 cycles was as follows: denaturation at 94^{o}C (30seconds), annealing at temperature according to Table 2.10 for 30 seconds, and extension at 72^{o}C (1 minute). After the completion of the cycling period, samples were electrophoresed on a 1% or 2% agarose gel, depending on the size of the fragment, to verify amplification (Section 2.7.1).

2.21.3. SNP genotyping by means of allele specific restriction enzyme analysis (ASREA)

Allele-specific restriction enzyme analysis (ASREA) was used in order to screen three cohorts of South African HCM families for selected *COMMD4* SNPs. Table 2.11 summarizes each polymorphism, with its respective restriction enzyme and the expected sizes of various alleles.

For each digest, a mixture was made up to a total volume of 30µl containing 20µl of PCR product, 2-3U of restriction enzyme (Table 2.11), 3µl 10X buffer according to the manufacturer's instructions, and the

appropriate volume of water. A no-enzyme control was also included as a control for the size of the undigested allele. The mixture was incubated overnight at 37°C and subsequently heat-inactivated at 65°C for 5 minutes.

Thereafter, 25µl of the digested sample was mixed with 2µl bromophenol blue loading dye and were electrophoresed on a 2.5% agarose gel along with a Promega 100bp molecular weight marker and the uncut control. Afterwards gels were visualized using the SYNGENE UV gel documentation system (SYNGENE, Synoptics Ltd Beacon House, Cambridge, UK).

Table 2.11. Characteristics of the selected polymorphisms in *COMMD4*

SNP	Alleles	HT value	Enzymes	Expected fragment sizes (bp)	
				Major allele	Minor allele
rs 3809547	C/G	0.404	Mbo II	225	176, 79
rs 11853141	C/T	0.475	Nhe I	356	234, 122
rs 11072542	A/G	0.468	Dra II	337	223, 114

Abbreviations: SNP - single nucleotide polymorphism, HT - heterozygosity, bp - base pairs

2.22. TAQMAN ASSAYS OF PUTATIVE COMMD4 INTERACTOR GENES

ABI Taqman allelic discrimination assays (Applied Biosystems, Foster City CA, USA) were used to genotype SNPs in genes encoding putative COMMD4-interactors in the 358 individuals.

Each Taqman SNP genotyping reaction consisted of forward and reverse primers for the amplification of the SNP polymorphic site region as well as two TaqMan probes for allele detection and the passive reference ROX (Perkin-Elmer). Each TaqMan probe has a fluorophore linked to the 5' end of the allele-specific probe (VIC for the Allele 1 probe and FAM for the Allele 2 probe), and a non-fluorescent quencher molecule (NFQ) linked to the 3′-end of each probe. When the probe is still intact, the quencher remains in close proximity to the fluorophore, thereby eliminating its signal. During PCR amplification, the probes anneal specifically to the complementary sequence, while during extension, it is degraded by the 5'-nuclease activity of the Taq DNA polymerase. This results in the separation of the fluorophore from the quencher, thereby generating an allele-specific detectable signal.

2.22.1. PCR GENOTYPING BY TAQMAN® SNP GENOTYPING ASSAYS

Briefly, a 5µl PCR cocktail was made consisting of 2.5µl Taqman® Universal PCR Mastermix, 0.25µl, 20x Taqman® SNP genotyping assay mix (Applied Biosystems assays-on-demand FAM™ dye- and VIC™ dye-labeled Taqman probes) mixed with the passive reference ROX (Perkin Elmer), 1µl genomic DNA template (20ng) and 2.25µl ddH_2O. Subsequently, the mixture was PCR-amplified in a 384-well plate using a GeneAmp PCR System 9700 (Applied Biosystems Inc, Foster City CA, USA). PCR-cycling parameters consisted of an initial hold at 50°C for 2 minutes then another hold at 95°C for 10 minutes, followed by 40

cycles at 92°C for 15 seconds and 60°C for 1 minute and 30 seconds. A last hold was set at 4°C for 7 minutes.

Following PCR amplification, the assays were analysed on an ABI Prism 7900HT Sequence Detection System (Applied Biosystems Inc, Foster City CA, USA) at the Central DNA Sequencing facility, University of Stellenbosch, Stellenbosch, South Africa. Allele assignment was performed using the Applied Biosystems Sequence Detection Software (SDS) version 2.3, whereby fluorescent signals was quantified and compared to determine the allelic content of each of the samples. The software employs the Applied Biosystems Maximum Likelihood Algorithm, which conducts a cluster analysis of the data based on the ratio of normalised reporter dye signal (the ratio of FAM™ to VIC™ dye) in each of the samples. All results were confirmed by visual inspection of the real-time PCR multicomponent analysis plots.

2.23. STATISTICS

Family-based statistical methods were used to assess association between gene variants and quantitative measures of hypertrophy indices within the South African HCM founder families.

Major hypertrophy indices that were focused on in the statistical analysis included maximal left wall thickness (mLVWT), left ventricular mass (LVM), maximal interventricular septum thickness (mIVST), maximal posterior wall thickness (mPWT), Maron score and CWT score. Echocardiographically-derived LVM, while frequently used to quantify hypertrophy in HCM, is known to be an inaccurate measure due to the variable nature of hypertrophy in HCM. Therefore, principle component analysis was further used to statistically define a composite score that best capture variability in LVH in the HCM cohort (Performed by Prof. Lize van der Merwe). The first of these components, namely Comp1 was included in the association analysis since it explains 76% of the variance in the dataset (the weighted sum of the16 wall thickness measurements) and shows significant heritability ($P = 0.018$, $H^2 = 0.38$) compared to LVM ($P = 0.028$, $H^2 = 0.29$) in the present cohort (Carstens *et al*., 2010).

Prior to association analysis, genotypic and phenotypic data were combined to generate a pedigree file for subsequent statistical analysis in PEDSTATS, Simwalk2, Haploview v. 4.1 and QTDT. Pedigree files were generated from the transformed dataset as specified by the QTDT tutorial (http://www.sph.umich.edu/csg/abecasis/QTDT/) for the subsequent analyses of all the quantitative hypertrophic traits. Simwalk2 was use to calculate identity-by-decent (IBD) files, which estimates IBD probabilities by the specific relationship between family members.

To test for association with the abovementioned hypertrophy traits, the quantitative transmission disequilibrium test (QTDT), was employed, which allows the tracing of transmitted alleles from heterozygous parent to offspring in multiple generations. Unaffected relatives functioned as “built in” controls, thus providing appropriately matched controls to reduce stratification. Since QTDT requires

heterozygous parents, offsprings of homozygous parents were uninformative, which resulted in some variants not being tested as the program only attempts to calculate allele effects when a minimum of 20 informative individuals are available.

For analysis, QTDT assumes that various trait values are normally distributed, since variance components analyses are sensitive to kurtosis and skewness in trait distribution. For the present study; this was not the case for some traits. Therefore as a standard procedure, each trait was transformed by Prof. Lize van der Merwe to approximate normality using quantile normalization as described by Pilia *et al*. (2006).

2.23.1 Descriptive Statistics

Pedstats v. 6.11 was implemented to validate input files and monitor Mendelian inheritance errors within families. In addition, deviation from Hardy-Weinberg equilibrium was also tested for each marker within unrelated individuals in the cohort, while graphic output files were generated of allele and genotype frequencies. The distribution and familial correlations of quantitative traits and covariates were also included into the graphic output file.

2.23.2 Linkage disequilibrium (LD) determination

Haploview v. 4.1 was used to determine linkage disequilibrium (LD) between SNPs within our study cohort. LD was measured in terms of D' values and each pair of SNPs was estimated as either being in strong LD, uncertain or incomplete LD.

2.23.3 Quantitative transmission disequilibrium test

As some of the pedigrees in the cohort were of Mixed Ancestry while others were of Caucasian descent, we considered that there was a possibility of population stratification within the cohort. Hence, QTDT was used to evaluate evidence for population stratification. Where evidence for population stratification was found, the orthogonal model as described by Abecasis *et al*. (2000), was implemented in QTDT, since it analyses the within-family component of association, which is robust to population stratification. To account for extended families, association modelling was also fitted with the variance components framework which models phenotypic similarities such as shared environment, polygenic and specific genes effects. Where no evidence for population stratification was found, the total model for association, which includes the within-family and between-family components were also evaluated, as it uses more of the available data.

Covariates, viz. systolic BP, diastolic BP, age, sex, BSA and HR were also incorporated in the model, as well as mutation status within the three founder mutation groups (A797T, R403W and R92W). This latter covariate was included to adjust for ancestral relatedness within each founder group and to adjust for the potential influence of the distinct HCM causal mutation on the hypertrophic phenotype variability.

Exact p-values for all QTDT significant associations were determined using Monte-Carlo permutation tests (McIntyre *et al*., 2000) and p-values < 0.05 were considered to be statistically significant. Effect sizes

obtained from analysis of transformed data cannot be interpreted in terms of the original measurement, thus the effect sizes of alleles showing significant association with hypertrophy traits, were estimated using the untransformed dataset.

CHAPTER 3: RESULTS

INDEX **PAGE**

3.1. IDENTIFICATION OF COMMD4 LIGANDS USING Y2H ANALYSIS 98

3.1.1. Y2H analysis of COMMD4 98

3.1.1.1. Generation of the COMMD4 bait construct 98

3.1.1.2. Quality control tests for the COMMD4 bait construct 98

3.1.1.3. Mating efficiency of AH109 transformed with COMMD4 bait construct 99

3.1.1.4. Y2H screening of pretransformed human cardiac cDNA library 100

3.2. PROTEIN-PROTEIN INTERACTION VERIFICATION ASSAYS 109

3.2.1. *In vitro* Co-immunoprecipitation 109

3.2.2. *In vivo* Co-immunoprecipitation 112

3.2.3. Live-cell 3D Co-localisation 115

3.3. FUNCTIONAL STUDIES 120

3.3.1. Domain/phosphorylation assay 120

3.3.2. si-RNA-mediated COMMD4 knockdown 122

3.4. COMPARISON OF Y2H DATA WITH VERIFICATION ASSAYS 124

3.5 FAMILY-BASED ASSOCIATION STUDY 125

3.5.1. Echocardiographic characteristics of the study cohorts 125

3.5.2. SNP selection 125

3.5.3. Genotyping 127

3.5.3.1. *COMMD4* ASREA 127

3.5.3.2. Taqman® Allele Discrimination 130

3.5.4. Statistical Analysis 130

3.5.4.1. Linkage Disequilibrium determination 132

3.5.4.2. Association test 135

3.5.4.2.1. *COMMD4* 136

3.5.4.2.2. *SNX3* 137

3.5.4.2.3. *DSCR3* 138

3.5.4.2.4. *LGMN* 140

3.5.4.2.5. *FBXL10* 140

3.5.4.2.6. Summary of association analysis 142

CHAPTER 3:

For the present study, the aim was two-fold: 1) to use COMMD4, a ligand of cMYBPC, as a bait in a yeast two-hybrid library screen to determine its distal ligands with a view to elucidating the function of COMMD4, and 2) to assess the genes encoding COMMD4 and its interactors as modifiers of hypertrophy in a panel of HCM founder families. Thus the data will be presented in this order.

3.1. IDENTIFICATION OF COMMD4 LIGANDS USING Y2H ANALYSIS

3.1.1. Y2H ANALYSIS OF COMMD4

3.1.1.1. Generation of the COMMD4 bait construct

PCR-amplification of the COMMD4 insert was successfully achieved using gene-specific primers as judged by gel electrophoresis. Afterwards, the amplified product was successfully cloned into the pGBKT7 cloning vector using the *Nde I* and *EcoRI* restriction enzyme sites for the 5' and 3' target sequence.

After the generation of the COMMD4 bait construct, automated sequencing was performed in order to ascertain that the reading frame of both bait and vector, as well as the integrity of the COMMD4 insert, was maintained. The sequence traces showed that both the integrity of the coding sequence and the reading frame were conserved throughout amplification and cloning.

3.1.1.2. Quality control tests for the COMMD4 bait construct

The COMMD4-bait construct was transformed into the MAT**a** yeast strain AH109, which is deficient for the *ADE2*, *HIS3*, *LEU2* and *TRP1* genes and are thus unable to grow on SD plates which lack the corresponding nutrients. Conversely, AH109 contains the *URA3* gene in its genetic make-up and are able to grow on SD plates lacking this nutrient. The COMMD4 bait construct transformed in AH109 was unable to grow on SD^{-Ade}, SD^{-Leu} and SD^{-His} but grew on SD^{-Ura} plates as well as SD^{-Trp} plates, because of the introduction of the functional *TRP1* gene which is encoded by the transformed pGBKT7 vector. This confirmed the phenotype of yeast strain AH109 (Table 3.1) and also showed that the COMMD4 bait construct did not autonomously activate any of the endogenous host reporter genes (*ADE2*, *HIS3*).

Table 3.1. Phenotype verification of yeast strain AH109 after COMMD4 bait construct transformation

Single dropout (SD) media	COMMD4-pGBKT7 in AH109
-Trp	++
-Ura	++
-His	-
-Leu	-
-Ade	-

Abbreviations: Ade – Adenine, COMMD4 – COMM domain containing protein 4, His – Histidine, Leu – Leucine, Trp – Tryptophan, Ura - Uracil

To assess the toxity of the COMMD4 bait construct on yeast strain AH109, growth curves were generated, which compared AH109 transformed with unrecombinant pGBKT7 to AH109 transformed with COMMD4-pGBKT7. Figure 3.1 shows that the slope of the growth curve representing AH109 transformed with the non-recombinant pGBKT7 was slightly higher than the growth curve representing the bait strain. Thus the AH109 strain containing COMMD4-pGBKT7 grew slower, which indicates that the COMMD4 bait construct was slightly toxic to yeast strain AH109. This problem was overcome by growing the AH109 yeast strain for subsequent mating and library screens on agar plates instead of in liquid cultures.

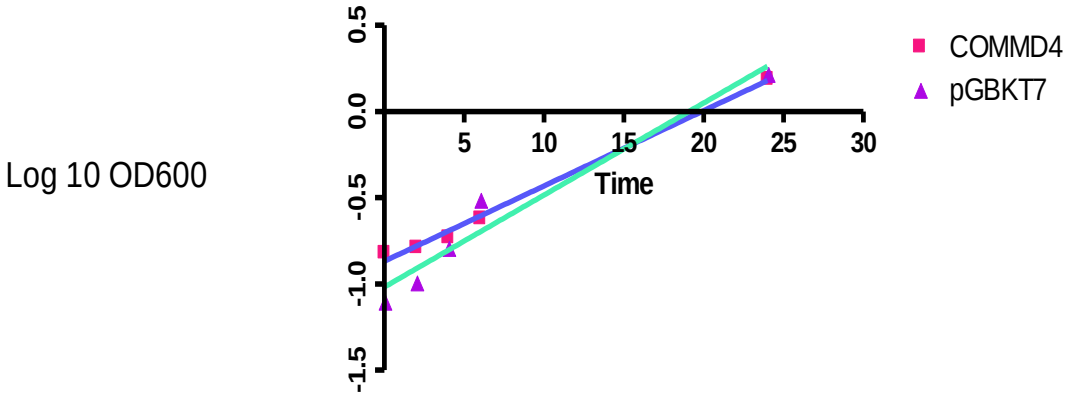


Figure 3.1. The growth curves of host yeast strain AH109 transformed with an non-recombinant pGBKT7 and COMMD4-pGBKT7. Growth rates of transformed pGBKT7-COMMD4 liquid cultures were compared to transformed non-recombinant pGBKT7 liquid cultures. The growth rate was estimated by calculating the slope of each of the curves. The slope of the blue line which represents the AH109 strain containing the COMMD4 bait construct was lower, indicating slower growth of yeast cells and thus slight toxicity on AH109.

3.1.1.3. Mating efficiency of AH109 transformed with COMMD4 bait construct

Small scale yeast matings were performed in order to assess whether the transformed COMMD4 bait construct affected the mating ability of yeast strain AH109. The COMMD4-pGBKT7-AH109 bait strain was mated with prey host strain, Y187 transformed with standard prey vector, pTD1.1. At the same time, a control mating was performed in which a control bait strain (pGBKT7-53-AH109) was mated with pTD1.1-Y187 strain. The mating efficiency of each experiment was calculated (Appendix II) and compared and results are represented in table 3.2. The results showed that the COMMD4 construct significantly enhanced the mating efficiency of yeast strain AH109, since the mating efficiency is much higher than that of the control experiment.

Table 3.2. Effect of COMMD4 bait construct on the mating efficiency of yeast strain AH109

Mating	Mating efficiency (%)
Control- pGBKT7-53-AH109 X pTD1.1-Y187	5.1
pGBKT7-COMMD4 -AH109 X pTD1.1-Y187	34

3.1.1.4 Y2H screening of pretransformed human cardiac cDNA library

A total of 3.6 X 10^5 yeast cDNA library clones were screened using COMMD4 as bait. After plating the library mixture on TDO plates, a total of 520 clones activated the *HIS3* reporter gene (Table 3.3; column A). Increasing the stringency of selection eliminated a further 150 clones, thus only 370 clones were able to activate both the *HIS3* and *ADE2* reporter genes on the QDO plates (Table 3.3; column B). These clones were further subjected to an X-α-galactosidase assay which yielded 123 clones which were able to activate the colourimetric *MEL1* reporter gene (Table 3.3; column C). Heterologous bait matings were performed on these 15 primary and 108 secondaries clones and a total of 28 clones, of which only 14 primary and 14 secondary clones showed specific interaction with COMMD4, and were subsequently sequenced (Table 3.4). Table 3.3 shows only scoring of the 28 clones, but the full scoring of clones is available in Appendix VII.

The physiological plausibility of these 28 putative interactors was evaluated using bioinformatic database mining (NCBI Genbank -http://www.ncbi.nlm.nih.gov or Ensemble- http://www.ensembl.org protein databases) to identify domains, localisation and function in order to prioritise interactors for subsequent verification studies (Table 3.5).

Four of the clones sequenced were discarded as putative interactors of COMMD4 based on their subcellular location (as determined by Proteome analyst [http://www.cs.ualberta.ca/~bioinfo/PA/Sub] and ESPpred [http://www.imtech.res.in/raghava/eslpred/), which was not compatible with COMMD4's localisation in the cytoplasm. A further fifteen clones were discarded as they encoded very short, physiologically irrelevant peptides due to a reading-frame shift of the prey inserts with respect to the GAL4 AD domain. In figure 3.2, clone 410, which encodes the protein, Legumain (LGMN) serves as a representative example to illustrate the *in silico* translation of a prey clone insert ORF, in the reading-frame as dictated by the GAL4-AD sequence.

Table 3.3. Scoring of the 28 primary and secondary clones which activated the nutritional (*HIS3* and *ADE2*) and colourimetric (*MEL1*) reporter genes by prey-COMMD4 interactions.

Colony nr	A Growth on TDO *HIS3* activation	B Growth on QDO *ADE2* activation	C X-α-Galactosidase assay (colour) *MEL1* activation
26	++++	++++	+++ (dark blue)
42	++++	+++	++ (medium blue)
46	++++	++++	+++ (medium blue)
52	++++	++++	+++ (medium blue)
53	++++	++++	+++ (medium blue)
136	++++	+++	+++(dark blue)
143	++++	++++	+++ (medium blue)
149	++++	++++	+++ (dark blue)
183	++++	++++	+++ (medium blue)
184	++++	++++	+++ (medium blue)
209	++++	++++	+++ (medium blue)
211	++++	++++	++ (dark blue)
217	++++	++++	+++ (medium blue)
234	++++	++++	++ (dark blue)
247	++++	+++	+++ (dark blue)
258	++++	++++	+++ (medium blue)
262	++++	++++	+++ (dark blue)
264	++++	++++	++ (medium blue)
281	++++	++++	+++ (medium blue)
288	++++	++++	+++ (dark blue)
335	++++	++++	+++ (dark blue)
339	++++	++++	+++ (dark blue)
342	++++	+++	+++ (medium blue)
345	++++	++++	+++ (medium blue)
383	++++	+++	++ (dark blue)
397	++++	++++	++ (dark blue)
410	++++	+++	++ (dark blue)
458	++++	++++	+++ (dark blue)

Colonies in red font represent the primary and those in black represent the secondary clones. TDO – solid media lacking Leu, Trp and His; QDO-solid media lacking Leu, Trp, His and Ade. Levels of colony growth on solid media: ++++ = robust, +++ = good, ++ = weak, - = no growth

Table 3.4. Heterologous bait matings of the 28 putative interactors which showed specific interaction with COMMD4.

Colony	pGBKT 7	COMMD4-pGBKT7	pGBKT7-53	Reelin
26	+	++++	+	-
42	+	++	+	+
46	+	+++	-	-
52	+	++	-	-
53	-	++	-	-
136	+	++	+	+
143	++	++++	-	+
149	+	++	+	-
183	-	++	-	+
184	-	++	-	+
209	-	+	-	-
211	+	+++	+	-
217	-	+++	-	-
234	+	++++	+	-
247	-	+++	+	-
258	-	++	-	+
262	+	+++	+	-
264	-	++	-	-
281	+	++	-	+
288	+	+++	+	++
335	-	+++	-	-
339	+	++	-	-
342	+	++	-	+
345	-	+	-	+
383	-	++++	+	+
397	+	++	++	-
410	-	++	+	-
458	+	++	-	-

Clones scored after 7 days on QDO plates. Colonies in red font represent the primary and those in black represent the secondary clones. QDO - solid media lacking Leu, Trp, His and Ade. Level of colony growth on solid media: ++++ = robust, +++ = good, ++ = weak, - = no growth

Table 3.5. Identification of the 28 COMMD4 putative interactor clones from the Y2H cardiac library cDNA library screen.

Genomic Hit			**Inframe protein Hit**			
Clone #	**Identity**	**Blast N Accession nr E-value**	**Blast P Accesion nr E-value**	**Insert size (aa)**	**Cellular localisation**	**Domains**
247, 264, 342	***Homo Sapiens* Actin 1, (ACTC1) cardiac**	NM_005159.4 (2e-143)	NP_005150.1 (2e-93)	193	Cytoplasm, cytoskeleton	Actin
217	***Homo Sapiens* Enolase1 (ENO1)**	NM_001428.2 (0.0)	AAA35698.1 (2e-13)	280	Cytoplasm, Nucleus	Enolase
143	***Homo Sapiens* Ankyrin repeat domain 1 (ANKRD1) cardiac**	NM_014391.2 (8e-178)	NP_055206.2 (5e-41)	121	Nucleus	5 ANK repeats
26	***Homo Sapiens* Sorting nexin 3 (SNX3)**	NM_003795.3 (5e-180)	BAB32649.1 (2e-69)	151	Cytoplasm	Phox (PX)
335	***Homo Sapiens* Down syndrome critical region 3 (DSCR3)**	NM_006052.1 (0.0)	NP_006043.1 (9e-20)	45	Nucleus	(Vps26)
410	***Homo Sapiens* Legumain (LGMN)**	NM_005606.5 (2e-159)	NP_005597.3 (2e-77)	139	Lysosome	-
46	***Homo Sapiens* F-box and Leucine rich repeat protein 10 (FBXL10)**	NM_032590.3 (4e-163)	NP_115979.3 (6e-21)	46	Nucleus	F-box domain JmjC domain Leucine Rich Repeat CXXC zinc finger
183	*Homo Sapiens* Guanine nucleotide binding protein, beta polypeptide 2-like 1 (GNB2L1 / RACK1)	NM_006098.4 (8e-155)	NP_006089.1 (1e-75)	242	Transmembrane	7 WD 40 repeats
209	*Homo Sapiens* Lysyl oxidase-like 4 (LOXL4)	NM_032211.6 (0.0)	NP_115587.6 (1e-135)	221	Cell Membrane	4 SRCR

Genomic Hit			Inframe protein Hit			
Clone #	**Identity**	**Blast N Accession nr E-value**	**Blast P Accesion nr E-value**	**Insert size (aa)**	**Cellular localisation**	**Domains**
383	*Homo Sapiens* Selenoprotein 1 (SEPP1)	NM_005410.2 (4e-94)	NP_005401.3 (8e-75)	138	Extracellular	SelP
258	*Homo Sapiens* Muskelin 1 (MKLN1)	NM_013255.3 (4e-86)	No significant similarity	39	Cytoplasm	6 kelsch motifs CTLH, LisH
458, 52	*Homo Sapiens* Chromosome 10 genomic contig - catenin, alpha 3 (CTNNA3)	NT_008583	No significant similarity	-	Actin cytoskeleton	-
149	*Homo Sapiens* Chromosome 9 genomic contig - Gelsolin (GSN)	NW_924573 (0.0)	No significant similarity	-	Actin cytoskeleton, cytosol, extracellular	6 Gelsolin-like repeats
136	*Homo Sapiens* Chromosome 7 genomic contig Zinc finger prot 680	NT_007758	No significant similarity	-	Nucleus	16 Zinc Fingers, 1 krab
53	*Homo sapiens* ras homolog gene family, member B (RHOB), mRNA	NM_004040.2 (0.0)	No significant similarity	-	Endosome, nucleus plasma membrane	-
281	*Homo Sapiens* Semaphorin 4C (SEMA4C)	NM_017789.3 (2e-116)	No significant similarity	-	Membrane	Immunoglobin , transmembrane, sema
345	*Homo sapiens* Chromosome 1 open reading frame 2 (C1orf2)	NM_006589.2 (0.0)	No significant similarity	-	Membrane	-
397	*Homo Sapiens* stearyl-CoA desaturase (delta-9-desaturase) (SCD),	NM_005063.4 (0.0)	No significant similarity	-	Endoplasmic reticulum, membrane	Histidine box
339	*Homo Sapiens* Chromosome 8 genomic contig - Dihydrodiol dehydrogenase (DDH2)	NT_007995.14 (0.0)	No significant similarity	-	Cytoplasm	-

Genomic Hit			Inframe protein Hit			
Clone #	**Identity**	**Blast N Accession nr E-value**	**Blast P Accesion nr E-value**	**Insert size (aa)**	**Cellular localisation**	**Domains**
288	*Homo Sapiens* Pyruvate dehydrogenase kinase, isozyme 4 (PDK4)	NM_002612.3 (0.0)	No significant similarity	-	Mitochondria	Histidine kinase
184	*Homo Sapiens* Syntaxin 7 (STX7)	NM_003569.1 (2e-97)	No significant similarity	-	Endosome, plasma membrane	T Snare coiled-coil
42	*Homo Sapiens* Prostaglandin D2 synthase (PTGDS)	NM_000954.5 (0.0)	No significant similarity	-	Extracellular, Intracellular, nucleus, Golgi membrane	-
211	*Homo sapiens*Mitochondrion, complete genome	NC_001807.4 (0.0)	No significant similarity	-	Mitochondrial	-
262	*Homo sapiens* Mitochondrion, complete genome	NC_001807.4 (0.0)	No significant similarity	-	Mitochondrial	-
234	*Homo sapiens* Mitochondrion, complete genome	NC_001807.4 (0.0)	No significant similarity	-	Mitochondrial	-

Clones highlighted in bold and grey (and with red font*) are those chosen for further verification assays. *Clone nr used for ACTC1 in further studies since it was represented by multiple clones

A) Translation of **clone 410_LGMN** (1-1053) using universal code

```
Total amino acid number: 336, MW=35201
Max ORF: 1-417, 139 AA, MW=16130

1         CCCCTACCTCCAGTCACACACCTTGACCTCACCCCCAGCCCTGATGTGCCTCTCACCATC
1          P  L  P  P  V  T  H  L  D  L  T  P  S  P  D  V  P  L  T  I

61        ATGAAAAGGAAACTGATGAACACCAATGATCTGGAGGAGTCCAGGCAGCTCACGGAGGAG
21         M  K  R  K  L  M  N  T  N  D  L  E  E  S  R  Q  L  T  E  E

121       ATCCAGCGGCATCTGGATGCCAGGCACCTCATTGAGAAGTCAGTGCGTAAGATCGTCTCC
41         I  Q  R  H  L  D  A  R  H  L  I  E  K  S  V  R  K  I  V  S

181       TTGCTGGCAGCGTCCGAGGCTGAGGTGGAGCAGCTCCTGTCCGAGAGAGCCCCGCTCACG
61         L  L  A  A  S  E  A  E  V  E  Q  L  L  S  E  R  A  P  L  T

241       GGGCACAGCTGCTACCCAGAGGCCCTGCTGCACTTCCGGACCCACTGCTTCAACTGGCAC
81         G  H  S  C  Y  P  E  A  L  L  H  F  R  T  H  C  F  N  W  H

301       TCCCCCACGTACGAGTATGCGTTGAGACATTTGTACGTGCTGGTCAACCTTTGTGAGAAG
101        S  P  T  Y  E  Y  A  L  R  H  L  Y  V  L  V  N  L  C  E  K

361       CCGTATCCACTTCACAGGATAAAATTGTCCATGGACCACGTGTGCCTTGGTCACTACTGA
121        P  Y  P  L  H  R  I  K  L  S  M  D  H  V  C  L  G  H  Y  *
```

B) Protein alignment of **LGMN** (NP_005597.3) reference sequence and the above translation of clone 410_LGMN

```
ORF       ------------------------------------------------------------
REF       MVWKVAVFLSVALGIGAVPIDDPEDGGKHWVVIVAGSNGWYNYRHQADACHAYQIIHRNG 60


ORF       ------------------------------------------------------------
REF       IPDEQIVVMMYDDIAYSEDNPTPGIVINRPNGTDVYQGVPKDYTGEDVTPQNFLAVLRGD 120


ORF       ------------------------------------------------------------
REF       AEAVKGIGSGKVLKSGPQDHVFIYFTDHGSTGILVFPNEDLHVKDLNETIHYMYKHKMYR 180


ORF       ------------------------------------------------------------
REF       KMVFYIEACESGSMMNHLPDNINVYATTAANPRESSYACYYDEKRSTYLGDWYSVNWMED 240


ORF       ------------------------------------------------------PLPPVT 6
REF       SDVEDLTKETLHKQYHLVKSHTNTSHVMQYGNKTISTMKVMQFQGMKRKASSPVPLPPVT 300
******

ORF       HLDLTPSPDVPLTIMKRKLMNTNDLEESRQLTEEIQRHLDARHLIEKSVRKIVSLLAASE 66
REF       HLDLTPSPDVPLTIMKRKLMNTNDLEESRQLTEEIQRHLDARHLIEKSVRKIVSLLAASE 359
************************************************************

ORF       AEVEQLLSERAPLTGHSCYPEALLHFRTHCFNWHSPTYEYALRHLYVLVNLCEKPYPLHR 126
REF       AEVEQLLSERAPLTGHSCYPEALLHFRTHCFNWHSPTYEYALRHLYVLVNLCEKPYPLHR 419
************************************************************

ORF       IKLSMDHVCLGHY* 139
REF       IKLSMDHVCLGHY* 432
*************
```

Figure 3.2. A representative example demonstrating the *in silico* translation of the ORF of the insert of prey clone 410 and protein alignment with the predicted protein reference sequence. A. Translation of in-frame nucleotide sequence of prey clone 410 using the DNAMAN™ software program (Lynnon Biosoft). **B**. Alignment of translated ORF sequences with protein reference sequence (REF) predicted by identity of prey insert nucleotide sequence using clustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

Table 3.5 shows the proteins identified as putative interactors of COMMD4, of which seven were prioritised for further analysis (in bold and highlighted in grey). The protein and domain structures of the seven prioritised prey proteins as well as the region of each prioritised prey encoded by the respective clone are illustrated in the linear schematic as shown in figure 3.3.

A brief description of these in-frame proteins, type, cellular compartment and possible functions as obtained via bioinformatics database mining, follows:

Three of the proteins, namely, α-enolase (ENO1), cardiac ankyrin repeat protein (ANKRD1) and cardiac actin (ACTC1) have been identified along with COMMD4 as ligands of the cMyBPC domains C1-C2 in a previous Y2H screens. ENO1 is an enzyme, which localizes to the nucleus and cytoplasm and functions in cellular metabolism (Jin *et al*., 2008). ANKRD1 is a nuclear transcription factor and localises to the I band, cytoplasm and nucleus where it may play a role in cardiac gene regulation (Zou *et al*., 1997).

Furthermore, four of the putative interactors identified seem to be involved in protein trafficking and protein degradation. The Down syndrome critical region 3 protein (DSCR3), resides in the nucleus and contains a Vps26 domain, which is homologous to a yeast vacuolar sorting protein 26. Sorting Nexin 3 (SNX3) localises to the cytoplasm and functions in the endosome-to-Golgi retrieval of proteins. Both SNX3 and DSCR are orthologs of each other and together form part of a retromer complex for their function in protein trafficking (Verges *et al*, 2007). The F-Box and leucine-rich repeat protein 10 (FBXL10), is a DNA-binding protein in the nucleus and forms part of an SCF complex that acts as an E3 ubiquitin ligase in the ubiquitin proteosome system (UPS) (Frescas *et al*, 2007). Lastly, Legumain (LGMN) localises to the late endosome, lysosome and cytoplasm. It contains the peptidase C domain and functions in lysosomal degradation (Chen *et al*., 1997). Taken together, these interactors implicate COMMD4 in protein trafficking and protein turnover.

A more detailed account of each of the respective preys is given in the next chapter (Discussion).

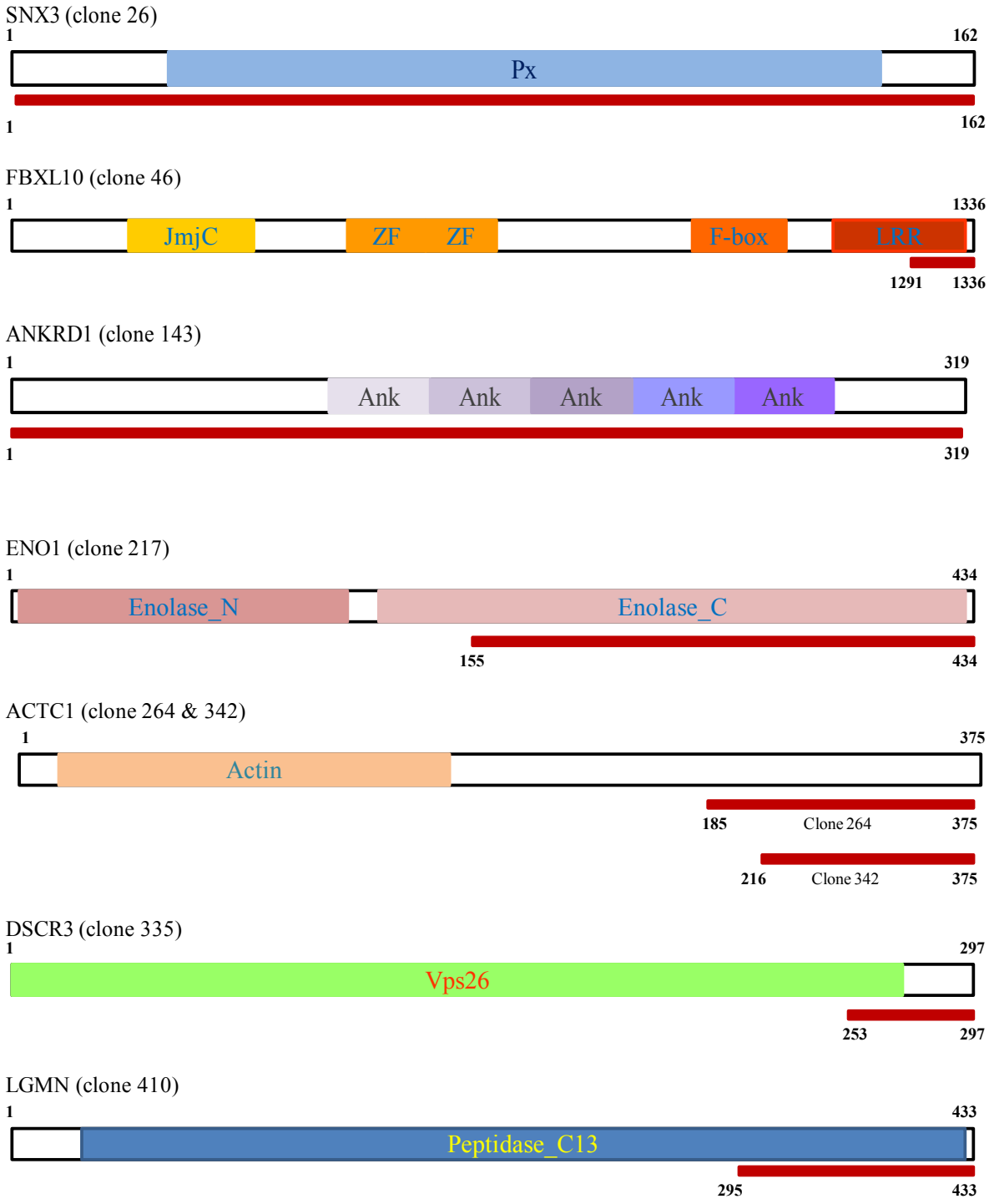


Figure.3.3. Linear schematic of protein and domain structure of the seven prioritised preys.
Different coloured blocks represent different domain structures in each respective prey protein. The numbers indicate the amino acid number and the red line below shows the fragment of prey protein encoded by the respective clone. ACTC1- cardiac Actin, Ank-Ankyrin domain, ANKRD1-Ankyrin repeat domain, DSCR3-Down Syndrome critical region 3/A, F-box-F-box motif, FBXL10-F-box and Leucine rich repeat protein, Px-phospholipid-binding motif, Sorting nexin 3, Vps26-vacuolar protein sorting, ZF-Zinc-Finger. *Diagram not drawn to scale*

3.2 PROTEIN-PROTEIN INTERACTION VERIFICATION ASSAYS

Following successful Y2H screening of a pre-transformed cardiac cDNA library using COMMD4 as "bait", protein-protein interactions of the seven prioritised clones with COMMD4 were assessed by additional, independent approaches, which include *in vitro* Co-IP analysis, *in vivo* Co-IP followed by Western blotting and live-cell 3D co-localisation with Z-stacking.

3.2.1 *IN VITRO* CO-IMMUNOPRECIPITATION

In vitro Co-IP was used to confirm bait and prey interactions in the absence of the GAL4 transcription factor domain. This was necessary, to ensure that interaction was not in some way mediated by the GAL4 domains.

Both the bait and prey inserts were successfully amplified to incorporate the T7 promoter and epitope-tags (cMyc-bait and HA-prey) for subsequent translation experiments. These fragments (COMMD4-bait and 7 preys) were successfully transcribed and translated and the actual size of the fusion proteins as observed on 15% SDS-PAGE gel (Figure 3.4) corresponded well with the *in silico*-predicted sizes of these protein fragments (Table 3.6). Some prey protein fragments, FBXL10 and DSCR3 repeatedly failed to be detected as bands on an SDS-PAGE gel and were thus excluded from subsequent *in vitro* Co-IP experiments.

Table 3.6. Predicted and approximate observed molecular weights of bait and prey fusion proteins used in Co-IP analysis

Cloned insert	Nr of predicted amino acids	Predicted MW (kDa)	Antibody tag and linker (kDa)	Size by electrophoresis (kDa)
Myc -COMMD4	199	21.7	1.4	~24
HA-SNX3	151	16.7	2.3	~19
HA-LGMN	139	16	2.3	~ 18
HA-ACTC1	193	21.8	2.3	~26
HA-ANKRD1	121	14.1	2.3	~16
HA-ENO1	280	30.4	2.3	~33
HA-FBXL10	46	5	2.3	-
HA-DSCR3	45	5	2.3	-

Electrophoresis were carried out on a 20% SDS-PAGE mini gel

Using this technique, a positive interaction was observed between COMMD4 and four of the seven preys. Figure 3.5 shows that SNX3, ACTC1, LGMN and ENO1 co-immunoprecipitated with COMMD4. Considering that equal volumes of prey was used in both the single IP and the Co-IP reactions, an intermediate interaction was observed between COMMD4, LGMN and ENO1, respectively. In comparison, a relatively strong interaction of COMMD4 with both ACTC1 and SNX3 was shown, whereas no interaction was observed between ANKRD1 and COMMD4. Interactions between COMMD4 with FBXL10 and DSCR3 could not be assessed using this method.

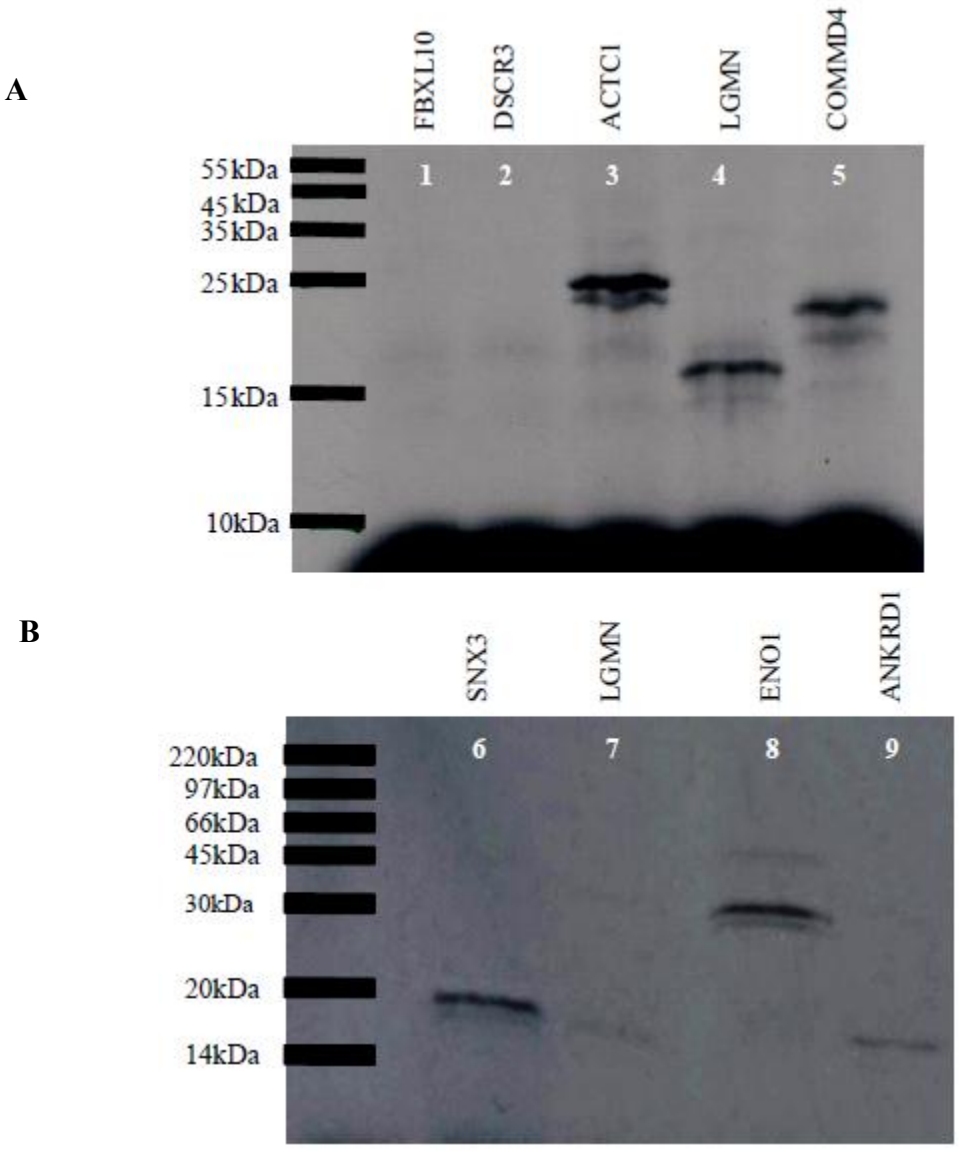


Figure 3.4. Autoradiographs of S^{35}-labeled methionine-labelled protein products from translation and transcription reactions. Lanes 3-9 shows respectively, translated bait and prey proteins, ACTC1, LGMN, COMMD4, SNX3, ENO1 and ANKRD1. In lane 1 and 2, the 5kDa translated FBXL10 and DSCR3 are not visible and may have been obscured by the dye front. The position of sizes from non-radioactive size markers, Spectra ™ Multicolor Broad Range Protein Ladder (Fermentas) **(A)**, and High-Range Rainbow Molecular Weight Markers (Amersham Biosciences) **(B)**, as transferred from dried polyacrylamide gel are shown in the marker lanes.

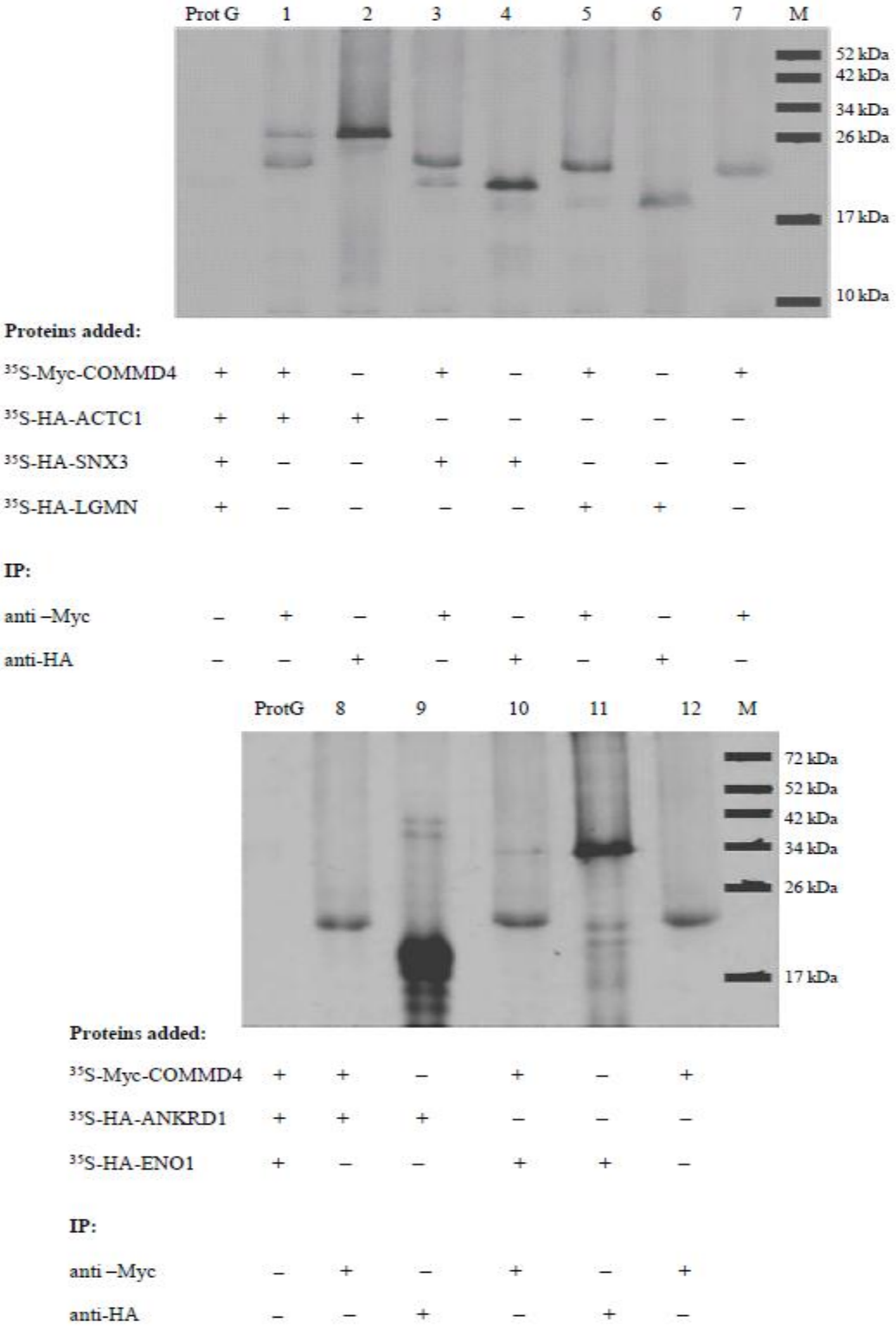


Figure 3.5. *In vitro* IP and Co-IP reactions of COMMD4 and each respective prey. Autoradiograph of 20% SDS-PAGE gel indicating immunoprecipitated Met-^{35}S-labelled *in vitro* transcribed/translated proteins and antibodies. COMMD4 co-immunoprecipitated with ACTC1 (lane 1), SNX3 (lane3), LGMN (lane 5) and ENO1 (lane 10). However, no interaction was detected between COMMD4 and ANKRD1 (lane 8). The protein G control lanes were clear, indicating that no non-specific occurred during pull-downs. The position of sizes from a non-radioactive size marker, Spectra ^{TM} Multicolor Broad Range Protein Ladder, (Fermentas), as transferred from dried polyacrylamide gel, is shown in the marker lane (M).

3.2.2. *IN VIVO* CO-IMMUNOPRECIPITATIONS

Since post-translational modification and protein folding may or may not occur properly in the coupled *in vitro* transcription-translation system, or indeed in the Y2H system, *in vivo* Co-IP experiments was used to confirm protein-protein interactions between COMMD4 and cMyBPC as well as its putative prey interactors within the cell. Where antibodies against proteins (COMMD4, DSCR3) were not available, or where commercial antibodies repeatedly failed to detect the correct size fragment (cMyBPC), fluorescently-tagged versions of these proteins were expressed in the relevant cells. Exogenously added EYFP-tagged COMMD4, pGFP-cMyBPC or DsRed-COMMD4 and DsRed-DSCR3 were respectively immunoprecipitated with an anti-YFP (JL8) or anti-DsRed antibody, while each of the prey interactors was immunoprecipitated with their respective antibody.

Interactions between exogenously expressed, EYFP-tagged COMMD4 or DsRed-COMMD4 and its ubiquitously expressed prey interactors (SNX3, LGMN, ENO1 and FBXL10) were investigated in HEK293 cells. In the case of ACTC1, cMyBPC and ANKRD1, whose expression is cardiac specific, H9C2 cardiomyocytes were used to determine bait:prey interactions.

Consistent with previous Y2H results, western blot data showed that COMMD4 co-immunoprecipitates with cMyBPC (Figure 3.6), confirming this interaction in a cellular environment. Interestingly, in the reciprocal co-immunoprecipitation experiment, COMMD4 pull-downs revealed only a band of about 50kD, rather than the band of ~140kD corresponding to GFP-cMyBPC (Figure 3.6), although this larger band was clearly detected in cells not transfected with DsRed-COMMD4. This suggests that cMyBPC undergoes enhanced degradation in the presence of additional COMMD4. The fact that this cMyBPC fragment can still be co-precipitated by COMMD4, which binds to the N-terminal region of cMyBPC, is congruent with previous findings that indicates that degradation of cMyBPC occurs from the C-terminal (Kulikovskaya *et al*., 2007).

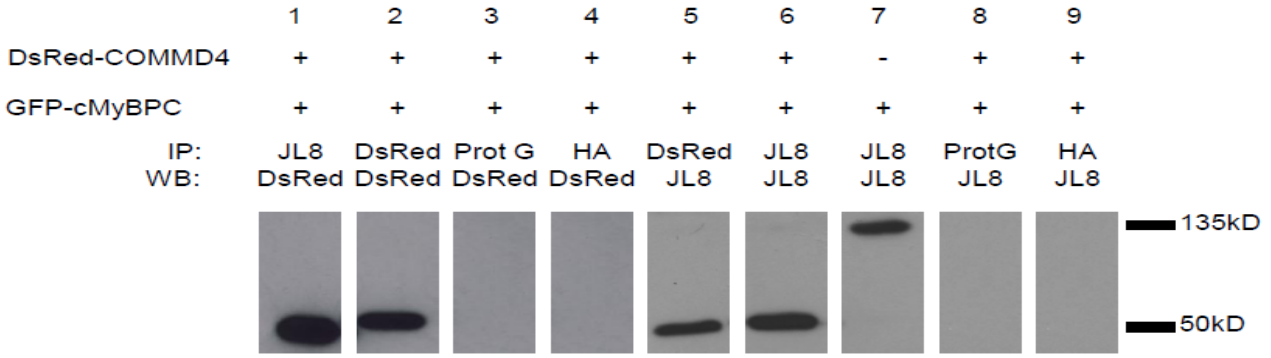
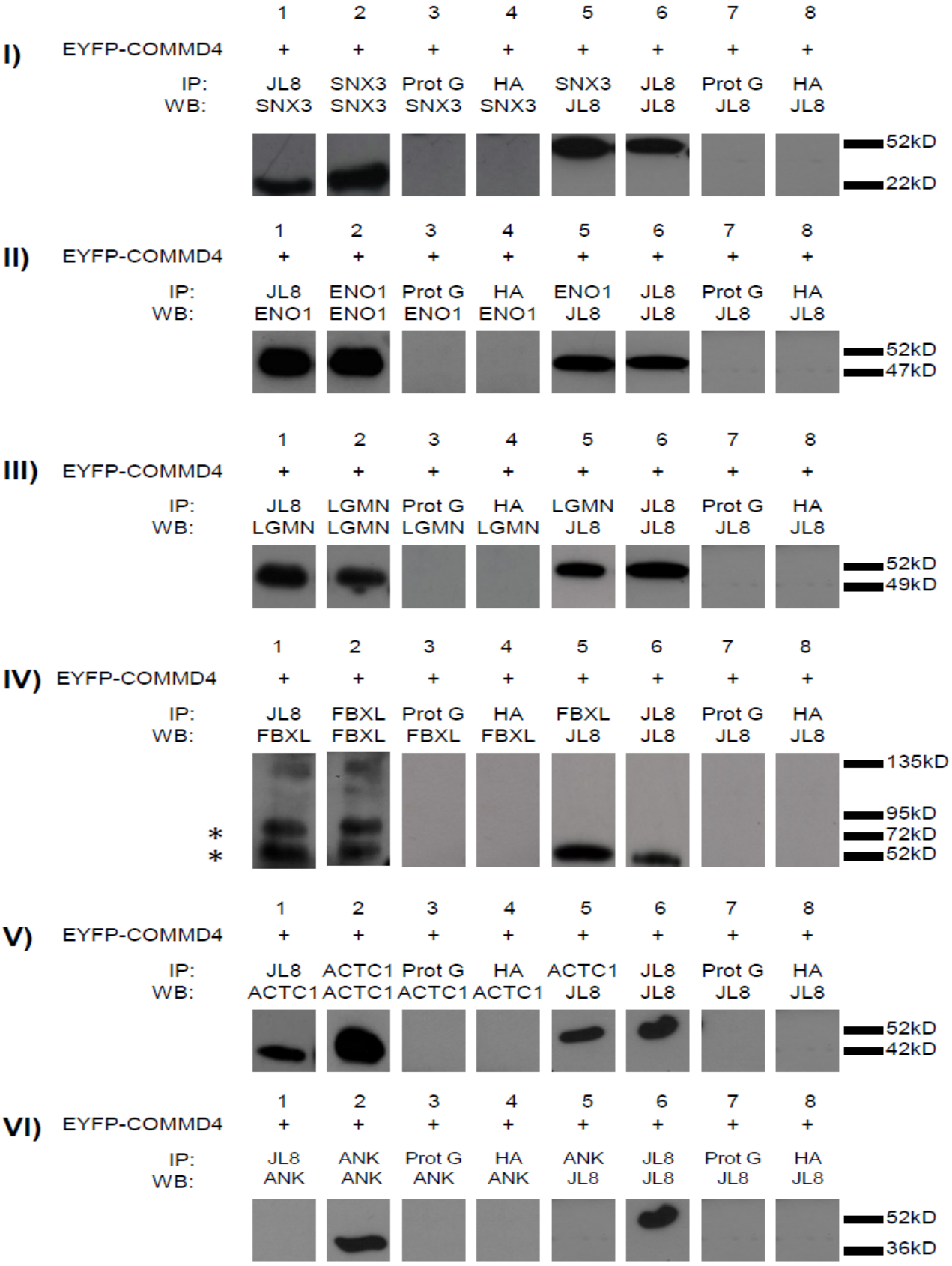


Figure 3.6. In vivo co-immunoprecipitation using H9C2 cell lysates, showing that GFP-cMyBPC specifically pulls down dsRed-tagged COMMD4. Cells were transfected with dsRed-COMMD4 and/or GFP-cMyBPC, as indicated above each lane. Pull-downs (IP) were performed using either the anti-GFP antibody (JL8) or the anti-dsRed antibody (DsRed), and western blots (WB) subsequently stained with either of these antibodies. Anti-haemaglutinin antibody (HA) was used as a negative control antibody (lanes 4 and 9), while a protein G control (Prot G) was employed to ensure that no non-specific binding occurred during pull-downs (lanes 3 and 8). Interestingly, in the reciprocal experiments, only a small band, rather than the band of ~140kD corresponding to GFP-cMyBPC, could be detected using the anti-GFP antibody (lanes 5 and 6), although this larger band was clearly detected in cells not transfected with dsRed-COMMD4 (lane 7). This suggests that cMyBPC undergoes enhanced degradation in the presence of additional COMMD4.

Moreover, western blot analysis of lysates from cells expressing YFP-labelled COMMD4 detected the endogenous proteins SNX3, LGMN, ENO1, FBXL10, ACTC1 and the exogenously expressed DSCR3, but not endogenous ANKRD1 (Figure 3.7). Conversely, when each putative prey protein was immunoprecipitated with their respective antibody, the COMMD4 protein was detected in all prey immunoprecipitates except the ANKRD1 immunoprecipitate (Figure 3.7). No detectable bands were observed in the protein G-agarose controls, indicating specificity of immunoprecipitation. Similarly, no bands were detected when the HA antibody was used in negative control immunoprecipitations reactions. These controls clearly demonstrate that the precipitations are not spurious but the result of physical association between the relevant proteins. These results demonstrate that COMMD4 interacts with cMyBPC, SNX3, LGMN, ENO1, FBXL10, DSCR3 and ACTC1.



Continue on next page

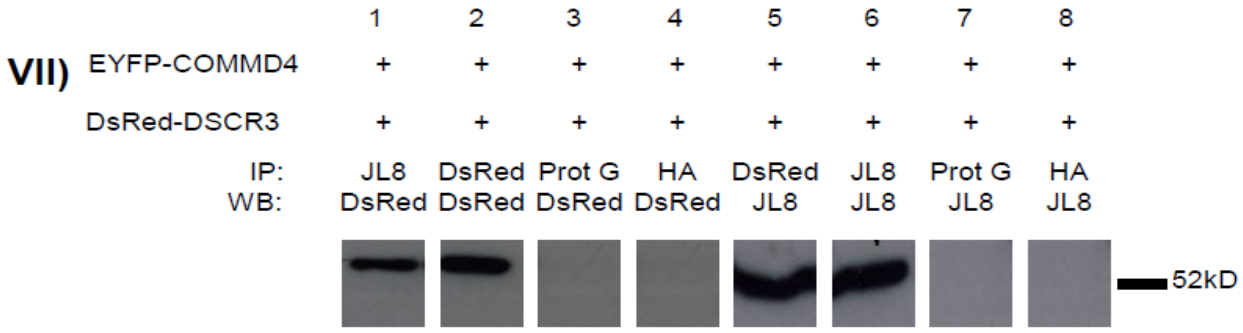


Figure 3.7. *In vivo* co-immunoprecipitation assays of EYFP-tagged COMMD4 and its Y2H-identified putative interactors. HEK293 or H9C2 cells were transfected with EYFP-COMMD4, or EYFP-COMMD4 and DsRed-DSCR3 (VII) as indicated above each lane. Pull-downs (IP) were performed using the anti-EYFP antibody (JL8), anti-dsRed antibody (DsRed), or prey protein-specific antibody as indicated, and Western blots (WB) subsequently stained with either of these antibodies, as indicated above each lane. Anti-haemaglutinin antibody (HA) was used as a negative control antibody (lanes 4 and 8), while a protein G control (Prot G) was employed to ensure that no non-specific binding occurred during pull-downs (lanes 3 and 7). EYFP-tagged COMMD4 could pull down endogenous SNX3 (I), ENO1 (II), LGMN (III), FBXL10 (IV), ACTC1 (V), as well as DsRed-tagged DSCR3 (VII), while these proteins also reciprocally immunoprecipitated EYFP-COMMD4 *in vivo*. However, no interaction was detected between EYFP-COMMD4 and endogenous ANKRD1 (VI). Both the protein G and negative control antibody lanes were clear, indicating that these precipitations were not spurious, but were the result of physical association between the relevant proteins.* denotes the extra bands seen for multiple alternatively-spliced transcripts, which are known to exist for FBXL10.

3.2.3. LIVE-CELL 3D CO-LOCALISATION

Live cell imaging was performed to assess protein-protein interactions between COMMD4 and its putative prey interactors within the 3D organisation of the cell and its subcellular compartments.

Using fluorescence microscopy, 3D co-localisation with Z-stacking further confirmed the results obtained by the pull-down assays. For this assay, either immunocytochemistry (Chapter 2, Section 2.16.3.1) or co-transfection of fluorescently-tagged proteins in H9C2 cells was used. A series of XY images were taken at different depths (Z), in order to map the entire volume of the H9C2 cell. A single frame of the Z-stack images is shown in figure 3.8, with images I-III representing a single colour channel, while image IV represent an overlay of the four colour channels used. Co-localisation with COMMD4 (Figure 3.8) was shown for SNX3, LGMN, ENO1, DSCR3, ACTC1 and FBXL10 but not for ANKRD1.

While monitoring the live cell fluorescence between EYFP-COMMD4 and GFP-SNX3, movement of these two proteins towards the nucleus was also observed. This interesting occurrence appeared to further substantiate SNX3 role in protein trafficking. Thus a movieclip (Appendix IX) was made and images in four different frames were captured and are shown in figure 3.9.

Further analysis of acquired images also revealed that the co-localisation of certain proteins frequently appears as punctate stains in the cytoplasm. As some of the verified COMMD4 interactors (SNX3, DSCR3

and LGMN) function in protein trafficking, this raised the question of whether this staining reflected co-localisation within subcellular structures such as endosomes. In order to address this question, co-localisation assays were performed for the fluorescently labelled COMMD4 and its confirmed interactors with Endo-DsRed, an endosomal marker, along with single transfections and appropriate controls. Co-localisation with Endo-DsRed was shown for COMMD4 (Figure 3.10); similarly, co-localisation was also observed for some of the respective putative interactors of COMMD4 (SNX3, LGMN and DSCR3) and the endosomal marker (Figure 3.10).

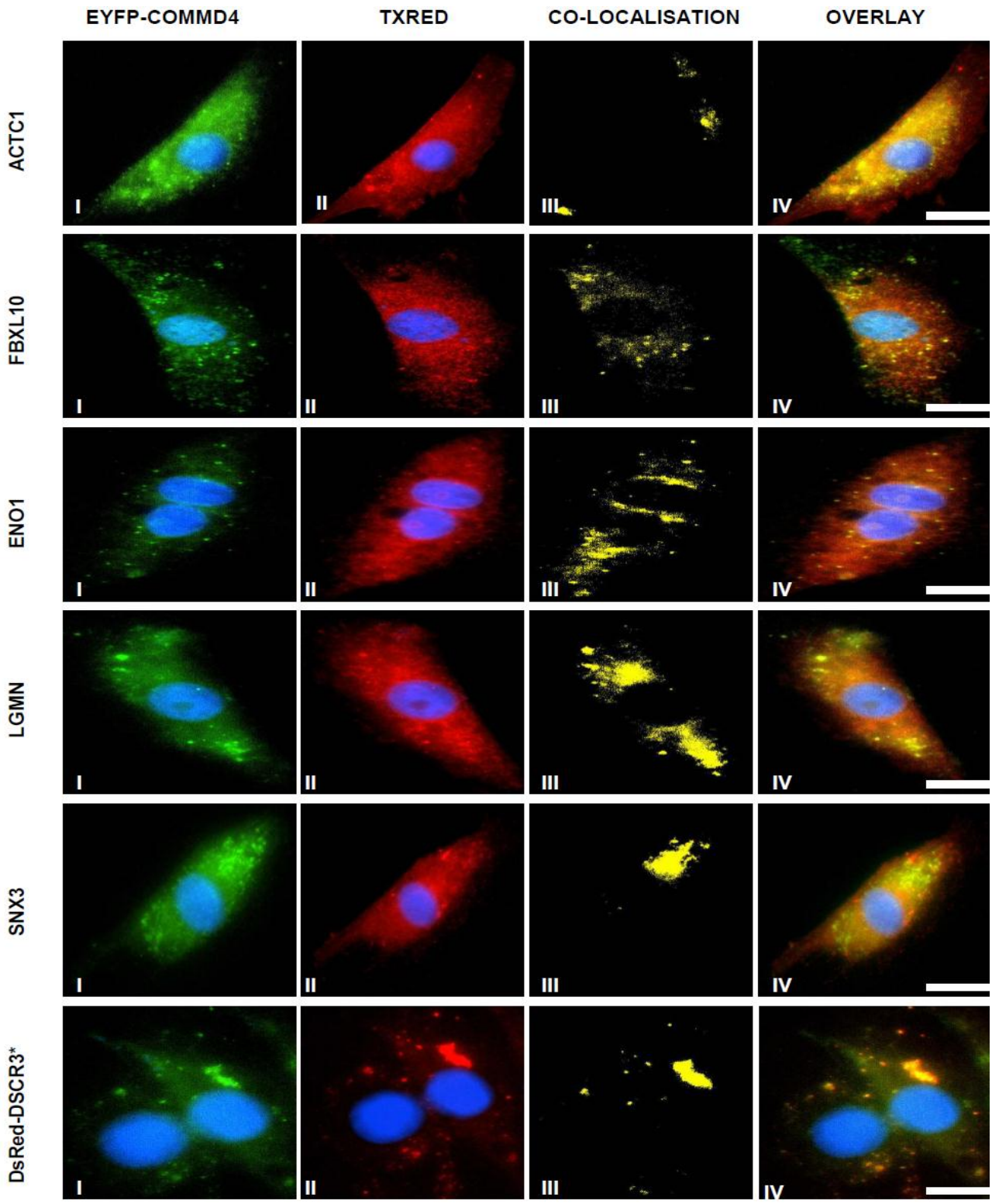


Figure 3.8. 3D Co-localisation of COMMD4 and respective preys identified in the Y2H library screen.
Representative images of fluorescence microscopy showing co-localisation of COMMD4 and its respective interactors in differentiated H9C2 cardiomyocytes. Each panel represents a single frame of the 25 images captured for the vertical Z-stack. The first three panels show a single colour channel, while the last panel shows an overlay of the three colour channels used, with the nuclei stained with DAPI for orientation purposes. Column (I) indicates the expression of EYFP-tagged COMMD4, artificially shown as green fluorescence. Column (II) indicates endogenous prey interactors stained with respective primary antibodies, followed by the relevant TxRed-conjugated secondary antibody (red fluorescence). * For DSCR3, 3D co-localisation was obtained using fluorescently-tagged proteins, as no antibodies against this protein were available. Column (III) shows co-localisation (yellow fluorescence) between EYFP-COMMD4 and each respective prey occurring in the cytoplasmic region of the cell. Scale bar: 0.02mm

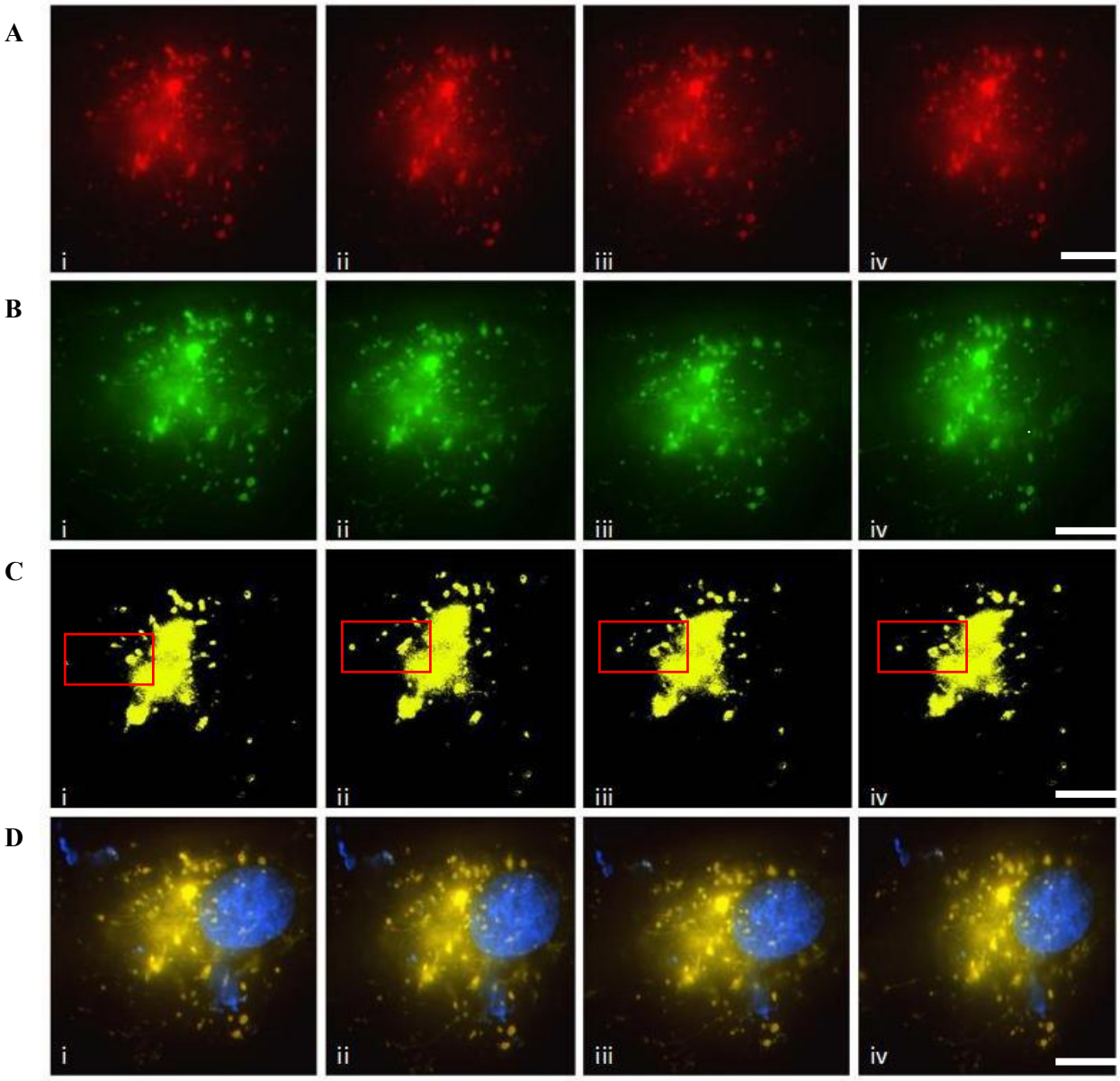


Figure 3.9. Live cell fluorescence imaging of COMMD4 and SNX3 showing different frame points from the movieclip. Each panel represents a single frame of the 25 images captured for the vertical Z-stack Row **(A)** shows differentiated H9C2 cell expressing EYFP-tagged COMMD4 constructs (red fluorescence). Row **(B)** shows the same transfected cell expressing GFP-SNX3 constructs (green fluorescence). Row **(C)** shows co-localisation between COMMD4 and SNX3 (yellow fluorescence). Row **(D)** shows overlay of images with nuclei stained with Hoechst H-33342. Images in rows (**i-iv**) show images captured at different frame points from the movie clip in order to demonstrate the movement of COMMD4 and SNX3 towards the nucleus. (red block highlights point of interest)**.** Scale bar: 0.02mm

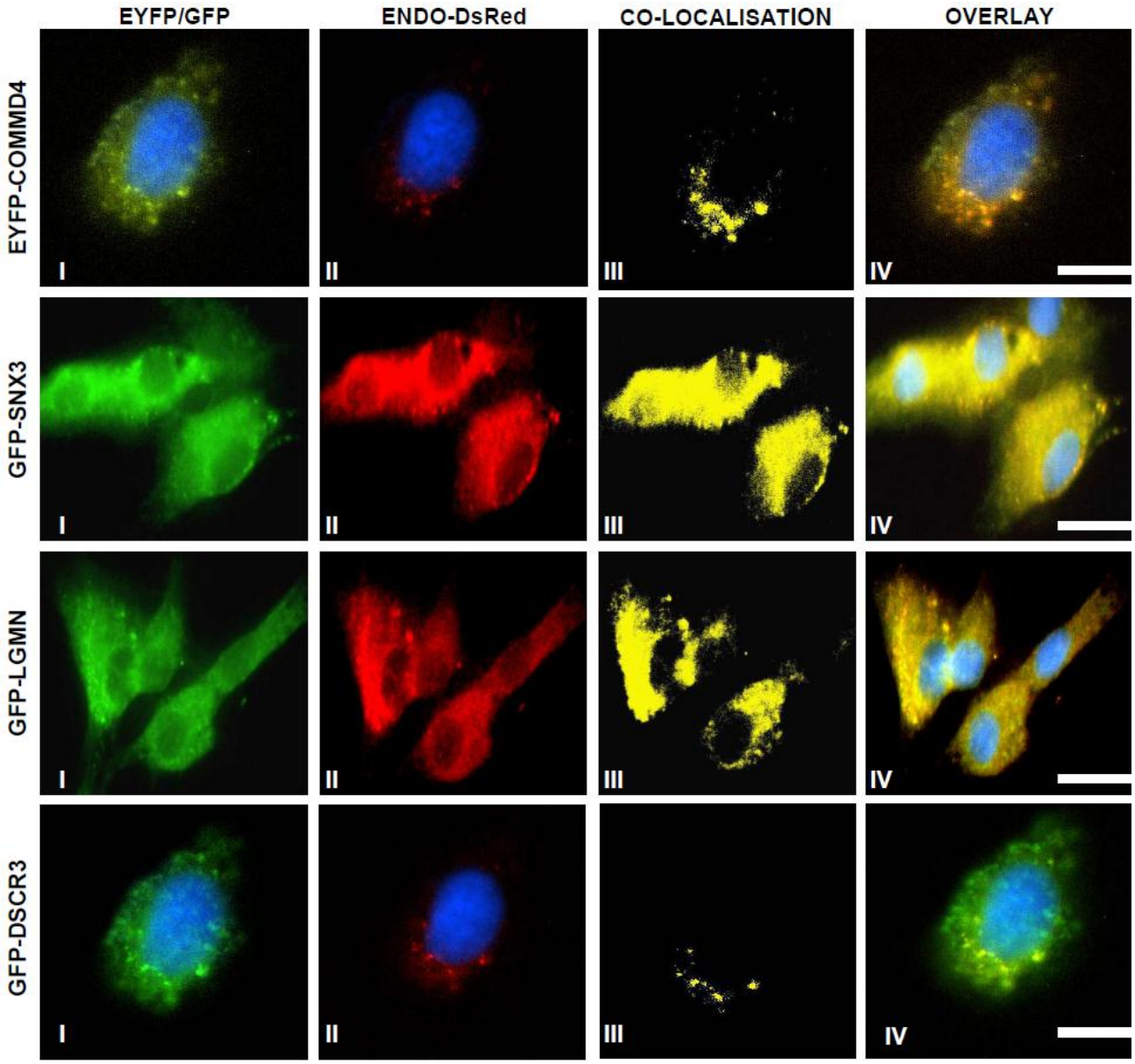


Figure 3.10. Live cell fluorescence imaging of COMMD4, SNX3, LGMN and DSCR3 with an endosomal marker. Representative image of live cell fluorescence microscopy showing co-localisation of EYFP-COMMD4, GFP-SNX3, GFP-LGMN and GFP-DSCR3 with the endosomal marker, RFP-Endo, in differentiated H9C2 cells. Cells in row 1 and 4 were tripled transfected with EYFP-COMMD4, GFP-DSCR3 and RFP-Endo, thus the same cell was used for imaging. Each panel represents a single frame of the 25 images captured for the vertical Z-stack. The first three panels show a single colour channel, while the last panel shows an overlay of the three colour channels used with the nuclei stained with Hoechst H-33342. Column (I) shows representative cells expressing either EYFP-COMMD4 (artificially coloured green in the image) or one of the GFP-tagged prey proteins (green fluorescence), while column (II) shows the same cells expressing DsRed-Endo (red fluorescence). Column (III) shows co-localisation (yellow fluorescence) between COMMD4 or its interactors and the endosomal marker. This data suggests that COMMD4, SNX3, LGMN and DSCR3 all share an endosomal localisation. Scale bar: 0.02mm

3.3. FUNCTIONAL STUDIES

3.3.1. Domain/phosphorylation interaction assay

Previously, Dr. Amsha Ramburan showed that COMMD4 associates with cMyBPC in a phosphorylation-dependant manner, based on her Y2H library screen, direct protein-interaction tests using mimics of all the phosphorylation states of the cMyBPC N-terminus, and quantitative β-galactosidase assays results. She found consistently more interaction between cMyBPC with the N-terminus in the mono or unphosphorylated states, compared to the triphosphorylated state (Figure3.11).

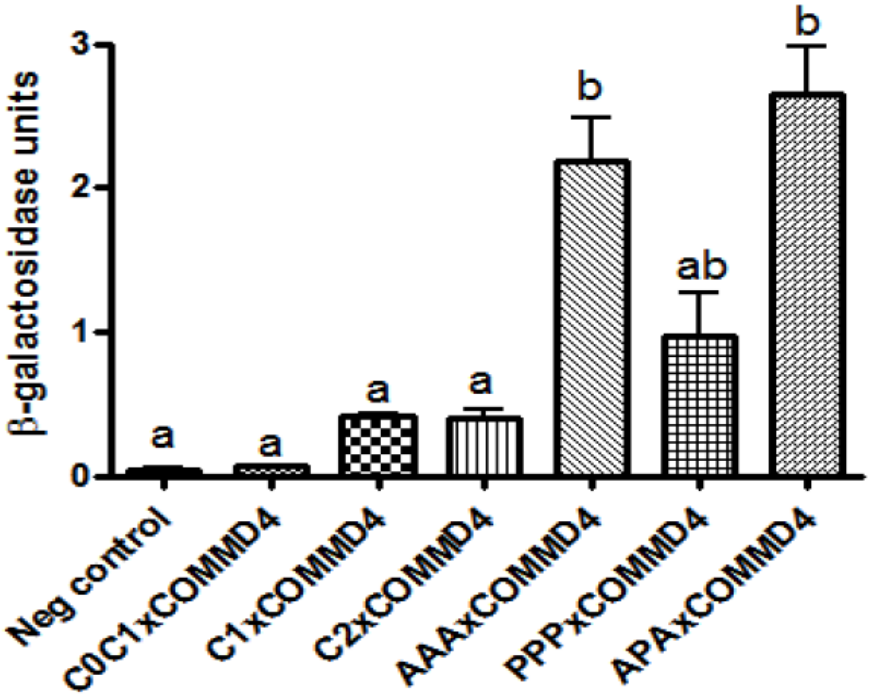


Figure 3.11 Phosphorylation-dependant association between COMMD4 and the cMyBPC-motif.
Bar-graph of quantitative β-galactosidase assay results, comparing the influence of the m-domain, in various phosphorylation states, on the interaction between cMyBPC C1-C2 and COMMD4 with that of the flanking C1 and C2 domains. One unit of β-galactosidase was defined as the amount of enzyme that hydrolyzes 1µmol of OPNG to *O*-nitrophenol and *D*-galactose per minute per cell. Diploid colonies expressing mono (AAA)- or unphosphorylated (PPP) mimics of C1-C2 fragments and COMMD4 generated significantly higher β-galactosidase levels compared to the negative control (non-recomninant pGBKT7), flanking domains C1, C2 and C0-C1, as well as the tri-phosphorylated mimic (PPP), suggesting that interaction between COMMD4 and cMyBPC occurs primarily with the m-domain in a mono- or unphosphorylated state. Combined data of three independent assays, each performed with triplicate samples, were normalized to the mean β-galactosidase levels of the negative control and are presented as Mean±S.E. Different lower case letters above data bars indicate statistically significant differences, while similar letters indicate non-significant differences; Differences were calculated by ANOVA with Bonferonni post-hoc testing; $p<0.05$ was considered statistically significant (A. Ramburan, Phd Thesis, 2008).

In order to assess this further in live mammalian cells, H9C2 cells expressing YFP-COMMD4 were incubated in medium containing 65mM $CaCl_2$ or in 65mM $CaCl_2$ followed by 0.1uM isoproterenol to induce mono- and triphosphorylation conditions (McClellan *et al*., 2001), respectively. Using immunofluorescense microscopy, 3D co-localisation with Z-stacking was performed in triplicate for each condition in two independent experiments and the mean data were pooled together for co-localisation quantification. Cells incubated in the $CaCl_2$-containing medium only, which induces the monophophorylated status (McClellan *et al*., 2001), consistently showed a high degree of co-localisation between COMMD4 and cMyBPC (Figure

3.12 AI) compared to cells treated with both $CaCl_2$ and isoproterenol, which induces the triphosphorylated state (Figure 3.12 AII); however, the differences did not reach statistical significance (Figure 3.12B).

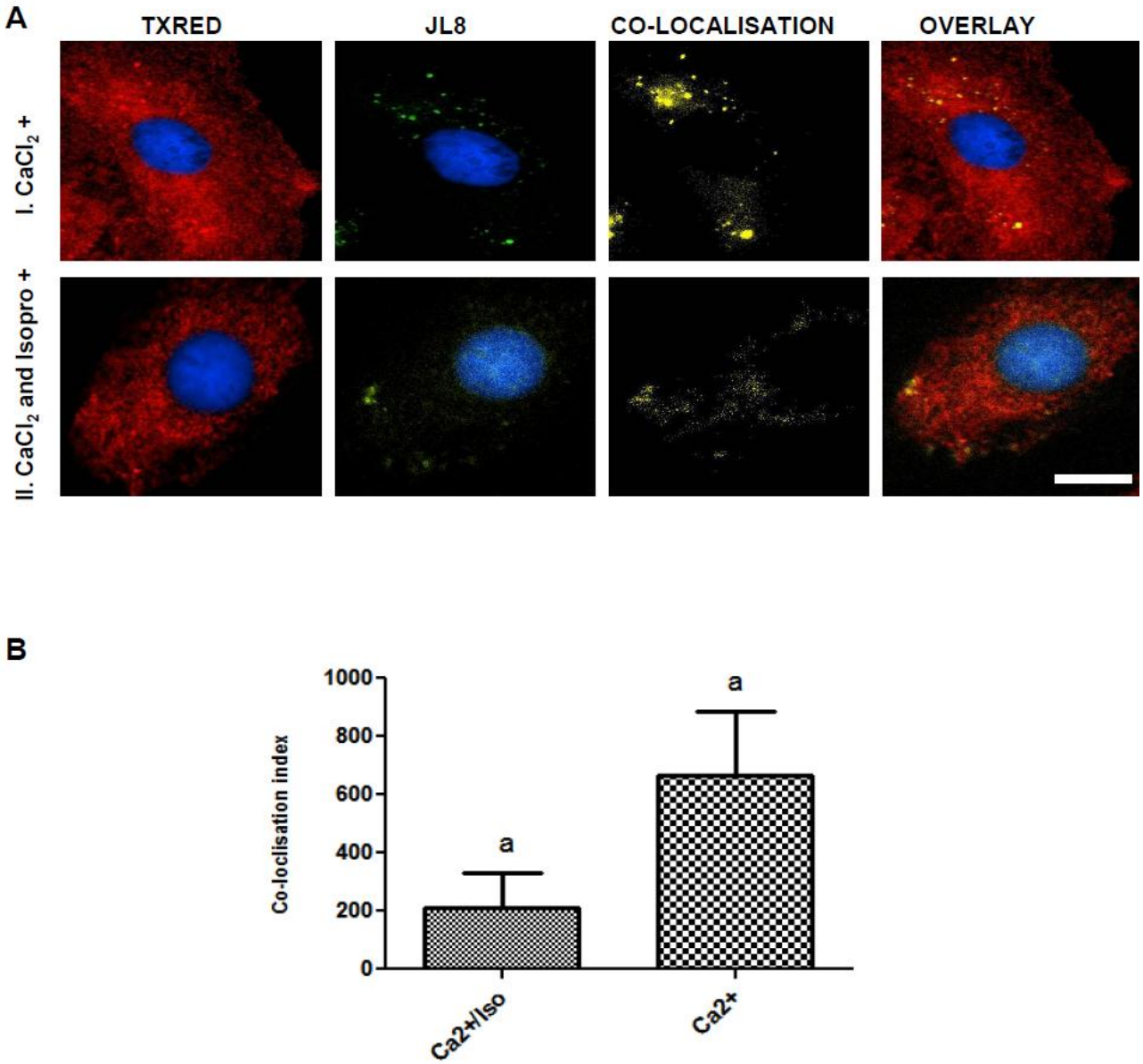


Figure 3.12. Effect of cMyBPC phosphorylation on the interaction between COMMD4 and the cMyBPC motif
H9C2 cells transfected with YFP-COMMD4 were treated with **(AI)** 2.5mM $CaCl_2$ or **(AII)** 2.5mM $CaCl_2$ and 0.1uM Isoproterenol representing mono- and triphosphorylation conditions, respectively**. A**: Representative image of live cell fluorescence microscopy showing that co-localisation of COMMD4 and cMyBPC decreases under conditions of tri-phosphorylation. Each panel represents a single frame of the 25 images captured for the vertical Z-stack. The first three panels shows a single colour channel, while the last panel shows an overlay of the three colour channels used and the nuclei stained with DAPI. Column (I) indicates endogenous cMyBPC stained its respective primary antibody followed by the relevant TxRed-conjugated secondary antibody (red fluorescence). Column (II) indicates the expression of GFP-tagged COMMD4 as green fluorescence. Column (III) shows co-localisation (yellow fluorescence) between GFP-COMMD4 and each respective prey. Scale bar: 0.02mm (**B)** A higher degree of co-localisation was consistently observed in cells treated with $CaCl_2$ only, compared to $CaCl_2$/Isoproterenol treated cells, although the p-value did not reach statistical significance. Data are presented as ±SEM from two independent experiments, each performed in triplicate. Similar letters above error bars indicate non-significance ($p>0.05$) upon unpaired t-testing.

3.3.2. si-RNA-mediated COMMD4 knockdown

As we had thus confirmed the presence of COMMD4 in the endosomes and shown its interactions with proteins involved in protein trafficking and turnover, we next investigated the functional significance of COMMD4 in relation to cMyBPC turnover. The effects of COMMD4 knockdown were investigated in H9C2 cells expressing cMyBPC-GFP and endo-DsRed using live cell fluorescence imaging. GFP-labelled cMyBPC was used to enable following the trafficking of the protein in live cells over time.

Two out of the four siRNAs optimised showed optimal knockdown of COMMD4 as confirmed via qRT-PCR and were subsequently used for downstream experiments. Expression levels of COMMD4 was decreased by 80-85% in H9C2 cells treated with COMMD4 siRNAs, compared to the expression levels in cells treated with a non-silencing control (Figure 3.13).

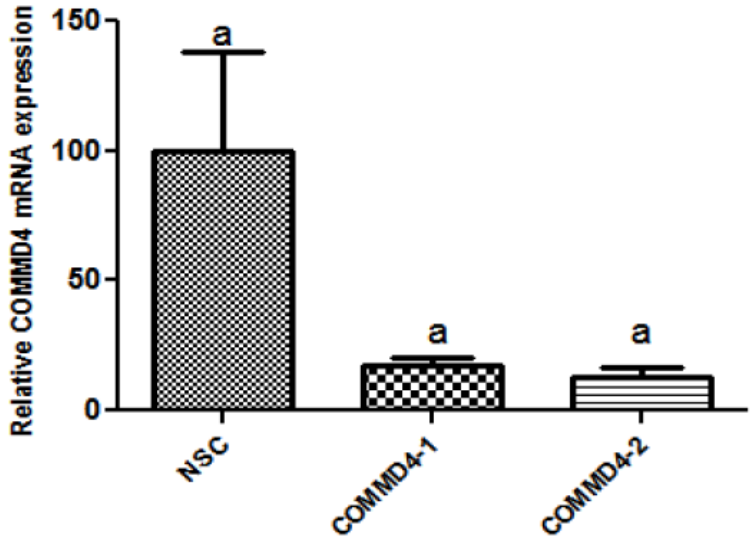


Figure 3.13. qRT-PCR analysis of COMMD4 mRNA expression levels. Bar graph demonstrating real-time quantification of COMMD4 cDNA, transcribed from RNA isolated from differentiated H9C2 cells treated with either 100nM COMMD4 siRNA for 24 hrs, or with a non-silencing control. COMMD4 mRNA levels were normalised according to the expression levels of rodent reference genes: Transferrin receptor (*TRFR*), Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and Heat Shock Protein (*1β-Hsp1β*). COMMD4 siRNA 2 *(Commd4*_Rn_LOC363068_2_HP) resulted in optimal knockdown of *COMMD4* (80-85%). Similar letters above error bars indicate non-significance ($p>0.05$) upon unpaired t-testing

In addition, GFP-cMyBPC co-localised more with Endo-DsRed in COMMD4 knock-down cells (Figure 5B), than in control cells ($p=0.039$, Figure 3.14B), suggesting that accumulation of GFP-MyBPC in the endosomes occurs as a result of the absence of COMMD4, rather than due to overexpression of an exogenous protein.

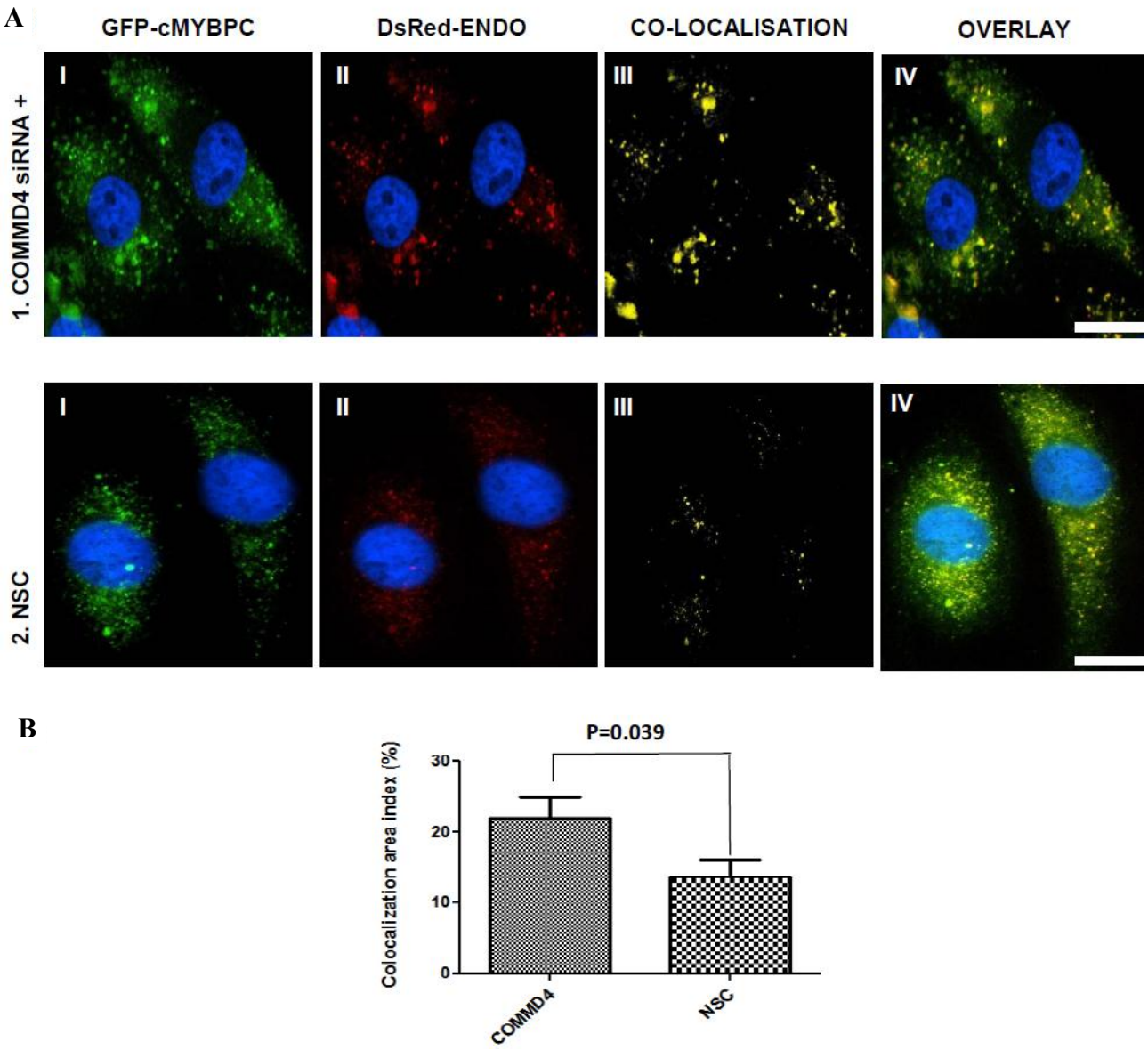


Figure 3.14. siRNA-mediated knockdown of COMMD4 in differentiated H9C2 cells expressing GFP-cMyBPC and RFP-Endo. A: Representative image of live cell fluorescence microscopy showing that co-localisation of cMyBPC and an endosomal marker in differentiated H9C2 cells increases in the (1) absence of COMMD4 (i.e. with COMMD4 knockdown) compared to the (2) presence of COMMD4 (non-silencing control, NSC). Column (I) shows H9C2 cells expressing GFP-cMyBPC (green fluorescence), while column (II**)** shows the same H9C2 cells expressing the endosomal marker, DsRed-Endo (red fluorescence). Column (III) shows co-localisation between cMyBPC and the endosomal marker (yellow fluorescence), while column (IV) shows an overlay of images, with nuclei stained with Hoechst H-33342 for orientation purposes. Scale bar: 0.02mm **B:** Quantification of the co-localisation signal shown in **A** shows that co-localisation of cMyBPC and the endosomal marker RFP-Endo, increased significantlyupon knockdown of COMMD4 ($p<0.05$). Change in co-localisation was calculated using the CellR software and presented as a false colour image. Colocalisation area index is calculated as the co-localisation area indexed to the sum of the EYFP and DsRed areas for each cell layer, and is given as %. Co-localisation data from three independent experiments was pooled together and the mean values ± standard error are portrayed.

Interestingly, knockdown of COMMD4 also affected cell growth and cell morphology (Data not shown). It appeared that knockdown of COMMD4 induced cell death, since a greater number of dead cells were observed between the siRNA treated cells compared to the control. Morphological differences were also observed between the two groups. The phenotype of the H9C2 cells treated with the NSC was thin, elongated and fused to form myotubes as expected. In comparison, only a few number of the COMMD4 siRNA treated

cells showed the elongated phenotype, the rest of the cells displayed a round and flat phenotype suggesting difficulties in differentiation.

3.4. COMPARISON OF Y2H DATA WITH VERIFICATION ASSAYS

To summarise, data from the Y2H assay was compared to data obtain from the verification assays. The interactions of COMMD4 with ACTC1, SNX3, LGMN and ENO1 were supported by all the methods employed. Moreover, the respective interactions of DSCR3 and FBXL10 with COMMD4 were supported by Y2H data, *in vivo* Co-IP experiments and 3D co-localisation data, but could not be assessed via *in vitro* Co-IP experiments. Lastly, despite the various verification assays used, no interaction but that observed in the original Y2H experiment was detected between ANKRD1 and COMMD4.

The data presented, provides convincing evidence that SNX3, LGMN, DSCR3, ENO1, ACTC1, FBXL10 have been identified as specific ligands of COMMD4 in this study. The majority of these proteins are involved in the protein trafficking/turn-over pathways. This, together with the fact that knockdown of COMMD4 leads to accumulation of cMyBPC within the endosomes, suggest that COMMD4 plays a crucial role in the turnover of cMyBPC and possibly other sarcomeric proteins. Since protein turnover has recently been implicated in cardiac hypertrophy development, COMMD4, and by implication, its interactors, were chosen as plausible candidate modifiers of sarcomeric function and contractility, and hence of cardiac hypertrophy and were investigated in clinically well-characterised cohorts of South African HCM families in which either one of three founder-mutations segregate.

3.5. FAMILY-BASED ASSOCIATION STUDY

A total of twenty-six SNPs in five genes (*COMMD4, SNX3, DSCR3, LGMN* and *FBXL10*) in the COMMD4 pathway, identified by Y2H analysis, were assessed for their potential role in hypertrophy development. Comparisons of allele frequencies between affected and unaffected individuals for each SNP tested will be discussed in the following sub sections.

3.5.1. ECHOCARDIOGRAPHIC CHARACTERISTICS OF THE STUDY COHORTS

Echocardiographic characteristics and hypertrophy covariate information of affected and unaffected individuals for the three South African HCM-founder mutations (seven $R92W_{TNNT2}$, three $R403W_{MYH7}$ and eleven $A797T_{MYH7}$) are shown in table 3.7. In this study cohort, 127 individuals had been identified as carriers of an HCM-causing founder-mutation.

3.5.2. SNP SELECTION

As described earlier (Section 3.4), *COMMD4* and its verified interactors *DSCR3, SNX3, LGMN* and *FBXL10* were prioritized as candidate hypertrophy genes. Intragenic SNPs spread throughout the genes were selected for use in the association studies, in order to best capture the spectrum of genetic variation within these candidate genes.

SNPs for each of these respective genes were genotyped using either ASREA (in the case of *COMMD4*) or validated Taqman® SNP genotyping assays (Applied Biosystems, Forster City CA, U.S.A), for the genes encoding putative COMMD4-interactors.

Table 3.7. Echocardiographic characteristics and hypertrophy covariate information of affected and unaffected individuals.

	A797T *MYH7*		R403W *MYH7*		R92W *TNNT2*	
	Affected	Unaffected	Affected	Unaffected	Affected	Unaffected
Total:	56	50	30	21	41	29
Males (%)	32 (57)	27 (54)	19 (63)	8 (38)	18 (44)	11 (38)
Age	43.5 (14-81)	36.5 (14-68)	39.5 (18-72)	35 (16-58)	37 (14-78)	33 (17-60)
BSA (m^2)	1.9 (1.3-2.5)	1.9 (1.5-2.3)	1.8 (1.3-2.3)	2.0 (1.5-2.3)	1.7 (1.3-2.2)	1.8 (1.3-2.1)
SBP (mm Hg)	120 (100-180)	120 (90-180)	122.5 (100-170)	125 (100-230)	120 (90-200)	120 (90-160)
DBP (mm Hg)	80 (60-100)	80 (60-120)	80 (60-110)	80 (60-110)	77.5 (60-100)	80 (60-90)
HR (bpm)	65 (44-120)	65 (50-95)	65 (52-96)	72 (53-90)	68 (45-110)	70 (45-83)
LVM (g)	200.8 (71.4-476.6)	133.3 (61.1-272)	173.9 (77.4-273.8)	141.9 (59.1-290.5)	133.7 (48.2-295.4)	113.7 (58.4-246.1)
IVSTmit (mm)	14 (8-34)	9.6 (5.6-20.5)	13.1 (7.2-27)	10.7 (6.6-16.1)	13.8 (6.3-27)	9.3 (6.8-16.5)
IVSTpap (mm)	15.4 (8-38.7)	9.6 (6.5-27)	13.4 (7.7-27.8)	10.5 (7-18.5)	13.7 (6.7-30)	8.6 (6.8-15.5)
IVSTapx (mm)	12.1 (8.3-37.4)	9.3 (5.6-20.6)	10.7 (7.2-24.9)	10.1 (6.8-12.6)	10.4 (6.5-25.6)	8.7 (6.3-11.9)
PWmit (mm)	9.1 (6-15.1)	8.6 (5.4-13)	8.8 (6-13.7)	8.8 (5.4-13.6)	8.1 (5.7-13.2)	7.5 (4.3-9.6)
PWpap (mm)	9.8 (6.3-16)	8.6 (5-15)	9.3 (6.5-16)	8.8 (5.2-11)	8.5 (6.6-12.7)	7.6 (5.6-10)
PWapx (mm)	9.6 (6.7-17.5)	8.3 (6.2-13.5)	10.1 (6.7-17.9)	9.3 (6-10.3)	9 (6.1-15.3)	7.8 (5.6-10.3)
Maron score	59 (40;70)	37 (35;40)	48 (35;60)	35 (32;39)	49 (41;58)	41 (37;47)
CWT score	196 (151;239)	139 (127;148)	159 (131;192)	131 (117;142)	172 (148;190)	153 (136;166)
Comp1	-3.36 (-8.26;4.72)	1.31 (-4.04;7.9)	-0.53 (-6.44;7.72)	2.53 (-1.86;7.19)	-1.61 (-6.21;5.16)	-0.30 (-2.65;7.88)

*Data summarised as median (Q_1, Q_3) Abbreviations: BSA-body surface area; DBP-diastolic blood pressure; HCM-hypertrophic cardiomyopathy; HR-heart rate; LVM-left ventricular mass; MC-HCM mutation carrier; IVST- interventricular septum thickness; IVSTmit-interventricular septum thickness at the mitral valve; IVSTpap-interventricular septum thickness at the papillary level; NC-non-carrier; PWapx-posterior wall thickness at the supra-apex level; PWmit-posterior wall thickness at the mitral valve; PWpap-posterior wall thickness at the papillary level; SBP-systolic blood pressure, CWT-cumulative wall thickness score and Comp1-principle component score1.

3.5.3. GENOTYPING

3.5.3.1 *COMMD4* ASREA

PCR amplification of the three *COMMD4* SNPs was successful and yielded fragments of expected sizes as shown in table 2.10 (Chapter 2, Section 2.19.3).

COMMD4 SNP, rs 3809547 (C/G), contains a invariant *MboII* restriction enzyme site, thus digestion yielded two fragments (176bp + 79bp) for the C allele, while the G allele caused the restriction site to be lost and only the 255bp fragment was seen in figure 3.15.

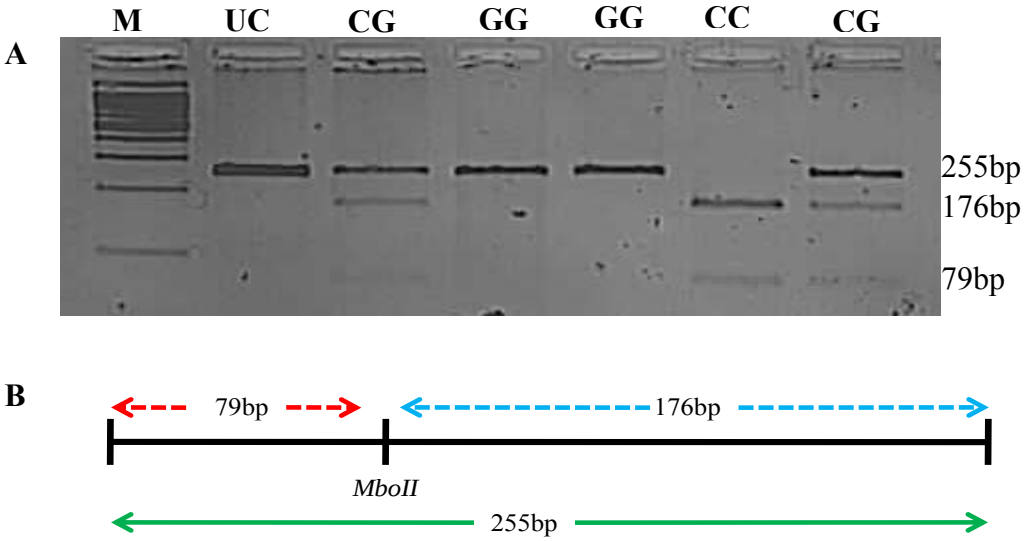


Figure 3.15. ASREA of *COMMD4*/rs 3809547 polymorphism (C/G). A. Representative 2% agarose gel showing the fragment sizes of *COMMD4*/ rs 3809547 amplicon following digestion with restriction enzyme, *MboII*. M represents a 100bp molecular marker, UC represent an undigested control. The 255bp fragment represents the G allele, whereas fragments 176bp and 79bp represent the C allele. The CG heterozygote is represented by all 3 fragments. **B.** Schematic representation of *COMMD4*/ rs 3809547 amplicon outlining the position of the *MboII* restriction site and the fragment sizes generated following digestion.

COMMD4 SNP, rs 11853141 (C/T) contained an invariant *NheI* restriction enzyme site, thus, digestion yielded two fragments (234bp + 122bp) for the C allele; the T allele caused the restriction site to be lost and yielded a fragment of 356bp (Figure 3.16).

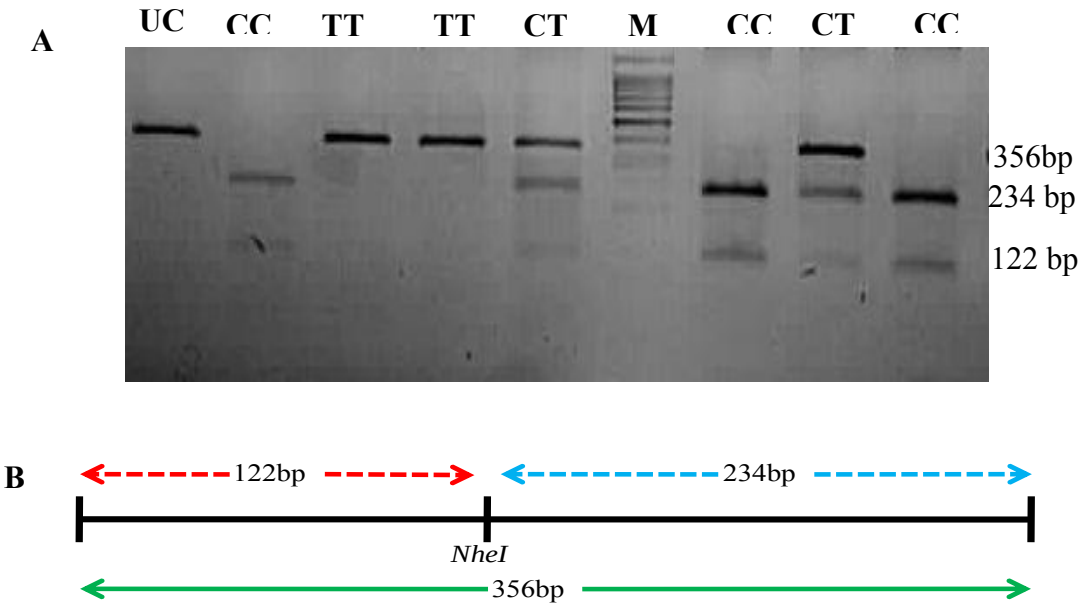


Figure 3.16. ASREA of *COMMD4*/rs 1185314 polymorphism (C/T). A. Representative 2% agarose gel showing the fragment sizes of *COMMD4*/rs 1185314 amplicon following digestion with restriction enzyme, *NheI*. M represents a 100bp molecular marker, UC represents an undigested control. The 356bp fragment represents the T allele, whereas fragments 234bp and 122bp represent the C allele. The CT heterozygote is represented by all 3 fragments. **B**. Schematic representation of *COMMD4*/rs 1185314 amplicon outlining the position of the *NheI* restriction site and the fragment sizes generated following digestion.

The last *COMMD4* SNP, rs 11072542 (A/G) contained an invariant *DraII* restriction enzyme site, thus digestion yielded two fragments (223bp + 114bp) for the G allele, while the T allele caused the restriction site to be abolished, thus, the product remained undigested and yielded a 337 bp product (Figure 3.17).

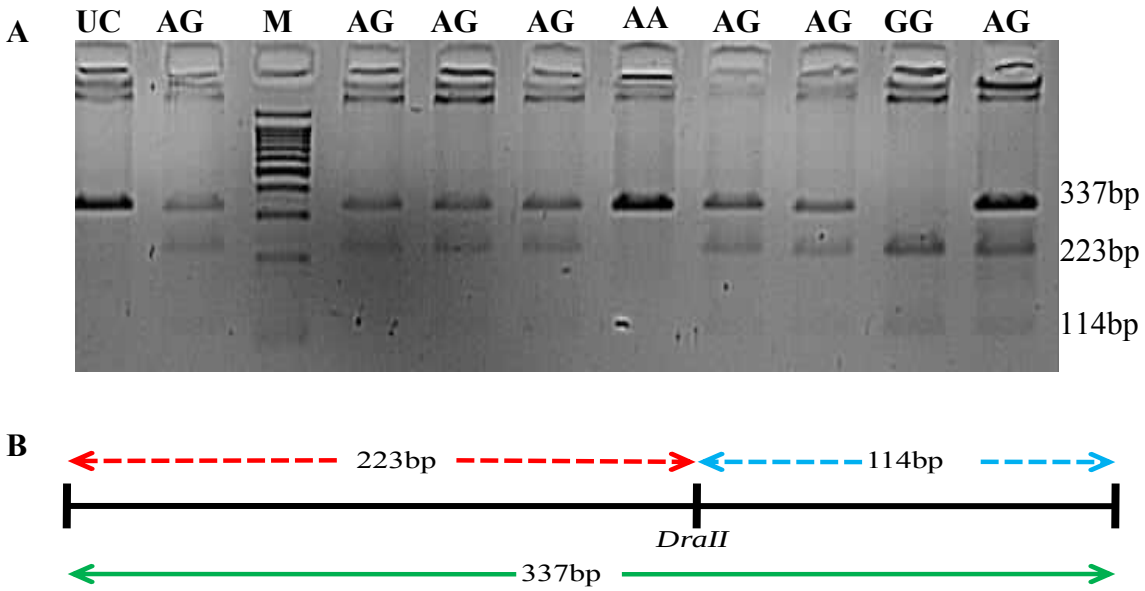


Figure 3.17. ASREA of *COMMD4*/ rs 11072542 polymorphism (A/G). A. Representative 2% agarose gel showing the fragment sizes of *COMMD4*/rs 11072542 amplicon following digestion with restriction enzyme, *DraII.* M represents a 100bp molecular marker, UC represents an undigested control, The 337bp fragment represents the A allele, whereas fragments 223bp and 114bp represents the G allele. The AG heterozygote is represented by all 3 fragments. **B**. Schematic representation of *COMMD4*/rs 11072542 amplicon outlining the position of the *DraII* restriction site and the fragment sizes generated following digestion.

3.5.3.2. TAQMAN® ALLELE DISCRIMINATION

TaqMan® PCR amplifications were successfully completed for all assays and no contamination was observed in the non-template controls. Figure 3.18 depicts a representative Taqman® genotype output that of *SNX3* rs579257, to illustrate the allelic discrimination plots produced during end-point analysis.

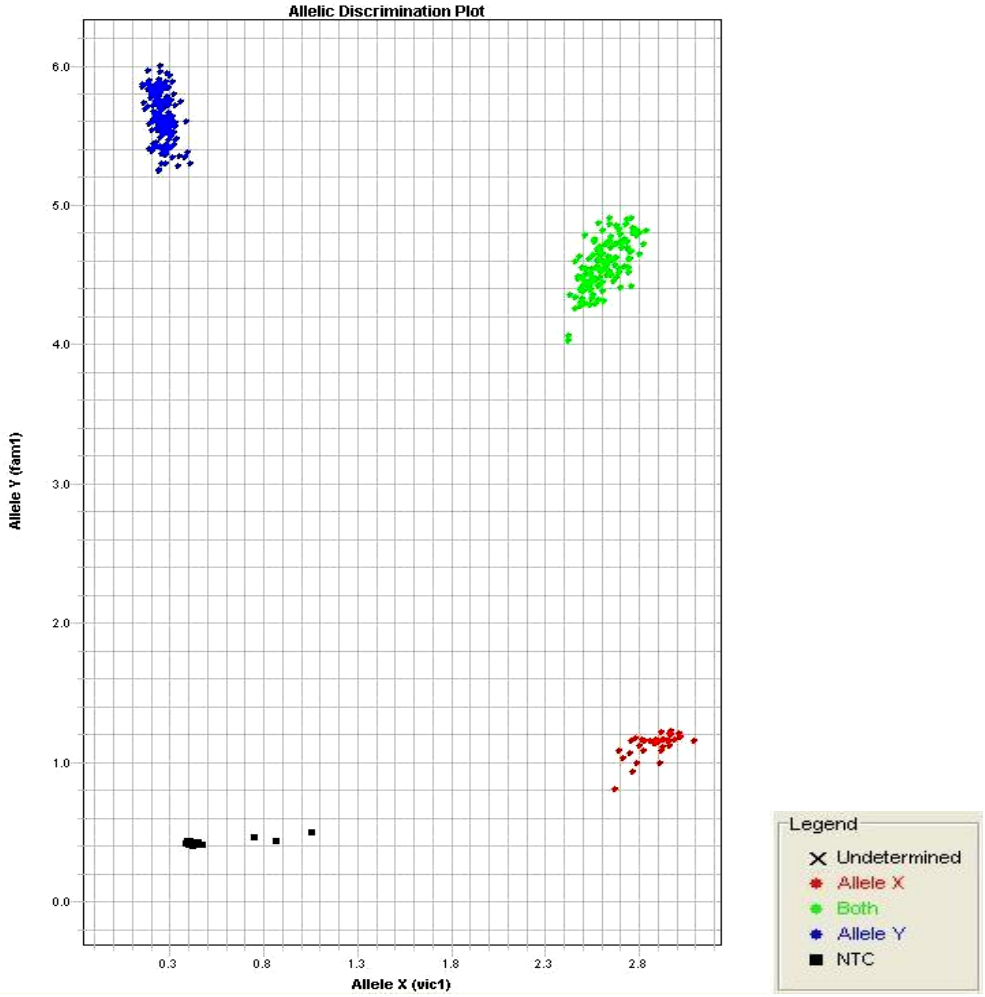


Figure 3.18. Representative TaqMan allelic discrimination plot. Taqman genotyping results of rs579257 in *SNX3* are shown; allele X represents the G-allele and allele Y represents the T-allele at this SNP.

3.5.4. STATISTICAL ANALYSIS

Pedstats was used to identify inconsistencies in Mendelian inheritance patterns for all genotypes. Those genotypes that were shown to be inconsistent were re-genotyped and included in the subsequent analysis. A Hardy-Weinberg Equilibrium (HWE) test among unrelated individuals within the cohort was performed for all SNPs and genotype frequencies for all the markers were in agreement with HWE.

Graphical outputs of allele frequencies for each marker generated by Pedstats are shown in Appendix VIII, which depicts the allele frequencies of each marker for the whole study cohort (blue) as well as for unrelated individuals (green). The X-axis represents the allele numbers, whereas the Y-axis represents the allele frequencies. Comparisons of minor allele frequencies of each SNP for each gene as observed in our population compared to populations used in HapMap are shown in table 3.8.

Table 3.8. Comparisons of minor allele frequencies observed in study cohort compared to other populations

		Minor allele frequencies			
	SNP	**Study cohort**	**CEU**	**YRI**	**JPT + HCB**
COMMD4	rs3809547	0.32 (C)	0.19 (C)	-	0.3 (C)
	rs11853141	0.34 (C)	0.49 (T)	0.18 (T)	0.46 (T)
	rs1107254	0.28 (G)	0.21 (G)	-	-
SNX3	rs579257	0.33 (G)	0.09 (G)	0.4 (T)	0.15 (G)
	rs628424	0.34 (A)	0.05 (A)	0.38 (C)	0.1 (A)
	rs588409	0.26 (C)	0.33 (C)	0.19 (C)	0.06 (C)
	rs635051	0.38 (C)	0.13 (C)	0.37 (C)	0.15 (C)
	rs9398166	0.36 (C)	0.1 (C)	0.36 (A)	0.15 (C)
FBXL10	rs12427382	0.27 (A)	0.37 (A)	0.09 (A)	0.33 (A)
	rs3751131	0.26 (A)	0.43(A)	0.09 (A)	0.36 (A)
	rs6489811	0.38 (A)	0.49 (A)	039 (A)	0.41(A)
	rs11065585	0.23 (C)	0.44 (C)	0.03 (C)	0.40 (C)
	rs4980986	0.26 (T)	0.48 (T)	0.02 (T)	0.42 (T)
	rs7307400	0.32 (G)	0.08 (G)	0.38 (A)	-
	rs10849893	0.33 (G)	0.38 (G)	-	-
	rs7313015	0.06 (T)	0.03 (T)	0.34 (T)	0.15 (T)
LGMN	rs2402189	0.30 (A)	0.45 (A)	0.22 (A)	0.26 (A)
	rs3783933	0.27 (T)	-	-	-
	rs4904980	0.35 (T)	-	-	-
	rs1242102	0.50 (C)	0.48 (T)	0.34 (C)	0.20 (C)
	rs1242095	0.48 (C)	0.23 (C)	0.26 (C)	0.35 (C)
DSCR3	rs3165	0.06 (C)	0.2 (C)	0.07 (C)	0.07 (C)
	rs2835680	0.34 (T)	032 (T)	0.44(T)	-
	rs9305614	0.38 (G)	0.41(G)	0.39(G)	0.49 (A)
	rs2835684	0.50 (G)	0.49(A)	-	-
	rs3787802	0.35 (A)	0.35(A)	0.22(A)	0.22(A)

Abbreviations: COMMD4–COMM domain containing protein 4, DSCR3-Down Syndrome critical region 3, FBXL10-F-Box Leucine rich repeat protein 10, MAF-minor allele frequency, SNX3–Sorting nexin 3, SNP-single nucleotide polymorphism, bp-base pair

3.5.4.1. LINKAGE DISEQUILIBRIUM DETERMINATION

The Haploview™ software package was used to analyse pairwise measurements of linkage disequilibrium (LD) between markers for each respective gene. As mentioned in section 2.23.2, LD was measured in terms of D' values and each pair of SNPs was estimated as either being in strong LD, uncertain or incomplete LD.

LD plots generated for the study cohort are shown in figures 3.19-3.23. Table 3.9 summarises the D prime (D') values within the blocks, which represents the degree of LD between markers for each gene.

Values of D' was specified by the following key: A red block represents a D' value close to or equal to 1 with a log of likelihood ratio (LOD) of ≠ 2, which represent perfect LD, blue also represents D'= 1, but with a LOD < 2, lighter shades represents decreasing D' values whereas white is indicative of low D' values which represents a degree of weak or incomplete LD.

Table 3.9. Summary of observed LD (shown by D') between relevant markers for each gene in the study cohort.

Gene	Marker1	Marker2	D'	Gene	Marker1	Marker2	D'
COMMD4	rs3809547	rs11853141	0.69	*FBXL10*	rs12427382	rs3751131	1.00
	rs3809547	rs11072542	0.92		rs12427382	rs6489811	1.00
	rs11853141	rs11072542	0.50		rs12427382	rs11065585	0.92
SNX3	rs579257	rs628424	0.87		rs12427382	rs4980986	0.67
	rs579257	rs588409	1.00		rs12427382	rs7307400	0.41
	rs579257	rs635051	1.00		rs12427382	rs10849893	0.64
	rs579257	rs9398166	1.00		rs12427382	rs7313015	1.00
	rs628424	rs588409	1.00		rs3751131	rs6489811	1.00
	rs628424	rs635051	1.00		rs3751131	rs11065585	0.92
	rs628424	rs9398166	1.00		rs3751131	rs4980986	0.72
	rs588409	rs635051	1.00		rs3751131	rs7307400	0.40
	rs588409	rs9398166	1.00		rs3751131	rs10849893	0.63
	rs635051	rs9398166	1.00		rs3751131	rs7313015	1.00
DSCR3	rs3165	rs2835680	0.08		rs6489811	rs11065585	1.00
	rs3165	rs9305614	0.23		rs6489811	rs4980986	0.56
	rs3165	rs2835684	1.00		rs6489811	rs7307400	0.89
	rs3165	rs3787802	1.00		rs6489811	rs10849893	0.91
	rs2835680	rs9305614	1.00		rs6489811	rs7313015	0.60
	rs2835680	rs2835684	0.10		rs11065585	rs4980986	1.00
	rs2835680	rs3787802	0.49		rs11065585	rs7307400	1.00
	rs9305614	rs2835684	0.48		rs11065585	rs10849893	0.46
	rs9305614	rs3787802	0.08		rs11065585	rs7313015	1.00
	rs2835684	rs3787802	1.00		rs4980986	rs7307400	1.00
LGMN	rs2402189	rs3783933	1.00		rs4980986	rs10849893	0.64
	rs2402189	rs4904980	0.68		rs4980986	rs7313015	1.00
	rs2402189	rs1242102	0.23		rs7307400	rs10849893	1.00
	rs2402189	rs1242095	0.28		rs7307400	rs7313015	1.00
	rs3783933	rs4904980	0.33		rs10849893	rs7313015	1.00
	rs3783933	rs1242102	0.26				
	rs3783933	rs1242095	0.65				
	rs4904980	rs1242102	0.42				
	rs4904980	rs1242095	0.04				
	rs1242102	rs1242095	0.32				

Abbreviations: COMMD4 – COMM domain containing protein 4, DSCR3 - Down Syndrome critical region 3, D' – measurement of LD, LD- Linkage disequlibrium, FBXL10- F-Box Leucine rich repeat protein 10, MAF- minor allele frequency, SNX3 – Sorting nexin 3, SNP - single nucleotide polymorphism, bp-base pair

Haploview analysis of *COMMD4* SNPs, showed that strong LD was observed between rs3809547 and rs11072542 with a D' value of 0.92. In comparison, incomplete LD was observed between rs3809547 and rs11853141 (D'= 0.69) and rs11853141 and rs11072542 (D'= 0.50) (Figure 3.19).

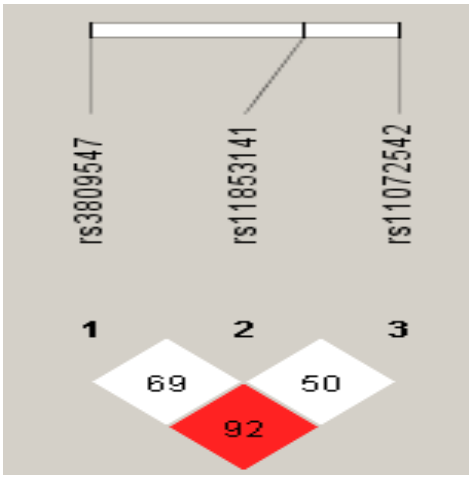


Figure 3.19. **Linkage disequilibrium plot of SNPs within *COMMD4*.** D' values within blocks represent the degree of LD. A blank red or light blue block represents a D' value of 1 and indicates complete LD. Lighter shades depicting low D' values is indicative of weak LD.

Strong LD was observed between the markers of SNX3, and haplotype analysis revealed an 8kb haplotype block between markers rs635051 and rs9398166 (Figure 3.20). D' values of 1 was observed between all markers, except between rs579257 and rs628424 (D'= 0.87).

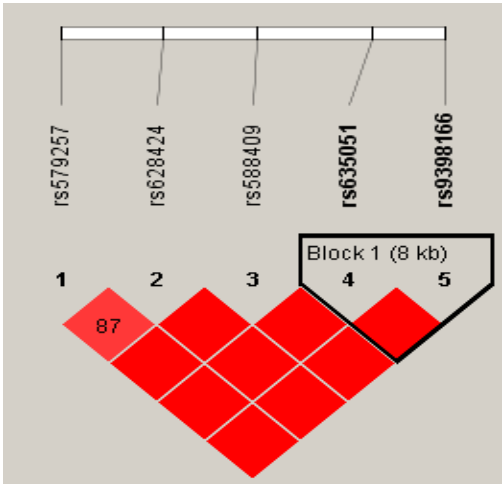


Figure 3.20. **Linkage disequilibrium plot block of SNPs within *SNX3*** D'values within blocks represents the degree of LD. A blank red or blue block represents a D' value of 1, indicating complete LD. Lighter shades depicting low D' values is indicative of weak LD.This LD plot shows a 8kb haplotype block indicative of complete LD between markers rs635051 and rs9398166 within this gene.

Figure 3.21 depicts the LD plot for *DCSR3*, which shows complete LD between rs2835680 and rs9305614 and between rs2835684 and rs3787802, each with a D'= 1.The light blue blocks also indicate complete LD (D'= 1) between markers rs3165 and rs2835680, and between rs3165 and rs3787802. Weak evidence of LD, i.e. incomplete LD, was observed between the rest of the SNPs within *DSCR3* as indicative of low D' values as observed in LD plot blocks.

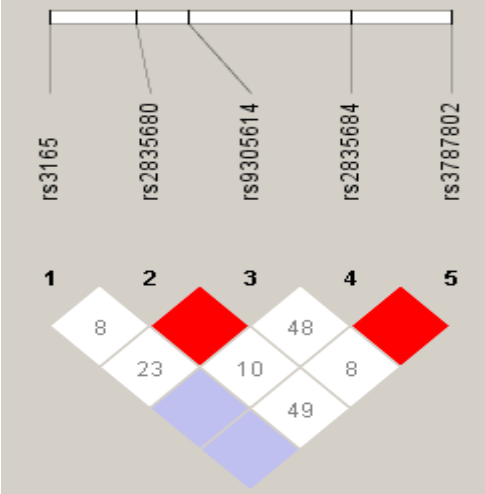


Figure 3.21. **Linkage disequilibrium plot of SNPs within *DSCR3.*** D'values within blocks represents the degree of LD. A blank red or blue block represents a D' value of 1, indicating complete LD. Lighter shades depicting low D' values is indicative of weak LD.

Haploview analysis of *LGMN* markers, shows complete LD between rs2402189 and rs3783933 with a D' value of 1. In contrast, incomplete LD was observed between the rest of the markers and their respective D' values is depicted in the LD plot blocks (Figure 3.22).

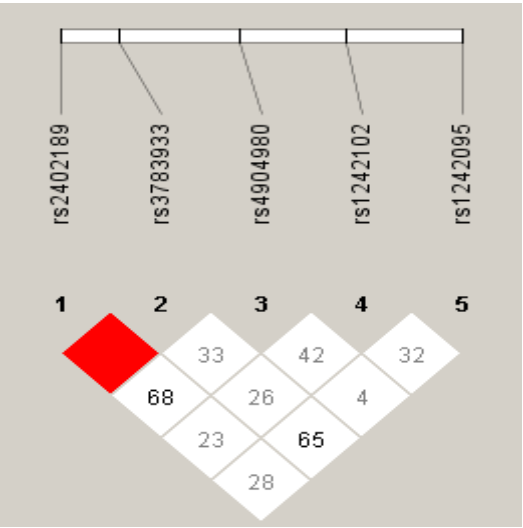


Figure 3.22. **Linkage disequilibrium plot of SNPs within *LGMN*.** D'values within blocks represents the degree of LD. A blank red or blue block represents a D' value of 1, indicating complete LD. Lighter shades depicting low D' values is indicative of weak LD.

Complete to strong LD was observed between most of the SNPs within the gene, *FBXL10*. In comparison, weak LD was observed between markers rs12427382 and rs7307400 (D'= 0.41) and rs10849893 (D'= 0.64); between rs3751131 and rs7307400 (D'= 0.40) and rs10849893 (D'= 0.63); between rs6489811 and rs4980986 (D'= 0.56) and rs7313015 (D'= 0.60); between rs11065585 and rs10849893 (D'= 0.46) and between rs4980986 and rs10849893 (D'= 0.64). In addition, hapotype analysis revealed 2 haplotype blocks of 9 and 10kb each, between rs12427382 and rs3751131 and between rs11065585 and rs4980986, respectively (Figure 3.23).

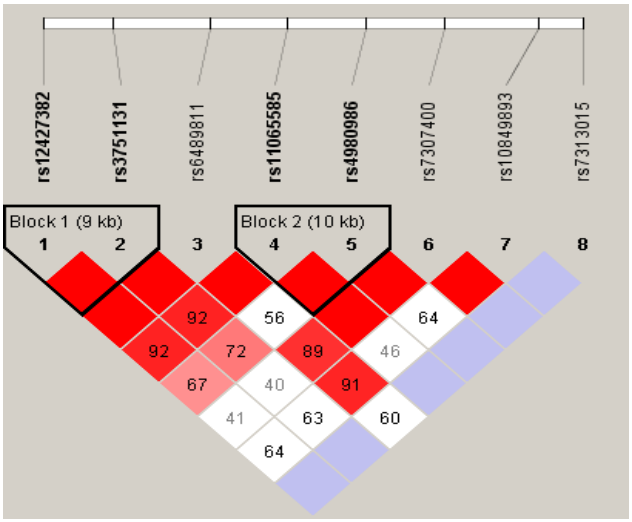


Figure 3.23. **Linkage disequilibrium plot and haplotype block of SNPs within *FBXL10*.**D'values within blocks represents the degree of LD. A blank red or blue block represents a D' value of 1, indicating complete LD. Lighter shades depicting low D' values is indicative of weak LD.This LD plot also shows 2 haplotype blocks of 9 and 10 kb each between markers, indicative of complete LD between the respective markers within this gene.

3.5.4.2. ASSOCIATION TEST

For all association analysis, a p-value less than 0.05 were considered significant. All markers of *COMMD4*, *SNX3*, *DSCR3, LGMN* and *FBXL10* were assessed under an additive model of association. Monte-Carlo permutation tests were performed to determine exact p-values for each marker within each gene. QTDT was first used to test for population stratification. Where population stratification was observed, the orthogonal model which employs within-family association and is therefore robust to population stratification was used. In the absence of evidence for population stratification, the total model, which employs both within and between family associations, and therefore utilizes more of the available data, was used. P-values of the population stratification tests are shown in tables 3.10, 3.12, 3.15, 3.17 and 3.19. P-values for association, derived using either the total model or the orthogonal model, as appropriate, are shown in tables 3.11, 3.13, 3.16, 3.18 and 3.20. In these tables, values between brackets indicate the number of informative genotyped individuals.

Furthermore, where association was found, non-quantile normalised (untransformed) data, adjusted for covariates, were analysed in order to obtain informative effect sizes in terms of the original measurement.

3.5.4.2.1 *COMMD4*

Table 3.10 and 11 summarizes the QTDT results between the three markers of *COMMD4* and the various hypertrophy indices. No significant evidence of population stratification was found. Consequently, the total model was used, but no association was observed.

Table 3.10 Results from population stratification analysis of COMMD4. Values in brackets indicate the number of informative individuals.

	COMMD4		
	rs3809547	**rs11853141**	**rs11072542**
LVM	0.536 (109/181)	0.663 (110/181)	0.633(100/181)
mIVST	0.880 (111/187)	0.555 (115/187)	0.616 (101/187)
mLVWT	0.852 (111/186)	0.975 (115/186)	0.569 (101/186)
mPWT	0.271 (111/186)	0.316 (115/186)	0.663 (101/186)
Maronscore	0.892 (108/179)	0.756 (108/179)	0.771 (98/179)
CWTscore	0.740 (107/172)	0.282 (195/172)	0.906 (96/172)
Comp1	0.729 (109/167)	0.644 (101/167)	0.926 (95/167)

Abbreviations:; Comp1-Principle component score; COMMD4-COMM domain containing protein 4;CWTscore-cumulative wall thickness score, LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness

Table 3.11. Results of *COMMD 4* association analysis using the total model.

P-values for association between rs3809547, rs11853141 and rs11072542 and hypertrophy indices. Values in brackets indicate the number of informative individuals.

	COMMD4		
	rs3809547	**rs11853141**	**rs11072542**
LVM	0.249 (181)	0.500 (181)	0.679 (181)
mIVST	0.111 (187)	0.532 (187)	0.690 (187)
mLVWT	0.175 (186)	0.559 (186)	0.816 (186)
mPWT	0.2235 (186)	0.395 (186)	0.752 (186)
Maronscore	0.103 (179)	0.504 (179)	0.709 (179)
CWTscore	0.279 (172)	0.5980 (172)	0.859 (172)
Comp1	0.178 (167)	0.830 (167)	0.942 (167)

Abbreviations:; Comp1-Principle component score; COMMD4-COMM domain containing protein 4;CWTscore-cumulative wall thickness score, LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness

3.5.4.2.2. *SNX3*

For *SNX3*, both the orthogonal and total model was employed for association analysis where appropriate as significant evidence was found for population stratification at various *SNX3* loci (Tables 3.12). Thus significant association (Table 3.13) following Monte Carlo permutation testing was found between rs579257 and the hypertrophy traits; mIVST (p-value=0.038), mLVWT (p-value=0.041) and the Maron-Spirito score (p-value=0.032). Association was found between rs635051 and the hypertrophy traits; LVM (p-value=0.025), mPWT (0.007) and the Comp1 score (p-value=0.037). Furthermore, marker rs9398166 were associated with hypertrophy trait, mPWT (p-value=0.012). In comparison, no evidence of association was observed between rs628424 and rs588409 and any of the hypertrophy traits.

Effect sizes of these significant associations are given in the original measurements as summarised in table 3.14. Interestingly, the effect-alleles of the various markers in *SNX3* for which association was found were shown to increase the hypertrophy indices. The G-allele of rs579257 increased the mIVST, mLVWT and Maron score respectively by 1.860mm, 1.713mm and 5.340mm. Likewise the C-allele of rs635051 increases LVM and mPWT respectively by 18.725g and 0.497mm. Furthermore, the C-allele of rs635051 also had an effect on comp1 and a value of 0.621 was obtained. However, since comp1 is derived from quantile normalised data, the effect size of this indice can not be deduced from the original cardiac measurements. Therefore, the effect size for comp1 was estimated from the CWT score which is the mean of the 16 untransformed cardiac wall thickness measurements (14.134mm). Thus a measurement of 0.883mm is attributed to each of the 16 measurements. In addition, the C-allele of rs9398166 was shown to increases the mPWT by 0.514mm.

Table 3.12. Results from population stratification analysis of SNX3. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals. Significant p-values are highlighted and indicated in bold red.

	SNX3				
	rs579257	**rs628424**	**rs588409**	**rs635051**	**rs9398166**
LVM	**0.002** (78/177)	**0.033** (88/177)	0.111 (95/175)	**0.003** (88/177)	**0.007** (90/176)
mIVST	**0.002** (80/183)	0.116 (91/183)	0.166 (99/181)	**0.022** (91/183)	**0.042** (93/181)
mLVWT	**0.004** (80/182)	0.156 (91/182)	0.381 (99/180)	**0.032** (91/182)	0.065 (93/180)
mPWT	**0.030** (80/182)	**0.034** (91/182)	0.311 (99/180)	**0.010** (91/182)	**0.014** (93/180)
Maronscore	**0.005** (75/175)	0.107 (86/175)	0.269 (94/173)	**0.022** (85/175)	**0.038** (87/173)
CWTscore	**0.018** (72/168)	**0.084** (83/168)	0.145 (91/166)	**0.013** (82/168)	**0.021** (84/166)
Comp1	**0.008** (70/163)	0.075 (79/163)	0.385 (90/161)	**0.011** (79/163)	**0.019** (81/161)

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score, LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness; SNX3-Sorting nexin 3.

Table 3.13. Results of SNX3 association analysis using the orthogonal and total model where appropriate. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals. Significant p-values are highlighted and indicated in bold red.

	SNX3				
	rs579257	**rs628424**	**rs588409**	**rs635051**	**rs9398166**
LVM	0.057 (78/177)	0.256 (88/177)	0.894 (175)	**0.025** (88/177)	0.104 (90/176)
mIVST	**0.038** (80/183)	0.248 (183)	0.959 (181)	0.258 (91/183)	0.465 (93/181)
mLVWT	**0.041** (80/182)	0.272 (182)	0.795 (180)	0.298 (91/182)	0.206 (180)
mPWT	0.094 (80/182)	0.052 (91/182)	0.834 (180)	**0.007** (91/182)	**0.012** (93/180)
Maronscore	**0.032** (75/175)	0.565 (175)	0.474 (173)	0.062 (85/173	0.151 (87/173)
CWTscore	0.158 (72/168)	0.308 (83/168)	0.655 (166)	0.070 (82/168)	0.177 (84/166)
Comp1	0.098 (70/163)	0.378 (163)	0.840 (161)	**0.037** (79/163)	0.099 (81/161)

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score, LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness; SNX3-Sorting nexin 3.

Table 3.14. Effect sizes for the significant associations between markers of SNX3 and hypertrophy indices. Effect sizes are given in terms of the original measurements for all indices except for comp1.

Gene	SNP	Effect allele	Hypertrophy trait	p-value	Effect size
SNX3	rs579257	G	mIVST	0.038	1.860mm
			mLVWT	0.041	1.713mm
			Maron score	0.032	5.340mm
	rs635051	C	LVM	0.025	18.725g
			mPWT	0.007	0.497mm
			Comp1	0.037	0.621
	rs9398166	C	mPWT	0.012	0.514mm

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score, LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness; SNX3-Sorting nexin 3.

3.5.4.2.3. *DSCR3*

Table 3.15 and 16 summarises the p-values obtained from the association analysis between markers of *DSCR3* and various hypertrophy traits. No significant association was observed for population stratification thus the total model was used. For this analysis, a significant association was observed between marker rs2835680 and the Comp1 score with a p-value of 0.045. Since the association was only observed with one hypertrophy trait, this association were thought to be spurious.

Table 3.15. Results from population stratification analysis of *DSCR3*. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals.

	DSCR3				
	rs3165	**rs2835680**	**rs9305614**	**rs2835684**	**rs3787802**
LVM	0.569 (30/177)	0.336 (93/175)	0.210 (104/172)	0.475 (121/177)	0.247 (117/177)
mIVST	0.730 (30/183)	0.442 (98/181)	0.332 (108/177)	0.541 (127/183)	0.263 (121/183)
mLVWT	0.715 (30/182)	0.521 (98/180)	0.524 (108/176)	0.775 (127/182)	0.522 (121/182)
mPWT	0.404 (30/182)	0.057 (98/180)	0.060 (108/176)	0.759 (127/182)	0.447 (121/182)
Maronscore	0.9029 (27/175)	0.348 (93/173)	0.235 (103/169)	0.465 (121/175)	0.192 (115/175)
CWTscore	0.701 (26/168)	0.079 (90/166)	0.055 (101/163)	0.209 (117/168)	0.091 (113/168)
Comp1	0.903 (24/163)	0.117 (88/161)	0.088 (97/158)	0.369 (113/163)	0.131 (110/163)

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score;DSCR3-Down syndrome critical region 3; LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness.

Table 3.16. Results of *DSCR3* association analysis using the total model.P-values for association between *DSCR3* SNPs and hypertrophy indices**.** Values in brackets indicate the number of informative individuals. Significant p-values are highlighted and indicated in bold red.

	DSCR3				
	rs3165	**rs2835680**	**rs9305614**	**rs2835684**	**rs3787802**
LVM	0.779 (177)	0.006 (175)	0.044 (172)	0.358 (177)	0.668 (177)
mIVST	0.947 (183)	0.142 (181)	0.384 (177)	0.709 (183)	0.937 (183)
mLVWT	0.854 (182)	0.243 (180)	0.861 (176)	0.918 (182)	0.789 (182)
mPWT	0.886 (182)	0.199 (180)	0.373 (176)	0.783 (182)	0.921 (182)
Maronscore	0.736 (175)	0.150 (173)	0.327 (169)	0.725 (175)	0.897 (175)
CWTscore	0.635 (168)	0.106 (166)	0.407 (163)	0.926 (168)	0.382 (168)
Comp1	0.666 (163)	**0.045** (161)	0.152 (158)	0.554 (163)	0.740 (163)

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score;DSCR3-Down syndrome critical region 3; LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness.

3.5.4.2.4. *LGMN*

For LGMN, testing for population stratification showed no significant association (Table 3.17). Subsequent association testing using the total model also revealed no significant associations between the markers of LGMN and hypertrophy indices (Table 3.18).

Table 3.17. Results from population stratification analysis of *LGMN*. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals.

	LGMN				
	rs2402189	**rs3783933**	**rs4904980**	**rs1242102**	**rs1242095**
LVM	0.487 (76/170)	0.643 (107/169)	0.211 (81/167)	0.462 (103/166)	0.936 (99/162)
mIVST	0.787 (77/176)	0.942 (111/175)	0.978 (85/173)	0.903 (106/172)	0.764 (103/167)
mLVWT	0.755 (77/175)	0.809 (111/174)	0.864 (85/172)	0.889 (106/171)	0.552 (103/166)
mPWT	0.994 (77/175)	0.885 (111/174)	0.928 (85/172)	0.710 (106/171)	0.486 (103/166)
Maronscore	0.987 (74/168)	0.486 (104/167)	0.925 (79/165)	0.627 (101/164)	0.749 (97/160)
CWTscore	0.707 (70/161)	0.635 (101/160)	0.726 (78/158)	0.614 (101/157)	0.656 (95/152)
Comp1	0.259 (68/158)	0.419 (97/156)	0.633 (77/155)	0.958 (98/154)	0.802 (93/150)

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score;LGMN-Legumain; LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness.

Table 3.18. Results of *LGMN* association analysis using the total model.P-values for association between *LGMN* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals.

	LGMN				
	rs2402189	**rs3783933**	**rs4904980**	**rs1242102**	**rs1242095**
LVM	0.656 (170)	0.513 (169)	0.352 (167)	0.269 (166)	0.826 (162)
mIVST	0.976 (176)	0.453 (175)	0.213 (173)	0.131 (172)	0.696 (167)
mLVWT	0.679 (175)	0.553 (174)	0.242 (172)	0.140 (171)	0.674 (166)
mPWT	0.894 (175)	0.402 (174)	0.568 (172)	0.526 (171)	0.678 (166)
Maronscore	0.707 (168)	0.240 (167)	0.376 (165)	0.497 (164)	0.960 (160)
CWTscore	0.789 (161)	0.348 (160)	0.366 (158)	0.551 (157)	0.705 (152)
Comp1	0.9003 (158)	0.4892 (156)	0.5708 (155)	0.5911 (154)	0.8761 (150)

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score;LGMN-Legumain; LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness.

3.5.4.2.5. *FBXL10*

For *FBXL10*, no significant association was found for its population stratification testing, thus the total model was fitted for association. However, no significance of association was found between any of the markers of *FBXL10* and the various hypertrophy traits. Since QTDT requires at least 30 informative individuals to enable association analysis, the marker, rs7313015 was not tested when population stratification was analysed. This marker of *FBXL10* lacked the sufficient number of informative genotyped individuals due to its low allele frequency in our study cohort.

Table 3.19. Results from population stratification analysis of *FBXL10*. P-values of association between *SNX3* SNPs and hypertrophy indices. Values in brackets indicate the number of informative individuals. P-values for marker rs7313015 are not shown since the lack of sufficient number of informative genotyped individuals to enable QTDT analysis

	FBXL10							
	rs12427382	**rs3751131**	**rs6489811**	**rs11065585**	**rs4980986**	**rs7307400**	**rs10849893**	**rs7313015**
LVM	0.228 (91/174)	0.137 (89/176)	0.557 (100/177)	0.233 (98/177)	0.175 (89/177)	0.066 (73/177)	0.415 (86/174)	Not tested* (19/175)
mIVST	0.249 (95/180)	0.139 (93/182)	0.736 (103/183)	0.288 (102/183)	0.141 (93/183)	0.444 (75/183)	0.844 (90/180)	
mLVWT	0.252 (95/179)	0.138 (93/181)	0.604 (103/182)	0.370(102/182)	0.281 (93/182)	0.478 (75/182)	0.940 (90/179)	
mPWT	0.988 (95/179)	0.971 (93/181)	0.435 (103/182)	0.829 (102/182)	0.752 (93/182)	0.684 (75/182)	0.641 (90/179)	
Maronscor	0.335 (90/173)	0.231 (87/174)	0.677 (98/175)	0.442 (96/175)	0.190 (87/175)	0.627 (73/175)	0.783 (84/172)	
CWTscore	0.464 (88/166)	0.326 (85/167)	0.348 (96/168)	0.512 (94/168)	0.062 (85/168)	0.427 (71/168)	0.488 (83/165)	
Comp1	0.374 (84/161)	0.318 (81/162)	0.812 (92/163)	0.384 (91/163)	0.139 (82/163)	0.635 (70/163)	0.764 (79/160)	

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score;FBXL10-F-box and leucine rich repeat protein 10; LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness

Table 3.20. Results of *FBXL10* association analysis using the total model.P-values for association between *FBXL10* SNPs and hypertrophy indices.

	FBXL10							
	rs12427382	**rs3751131**	**rs6489811**	**rs11065585**	**rs4980986**	**rs7307400**	**rs10849893**	**rs7313015**
LVM	0.410 (174)	0.163 (176)	0.574 (177)	0.781 (177)	0.287 (177)	0.254 (177	0.556 (174)	0.842 (170)
mIVST	0.919 (180)	0.877 (182)	0.462 (183)	0.569 (183)	0.933 (183)	0.625 (183)	0.828 (180)	0.583 (176)
mLVWT	0.656 (179)	0.553 (181)	0.609 (182)	0.865 (182)	0.810 (182)	0.563 (182)	0.583 (179)	0.499 (175)
mPWT	0.670 (179)	0.421 (181)	0.501 (182)	0.656 (182)	0.175 (182)	0.075 (182)	0.486 (179)	0.868 (175)
Maronscor	0.989 (173)	0.857 (174)	0.291 (175)	0.521 (175)	0.995 (175)	0.515 (175)	0.939 (172)	0.680 (168)
CWTscore	0.966 (166)	0.782 (167)	0.204 (168)	0.547 (168)	0.956 (168)	0.245 (168)	0.934 (165)	0.274 (161)
Comp1	0.682 (161)	0.482 (162)	0.202 (163)	0.987 (163)	0.510 (163)	0.100 (163)	0.666 (160)	0.282 (156)

Abbreviations:; Comp1-Principle component score;CWTscore-cumulative wall thickness score;FBXL10-F-box and leucine rich repeat protein 10; LVM-left ventricular mass; Maron-Spirito score-as defined by Spirito and Maron (1990); mIVST-maximal interventricular septum thickness; mLVWT-maximal left ventricular wall thickness; mPWT-maximal posterior wall thickness

3.5.4.2.6. Summary of association analysis

Taken together, except for markers observed in *SNX3* and *DSCR3,* the majority of the markers screened in our South African HCM cohort showed no significant associations with any hypertrophy indices. However, for *DSCR3*, the association was assumed to be spurious since a significant association was only found with one hypertrophy indice. In contrast, for *SNX3,* three of the five markers screened, showed significant association with various hypertrophy indices and suggest a possible role for *SNX3* in the development of cardiac hypertrophy. Thus, the possible implications of these associations of *SNX3* are discussed further in the next chapter.

CHAPTER 4: DISCUSSION

INDEX **PAGE**

4.1.VERIFICATION OF THE PHOSPHORYLATION-DEPENDENT ASSOCIATION BETWEEN COMMD4 AND cMyBPC 144

4.2. IDENTIFICATION OF COMMD4 INTERACTORS VIA YEAST-TWO-HYBRIDANALYSIS 145

4.2.1. Preys excluded from further analysis 145

4.2.2. Preys excluded from study 146

4.2.3. Preys identified as putative COMMD4 ligands 146

4.2.3.1. Sorting nexin 3 (SNX3) 147

4.2.3.2. Down syndrome critical region 3 (DSCR3) 150

4.2.3.3. Legumain (LGMN) 151

4.2.3.4. F-box and Leucine rich repeat protein 10 (FBXL10) 152

4.2.3.5. α-enolase (ENO1) 154

4.2.3.6. Cardiac Actin (ACTC1) 156

4.2.4. Limitations of Y2H 157

4.2.5. Limitations of verification assays 158

4.2.5.1. *In vitro* Co-IP 158

4.2.5.2. *In vivo* Co-IP 159

4.2.5.3. 3D co-localisation 159

4.3. POSSIBLE MECHANISMS OF ACTION OF COMMD4 IN cMYBPC FUNCTIONING 161

4.3.1. Other possible effects of increased COMMD4 levels 168

4.3.2. COMMD4 and cMyBPC turnover 168

4.4. FAMILY-BASED ASSOCIATION STUDIES 170

4.4.1. *SNX3* 171

4.4.2. Limitations of association studies 173

4.4.2.1. Population stratification 173

4.4.2.2. Phenotypic heterogeneity 174

4.4.2.3. Statistical power and sample size 174

4.4.2.4. Confounding variables 175

4.4.2.5. Multiple testing 175

4.5. OVERVIEW AND FUTURE DIRECTIONS 176

4.6. CONCLUSION 177

CHAPTER FOUR: DISCUSSION

This thesis focussed on COMMD4, a protein of unknown function, which was previously identified as a novel ligand of the N-terminal region of cMyBPC (Ramburan A, Phd Thesis, 2008). COMMD4 is a little understood member of the COMMD family of proteins and has previously been linked to the development of cardiac hypertrophy. cMyBPC, in turn, is a sarcomeric protein with dual function involving the structural assembly and stability of the thick filament, as well as the regulation of cardiac contractility (Section 1.2.4). Key to cMyBPC-mediated regulation of cardiac contractility is the interaction of myosin S2 (Gruen *et al*., 1999), as well as thin filament proteins (Whitten *et al*., 2008; Shaffer *et al*., 2009), with the cMyBP-C motif, in a manner dependent on the dynamic phosphorylation of the motif by CaMK and PKA in response to β-adrenergic stimulation.

In the abovementioned study during which COMMD4 was identified as a novel MyBPC-interacting protein, a study conducted by A. Ramburan (2008) to further investigate the role of the N-terminal domains of cMyBPC in the regulation of cardiac contractility, it was noted that the interaction between cMyBPC and COMMD4 was also dependent on the phosphorylation status of the MyBPC-motif of cMyBPC (Ramburan A, PhD Thesis, 2008).

Given the phosphorylation-dependent interaction of COMMD4 with cMYBPC and the previously described association of COMMD proteins with cardiac hypertrophy, we hypothesised that COMMD4 may play a crucial role in the both the regulation of cardiac contractility and that variation in its corresponding gene may play a role in the development of cardiac hypertrophy. Therefore, the present investigation sought to elucidate the functions of COMMD4 by identifying COMMD4-interacting proteins using Y2H analysis and by subsequently confirming these interactions via various biochemical verification assays. Additionally, the previously identified COMMD4/cMyBPC interaction was further investigated in order to explore its potential role in the regulation of cardiac contractility. Furthermore, variations in the *COMMD4* gene as well as the genes corresponding to certain verified COMMD4 interactors were subsequently assessed for association with hypertrophy parameters in a cohort of South African HCM families bearing founder mutations, to assess the role of this pathway in hypertrophy development.

4.1. VERIFICATION OF THE PHOSPHORYLATION-DEPENDENT ASSOCIATION BETWEEN COMMD4 AND cMyBPC

Previously, the interaction between COMMD4 and the N-terminal cMyBPC APA bait (which encompasses the C1-C2 region of cMyBPC, and mimics the MyBPC motif in the monophosphorylated state) was identified using Y2H analysis, and was verified using 3D co-localisation assays in live cells. Subsequent quantitative β-galactosidase assays further showed that the interaction was significantly stronger when the MyBPC motif was in the mono- or unphosphorylated states, rather than in the trisphosphorylated state ($p<0.01$), and that the interaction was primarily mediated by the cMyBPC motif and not by the flanking Ig domains (Figure 3.11; Section 3.3.1).

In the present study, we further investigated the reported interaction. We confirmed the interaction between COMMD4 and cMyBPC using a pull-down assay (*in vivo* immunoprecipitation, Figure 3.6, Section 3.2.2). Furthermore, using the domain/phosphorylation assay, the data showed a trend that the interaction is consistently and visibly reduced by phosphorylation of cMyBPC in response to adrenergic stimulation (Figure 3.11 and 3.12), although, in these live cells, the quantitative effect of phosphorylation on co-localisation did not reach significance. The latter may be attributable to the low sample number (n=3) and repeat experiments (N=2) used, however, due to time-constraints, the matter was not further pursued.

Interestingly, in the library screen performed in the present study, in which COMMD4 was used as bait, cMyBPC was not identified as a COMMD4 interactor. The most likely reason for this is that the region of COMMD4-interaction with cMyBPC is the N-terminally located cMyBPC motif, and it is well known that the N-terminal regions of large proteins are extremely poorly represented, if at all present, in oligo dT-primed cDNA libraries (Adler *et al*., 1997).

4.2. IDENTIFICATION OF COMMD4 INTERACTORS VIA YEAST-TWO-HYBRID ANALYSIS

Using COMMD4 as bait, a total of $3.6x10^5$ cardiac cDNA yeast library clones were screened; 123 of these clones were able to activate the *HIS3*, *ADE2* and *MEL1* reporter genes. Following heterologous mating experiments, 28 of the 123 clones interacted specifically with COMMD4 (Table 3.4; Section 3.1.1.4).

4.2.1. Preys excluded from further analysis

After evaluating the physiological plausibility of the 28 clones, 19 were discarded from further analysis for the following reasons:

- No significant protein matches

  Even though all 28 clones had significant DNA matches in both NCBI Genbank (http://www.ncbi.nlm.nih.gov) and Ensemble databases (http://www.ensembl.org), 15 clones were discarded as they were not in-frame according to the reading frame dictated by the GAL4 activation domain. One reason for the lack of significant protein matches, in spite of significant DNA matches, is that in classical two-hybrid library constructions such as the Clontech library used in this study, which are derived from oligo-dT primed cDNA, only one out of six of all cloned inserts are in-frame with the transcription factor activation domain (van Criekinge and Beyaert, 1999).

- Incompatible cellular compartments

  Knowledge of a protein's localisation is critical to understanding its function. This is of particular relevance when studying and attempting to verify protein-protein interactions, given that proteins localised in separate cellular compartments are physically unable to interact with one another. For proteins where the sub-cellular localisation has not yet been experimentally determined, publicly available sub-cellular localisation prediction programs have been developed to predict sub-cellular localisation based on amino acid sequence

information (Nakai and Kanehisa, 1992; Horton and Nakai, 1997; Reinhardt and Hubbard, 1998; Hua and Sun, 2001; Emanuelssen, *et al.,* 2000; Nair and Rost, 2002). These subcellular prediction programs do, however, have two major limitations. Firstly, they have limited coverage (the number of sub-cellular regions supported by the prediction software are few) and secondly, they have limited accuracy for predicting sub-cellular localisation (Lu *et al.,* 2003). Given these short-comings, two subcellular prediction programs; viz Proteome analyst (http://www.cs.ualberta.ca/~bioinfo/PA/Sub) and ESPpred (http://www.imtech.res.in/raghava/eslpred/), were used to improve prediction validity when predicting subcellular localisation of putative COMMD4 interactors identified in the present investigation.

Although initial reports have classified COMMD proteins as cytosolic proteins, more recent findings have shown that these proteins are also found in the nucleus, perinuclear organelles, lysosomes and endosomes (Klomp *et al*., 2003). Therefore, prey proteins known or predicted to be located in the nucleus, lysosomes and endosomes were also considered as physiologically plausible interactors, but those that were known to or predicted to localise to the mitochondria, endoplasmic reticulum or the extracellular matrix were excluded from further analysis.

4.2.2. Preys excluded from study

Following the exclusion of 19 clones based on reasons expounded in section 4.2.1., seven clones (Table 3.5; Section 3.1.1.4), of which one, ACTC1, was repeatedly identified, were retained as plausible COMMD4 interactors for further analysis. However, after the various verification analyses, the number of putative positive prey clones was further reduced to six: direct interaction of ANKRD1 with COMMD4 could not be confirmed using any of the verification techniques, suggesting that it was a yeast two-hybrid false positive result.

4.2.3. Preys identified as putative COMMD4 ligands

Two of the six verified COMMD4 interactors, ACTC1 and ENO1, have previously been identified, along with COMMD4, as interactors of the C1-C2 region of cMyBPC, during various Y2H library screens conducted in our laboratory (A, Ramburan, Phd Thesis, 2008). In another Y2H library screen, COMMD4, along with SNX3, ENO1 and ENO3, was identified as interactors of the phosphodiesterase-4D-interacting protein (PDE4DIP), myomegalin, another cMyBPC interacting protein (Uys G, Phd Thesis, 2010). The fact that these interactors from three different screens overlap suggests that these proteins are part of a sarcomeric protein complex. However, these results may also reflect isolated interactions with COMMD4.

Interestingly, four of the six verified COMMD4 interactors (SNX3, DSCR3, LGMN and FBXL10) have been shown to be involved in protein trafficking and protein degradation, suggesting a role for COMMD4 in the protein turnover pathway.

The functions of each of these interactors will now be discussed in the following sections, before a synthesis of the results will be presented.

4.2.3.1. Sorting nexin 3 (SNX3)

Sorting nexin 3 (SNX3) belongs to the sorting nexin (SNX) family of proteins that is involved in endocytosis and protein trafficking through the endosomal pathway (Worby and Dixon, 2002). These proteins characteristically contain a phospholipid-binding motif (PX) (Figure 3.3) domain, which has been found to play important roles in a number of cellular processes including endocytic protein trafficking, protein sorting and lipid modification, cell signalling, as well as serving as an interface for protein-protein interactions (Section 1.5) (Sato *et al*., 2001; Carlton *et al*., 2005).

Sorting nexins are divided into three groups based on their structural organisation; the latter also helps to determine their phospholipid-binding preference with regard to the endosomal membrane.These include SNX^{PX}, which consists of PX domains and minimal flanking regions, $SNX^{PX\text{-}BAR}$, which, in addition to the PX domain, also has a C-terminal coiled-coil domain that is related to the BAR domain (a domain involved in protein dimerization and membrane dynamics), and $SNX^{PX\text{-}other}$, which contains other flanking regions in addition to the PX domain (Seet and Hong, 2006). SNX3, along with SNX10, SNX12, SNX22 and SNX24, belong to the SNX^{PX} group of SNXs and is therefore thought unlikely to form heteromeric complexes with other family members because of its lack of coiled-coil domains (Haft *et al*., 1998). Unlike the other SNXs, which are predominantly associated with the cell membranes, SNX3 has been found to localise mainly in the cytosol and is highly expressed in the heart, skeletal muscle and spleen (Haft *et al*., 1998; Vervoort *et al*., 2002).

SNX3 exclusively interacts with phosphatidylinositol 3-phosphate (PtdIns3P), a lipid enriched in endosomal membranes. The latter further indicates SNX3 cellular localisation in the early and perinuclear endosomes (Xu *et al*., 2001). Similarly, SNX3 co-localisation with an early endosome-associated (EEA1) protein, a marker for early endosomes as well as its co-localisation with the transferrin receptor (TFR), confirmed its presence in the tubular-vesicular structures of the early endosomes and the recycling endosomes (Xu *et al*., 2001; Mizutani *et al*., 2009). This data thus suggests the involvement of SNX3 in the recycling of proteins either to the plasma membrane, TGN or to the lysosomes.

Similary, findings involving the yeast ortholog of human SNX3, Grd19p, confirmed an association of SNX3 with the late Golgi proteins (Haft *et al*., 1998; Brown *et al*., 1995). Moreover, Grd19p has been shown to interact with the retromer complex, where it functions as the cargo-specific adaptor, as it binds to the recycling signal in the cargo protein (e.g. iron transporter, Fet3-Ftr1) and the retromer (Strochlic *et al*., 2007; Strochlic *et al*., 2008). Thus, this association with the retromer complex, which localises to the early endosome and mediates the recycling of sorting receptors, further supports SNX3's endosomal localisation and proposed involvement in the recycling of proteins.

In addition to endosomal sorting, SNX3 was also shown to be involved in the formation of multivesicular bodies (MVB) through its interaction with PtdIns3P. SiRNA knockdown studies showed that SNX3 knockdown inhibits the invagination process involved in the formation of MVB's, while overexpression of SNX3 resulted in the accumulation of intraluminal vesicles (Pons *et al*., 2008). Moreover, knockdown of Hrs (Section 1.5.1.2), which functions upstream of SNX3 in sorting of receptors to the lysosome, was shown to reduce SNX3 expression (Pons *et al*., 2008). These findings, therefore, show that Hrs and SNX3 function together in the lysosomal pathway, but whether SNX3 directly associates with the ESCRT machinery remains to be elucidated (Section 1.5.1.2).

SNX3 also interacts with the de-ubiquitylating enzyme, ubiquitin-specific protease 10 (USP10), which was shown to reduce SNX3 proteosomal degradation and stabilise its expression (Boulkroun *et al*., 2008). Moreover, vasopressin-induced USP10 interaction with SNX3was also found to indirectly increase ENaC cell surface expression (Section 1.3.4.3.). Consequently, based on the involvement of SNX3 in early–to recycling endosome transport, it was suggested that the SNX3/ USP10 interaction either promotes the export of ENaC to the cell surface, or alternatively, SNX3 may act as an adapter for ENaC and promote its recycling from the endosomes to the cell surface via the retromer complex (Boulkroun *et al*., 2008).

Taken together, the above findings clearly demonstrate the important roles of SNX3 in sorting and recycling of proteins at the early endosomes, via retromer-dependent trafficking, or to the lysosome, via the ESCRT machinery. However, the manner in which SNX3 exerts its function in the latter still remains unclear.

In the present study, the Y2H screen showed that COMMD4 interacted with SNX3. This interaction was subsequently confirmed by each of the verification assays, suggesting that the proposed interaction between COMMD4 and SNX3 does occur *in vivo*. More importantly, time-lapsed images of the colocalisation between EYFP-COMMD4 and GFP-SNX3 (Figure 3.9; Section 3.2.3), together with our observation that COMMD4 and SNX3 colocalised within endocytic vesicles, suggest that the interaction between COMMD4 and SNX3 is important for endocytic protein trafficking. In this study, we also provide the first evidence for endosomal trafficking of COMMD4 through its interaction with SNX3.

We know from our own investigations and from the literature that COMMD4, similar to COMMD1, seems to act as a facilitator in protein complexes and seems to promote the ubiquitination and subsequent degradation of target proteins by stabilising the interaction between the target protein and E3 ubiquitin ligases (Burstein *et al*., 2005; de Bie *et al*., 2006; Maine *et al*., 2007). Therefore, this scenario might also be applied to its role in the function of SNX3. Given the fact that SNX3 is important in recycling and sorting of intracellular proteins, tight regulation of its intracellular levels is needed. Thus the most plausible role for COMMD4 in this context may be to promote the ubiquitination and degradation of SNX3, thus regulating SNX3 turnover.

Interestingly, both SNX3 and COMMD4 were also identified, in another library screen performed in our laboratory, as interactors of PDE4DIP, also known as myomegalin (Uys G, Phd Thesis, 2010). Myomegalin has also been identified as a cMyBPC-interacting protein (Ramburan A, Phd Thesis, 2008) and functions as an A-kinase anchoring protein (AKAP) that anchors PKA to cMyBPC and in turn facilitates its phosphorylation. Although each of these interactions may occur independently of each other, another interpretation of the above-mentioned interactions may be that these proteins are involved in the formation of multi-protein complexes, most likely during sarcomeric disassembly. Both COMMD4 and PDE4DIP can bind to cMYBPC in its unphosphorylated state. Since this state is also associated with the degradation of cMyBPC, one can speculate that SNX3 associates with this multi-protein complex when MyBPC is released from the sarcomere for degradation via the UPS. SNX3 may aid in the trafficking of this complex to the proteasome; however, as there is currently no direct evidence linking SNX3 to the proteasome, this proposal remains speculative.

Furthermore, the involvement of COMMD4 in the turnover of SNX3 might have indirect consequences on SNX3 target proteins, for example, on the expression of ENaC. As mentioned above, the stable expression of SNX3 is indirectly linked to increased ENaC cell surface expression, thus a regulatory effect of COMMD4 on SNX3 protein turnover, as proposed, would in turn reduce cell surface ENaC expression, and so inhibit overall ENaC activity in an indirect manner. Since SNX3 is a short-lived protein, this would be consistent with the proposed functions of COMMD4 based on its interactions with E3 ubiquitin ligases. However, whether COMMD4 plays a more direct role in ENaC trafficking remains speculative for now. Of interest though is the reported functional overlap between SNX3 and COMMD1 with regards to their roles in the regulation of ENaC activity, since both COMMD1 and SNX3 interact with ENaC subunits (Boulkroun *et al*., 2008; Ke *et al*., 2010). Via its COMMD domain, COMMD1 forms a multimeric complex with Nedd4.2, an E3 ubiquitin ligase, and facilitates the ubiquitination, endocytosis and degradation of ENaCs, and subsequently downregulates the cell surface expression of ENaCs (Ke *et al*., 2010). It seems plausible that COMMD4, as another COMMD family member, could fulfil a similar role to COMMD1, given the interaction between SNX3 and COMMD4 shown in the present study. Since the expression of these sodium channels is dependent on the trafficking of these proteins via the endosomal/recycling pathway, this scenario with SNX3 in complex with COMMD4 is feasible. Upon downregulation of the ENaCs, SNX3 might be associated with the trafficking of the target protein, while COMMD4 might facilitate the ubiquitination, which is needed for endocytosis and its subsequent sorting to the lysosome.

Furthermore, both COMMD1 and SNX3 interact with specific phospholipids. While COMMD1 binds to PtdIns 4.5P through its COMM domain, SNX3 binds to PtdIns 3P through its PX domain. The preference for binding to specific phospholipids dictates where each functions within the endosomal pathway. Based on the capacity of COMMD1 for binding PtdIns 4.5P and its co-localisation with EEA1 and ESCRTIII through its COMM domain (Section 1.3.4.5.1) (Burkhead *et al*., 2009), one can speculate that a similar role might exist

for COMMD 4 in association with SNX3 in the endosomal trafficking/recycling/degradation pathway. Although the phospholipid binding domain, and thus where COMMD4 might localise within the endosomal pathway, has not been defined for COMMD4, one might speculate that COMMD4 might also be localised to the ESCRT complex, based on this evidence from COMMD1. If this is the case, then it could most likely explain an indirect association of SNX3 with the ESCRT complex through its interaction with COMMD4, as a direct association between SNX3 and the ESCRT machinery has not been established yet. This association would also provide a plausible explanation for the role of SNX3in MVB formation, since SNX3 has been shown to be associated with the early endosome (Xu *et al*., 2001).

4.2.3.2. Down syndrome critical region 3 (DSCR3)

The Down syndrome critical region 3 gene (*DSCR3/DSCRA*) was identified in the region of chromosome 21 that is associated with partial trisomy of chromosome 21 in Down's syndrome (Nakamura *et al*., 1997). DSCR3 is expressed in a variety of tissues, including the heart, brain, kidney and skeletal muscle and has been found to localise within the nucleus of the cell (Hu *et al*., 2006). Since COMMD4 is also located in the nucleus, an interaction between DSCR3 and COMMD4 is feasible. While DSCR3 might be involved in the pathogenesis of Down syndrome, its function still remains unknown. DSCR3 has, however, been shown to have significant homology to the mouse *Hβ58* gene, which functions in embryogenesis. More interestingly, the yeast homologue of *Hβ58*, *PEP8*, just like SNX3, was shown to be involved in the protein sorting pathway (Nakamura *et al*., 1997). This together with the domain structure of DSCR3 further implicates DSCR3 in protein trafficking.

DSCR3 is a member of the Vps family of proteins by virtue of containing a Vps26 domain, a domain that has been implicated in protein trafficking (Haft *et al*., 2000). Furthermore, despite the lack of sequence homology amongst Vps proteins, this family of proteins has been shown to be similar to the arrestin protein family that functions in cell signalling and protein trafficking (Gurevich and Gurevich, 2006). As described in section 1.5.1.1, Vps26, along with Vps35, Vps29 and SNXs 1, 2 and 3, forms part of a retromer complex involved in the endosome-to-Golgi retrieval of proteins (Haft *et al*., 2000). Interestingly, depletion of Vps26 by siRNA knock-out caused the retromer to be inactive, which resulted in significant defects in the retrograde transport of cargo. This further suggests the importance of Vps26 as a structural scaffold in the cargo-recognition complex of the retromer (Seaman, 2005).

In the present study, the interaction between COMMD4 and DSCR3 identified by Y2H analysis was verified both by *in vivo* Co-IP experiments and 3D co-localisation assays. The 3D co-localisation further suggest that DSCR3 localised within endosomes (Figure 3.10; Section 3.2.3), giving further credence to its possible role in protein trafficking.

It is reasonable to assume that COMMD4 might regulate the turnover of DSCR3, based on the ability of COMMD proteins to promote ubiquitination. Turnover of DSCR3 would have an effect on the regulation of DSCR3 target proteins; however other interactors of DSCR3 have not yet been reported.

Alternatively, both DSCR3, based on its Vps26 domain, and SNX3, as mentioned in section 4.2.3.1, form part of a retromer complex that mediates endosome-to-Golgi trafficking of various proteins (Section 1.5.1.1). Moreover, DSRC3, based on its Vps26 domain, together with Vps35 and Vps29, make up the substrate-recognition complex of the retromer complex. Therefore, it is tempting to speculate that, because of its interactions with DSCR3 and SNX3, COMMD4 may also form part of the retromer complex and in turn help regulate the turnover of the retromer complex. However, at this point, this remains speculative and would require further investigations.

4.2.3.3. Legumain (LGMN)

Legumain (LGMN), or asparaginyl endopeptidase, a member of the caspase family of proteins, is a cystein protease, which is expressed in a variety of tissues, but is particularly abundant in the kidney, heart and placenta (Tanaka *et al*., 1996). Moreover, LGMN functions in intracellular degradation within the lysosome (Section 1.5.1.2 and 1.5.2) and specifically hydrolyses internal asparaginyl bonds of a protein (Ishii, 1994; Chen *et al*., 1997, Dickinson, 2002). This protein is activated by the acidic pH in the lysosome, and requires the removal of its C-and N-terminal peptides through autocatalysis to exert its function. However, just what triggers autocatalysis is poorly understood (Li, *et al*., 2003). Interestingly, its activation also controls the activation and processing of downstream substrates, including lysosomal proteases. For example, in LGMN-deficient mice, the processing of the lysosomal proteases is defective, resulting in the accumulation of these proteases and enlarged lysosomes (Shirahama-Noda *et al*., 2003; Chan *et al*., 2009).

Moreover, LGMN inhibits the activity of another family of proteases, viz. calpains, which are papain-related enzymes (Chen *et al*., 1997). Calpains are Ca^{2+} regulated proteases that prefer cytoskeletal and myofibrillar proteins as substrates (Dickinson, 2002) (Section 1.4.2). In addition, LGMN can either abolish protein function by digestion, or digest a pro-protein into its principal component, thereby activating it (Dickinson, 2002). For example, LGMN slowly degrades vitamin D-binding protein (DBP) in order to release vitamin D in the cytosol (Yamane *et al*., 2002). Examples of target proteins that are degraded and inactivated by LGMN include annexin II, a cell surface receptor (Yamane *et al*., 2002), as well as macro molecules such as metalloproteinases (MMP-2), cathepsins H, B, L and fibronectin (Chen *et al*., 2000; Shirahama-Noda *et al*., 2003; Morita *et al*., 2007).

Interestingly LGMN also been implicated in the development of atherosclerosis, in tumor growth/metastasis and in various pathological conditions (Clerin *et al*., 2008; Murthy *et al*., 2005). High levels of LGMN were observed in a mouse model of spontaneous atherosclerosis (Clerin *et al*., 2008) as well as several types of tumors. It has been suggested that the role of LGMN in the control of ECM remodelling contributes to its involvement in these various conditions (Morita *et al*., 2007). LGMN plays a role in ECM turnover through its interaction and degradation of fibronectin, which is thought to be mediated through LGMN's regulatory effects on MMP-2 (Morita *et al*., 2007). Fibronectin, which is the main component of ECM, is important in the stabilisation of the ECM (Sottile and Hocking, 2002). In LGMN-deficient mice, fibronectin accumulated

within the kidneys, which is in turn, was associated with increased renal interstitial fibrosis (Morita *et al*., 2007). Of interested to us is that elevated levels of metalloproteinases, in particular MMP-2 was found to be associated with LV remodelling and poor prognosis in HCM patients (Cambronero *et al*., 2009; Lombard *et al*., 2003). Since these metalloproteinases are involved in myocardial collagen turnover, they were established as biomarkers for HCM pathophysiology. Based on LGMN's involvement in ECM remodelling, LGMN itself might also be implicated in the pathophysiology of HCM.

For the present study, the interaction between COMMD4 and LGMN observed in the Y2H screen was confirmed by all the verification assays. Moreover, we were able to show that LGMN and COMMD4 co-localise to the cytoplasm as well as the endosomes (Figure 3.8 and 3.10; Section 3.2.3). Based on this interaction, we propose that COMMD4 functions in the regulation of LGMN turnover, based on the role of COMMD4 in ubiquitin modification. Thus, LGMN turnover may, in turn, affect the regulation of downstream LGMN substrates, such as fibronectin and MMP2, for example. If one were to consider this scenario, one can speculate that when COMMD4 expression is high, this would result in decreased levels of LGMN. Since LGMN is responsible for the degradation of fibronectin and MMP2, lower levels of LGMN would in turn lead to accumulation of fibronectin and subsequent interstitial fibrosis. Moreover, decreased levels of LGMN would attentuate the inhibitory effect of LGMN on calpains which in turn would iniatiate or contribute to the untimely degradation of sarcomeric proteins from the myofibril

These data therefore suggest a role for COMMD4 in the regulation of cystein proteases expression, which include calpain.

4.2.3.4. F-box and Leucine rich repeat protein 10 (FBXL10)

The F-box family of proteins consists of a large family of eukaryotic proteins characterised by a ~40 amino acid motif, the F-box motif (Bai *et al*., 1996; Jin *et al*., 2004). In addition to their F-box domains, these proteins contain WD40-domains, leucine-rich repeats (LRR) or other protein-protein interaction motifs. Based on these additional domains, the F-box family can be divided into three groups. These are the Fbxws containing WD40 domains, Fbxls containing LRR, and Fbxos containing "other" motifs (Cenciarelli *et al*., 1999; Jin *et al*., 2004). Furthermore, many F-box proteins exist as multiple splicing variants (Jin *et al*., 2004).

FBXL10/JHDM1B, which belongs to the Fbxls group, is the largest member of the family of F-box proteins. The F-box motif and the eight LRRs at the FBXL10 C-terminal are the domains involved in the role of this protein as part of the SCF E3 ubiquitin ligase (Figure 3.3; Section 3.1.1.4) (Jin *et al*., 2004), while the JmjC domain, a DNA-binding CxxC zinc-finger, and a PHD zinc finger at its N-terminal mediate the effects of FBXL10 on gene expression (Takeuchi *et al*., 2006; Frescas *et al*., 2007). The JmjC domain is common to members of the jumonji (JMJ) family of transcriptional repressors that regulate chromatin and gene expression, while both ZF domains are DNA-binding motifs (Takeuchi *et al*., 2006). The CxxC zinc-finger

on the other hand recognises and binds unmethylated CpG-rich regions, while the PHD finger recognises lysines on histones (Aasland *et al*., 1995; Frescas *et al*., 2007).

FBXL10 and FBXL11/JHDM1A have been reported to function as histone lysine-demethylases (enzymes that remove methyl groups from lysines in histones) and are thought to be involved in histone modification. Both these F-box proteins target lysine36 on histone3 (H3K36) for demethylation and their JmjC domain has been shown to exhibit demethylase activity (Tsukada *et al*., 2006). More recently, *in vivo* studies also implicated FBXL10 in the demethylation of trimethylated histone H3 lysine 4 (H3K4me3) (Koyama-Nasu *et al*, 2007).

Additionally, there has been a certain amount of controversy surrounding the biological functions of FBXL10 with regards to its effect on gene expression. Some studies have reported that FBXL10 is a proto-oncogene, while others have shown that it acts as tumor suppressor (Frescas *et al*., 2007; Koyama-Nasu *et al*., 2007; He *et al*., 2008). Most recently, FBXL10 was shown to regulate cell proliferation and senescence in a p53 pathway-dependent manner. Knockdown of FBXL10 resulted in a significant decrease in cell proliferation, an effect that was shown to be mediated through the repression of $p15^{Ink4b}$, one of three cell cycle inhibitors on the *Ink4a-ARF-Ink4b* locus, by FBXL10 (He *et al*., 2008). In contrast, FBXL10 has also been reported to act as a negative regulator of c-Jun and ribosomal RNA (rRNA) genes (Frescas *et al*., 2007; Koyama-Nasu *et al*., 2007). It was shown that the CxxC zinc-finger domain of FBXL10 binds to the unmethylated CpG islands within the c-Jun promoter and tethers transcriptional repressor complexes, such as Sin3A and histone deacetylase (Silverstein and Ekwall, 2005). In addition to its down-regulation of c-Jun, FBXL10 was also implicated in the control of cell death and cell cycle progression, since knockdown of FBXL10 caused abnormal cell-cycle progression and increased UV-induced cell death (Koyama-Nasu *et al*., 2007). Similarly, FBXL10 was also shown to bind the transcribed region of ribosomal genes and, in so doing, repress their transcription (Frescas *et al*., 2007). The repression was mediated via the JmjC domain, which is required, in this case, for the demethylation of H3K4me3 rather than H3K36. Moreover, it was suggested that FBXL10-mediated modulation of rRNA gene expression has an effect on cell growth and proliferation, since FBXL10-depleted cells proliferated faster than control cells (Frescas *et al*., 2007).

While the F-box proteins function in cell death, cell differentiation as well as transcription, their most important function is to provide substrate specificity for E3 RING ligases. F-box proteins form part of the substrate-recognition component of the SCF E3 ubiquitin ligase complex and are thought to be involved in the controlled degradation of cellular regulatory proteins. The SCF complex transports ubiquitin-conjugating enzymes to target substrates that are specifically recruited by the different F-box proteins (Section 1.4.3.3) (Powell, 2006). Studies have shown that FBXL10 binds to Skp1 (Fbl1) via its F-box motif, but not to Cul1 (Jin *et al*., 2004). In addition, the LRR domain, domains which are involved in protein-protein interactions, serves as the substrate-binding domain within FBXL10 (Jackson *et al*., 2000). Furthermore, F-box proteins are also unstable, so once substrates are released for degradation, the F-box proteins can be self-ubiquitinated

via the SCF E3 ubiquitin ligase complex (Willems *et al*., 2004). Subsequent degradation of the F-box protein in turn would attenuate the activity of the SCF E3 ubiquitin ligase complex.

For the present study, the putative interaction between COMMD4 and FBXL10 was verified by the *in vivo* Co-IP and 3D co-localisation assays. Although the *in vitro* Co-IP assayswere inconclusive since the *in vitro* translated product was too small and fell outside of the detection limit for the SDS-PAGE system, despite using various SDS gel concentrations as well as loading dye containing Ficoll Orange to ensure that the dye front does not obscure any visible bands interaction between COMMD4 and FBXL10 were confirmed in a cellular environment using *in vivo* Co-IP. Even though FBXL10 has been reported to localise within the nucleus, the co-localisation experiment conducted as part of the present investigation suggests that this protein also localises to both the nucleus and the cytosol. This finding is, however, not unexpected since F-box proteins form part of a SCF E3 ubiquitin ligase complex, which is known to function in both the nucleus as well as the cytosol (Patton *et al*., 1998; Cenciarelli *et al*., 1999). Furthermore, based on its homology to COMMD1, COMMD4 could aid in the export of proteins, such as FBXL10, from the nucleus to the cytosol.

Moreover, since the the F-box protein of the SCF complex is responsible for substrate recognition and recruitment to the SCF complex, we propose that COMMD4 aids in the recognition and recruitment of target proteins by facilitating protein complex formation between target proteins and the F-box protein, linking the target to the rest of the SCF E3 ubiquitin ligase. Interaction between COMMD4 and FBXL10 could provide added stabilisation, and could help maintain the required angle of orientation (Cardozo and Pagano, 2004) (Section 1.4.3.3) between the F-box and the E2 enzyme, ensuring the optimum distance for efficient ubiquitin transfer, and promoting degradation of the target protein. This notion is given further support from studies that have shown that COMMD 1 interacts with several E3 ubiquitin ligase subunits including Cullins 1, 2, 3, and 5, Rbx1 as well as XIAP proteins via its COMMD domain (Maine *et al*., 2007).

Given that FBXL10 has also been implicated in transcriptional regulation (Frescas *et al*., 2007; Koyama-Nasu et al., 2007), one could speculate that COMMD4 may also play an indirect role in the regulation of FBXL10 targeted proteins. In fact, COMMD4, together with COMMD1, has been shown to inhibit the NFκB pathway by being recruited to chromatin of the κB-responsive promoter and destabilising the interaction between the NFκB subunit, RelA and chromatin (Burstein *et al*., 2005; de Bie *et al*., 2006) (Section 1.3.4.1.3). However, whether the interaction between COMMD4 and FBXL10 is able to regulate transcription in a similar manner remains unresolved.

4.2.3.5. α-enolase (ENO1)

Enolase (2-phospho-D-glycerate hydrolase) is a glycolytic enzyme expressed in tissues with high energy requirements, such as the brain and striated muscle, in order to produce ATP. During glycolysis, this cytosolic enzyme, which requires magnesium or other divalent metal ions for optimal activity, catalyzes the conversion of 2-phospho-D-glycerate (PGA) to phosphoenolpyruvate (PEP) (Pancholi, 2001).

Enolase is active as dimers comprising three different subunits, α, β and γ, which are encoded by *ENO1*, *ENO3* and *ENO2*, respectively (Fougerousse *et al*., 2001). The expression of these genes is developmentally regulated as well as being tissue-specific. Alpha-enolase (the αα dimer), which is ubiquitously expressed in most aldult tissues, is the major isoform expressed during embryonic development. However, during the fetal-to-postnatal transition, the expression of α-enolase decreases, while the expression of the β and γ isoforms increases in striated muscle (αβ and ββ dimers) and the brain (αγ and γγ dimers), respectively (Merkulova *et al* 1997; Merkulova *et al* 2000). It was shown via the metabolic adaptations of the fetal and postnatal heart that the transition from a hypoxic uterine environment to normoxic conditions may account for the observed decrease in α-enolase expression (Nau *et al*., 2002).

In striated muscle, α- and β-enolase are located at the M-band and Z-disk of the sarcomere. In addition, both isoforms were also shown to associate with other glycolytic enzymes (i.e. creatine kinase and pyruvate kinase), with the cytoskeletal component F-actin, as well as tubulin and microtubules (during myogenesis) and with components of the sarcomeric contractile apparatus (viz. troponin), with high affinity (Merkulova *et al* 1997; Keller *et al*., 2007; Pancholi, 2001). Given its role in glycolysis, it is tempting to suggest that enolase forms part of a functional complex responsible for availability of ATP during energy-consuming processes, such as myogenesis and possibly muscle contraction (Keller *et al*., 2007).

The expression of both α- and β-enolase is upregulated when cells are exposed to hypoxic conditions (Merkulova *et al*., 2000); however, the hypertrophic heart, similar to the embryonic heart, is characterised by high levels of α-enolase rather than β-enolase (Keller *et al*., 1995; Jin *et al*., 2008). Since hypertrophic hearts require more ATP and oxygen for effective cardiac function, it was suggested that the elevation in α-enolase expression might be the result of increased oxygen demand (Mizukami *et al*., 2004). Thus α-enolase seems to offers protection, by increasing glycolysis, under low oxygen conditions present in myocardial hypertrophy (Zhu *et al*., 2009). Furthermore, α-enolase was also shown to improve the contractility of the cardiomyocytes, which had been impaired by ischemic hypoxia (Mizukami *et al*., 2004).

In the present study, the Y2H interaction between COMMD4 and ENO1 was confirmed by all the verification assays employed. Moreover, 3D co-localisation further confirmed that these two proteins share the same subcellular space, suggesting that the interaction is physiologically relevant. In addition, ENO1, along with COMMD4 and ENO3, was also identified in an additional library screen, performed in our laboratory, as an interactor of cMyBPC (Ramburan A, Phd Thesis, 2008) as well as as a ligand of another interactor of the N-terminus of cMyBPC, PDE4DIP (Uys G, Phd Thesis, 2010). Based on these observations and the fact that both α- and β-enolase are known to locate to the sarcomere, although at the M-band and Z-disk (Foucault *et al*., 1999; Merkulova *et al*., 2000), these interactions might reflect complexes that occur within the sarcomere. The fact that both ENO1 and ENO3 were identified as interactors of cMYBPC further also testifies to enolase's activity as a dimer. Thus, the anchoring of the αβ heterodimer to cMyBPC within

the sarcomere may provide energy via glycolysis. However, the added binding of COMMD4, which is most likely involved in the turnover of cMyBPC once it is released from the sarcomere, also offers the possibility that ENO1 might provide the energy required for cMyBPC degradation, as this process is ATP-dependent.

However, as expounded for the previous preys, the most plausible role for COMMD4 in the COMMD4/ENO1 interaction is in the regulation of ENO1 turnover, through its ability to act as a facilitator in protein complexes: thus it may promote the ubiquitination and subsequent degradation of ENO1. Since the expression of the ENO1 isoform is tightly regulated, it is plausible that COMMD4 might be involved in this regulation. In addition, COMMD4 might indirectly affect other proteins of the ENO1 pathway. Given the role of ENO1 in glycolysis, one can speculate that downregulation of ENO1 expression through its interaction with COMMD4 could have dentrimental effects on energy-consuming processes such as muscle contraction since the required ATP needed would be limited. This in turn would affect proper cardiac function.

4.2.3.6. Cardiac Actin (ACTC1)

As described in section 1.2.1, cardiac actin, along with myosin, constitute the major contractile proteins of the sarcomere; the gene encoding this sarcomeric protein has also previously been implicated in HCM (Mogensen *et al*. 1999). Actin itself is a ubiquitous protein, and its various isoforms function in a variety of cellular processes, such as muscle contraction, cell motility, endo- and exocytosis, cytokinesis and maintenance of the cytoskeleton (Au, 2004). Six isoforms of actin exist, which are developmentally regulated in a tissue-specific manner and thus reflect the different isoform-specific functions. In the present study, our interest was focussed on cardiac α-actin, which is the main actin isoform in the adult heart and which functions in muscle development and contractility (Bertola *et al*., 2008; Tondeleir *et al*., 2009) due to the observed interaction between this protein and COMMD4. Changes in the expression pattern of actin isoforms are associated with pathologies such as cardiac hypertrophy.

With regards to its role in cMyBPC functioning, several studies to date have shown an interaction between actin and the N-terminal regions of cMyBPC (Herron *et al*., 2006; Whitten *et al*., 2008; Shaffer *et al*., 2009). Although the exact cMyBPC region involved is still uncertain, findings have demonstrated that there is more than one actin binding site. This includes the PA-rich region between domain C0 and C1, domains C1-C2 and the cMyBPC motif (Sections 1.2.4.2.2 and 1.2.4.2.3).

For the present study, the interaction between ACTC1 and COMMD4 were confirmed by all the verification assays**.** In addition, ACTC1 was the only prey that was identified in multiple clones (three times), which further suggests that this particular interaction is more likely to be biologically significant, rather than a mere two-hybrid artefact (Parrish *et al*., 2006). These findings, together with the fact that ACTC1 has been previously identified as an interactor of the C1-C2 region of cMyBPC, where COMMD4 also binds, make this interaction between COMMD4 and ACTC1 intrigueing.

Since COMMD4 facilitates protein complexes and promotes ubiquitination through its interactions with E3 ubiquitin ligase subunits, one can assume that COMMD4 might also be involved in the turnover of actin. However, the E3 ubiquitin ligases involved in actin turnover is still unclear. According to Cohen *et al*, the thin filaments are degraded in a distinct mechanism that does not involve known E3 ubiquitin ligases such as MurF1 or Atrogin 1 (Cohen *et al*., 2009). Therefore, the exact role of COMMD4 with regards to actin turnover and the ubiquitin machinery would warrant further independent investigations.

4.2.4. Limitations of Y2H

While the Y2H system is a well-known and popular technique to map protein-protein interactions *in vivo*, it has its limitations. One such limitation that should be considered is whether the right library was used and how the library was constructed, as this could have an affect on the results obtained. For the present study, a pre-transformed human MATCHMAKER cardiac cDNA library from Clontech, generated via Xho1-$(dT)_{15}$ priming was used for the Y2H library screen. In such libraries, only one out of six of all cloned inserts in oligo dT-primed cDNA libraries are in-frame with the GAL4 AD domain (van Criekinge and Beyaert, 1999); this explains the reason why numerous clones were discarded (Section 4.2.1). These clone inserts were out of frame with respect to the GAL4 AD domain and resulted in no protein matches upon translation. Moreover, the N-terminal regions of large proteins are also poorly represented in oligo dT-primed cDNA libraries: the average insert size of these libraries is 2kb, with a range between 0.4 and 4.0kb, (Adler *et al*., 1997). This explains why cMyBPC was not identified as a putative interactor of COMMD4.

Furthermore, the stringent use of heterologous bait matings to eliminate false positives could lead to some true interactors not being identified. This may explain the fact that COMMD4 itself and probably any other COMMD proteins could have been missed. One of the bait controls used in the present study, which is provided by the manufacturer of the MATCHMAKER cDNA library (CLONTECH), encodes murine p53, a known target of NF-κb (Ryan *et al*., 2000). The fact that COMMD4 and several COMMD proteins interact with NF-κB and regulate its activity (Section 1.3.4.1.2), could result in non-specific binding to this bait, most likely due to p53-NFκB interactions, and thus could explain the absence of these proteins in the list of COMMD4 preys identified.

For Y2H, interacting proteins have to be able to localize to the nucleus to facilitate the activation of transcription of each of the reporter genes (Mcalister-Henn *et al.,* 1999). This may prevent some proteins, such as transmembrane proteins, from interacting appropriately. For the present study, this was not a problem as COMMD4; similar to COMMD1, is capable of transport between the nucleus and the cytosol.
Furthermore, interactions that depend on posttranslational modifications (such as phosphorylation) may not be detected in a Y2H system, as these modifications either do not occur, or occur inappropriately, in yeast cells (Golemis, 1999). Usually this problem can be circumvented by co-expressing the enzymes required for posttranslational modifications in the yeast cell. However, at the time of the Y2H library screen, very little

information was available regarding COMMD4 and especially its regulation, and hence no attempt was made to account for posttranslational modification; it remains possible that this may have resulted in the loss of potential putative interactors.

In some cases, fusion proteins expressed in yeast cells may either be toxic to the yeast and may interfere with the functions of proteins that are essential for the proper functioning of the Y2H system or inhibit growth (Van Criekinge and Beyaert, 1999). In the present study, the COMMD4 bait construct was slightly toxic to the yeast strain AH109 and caused the AH109 strain containing COMMD4-pGBKT7 to grow slower in liquid cultures compared to controls. However, this problem was overcome by growing the AH109 yeast strain for subsequent mating and library screens on agar plates instead of in liquid cultures.

Another important limiting factor of Y2H is that it produces a large number of false positives (Vidalain *et al*., 2004), such as the case with ANKRD1 (Section 4.2.2), in the present study. However, the number of false positives, interactions that occur due to auto-activation of reporter genes by bait or/and prey constructs and alterations in yeast physiology, was reduced by rigorous quality control tests of multiple reporter gene activationand heterologous matings (Gietz *et al*., 1997).

Thus, wherever possible, steps were taken to reduce the impact of the limitations that typically haunt Y2H studies.

4.2.5. Limitations of verification assays

4.2.5.1. I*n vitro* Co-IP

Co-IP analysis is one of the classical biochemical methods for confirming protein-protein interactions. In the present study, both *in vitro* and *in vivo* co-IP assays were used for verification of the Y2H analysis. A limitation of the *in vitro* method is that the protein-protein interactions detected may, in fact, not be representative of true physiological interactions that occur *in vivo*. The reason for this is that in an *in vitro* Co-IP system, like the one used in the present investigation, proteins that normally are unable to interact with each other because of their cellular localisations, may now do so. However, in the present study, this problem was overcome by only testing proteins that occur in cellular compartments where COMMD4 has been reported to localise. Finally, in *in vitro* systems, peptides may not be able to fold into their proper, physiologically active conformations, while any post translational mofication required for proper functioning may be be absent. Thus, *in vivo* Co-IP experiments using appropriate cell lines, as was done in this study, was another way to circumvent this limitation.

A further limitation of the *in vitro* Co-IP system is the fact that some of the peptides under investigation may bind the protein G agarose beads non-specifically. This may cause these proteins to precipitate because of their binding to the protein G, rather than via their binding partner and, in so doing, yield a false positive precipitation result. However, to overcome this limitation, in the present investigation, the protein G agarose beads were pre-washed multiple times.

In the present study, *in vitro* Co-IP analysis confirmed the interactions of COMMD4 with SNX3, ACTC1, LGMN and ENO1, in the absence of the Y2H GAL4 domains, although the latter interaction was very weak. In contrast, no interaction was observed between COMMD4 and ANKRD1, while the interaction of COMMD4 with FBXL10 and DSCR3 could not be assessed as the molecular weights of the peptides encoded by their respective clones were too low and therefore fell outside the detection limit of the SDS-PAGE system used. Different concentration of SDS-PAGE gels were used, and other loading dye (Ficoll Orange) in which the dye front might not obscure any visible bands were tested; however no visible bands were detected. It is possible that the small fragments were unstable and degraded very quickly in the *in vitro* system. Therefore, due to time constraints, in stead of expending more effort on optimizing the *in vitro* Co-IP conditions, the interaction of COMMD4 with FBXL10 and DSCR3 was further investigated using *in vivo* Co-IP.

4.2.5.2. *In vivo* Co-IP

While the *in vivo* Co-IP assay is an improvement over the *in vitro* assay, in that it allows for the examination of protein-protein interaction in a cellular environment, this method is not without its flaws. In the present study, cell lysates expressing endogenous and/or exogenously-tagged proteins (as is the case for COMMD4, DSCR3 and cMyBPC) were used for the *in vivo* experiments. Consistent with Y2H results, interactions between exogenously expressed COMMD4 and cMyBPC, LGMN, SNX3, ACTC1, FBXL10, ENO1 and DSCR3 was observed. In contrast, no interaction was observed between COMMD4 and the endogenous ANKRD1 protein.

A general problem that is usually associated with *in vivo* Co-IP assay is the high background signals that arises when an antibody from the same species (cross-reactivity) is used for IP and subsequent WB analysis. However, in order to reduce the background signal, this problem can be resolved by careful selection of the antibodies to be used in pull-down assay or using HRP-conjugated protein A or G, rather than HRP-conjugated secondary antibodies for WB analysis. In the present study, antibodies were chosen in such a way to avoid cross-reactivity. In addition, the specificity and affinity of the antibody could also be attenuated by the 3D structure of the tagged proteins and the nature of protein complexes (Zhou and Veenstra, 2007). With *in vivo* Co-IP, one also has to take into account that it allows only detection of stable interactions and does not indicate whether the interaction is direct. Another factor that should be considered when using cell lysates is that cell lysis mixes compartments and brings proteins that are normally excluded due to their localisation in different cellular compartment, together. Therefore, 3D *in vivo* co-localisation was used to show that proteins existed in the same subcellular compartment, and thus that their interaction could be considered to be physiological relevant.

4.2.5.3. 3D co-localisation

In the present study, 3D *in vivo* co-localisation in live or fixed cells was used to determine whether proteins share the same subcellular environment, thereby further supporting the physiological relevance of the

interaction as confirmed via Co-IP. It is important to note that for co-localisation, 3D Z-stacking was performed in order to map the entire volume of the cell, creating a 3D image of the sample under investigation, rather than simply using a two-dimensional overlay of fluorescence from the proteins under investigation; thus the technique used was able to establish whether two proteins actually share a subcellular compartment, rather than just appear to do so in a two-dimensional view.

We were able to show that COMMD4 shares the same subcellular space as ACTC1, SNX3, DSCR3, ENO1, and LGMN1 in live cells, but does not do so with ANKRD1. Given the fact that the GFP and YFP fluorophores were used, spectral unmixing was necessary to differentiate between the overlapped emission spectra of these fluorophores (Section 2.16.3.2). However, with spectral unmixing, the co-localisation signal might decrease; therefore the co-localisation assays for each of these proteins with COMMD4 were repeated, this time using immunocytochemistry with fluorophores with non-overlapping spectra. Similar results were observed for both approaches.

For the functional studies (domain/phosphorylation interaction assay and the effect of COMMD4 knockdown), the co-localisation signal was quantified; therefore co-localisation was performed in triplicate in two independent experiments to overcome problems that might arise from differences in protein expression of transfected cells.

Despite the advances in live cell imaging, several major problems are associated with this technique. As mentioned earlier, signal bleeding due to overlapping spectral properties, as is the case when the combination of GFP and YFP fluorophores are used, might influence accurate co-localisation analysis (Bolte and Cordelieres, 2006). However, this problem can be addressed by using fluorescent proteins that are spectrally separate from each other and this highlights the importance of choosing the right fluorophore. However, at the time of the study, the GFP and YFP-fluorophores vectors were the only ones conveniently available in the laboratory. In time, the availability of required primary antibodies against some of the proteins of interest as well as fluorescently-tagged secondary antibodies made immunocytochemistry an alternative possibility to re-confirm the co-localistion results. The latter was simple and quicker to use than the recloning of proteins into non-overlapping fluorophore vectors.

Furthermore, co-localisation can also be limited by the optical resolution of the fluorescence microscope, which is 200nm (Day and Schaufele, 2005); because of this, substantial distances might actually separate proteins that appear to be co-localised. Thus, in the present study, vertical Z-stack images were acquired at 260nm intervals, creating a 3D image of the total volume of the cell, and hence co-localisation results were taken as evidence of proteins sharing a similar space. These data were seen as support for physiological relevance of a given interaction also verified by *in vivo* Co-IP, rather than stand-alone evidence for protein-protein interactions, given that even in any image of 260nm, there may still be sufficient distance between molecules to prevent their interaction.

Another problem of live-cell imaging is the heterogeneity that arises from differences in protein expression level, particularly when using transiently transfected cells (Voss *et al*., 2005), as one batch of cells might have had a higher transfection efficiency than the next. This might pose a problem for subsequent analysis, as a single image does not adequately represent a population of cells that express the tagged protein. However, this problem can be overcome by using cloned cell lines that stably express the fluorescently-tagged protein. Alternatively, as was done in the present study, co-localisation assessments can be repeated using a number of cells, in a number of independent experiments (in this study, three cells were assessed in two or three independent experiments). Where cells do not have to be studied while alive, immunocytochemistry can also circumvent this problem, as was also employed in this study.

Furthermore, fluorescence-based approaches also suffer from cellular autofluorescence, whether arising from fluorophores or background staining by antibodies (Piehler, 2005). However, in the present study, this was not a problem since cell specimens were carefully handled and the proper controls were also used. In addition, both primary and secondary antibodies used were tested for cross-reactivity and non-specific background staining, and non-transfected cells were used as controls.

4.3. POSSIBLE MECHANISMS OF ACTION OF COMMD4 IN cMYBPC FUNCTIONING

In summary, it is clear that the COMMD4-interacting proteins, identified by Y2H and confirmed as interactors via the various adjunctive techniques, have established roles in the protein trafficking and degradation pathway. Given the functions of these prey proteins, and given that COMMD1, which is also ubiquitously expressed, like COMMD4, has also been implicated in protein trafficking and in the regulation of protein stability via ubiquitin modification, it seems plausible that COMMD4 fulfills a similar role to COMMD1 in these processes. Thus COMMD4 seems to facilitate the assembly of protein complexes with the E3 ubiquitin ligase and help to stabilise the target substrate for subsequent ubiquitin conjugation.

Since COMMD4 has been identified as a novel interactor of the cMyBPC motif, the question arises as to how COMMD4 fits into the picture with regards to cMyBPC functioning.

As described in section 1.4, the assembly and turnover of the cardiac sarcomere is important for cardiac performance and requires the integrated involvement of chaperone proteins, calpains and the UPS system. While the chaperones aid in the prevention of aggregate formation and help sort misfolded proteins to the proteasome (Westhoff *et al*., 2005); the calpain system mediates the dissociation of sarcomeric proteins from the assembled myofibril before it can be degraded via the UPS. The UPS recognises and degrades 60% of muscle proteins, via the N-end rule mechanism, except for myosin and actin (Solomon *et al*., 1998; Lecker *et al*., 1999). Thus, muscle proteins available for proteolysis would undergo N-terminal modifications in order to create or expose a functional degradation signal (N-degron) consisting of a destabilizing N-terminal residue and an internal lysine residue within the substrate; however the exact mechanism involved in these

processes as well as the regulation involved in scheduled degradation of muscle proteins is still unclear (Solomon *et al*., 1998) (Section 1.4.3.4).

As MyBPC is a short-lived protein, and subjected to scheduled degradation, we speculate that the N-terminal domains of cMyBPC, most likely the cMyBPC motif might house an internal lysine that forms part of the N-degron. These observations are made based on the N-terminal degradation mechanism involved in sarcomeric protein degradation as well as findings from Bahrudin *et al* (2008). The latter study showed that, compared to four other mutations, only the missense mutation Glu334Lys within the cMyBPC motif had an affect on the stability of the cMyBPC mutant protein and caused UPS impairment. This is most likely due the introduction of the lysine group, which would allow for recognition and subsequent ubiquitination of the mutant protein via an E3 ubiquitin ligase (Bahrudin *et al*., 2008; Wang *et al*., 2008).

Given the likely involvement of COMMD4 in protein trafficking and turnover, the following proposal of the involvement of COMMD4 in cMyBPC turnover can be made:

cMyBPC is a known calpain substrate, therefore enhanced calpain activity would mediate the dissociation of cMyBPC from the assembled myofibril prior its degradation via the UPS (Huang and Forsberg, 1998). When unphosphorylated WT-cMyBPC is released from the sarcomere, such as during ischemia (Decker *et al*., 2005; Sadayappan *et al*., 2006) (Section 1.2.4.2.5), COMMD4 binds to the cMyBPC motif (Figure 4.1A). COMMD4, as we showed in this study, also binds to FBXL10; this binding may direct FBXL10 to cMyBPC, although the possibility of a direct binding of FBXL10 to cMyBPC cannot be excluded, as atrogin 1, another F-box protein, is known to bind WT-cMyBPC directly (Mearini *et al*., 2009). FBXL10 would most likely associate with cMyBPC via its LRR substrate binding domain, and would also bind to the anchor protein, Skp1, via its F-box motif (Jin *et al*., 2004) (Figure 4.1B).

Furthermore, the SCF subunit, cullin, functions as a scaffold that would bind Skp1 via its N-terminal and bind the RING finger protein, Rbx1, which in turn would recruit the E2 enzyme via its C-terminal (Section 1.4.3.3) (Figure 4.1B). Prior to the assembly of Skp1-FBXL10-WT-cMyBPC to the cullin/Rbx1 subunits, the latter complex would first be neddylated via Nedd8 to be activated. Neddylation would allow the dissociation of the sequestration factor CAND1, which in turn would allow binding of WT-cMyBPC-FBXL10-Skp1 to the cullin-Rbx1 complex as well as binding of the E2 enzyme to Rbx1 (Skaar and Pagano, 2009). Neddylation would also allow Rbx1 via its flexible linker to be released from its tight association with cullin (Figure 4.1 C).

Thus: once FBXL10 is bound to WT-cMyBPC, it in turn would recruit the targeted cMyBPC to the SCF E3 ubiquitin ligase complex upon its activation. Once assembled and activated, the SCF E3 ubiquitin ligase complex, which is known as “N-terminal readers”, would then mediate transfer of activated ubiquitin from the E2 ubiquitin-conjugating enzyme to a lysine on cMyBPC via the Rbx1 flexible linker (Section 1.4.3.3).

According to structural modelling, the different subunits of the SCF E3 ubiquitin complex is orientated in a single rigid C-shaped superstructure with a conserved distance of ~50Å between the substrate binding surface of the F-box protein and the E2 enzyme (Cardozo and Pagano, 2004; Skaar and Pagano, 2009). Thus, the flexible Rbx1 linker would be able to bridge the conserved gap and allow subsequent ubiquitin transfer to the substrate (Skaar and Pagano, 2009) (Figure 4.1 C). In addition, the F-box protein must also be appropriately docked to Skp1 and the substrate must be stably tethered for subsequent ubiquitin conjugation. Thus, any slight disturbances in the orientation at one end of the complex would have downstream consequence on the other end of the complex which might affect the mechanism of ubiquitin ligation (Cardozo and Pagano, 2004). In addition, F-box proteins on their own are unstable and once a substrate is released, these F-box proteins can be self ubiquitinated via the SCF E3 ubiquitin ligase itself, which, in turn, would attenuate the activity of the SCF E3 ubiquitin ligase. The mechanism involved with regard to the self-ubiquitination is unclear; however one can assume that COMMD4 might also play a role. In addition one can assume that unstable tethering of substrates would result in similar consequences.

As COMMD4 binds cMyBPC and the F-box protein within the SCF E3 ubiquitin complex, COMMD4 may aid in stabilising the interaction between WT-cMyBPC and the E3 ligase (Figure 4.1C). This stabilising effect would aid in the orientation of the FBXL10 and WT-cMyBPC with regards to the E2 enzyme such that the appropriate distance that is necessary for effective ubiquitin conjugation is attained (Cardozo and Pagano, 2004; Skaar and Pagano, 2009). It is also possible that COMMD4 binds the cullin-Rbx 1 complex within the SCF E3 ubiquitin ligase complex, just as COMMD1 does; even though cullin and/ or Rbx 1 was not identified as a ligand of COMMD4 in the current study, it is possible that these proteins could have been excluded by over-stringent elimination of putative interactors during the heterologous bait-mating step.

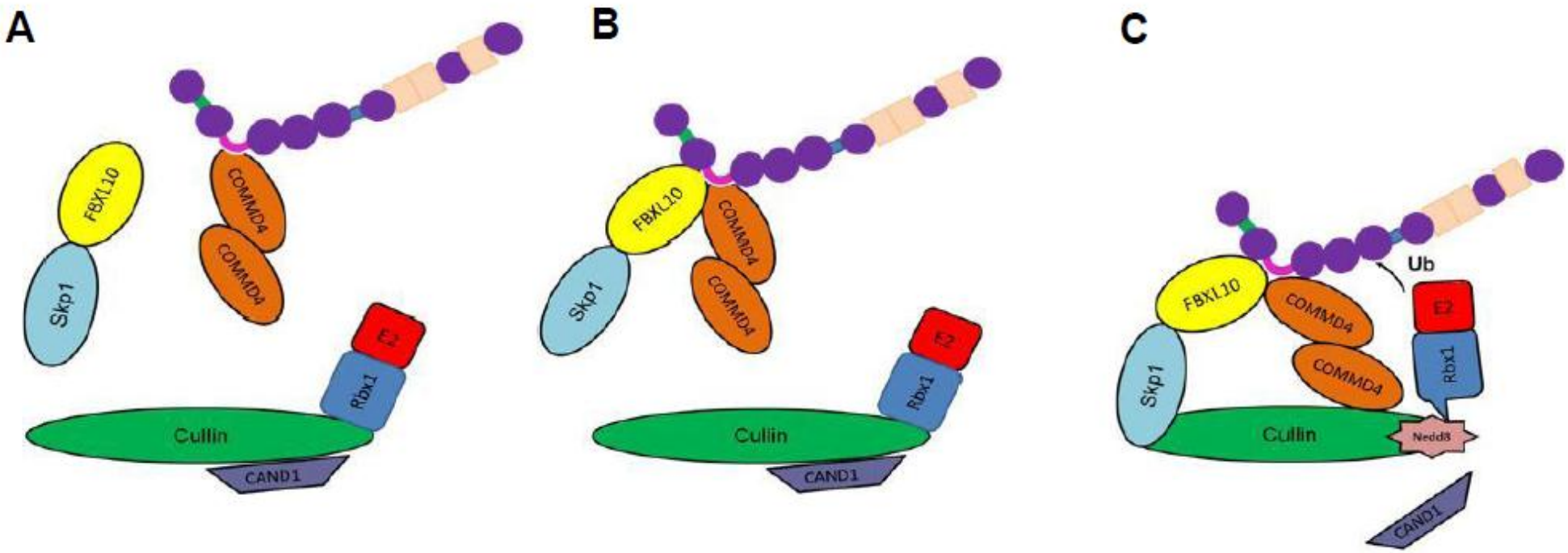


Figure 4.1. Schematic depicting the proposed role of COMMD4 in the transfer and stabilisation of cMyBPC within the SCF E3 ubiquitin ligase complex. I.A. COMMD4 binds to the m-domain of dephosphorylated cMyBPC, which had been released from the sarcomere. **B**. COMMD4 recruits F-box protein (FBXL10), which possibly also binds to cMyBPC via its LRR domains, and which binds Skp1 via its F-box motif. Meanwhile the scaffold protein of the eventual SCF E3 ubiquitin ligase complex, Cullin, binds to the RING protein, Rbx1, which, in turn, recruits the E2 enzyme; this complex would remain inactivated due to the binding of the sequestering factor, CAND1. **C**. Upon neddylation via Nedd8, CAND1 dissociates and the full SCF E3 ubiquitin ligase assembles and becomes activated, involving release of Rbx1 on a flexible linker, to facilitate transfer of activated ubiquitin to the cMyBPC substrate. COMMD4, by binding FBXL10, the MyBPC substrate, and Cullin, stabilises the complex and facilitates presentation of the target at the angle and distance required for efficient ubiquitin transfer by the E2 enzyme. Modified from van de Sluis *et al*., 2007

Following the initial mono-ubiquitination event, a poly-ubiquitin chain of at least four, sequentially-attached, ubiquitins are necessary for subsequent target-protein recognition, after which the protein may be shunted towards recycling, possibly by the involvement of SNX3 and DSCR3 or towards either proteasomal or lysosomal degradation (Sarikas *et al*., 2005; Mukhopadhyay, 2007). Escorting target proteins to the proteasome can be achieved either by direct association of the E3 ubiquitin protein ligase to the proteosome or via adaptor proteins that associate with poly-ubiquitinated proteins and specific proteasomal subunits (Hartmann-Petersen *et al*., 2003). Interestingly, an interaction between COMMD1 and the 20S proteasome has been observed (van de Sluis *et al*., 2009), therefore one can speculate that COMMD4, similar to COMMD1 may also play a role in targeting substrates to the proteosome as it binds subunits of the SCF E3 ubiquitin ligase and most likely the proteasome as well (Section 1.3.4.1.2 and Section 1.3.4.4.3).

In recent years, protein turnover has been implicated in cardiac hypertrophy development: Sarikas *et al* (2005) proposed the involvement of the UPS and to a lesser extent, the lysosome, in cMyBPC turnover, and in particular, impairment of the UPS plays a role in cMyBPC-derived HCM. Briefly, their paper demonstrated that an 80%-truncated protein, which lacked most of the cMyBPC molecule C-terminal from the cMyBPC-motif (M7t, expressing only domain C0 and partial domain C1), formed ubiquitinated aggregates in the cytoplasmic as well as peri- and intranuclear regions. In some cells, this M7t protein mis-incorporated into the Z-disk, probably due to the remaining actin binding site in C0 and/or the PA region. In contrast, a 3%-truncation, which left the cMyBPC-motif intact (M6t, expressing most of the cMyBPC molecule, with only a partial deletion of domain C10), did not form such aggregates and showed diffuse A-band incorporation. The aggregates composed of the 80%-truncated protein were ubiquitin-positive and were thus marked for degradation, but somehow could not be processed, which further indicated UPS impairment. Interestingly, an increased rate of apoptosis was also observed in cells expressing the M7t protein compared to WT cMyBPC. Although the authors speculated that the mutant cMyBPC provided an unusually strong degradation signal and effectively competed with other UPS substrates, this does not explain the mechanism by which these aggregates induced UPS impairment.

In line with these findings, in our study, we show that cMyBPC accumulates in the endosome after RNAi-mediated COMMD4 knockdown (Figure 3.1.4; Section 3.3.2). In addition, knockdown of COMMD4 also affected cell growth and morphology and seemed to induce apoptosis, since increased dead cells were observed in COMMD4-knockdown cells, compared to control cells in which COMMD4 expression was unimpeded. We therefore speculate that the data presented by Sarikas *et al* (2005) could be explained by the fact that the 80%-truncated cMyBPC lacked the ability to bind COMMD4, whereas the 3%-truncated version retained this ability as the cMyBPC motif remained intact (Figure 4.2 I and II). In M7t, the lack of COMMD4 binding may result in an unstable interaction between the 80% truncated protein and the SCF complex via FBXL10, such that correct orientation as well as the distance needed for effective ubiquitin conjugation would be obtained only occasionally by random molecular movement. Thus, some of these very short proteins in the Sarikas *et al.* study may have gone through the first step of being marked for processing,

but were not likely to be systematically poly-ubiquinated, because of the absence of the stabilising influence of a COMMD4 attachment to the truncated cMyBPC protein (Figure 4.2.II). In contrast, the longer truncation mutants, which had retained the ability to bind COMMD4, would be ubiquitinated in the usual way (Figure 4.2.I). As poly-ubiquination is necessary for degradation by the UPS (Mukhopadhyay, 2007), these mono-ubiquinated, truncated cMyBPC cannot be degraded by the UPS, nor can they be recycled via SNX3 and DSCR3 in the absence of COMMD4-binding, and thus accumulate as aggregates (Weekes *et al*., 2003; Zheng et al., 2010; Portbury *et al*., 2011).

Accumulation of these ubiquitinated aggregates poses a problem as it is deleterious to the cardiomyocytes.. These aggregates impair the UPS by blocking the proteasome either due to size constraints, or by depletion of free ubiquitin and consequential inactivation of UPS components, or by other as yet unknown mechanism (Dantuma *et al*., 2006, Salomons *et al*., 2009; Verhoef *et al*., 2009; Willis *et al*., 2009). Upon UPS impairment, autophagy is increasingly activated (Tannous *et al*., 2008), which leads to lysosomal-mediated degradation.In our study, accumulation of cMyBPC in the endosomes after knockdown of COMMD4 (Figure3.14; Section 3.3.2) may reflect increased shunting of the protein towards lysosomal degradation and autophagy, possibly because the absence of COMMD4 may lead to decreased SNX3 and DSCR3-mediated recycling.

Given the fact that the UPS tightly controls the regulation of proteins involved in intracellular signalling and transcription, Sarikas *et al* (2005) speculated that impairment of the UPS leads to a deficit in protein turnover. This in turn results in increasing levels of, and continued signalling by, proteins involved in cellular hypertrophic signalling or apoptosis, such as human growth hormone receptor and β-adrenergic receptorsand their ligands, various transcription factors, p53 and calcineurin A. In addition to the UPS impairment, COMMD4 would itself not be degraded, and so would also be available in excess in the cytosol, which could have detrimental effects on the cardiac cell, as detailed below.

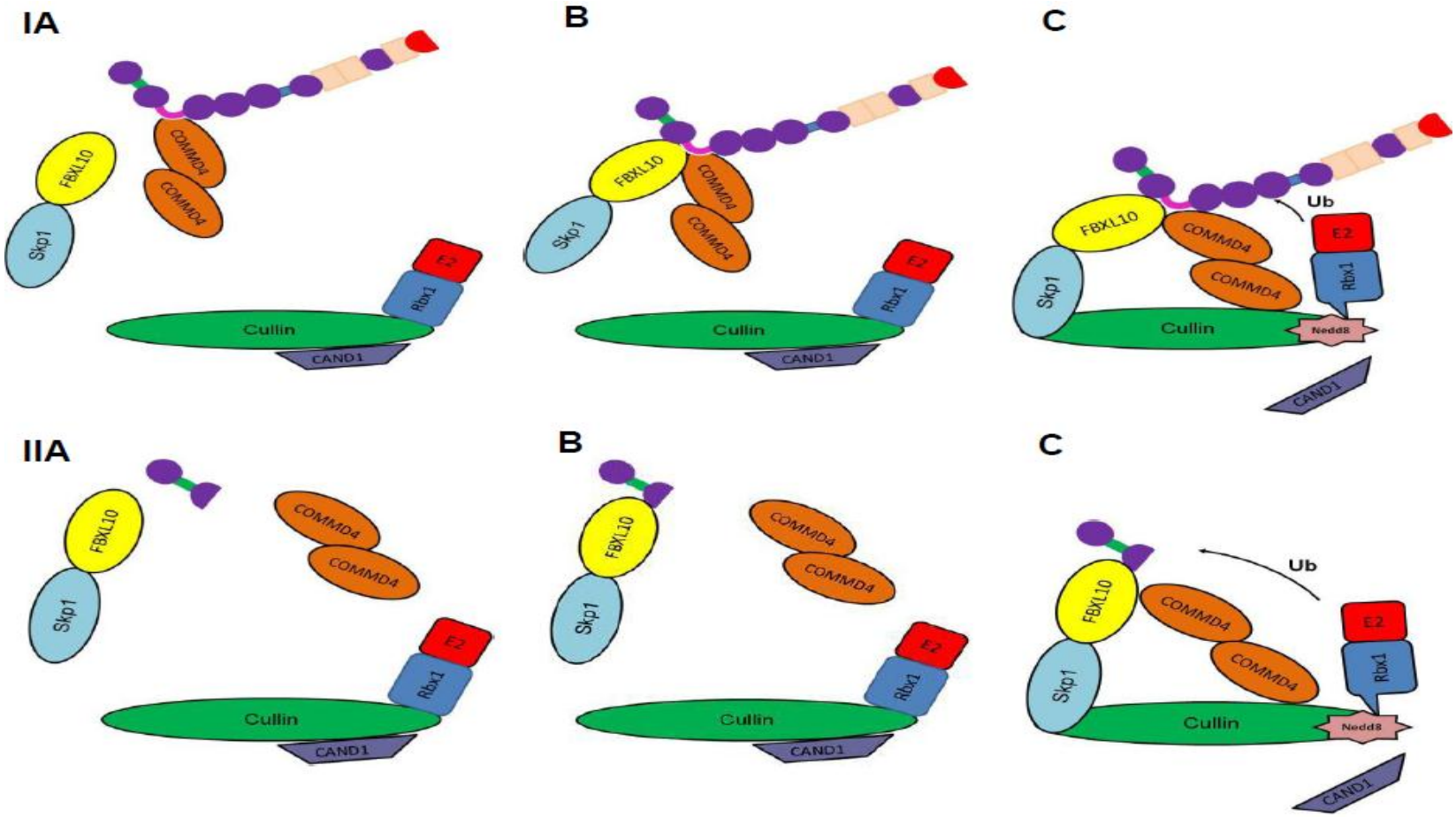


Figure 4.2. Schematic model depicting the proposed role of COMMD4 in the transfer and stabilisation of the 3% truncated cMyBPC (I) as well as lack of proper stabilisation of the 80% truncated cMyBPC (II). I. Just as with wild-type cMyBPC in figure 4.1, ubiquitin transfer can occur unimpeded with the 3% truncated cMyBPC that retains the m-domain (M6t from Sarikas *et al., 2005.*).**II. A.** The 80% truncated cMyBPC (M7t from Sarikas *et al*.,2005), which is comprised of only domain C0 and partial domain C1, is unable to bind COMMD4 due to the absence of the m-domain, although (**B.**) FBXL10 may still be able to bind the truncated protein. **C.** Without stable tethering of the cMyBPC fragment to FBXL10 and the rest of the SCF complex via COMMD4, the requisite conserved orientation of the target with regard to the rest of the SCF E3 subunits is disturbed, which affects the efficiency of ubiquitin transfer. Due to molecular movement, the severely truncated cMyBPC might randomly present in the correct orientation, allowing for mono-ubiquitination, but not for structured poly-ubiquitination. This in itself may tie up the UPS, impairing the system by effectively depleting its components from the cytoplasm through the formation of aggregates. Modified from van de Sluis *et al*., 2007

4.3.1. Other possible effect of increased COMMD4 levels

Increased levels of COMMD4, such as my result from an impeded UPS in the presence of severe cMyBPC truncation mutations, would result in increased efflux of copper from the cell and would lead to copper depletion (Section 1.3.4.2.2). In this respect, it is interesting that studies in humans and animals have also previously shown the role of dietary copper depletion in cardiac hypertrophy development, although the mechanism for this is not completely understood (Elsherif *et al*., 2004;Freisinger *et al*., 2004). Thus, it is tempting to consider that this mechanism could contribute to the cardiac hypertrophy seen in HCM associated with cMyBPC truncations Conversely, as it is evident that intracellular copper levels are associated with the structural stability of the COMMD4; one could further speculate that this in turn might have an effect on the turnover of the remaining wild-type cMyBPC in heterozygous patients, and thus further affect muscle structure and function.

Alternatively, based on the inhibitory effect of COMMD proteins on NFκB, the increased levels of COMMD4 could also mitigate against the development of hypertrophy (Section 1.3.4.1.2.). Both *in vitro* and *in vivo* studies have shown that NFκB activation is required for cardiac hypertrophy development and that this hypertrophic response is attenuated following inhibition of NFκB activation (Purcell *et al*., 2001; Li *et al*., 2006). However, the inhibition of NFκB via COMMD proteins might not be adequate to inhibit this pathway completely, given the fact that NFκB could also be activated via a number of other signalling pathways, which include the PI3K/Akt/mTOR pathway, JAK-STAT-mediated signalling pathway and calcium/NFAT, all of which also have been implicated in the cardiac hypertrophic response pathway (Jones *et al*., 2003; Li *et al*., 2006). Thus, the discovery that COMMD4 is involved in cMyBPC turnover opens up a new area for investigation into the regulation of cMyBPC-associated hypertrophy pathways.

4.3.2. COMMD4 and cMyBPC turnover

As mentioned in section 1.4, myofibrillar proteins in the sarcomere are in a constant flux between degradation and resynthesis, in order to adapt the muscle to physiological stresses and requirements; this process, which is a standard part of normal sarcomeric functioning, has to happen without affecting contractile function. However, despite the importance of the UPS in the turnover of sarcomeric proteins, little is known about the mechanisms that underlies the functional association between the UPS and sarcomeric proteins.

In this study, the turnover of cMyBPC was of particular interest. As mentioned in section 1.2.4.2.5, under normal physiological conditions, cMyBPC exists in two different forms depending on its phosphorylation status. The phosphorylated form is associated with greater thick filament stability and is protected against proteolysis, while the unstable, dephosphorylated form is associated with lower thick filament stability and is easily proteolysed (Kulikovskaya *et al*., 2007). Interestingly, most of the proteolytic fragments of the dephosphorylated form stained with the antibody, anti-C0-C2 but not with anti-C5 or anti-C8C9. This suggests that cleavage of cMyBPC by calpains into large polypeptide fragments suitable for proteosomal

degradation occurs from the C-terminal region, while the N-terminal region at first remains intact. Sequencing confirms that the N-terminal region of the dephosphorylated form was still intact, apart from the replacement of an arginine residue for lysine (Kulikovskaya *et al*., 2007). This lysine residue within the N-degron, would functions as an acceptor the poly-ubiquitin chain which would make this dephosphorylated form more susceptible to proteolysis (Tasaki et al., 2007) (Section 1.4.3.4). Thus, the complete release of cMyBPC from the myofibril and its subsequent proteolysis, involving the UPS would result in the disassembly of myofibrillar thick filament (Kulikovskaya *et al*., 2007; Goll *et al*., 2008). However, just what signals are integral to the scheduled degradation of cMyBPC, remain unknown; although N-degrons are thought to be important for scheduled degradation, it is not clear what triggers calpains to cleave cMyBPC and how that ties in with scheduled degradation.

The importance of the N-terminal region in the regulation of scheduled degradation of cMyBPC is further highlighted in a study by Bahrudin *et al*, which reported the first missense mutation (Glu334Ly), located in the cMyBPC motif to influence the stability of cMyBPC and cause UPS impairment (Bahrudin *et al*., 2008). The change from a negatively charged glutamic acid to a positively charge lysine was shown to affect the stability of the mutant protein and the authors speculated that the introduction of the lysine group allowed the poly-ubiquitination of the mutant protein. This gives added support for the role of the N-terminal region of cMyBPC for ubiquitin modification and thus for the regulation of cMyBPC protein turnover. Furthermore, this also highlights the importance of the cMyBPC motif as part of the N-degron, as mutations that result in modification of the motif or causes its truncation results in the impairment of the UPS due to ubiquitinated aggregates (Sarikas *et al*., 2005; Bahrudin *et al*., 2008). Furthermore, as there are numerous point-mutations in cMyBPC, including mutations in the C1-C2 region which do not introduce lysine residues, for which the patho-mechanisms are still unknown, our results suggest that one possible mechanism may be by affecting the binding of COMMD4, and thereby affecting cMyBPC ubiquitination and turnover.

As phosphorylation of cMyBPC protects it against proteolysis, one could speculate further that the cMyBPC motif houses the internal lysine that forms part of the N-degron and that phosphorylation of the cMyBPC motif blocks certain E3 ubiquitin ligases from recognising the ubiquitination signal and protects against degradation. Thus, upon release of dephosphorylated cMyBPC from the sarcomere, it is likely that COMMD4 together with the SCF E3 ubiquitin ligases ubiquitinate the N-terminal of cMyBPC, and marks it for degradation via the UPS. TheY2H data from Ramburan, A (2008), previous β-galactosidase assays and the data from the domain/phosphorylation assay have shown that COMMD4 is able to bind the unphosphorylated and monophopshorylated cMyBPC motif, but not the fully phosphorylated cMyBPC. This may at least partially explain the finding that fully phosphorylated cMyBPC is stable and resistant to proteolysis. As mentioned in section 1.4.2, other E3 ubiquitin ligases such as MuRF-1 also targets cMyBPC, under stress conditions such as atrophy. Furthermore, MuRF-1, unlike the SCF E3 ubiquitin ligases, is a single subunit E3 ubiquitin ligase and can ubiquinate cMyBPC directly even when it's still associated within the myofibril (Cohen *et al*., 2009). However the direct binding with cMyBPC within the myofibril has only

been observed during atrophying conditions, even though MuRF1 has been implicated in regulating myofibrillar proteolysis in both cardiac atrophy and hypertrophy (Adams *et al*., 2008). Whether MuRF-1 binds to specific regions on cMyBPC is also still unknown. In addition, whether the same mechanism that stimulate degradation are involved during normal or stress conditions, are also still unclear.

Taken together; it is evident that from the abovementioned findings that turnover of cMyBPC is a regular process that occurs under normal physiological conditions and not just under pathological conditions. In addition to cMyBPC phosphorylation, it seems that its ubiquitination and deubiquitination adds another regulatory step with regard to its function in the sarcomere. Furthermore, studies have also shown that de-ubiquinating enzymes, such as USP25 interacts with other sarcomeric proteins (actin, filamin C and sscMyBPC) (Bosch-Comas *et al*., 2006) and provides an additional mechanism for regulating their turnover; therefore sarcomeric protein turnover is a standard and necessary part of normal sarcomeric functioning. Since protein turnover, is a highly regulated process, any alteration to this could therefore have detrimental effects.

4.4. FAMILY-BASED ASSOCIATION STUDIES

As described in section 4.2, evidence from our Y2H studies suggested that COMMD4, a ligand of cMyBPC, and its interactors may play a vital role in the turnover of cMyBPC and possibly other sarcomeric proteins such as actin. Given the fact that sarcomeric protein turnover has recently been implicated in hypertrophy development, COMMD4 and those of its interactors that are involved in the trafficking/turnover pathway were chosen as plausible candidate modifiers of sarcomeric function and contractility. Thus, for the present study, a total of twenty-six SNPs in five genes, which include *COMMD4*, *SNX3*, *DSCR3*, *LGMN* and *FBXL10* were assessed for their potential role in hypertrophy development in a family-based association study using cohorts of South African HCM families.

At the beginning of the study, ASREA was employed (Section 3.5.3.1) for the genotyping of *COMMD4* as this PCR-based method was simple and cheap. SNPs were selected from public domain databases based on high heterozygosity values, validation status and whether the nucleotide variant affected a restriction enzyme site. However, during the course of this study, improved infrastructure became available, and thus, ABI Taqman validated SNP genotyping assays were employed to genotype SNPs in genes encoding COMMD4-interactors. These SNPs were selected using the SNPbrowser™ Software at an even spacing of 0.5 LDUs in the CEU and YRI HapMap populations. The abovementioned populations were used since the LD structure of our own South African population is still uncharacterised and these populations represented the HapMap populations likely closest to our own population.

For subsequent association analysis, QTDT was employed to test for association of any of the SNPs with major hypertrophy indices. As some of the families in our cohort are derived from South African Caucasian and others from Mixed Ancestry descent, we ran population stratification tests to establish whether the total

model could be used, or whether the orthogonal model, which is robust to population stratification, should be employed. No evidence of population stratification was noted for *COMMD4*, *DSCR3*, *LGMN* and *FBXL10,* therefore the total model, which uses all available data, was employed to test for association. In contrast, significant evidence for population stratification was found at some *SNX3* loci, thus either the orthogonal or total model, where appropriate, were employed for association analysis of SNPs in *SNX3*. No evidence of association was observed between markers of *COMMD4*, *LGMN*, *FBXL10* and the various hypertrophy indices. Lack of association is not evidence of absence of association; thus, replication of this study using different populations would be required to ascertain the involvement of these genes in the hypertrophy phenotype. For *DSCR3*, a significant association was observed between marker rs2835680 and the Comp1 score with a p-value of 0.045. However, considering that this association was not observed with rs9301564, which was in perfect LD with rs2835680 in the study population, this association was considered inconclusive, and further studies either in other populations, or with additional markers, should be considered. In contrast, strong evidence of association was observed in our population for various markers in *SNX3*. These results thus implicate *SNX3* as a candidate modifier in the development of HCM.

4.4.1. *SNX3*

As described earlier in section 4.2.3.1, *SNX3* encodes a protein that forms part of a protein family that is involved in intracellular trafficking. Moreover, our Y2H data has implicated this protein, together with COMMD4, in the trafficking and turnover of cMyBPC and maybe other sarcomeric proteins.

In addition to this, our association study suggests for the first time that *SNX3* might also play a role in the development of hypertrophy, independent of BP, the primary causal HCM mutation and other known hypertrophy covariates. Of the five *SNX3* markers screened in the present HCM cohort; significant association with major hypertrophy indices was found for markers rs579257, rs635051, rs9398166. In contrast, no evidence of associations was found for *SNX3* markers rs628424 and rs588409 and any of the various hypertrophy indices. Even though this appeared curious, given the strong LD that was observed between these markers and the hypertrophy-associated *SNX3* markers,, haplotype analysis revealed an 8kb haplotype block only between markers rs635051 and rs9398166 (Figure 3.20), which could explain the lack of evidence for association of markers rs628424 and rs588409 with hypertrophy. The incomplete LD between rs579257 and the non-associated markers could further explain this phenomenon.

The G-allele of rs579257 was associated with an increase in mIVST (1.860mm), mLVWT (1.713mm) and the Maron score (5.340mm); the C-allele of rs9398166 was shown to increase mPWT (0.514mm), while the C-allele of rs635051 was shown to increase LVM (18.725g), mPWT (0.497mm) and Comp1(14.134mm). The latter is a composite hypertrophy score that takes into account the extent and distribution of hypertrophy, and is more heritable than LVM within this HCM family-cohort (as described in section 2.23). Since Comp1 is derived by principle component analysis, the effect size of this trait, unlike those for the other indices which showed significant associations, cannot be deduced from the original cardiac measurements (Carstens

et al., 2010). Therefore, the effect size for Comp1 was estimated from the CWT score, which is the mean of the 16 untransformed cardiac wall thickness measurements, as described before (Carstens *et al*., 2010). Although these effect sizes are modest, this seems to be typical for hypertrophy modifier genes; and probably reflect a synergistic and cumulative effect on the increased risk for hypertrophy development (Mayosi *et al*., 2003; van der Merwe *et al*., 2008).

However, the biological mechanisms by which these *SNX3* variants may lead to hypertrophy development have yet to be identified. All the SNPs that showed association are intronic and do not have any obvious functional roles, thus any explanation for the described associations remains speculative: Firstly, intronic SNPs might have a functional effect depending on whether the SNP in question influences gene splicing by creating or causing loss of splice donor or acceptor sites. However, this possibility was excluded since sequence analysis using the NetGene2 internet-based neural network (http://www.cbs.dtu.dk/services/NetGene2/) revealed that these polymorphisms neither create any cryptic splice sites nor have any effect on either donor or acceptor splice sites. Alternatively, it is quite possible that the associations observed are due to LD between the associated SNPs and a nearby causal variants, either within *SNX3* itself or within an adjacent gene, that has an effect on hypertrophy development. Since the LD distribution and genetic variation of *SNX3* is unknown in our population, the extent of genetic variation and LD distribution was therefore analysed in the CEU and YRI populations as they are currently the best possible representation of our own study population. Two non-synonymous SNPs were observed in SNX3, however LD analysis of *SNX3* showed that the hypertrophy-associated markers rs579257, rs635051 and rs9398166 are not in LD with these two non-synonymous SNPs within *SNX3* itself. Therefore, it is possible that either the LD structure in our population is different, or that these hypertrophy-associated markers might be in LD with another unknown functional variant in *SNX3*, or with functional variants on adjacent genes that are involved in hypertrophy development.

Another interesting possible mechanism that might underlie the association with *SNX3* in our HCM family-cohort might be deduced from previous reports from our laboratory involving RAAS. Interestingly, significant associations in our present HCM cohort have been found with hypertrophy traits and various markers in the gene encoding for the ENaC β-subunit. Moreover, gene-gene interaction analysis revealed that significant interaction between variants in Renin and the ENaC β-subunit was found to influence various hypertrophy traits (Carstens, MSc, 2009). Similarly, one could speculate that *SNX3* might have an indirect influence on hypertrophy development through gene-gene interactions between variants in the genes implicated in the RAAS system, or other genes. Importantly, as shown earlier in discussions of the protein SNX3 (Section 4.2.3.1), a connection with the ENaC subunits and SNX3 is supported, which further suggests a role for *SNX3* in a common pathway that influences hypertrophy development.

Thus, in summary, our preliminary results indicate that various loci in the *SNX3* gene might be associated with HCM. To the author's knowledge, this is the first report of association between genetic variation in

SNX3 and hypertrophy development. However, further studies of SNX3 in other HCM populations as well as in other conditions in which hypertrophy is a feature such as hypertension, diabetes, etc. will be required to confirm these findings reported here.

4.4.2. Limitations of association studies

Although the results obtained are very interesting, one should keep in mind that, with any statistical analysis of association, the association observed may still be false due to a number of confouding factors such as population stratification, phenotypic heterogeneity, statistical power and sample size and other confounding variables, which are discussed in the following sections. Furthermore, the fact that no association was observed with some genes does not necessarily mean that these genes are completely uninvolved with the condition in question, as association may occur with the same variants in other populations, and even more so when the variants investigated are not functional and association relies on LD between the studied SNP and an unknown functional variant. Thus, it is acknowledged that caution in the interpretation of results is advisable, since association studies in general remain limited in the absence of functional studies.

4.4.2.1. Population stratification

Concern regarding population stratification in association studies arises when the populations under investigation are comprised of several subpopulations and the allele frequency at a particular SNP differs in each subpopulation (Pritchard and Donnely, 2001). However, significant population-stratification has also been reported within apparently homogeneous populations, such as the North American population of European ancestry (Campbell *et al*., 2005). Thus, the issue of population stratification in any association study still needs to be addressed in order to distinguish between true and false associations, since spurious association can arise from undetected population substructure (Lange *et al*., 2008). Not only would this confound association, since it tends to cancel out true signals of genetic effects, but it also might compromise the interpretation of these results. Nevertheless, a family-based association study design offers protection against false positives due to population-stratification, since it has the advantage of a common genetic background, as transmitted alleles are matched with non-transmitted alleles from the same family, although statistical power is reduced (Evangelou *et al*., 2006). Furthermore, consistant replication of association studies in other populations also relieves the concern about population stratification.

We addressed this possible problem by first analysing our samples for population stratification prior to association analyses, as our cohort contains families derived from South African Caucasian and Mixed Ancestry descent. The latter is an ethnic admixture group, also called the Coloured subpopulation, which originated by the intermingling of several subpopulations such as the European immigrants, the local Khoi-Khoi and San peoples as well as native black Africans and later, slaves imported from Malaysia and other East Indian countries. Thus, depending on whether population-stratification was observed or not, the within-family (orthogonal model) or total model was used for association analysis, as appropriate.

4.4.2.2. Phenotypic heterogeneity

Defining the phenotype is also of utmost importance in association studies, especially when attempting to replicate a previous study. Misclassification of phenotypes, and unmeasured or poorly measured covariate factors can all contribute to reduce the power and increases the background ‚noise' of a study. Therefore, all relevant measurements to the disease phenotype must be taken into account. However, as mentioned in section 1.1.2, cardiac hypertrophy in HCM patients poses a problem as it differs between individuals not only with regards to the degree, but also the distribution of hypertrophy. This makes it difficult to differentiate a single echocardiographic measurement that accurately accounts for the total extent of hypertrophy in all patients. Previously, echocardiographic LVM was frequently used to quantify total left ventricular hypertrophy in HCM; however, it is known to be an inaccurate measure due to the variable distribution of hypertrophy in HCM, which negates the assumption of symmetrical geometry that underlies the calculation of LVM derived by echocardiography. An alternative would be to use MRI derived LVM; however cardiac MRI is not a widely available clinical modality, is expensive when it is, and cannot always be used on all patients due to contra-indications.

Therefore, in this study, single measurements as well as composite scores that encompass a number of hypertrophy indices were utilized in order to best describe the extent and degree of hypertrophy. As described in section 2.19.2, in the present study, subjects were clinically assessed via 2D and M-mode echocardiography by a single, experienced echocardiographer and 16 wall thickness measurements, in total, were taken at three different levels of the heart, from which a number of composite scores were derived.

4.4.2.3. Statistical power and sample size

For association studies, power is an important aspect of the study design as it reflects to the likelihood of the study to detect a true association or the lack thereof (Gordon *et al*., 2005). The power of a study is dependent on the size of the genetic effect, disease-allele frequency as well as sample size, however it could be compromised by genotyping errors as well as errors in phenotype misclassification (Kang *et al*., 2004, Gordon and Finch, 2005 and Laird and Lange, 2006). As mentioned in section 4.4.2.2, phenotype misclassification was reduced by the use of multiple single measurements, taken by a single experienced echocardiographer, as well as by composite scores derived from these measurements. In addition, genotyping errors were reduced by verifying input files and monitor Mendelian inheritance errors within families via Pedstats.

Non-normality of traits can also increase type I error rates and decrease the power of the study (Diao and Lin, 2006). For the present study, as a standard procedure, trait values were transformed to achieve normality (Section 2.23.1), given the fact that QTDT assumes normality of trait values. The latter also helped in the effort to minimise the effects that could influence the power of the study.

In addition to dealing with other factors that lead to increased error rates, the sample size of a study must be large enough in order to avoid type I error. One drawback of using family-based association studies relates to the calculation of statistical power, as no comprehensive method exist for calculating the power of this type of association study. Difficulties in accumulating large numbers of well-characterised families contribute to this disadvantage (Evangelou *et al*., 2006). For the above reasons, and as the present study is explorative, no formal calculations was performed to attempt to estimate the power of the present study, although every effort was made to minimise other confounding factors that may have decreased the power of the study. Moreover, this is the largest reported family cohort with founder mutations yet to be used for such a study in HCM, and every family-member willing to participate in the study was included; thus the sample size was at its maximum possible. For the majority of the association analysis, the total model was employed (when no population stratification was observed between markers and the various hypertrophy traits); in these analyses, every individual in the HCM co-hort was included. In contrast, for *SNX3*, population stratification was detected, and hence the orthogonal model, which utilises the within-family component of QTDT, was used, during which some individuals are excluded from analyses.

4.4.2.4. Confounding variables

When an association is observed between a variant and a specific phenotypic trait, it is essential to ensure that the effect is due to the effect of the variant on the trait, and not due to the effect of the variant on covariates. If no correction is made for these variables, it would result in an increase in type I error rates. Therefore, in the present study, all analyses were adjusted for known hypertrophy covariates, which include BP, HR, BSA, age, sex and the primary disease-causing mutation.

4.4.2.5. Multiple testing

When performing association tests, the issue of multiple testing always arises when different independent statistical tests are conducted (Laird and Lange, 2006). As a result of these multiple tests, the probability of obtaining false positives also increases. Therefore, it is of utmost importance to correct for multiple testing, whereby the p-value for each test performed is adjusted to keep the overall error rate less than or equal to the significance level as specified by the user, thus reducing the false positive rate. However, as a trade-off, the rate of false negatives (viz. those that are found to be non-significant are truly non-significant) is very high (Hattersley and McCarthy, 2006). In most cases, Bonferonni correction is applied; this test assumes that all tests performed are independent and each p-value is then multiplied by the number of tests performed. However, Bonferonni correction is generally considered to be too conservative, and is not applicable when SNPs are tightly linked, thus as a stringent alternative, permutation testing is usually applied. Permutation testing is not a form of multiple testing corrections; however it maintains the overall rate of false positives without inflating the rate of false negatives (Balding, 2006).

In the present study, many variants were tested against different hypertrophy indices, however these test are not truly independent tests given the fact that the LVH measurements are highly correlated and that LD

exists between several of the SNPs: therefore the Bonferonni correction would be inappropriate (Gao *et al*., 2008). Thus, for the present study, Monte Carlo permutation testing was used to determine exact p-values: during this process, genotype data are retained unchanged, while phenotype data are randomised during permutations. An added advantage of this method is that it also accounts for correlations between SNP's due to LD (McIntyre *et al*., 2000, Diao and Lin, 2006 and Sullivan, 2007).

4.5. OVERVIEW AND FUTURE DIRECTIONS

The molecular mechanism whereby *MYBPC3* mutations cause HCM is still poorly understood. However, evidence to date suggests that both the haplo-insufficiency mechanisms as well as the impairment of the UPS contribute to the hypertrophy phenotype (Flashman, *et al*., 2004; Sarikas *et al*., 2005). These findings clearly highlight the importance of the UPS in at least *MYBPC3*-derived HCM. However, the exact way that this pathway contributes to the disease phenotype on a molecular level is still unclear.

The aim of the present study was to dissect the role of COMMD4 in the context of cMyBPC functioning. The detected interactions implicate COMMD4 in both protein trafficking and the protein degradation pathway. More specifically, COMMD4 seems to facilitate and stabilise the interaction between cMyBPC and the E3 ubiquitin ligase for efficient ubiquitin transfer. Thus, any alterations that would affect ubiquitin conjugation of cMyBPC in a timely manner would lead to accumulation of ubiquitin-positive aggregates, which in turn would impair the UPS via overusage or blocking. Thus, our data, as deduced from our COMMD4 knockdown studies taken in conjunction with the findings of Sarika *et al* (2005), suggest a role for COMMD4 in cMyBPC turnover, whether under normal physiological or pathological conditions. The results from the present study suggest that alterations of normal cMyBPC turnover in the absence of COMMD4 may be the cause of ubiquitin-positive aggregate formation due to inefficient ubiquitin conjugation or, alternatively, inefficient transfer of ubiquitinated substrate to the proteasome. This, in turn, would result in the impairment of the UPS with detrimental consequences. However, to fully understand the exact nature of the underlying mechanisms involved in the function of COMMD4 and other UPS machinery in the context of cMyBPC function and turnover would warrant further independent investigations. In addition, our findings suggest that the role of COMMD4 in turnover of cardiac actin should be further investigated.

Thus, the present study has added to the little information currently available on sarcomeric protein turnover. However, this is only a starting point and would require further investigations in order to provide clarity regarding the complexity of this process (Wang *et al*., 2008; Willis *et al*., 2009). Although we answered some questions regarding the turnover of cMyBPC; this study generated additional, unexpected questions that create focus points for further research. These would involve determining how exactly the cMyBPC substrate is recognised for turnover, what signals are involved in scheduled degradation of this sarcomeric protein, and whether there are other E3 ubiquitin ligases as well as other intermediate proteins that are involved in the ubiquitin-conjugation process as well as in the transfer of ubiquitinated substrates to the

proteasome. Furthermore, these questions apply not only to cMyBPC turnover, but also to the turnover of other sarcomeric proteins.

Thus in future, it should be investigated whether COMMD4 has an effect on the ubiquitination of both cMyBPC and ACTC1; future studies could involve an *in vivo* ubiquination assay using H9C2 cardiomyocytes with and without knockdown of COMMD4. H9C2 cells could be transfected with an HA-tagged ubiquitin plasmid construct and respectively with GFP-tagged cMyBPC or DsRed-tagged ACTC1. *In vivo* IP and Co-IP experiments could be performed to investigate the respective interaction between both GFP-cMyBPC and DsRed-tagged ACTC1 and HA-Ub, followed by WB analysis to detect any changes in the expression of these proteins in the presence and absence of COMMD4. In addition, the direct interactions of COMMD4 with the Cullin1/ Rbx1 complex should also be investigated. Furthermore, deletion mapping of FBXL10 could be performed in order to determine which domains mediate the interaction with COMMD4, while direct protein-protein interaction assays could also be performed to ascertain the binding of FBXL10 to its proposed target proteins, particularly cMyBPC and cardiac actin. In addition, it could be determined whether COMMD4 binds to any proteasomal subunits, to see if COMMD4 might be involve in the transfer of ubiquitinated cMyBPC to the proteasome. All of these protein protein interactions could be determined via the combination of immunochemical capture and mass-spectrometry which is currently the method of choice for protein-protein interactions due to its sensitivity, specificity and accuracy. This method also allows for the distinguishment between specific protein interactors from those non-specific background proteins (Kaake, Wang and Huang, 2010).

Furthermore, as association analysis implicated *SNX3* as a modifier of hypertrophy development, and while the present study and other studies in this laboratory has shown it to be an interactor of cMyBPC itself, it is possible that mutations in *SNX3* could also be HCM-causing. Novel disease-causing genes have previously been identified in interactors of known disease-causing proteins in other conditions (Mohler *et al*., 2003; Vatta *et al*., 2006). As part of an ongoing study in collaboration with Dr. Michael Christiansen, *SNX3* is being screened for mutations using bidirectional sequencing in HCM probands from South Africa, Denmark and England.

4.6. CONCLUSION

Thus, to conclude, the present study had identified eight putative interactors of COMMD4, using Y2H library screening. Six of these interactors, namely SNX3, ENO1, LGMN, DSCR3, FBXL10 and ACTC1, were confirmed as true and novel interactors of COMMD4 via additional verification assays. Our results demonstrate that COMMD4, a little understood member of the COMM family of proteins, binds to the cMyBPC motif in a phosphorylation-dependent manner. We delineated the protein interactions of COMMD4, and for the first time implicated this protein in protein trafficking and protein turnover, thus establishing a novel link between the sarcomere, HCM and the protein turnover machinery in COMMD4. In addition, identification of *SNX*3 has added to the compilation of hypertrophy modifier genes, and may, along

with the identification of further hypertrophy modifier genes, contribute to the existing knowledge of the HCM patho-aetiology with regards to their role in modulating hypertrophy development.

APPENDIX I

1. BACTERIAL MEDIA

LB-Media

Bacto tryptone	5g
Yeast Extract	2.5g
NaCl	5g

Sterile water to a final volume of 500ml. Autoclave and add the appropriate antibiotic (Ampicillin 50mg/ml, Kanamycin 50mg/ml or Zeocin™ 100mg/ml) to media when temperature of ≈55^{0}C is reached.

LB-Agar Plates

Bacto tryptone	5g
Yeast Extract	2.5g
NaCl	5g
Bacto agar	8g

Sterile water to a final volume of 500ml. Autoclave and add the appropriate antibiotic (Ampicillin 50mg/ml, Kanamycin 50mg/ml or Zeocin™ 100mg/ml) to media when temperature of ~55^{0}C is reached, prior to pouring ~20, 90mm plates. These plates were then allowed to set for 2-5 hours and stored at room temperature for up to three weeks.

2. YEAST MEDIA

YPDA Media

Difco peptone	10g
Yeast Extract	10g
Glucose	10g
L-adenine hemisulphate (0.2% stock solution)	7.5ml

Add sterile water to a final volume of 500ml. Autoclave at 121^{0}C for 15 minutes.

YPDA Agar Plates

YPDA media

Difco peptone	10g
Yeast Extract	10g
Bacto agar	10g
Glucose	10g
L-adenine hemisulphate (0.2% stock solution)	7.5ml

Add sterile water to a final volume of 500ml and autoclave at 121^{0}C for 15 minutes. Allow to cool to a temperature of ~55^{0}C, before pouring ~20, 90mm plates. These plates were then allowed to set for 2-5 hours and stored at room temperature for up to three weeks.

SD^{-Trp} Media

Glucose	12g
Yeast nitrogen base without amino acids	4g
SD^{-Trp} amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^{0}C for 15 minutes.

SD^{-Trp} Agar Plates

Glucose	12g
Bacto agar	12g
Yeast nitrogen base without amino acids	4g
SD^{-Trp} amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^{0}C for 15 minutes. Allow to cool to a temperature of ~55^{0}C, before pouring ~20, 90mm plates. These plates were then allowed to set for 2-5 hours and stored at room temperature for up to three weeks.

SD^{-Leu} Media

Glucose	12g
Yeast nitrogen base without amino acids	4g
SD^{-Leu} amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^{0}C for 15 minutes.

SD^{-Leu} Agar Plates

Glucose	12g
Bacto agar	12g
Yeast nitrogen base without amino acids	4g
SD^{-Leu} amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^{0}C for 15 minutes. Allow to cool to a temperature of ~55^{0}C, before pouring ~20, 90mm plates. These plates were then allowed to set for 2-5 hours and stored at room temperature for up to three weeks.

SD$^{-Leu-Trp}$ Media

Glucose	12g
Yeast nitrogen base without amino acids	4g
SD$^{-Leu-Trp}$ amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^{0}C for 15 minutes.

SD$^{-Leu-Trp}$ Agar Plates

Glucose	12g
Bacto agar	12g
Yeast nitrogen base without amino acids	4g
SD$^{-Leu-Trp}$ amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^{0}C for 15 minutes. Allow to cool to a temperature of ~55^{0}C, before pouring ~20, 90mm plates. These plates were then allowed to set for 2-5 hours and stored at room temperature for up to three weeks.

TDO Media

Glucose	12g
Yeast nitrogen base without amino acids	4g
SD$^{-Leu-Trp-His}$ amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^{0}C for 15 minutes.

TDO Agar Plates

Glucose	12g
Bacto agar	12g
Yeast nitrogen base without amino acids	4g
SD$^{-Leu-Trp-His}$ amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^{0}C for 15 minutes. Allow to cool to a temperature of ~55^{0}C, before pouring ~20, 90mm plates. These plates were then allowed to set for 2-5 hours and stored at room temperature for up to three weeks.

QDO Media

Glucose	12g

Yeast nitrogen base without amino acids	4g
$SD^{-Leu-Trp-His-Ade}$ amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^0C for 15 minutes.

QDO Agar Plates

Glucose	12g
Bacto agar	12g
Yeast nitrogen base without amino acids	4g
$SD^{-Leu-Trp-His-Ade}$ amino acid supplement	0.4g

Add sterile water to a final volume of 600ml. Adjust pH to 5.8 and autoclave at 121^0C for 15 minutes. Allow to cool to a temperature of ~55^0C, before pouring ~20, 90mm plates. These plates were then allowed to set for 2-5 hours and stored at room temperature for up to three weeks.

X-α-Galactosidase Solution (5mg/ml)

X-α-Galactosidase	25mg
Dimethylformamide	1ml

Make a 25mg/ml stock. Dilute with dimethylformamide to 5mg/ml working solution.

3. EUKARYOTIC CELL CULTURE MEDIA

Complete Growth Media (for H9C2 cells)

DMEM (4.5g/L glucose, with L-glutamine)	178ml
Fetal bovine serum	20ml
Pen-strep	2ml

Pre-warm to 37^0C before use.

Complete Growth Media (for HEK293 cells)

DMEM (4.5g/L glucose, with L-glutamine)	89ml
HAMS-F12	89ml
Fetal bovine serum	20ml
Pen-strep	2ml

Pre-warm to 37^0C before use.

Serum-Free Media

DMEM (4.5g/L glucose, with L-glutamine)	100ml

Pre-warm to 37^{0}C before use.

Differentiating Growth Media

DMEM (4.5g/L glucose, with L-glutamine)	196ml
Horse serum	2ml
Pen-strep	2ml

Pre-warm to 37^{0}C before use.

4. YEAST TRANSFORMATION REAGENTS

1M LiAc

LiAC	5.1g

Make up to 50ml with sterile water.

100mM LiAc

1M LiAC	5ml

Make up to 50ml with sterile water

50% PEG 4000

PEG 4000	25g

Make up to 50ml with sterile water

5. YEAST PLASMID PURIFICATION SOLUTIONS

Yeast Lysis Buffer (Smash and Grab buffer)

SDS	1%
Triton X-100	2%
NaCl	100mM
Tris, pH 8	10mM
EDTA, pH8	1mM

6. ELECTROPHORESIS STOCK SOLUTIONS

10% Ammonium Persulphate

Ammonium persulphate (Merck)	10g
Sterile water	10ml

Mix well and stock at 4^{0}C

20X SB Stock Solution

di-Sodium tetraborate decahydrate	38.14g
Sterile water to a final volume of 1 litre	

1X SB Solution

20X SB solution	50ml
Sterile water to a final volume of 1 litre.	

SDS-PAGE Resolving Gel Buffer (4X)

Tris base	109.2g
ddH_2O	330ml
10% SDS	24ml
pH to 8.8 using 1M HCl. Make up to 600ml using ddH_2O	

SDS-PAGE Stacking Gel Buffer (4X)

Tris base	36.3g
ddH_2O	330ml
10%SDS	24ml
pH to 6.8 using 1M HCl. Make up to 600ml using ddH_2O	

10X SDS-PAGE Running Buffer

Tris base	30g
Glycine	144g
10% SDS	100ml
Add ddH_2O to a final volume of 1L	

1X SDS-PAGE Running Buffer

10X SDS-PAGE running buffer	100ml
Sterile water to a final volume of 1 litre	

7. GELS

1% Agarose Gel

Agarose	1g
SB buffer (1X)	100ml

Microwave for 1 minute on maximum power and add 5µl ethidium bromide (10mg/ml) when temperature of $\approx 55^0$C is reached.

2% Agarose Gel

Agarose	2g
SB buffer (1X)	100ml

Microwave for 1 minute on maximum power and add 5µl ethidium bromide (10mg/ml) when temperature of ≈55^{0}C is reached.

15% SDS-PAGE Stacking Gel

Sterile water	2.56ml
10% SDS	40µl
SDS-PAGE stacking gel buffer (4X)	1ml
Acrylamide (40%)	390µl
APS (10%)	30µl
TEMED	6µl

Makes two gels for the Bio-Rad Mini gel apparatus system (Bio-Rad Laboratories, Hercules, CA, USA).

15% SDS-PAGE Resolving Gel

Sterile water	4.82ml
10% SDS	100µl
SDS-PAGE resolving gel buffer (4X)	1.25ml
Acrylamide (40%)	3.75ml
APS (10%)	80µl
TEMED	6µl

Makes two gels for the Bio-Rad Mini gel apparatus system (Bio-Rad Laboratories, Hercules, CA, USA).

20% SDS-PAGE Stacking Gel

Sterile water	8.8ml
SDS-PAGE stacking gel buffer (4X)	3.2ml
Acrylamide (40%)	950µl
APS (10%)	134µl
TEMED	13.4µl

Makes two gels for the Modular Vertical Electrophoresis system (VS20CBS, Cleaver Scientific, Warwickshire, United Kingdom).

20% SDS-PAGE Resolving Gel

Sterile water	8.8ml
SDS-PAGE resolving gel buffer (4X)	8.8ml
Acrylamide (40%)	17.6ml
APS (10%)	350µl
TEMED	35µl

Makes two gels for the Modular Vertical Electrophoresis system (VS20CBS, Cleaver Scientific, Warwickshire, United Kingdom).

8. LOADING DYES

Ethidium Bromide Stock (10mg/ml)

Ethidium bromide (Sigma)	500mg
Sterile water	50ml

Stir well on magnetic stirrer for 4 hours and store in a dark container at 4^0C

Bromophenol Blue Loading Dye

Bromophenol blue	0.1% (w/v)

Sterile water to a final volume of 100ml. Store at 4^0C

SDS Loading Dye

1M Tris-HCl (pH6.8)	50mM
DTT	100mM
SDS	2%
Bromophenol blue	0.1%
Glycerol	10%

9. MOLECULAR SIZE MARKERS

λPst Molecular Size Marker

Bacteriophage Lambda genomic DNA	200µl
*Pst*I Restriction enzyme	3µl (30U)
Promega buffer H	30µl

Sterile water to a final volume of 300µl

Incubate at 37^0C for 3 hours, heat inactivate enzyme at 65^0C for 10 minutes. Use 1µl of digestion product on polyacrylamide gels, or 3µl on ethidium bromide stained agarose gels.

Fragment sizes:

11497bp, 5077bp, 4507bp, 2838bp, 2560bp, 2459bp, 2443bp, 2140bp, 1986bp, 1700bp, 1159bp, 1093bp, 805bp, 514bp, 468bp, 448bp, 339bp, 264bp, 247bp, 211bp, 200bp, 164bp, 150bp, 94bp, 87bp, 72bp

10. *IN VITRO* CO-IP REAGENTS

Co-IP Buffer

1M Tris (pH-7.5)	400µl
5M NaCl	60µl
1M DTT	20µl
2ng/ml Aprotinin	50µl
50mM PMSF	200µl
Tween 20	20µl

Make up to 20 ml using sterile water.

Pre-washed Protein G Agarose

A volume of Co-IP buffer (double the volume of the Protein G agarose) was used to wash the Protein G agarose (10µl/reaction) by centrifugation for 30 seconds. Following centrifugation, the supernatant was discarded and the pellet resuspended (double volume of Co-IP buffer) before another round of centrifugation. This procedure was repeated five times. After the final centrifugation step, the pellet was resuspended in a volume of Co-IP buffer representing the initial volume of Protein G aliquoted.

TBST (pH 7.6)

5M NaCl	30ml
1M Tris-HCl (pH 7.6)	20ml
Tween 20	1ml

Make up to 1L using sterile water.

11. *IN VIVO* CO-IP AND WB REAGENTS

20 ml Lysis Buffer (LyB)

Stock solution:

1M HEPES (pH-7.5)	1ml
5M NaCl	0.4 ml
0.5 M EDTA	0.4 ml
1% Triton X	0.2 ml
10mM NaPPi	8ml
1M Na3VO4	0.04ml

Add ddH2O to a final volume of 20ml and store at -20^{o}C. Prior cell lysis add 400ul PMSF (50mM) and1 complete protease inhibitor cocktail tablet.

Optional: Can add 1ml LiCl (10mM) to lysis buffer during wash steps. Help reduce non-specific banding.

2X Sample buffer

25% Glycerol	25ml
0.01% Bromophenol blue	10mg
50mM DTT	80mg

Make up in 100mM Tris buffer, pH 6.8 to a final volume of 100ml. Store at -20^{o}C.

Transfer Buffer

25mM Tris Base	3.03 g
192mM Glycine	14.4g
20% Methanol	200ml

Make up to 1L using sterile water.

Bradford Reagent

Coomasie Brilliant Blue G-250	100mg
Phosphoric acid (85%w/v)	100ml
95% Ethanol	50ml

Make up to 1L using sterile water. Filter and store in dark or foil covered bottle at RT.

12. SOLUTIONS USED FOR THE PRODUCTION OF BACTERIAL COMPETENT CELLS

CAP-Buffer

$CaCl_2$	2.21g
Glycerol	37.5ml
PIPES	0.76g

Sterile water to a final volume of 250ml. Adjust pH to 7.0 and store at 4^0C.

13. SOLUTIONS USED IN THE DOMAIN/PHOSPHORYLATION ASSAY

10X Isoproterenol

Isoproterenol	2.48mg

Add 100ml sterile water for 1000X stock. Dilute further 1:100 with sterile water for 10x working solution. Make up as required as it cannot be stored in solution.

65mM $CaCl_2$

Calcium chloride	478mg

Add 50ml sterile water for a 10X working solution.

14. OTHER SOLUTIONS USED

0.5 M EDTA pH 8.0

$Na_2EDTA.2H_2O$	186.1 g
dd H_2O	800 ml

Bring pH to 8 with NaOH pellets (about 20 g); the EDTA will not dissolve until the pH is about right. Add dH_2O to a final volume of 1 liter and autoclave.

PBS

NaCl	8g
KCl	0.2g
$Na_2HPO_4.2H_20$	1.42g
KH_2PO_4	0.2g

Add ddH2O to a final volume of 1 litre.

APPENDIX II

CALCULATING YEAST MATING EFFICIENCIES (Calculations based on Clontech Manual)

Count number of colonies on all plates with 30-300 colonies after 4 days

$$\#\text{colony forming units (cfu)/ml} = \frac{\text{cfu} \times 1000\mu\text{l/ml}}{\text{volume plated } (\mu\text{l}) \times \text{dilution factor}}$$

1. Number of cfu/ml on SD^{-Leu} plates = viability of prey partner
2. Number of cfu/ml on SD^{-Trp} plates = viability of bait partner
3. Number of cfu/ml on $SD^{-Leu-Trp}$ plates = viability of diploids
4. Lowest Number of cfu/ml of SD^{-Leu} or SD^{-Trp} plates indicate limiting partner

$$5.\ \text{Mating efficiency} = \frac{\#\text{cfu/ml of diploids} \times 100}{\#\text{cfu/ml of limiting partner}}$$

Library titre

Count number of colonies on all plates with 30-300 colonies after 4 days

$$\#\text{cfu/ml} = \frac{\#\text{colonies}}{\text{plating volume(ml)} \times \text{dilution factor}}$$

\# colonies clones screened= # cfu/ml x final resuspension volume

Haemocytometric cell count

A cell count using, a Neubauer haemocytometer (Superior, Berlin, Germany), was required to determine the titre of bait cultures used in the library mating experiments, as well as to determine the number of H9C2/HEK293 cells to be seeded prior transfections. Prior to counting, both the surface of the chamber and the glass cover slip was cleaned with alcohol. The glass cover slip was then placed over the counting surface; a dilution (1:100) of the sample prepared and an aliquot applied into one of the V-shaped wells. Once the area under the coverslip filled with the sample through capillary action, the counting chamber was subsequently placed on a microscope (Nikon TMS, Nikon Instruments., New York, USA) stage and the counting area was brought into focus

under low magnification. The number of cells in the large central quadrant of the haemocytometer was counted and this value was used to calculate the number of cells per millilitre using the following formula:

Number of cells/ml= number of cells x dilution factor x 10^4 (a constant used because the depth of the haemocytometer is 0.1mm)

APPENDIX III

METHODOLOGY AND SOLUTIONS USED FOR THE EXTRACTION OF DNA FROM BLOOD

Cell lysis buffer

0.32M Sucrose	109.5g
1%Triton X-100	10ml
1M $MgCl_2$	5ml
10Mm Tris-HCL pH 8.0	10ml

Add ddH20 to a final solution of 1litre

3M sodium-acetate

$NaAc.3H_2O$	40.81g
ddH_2O	50ml

Adjust pH to 5.2 with glacial acetic acid and adjust volume to 100 ml with ddH_2O

Phenol/Cloroform

Phenol, saturated with 1X TE	50ml
Chloroform	48ml
8-Hydroxiquinoline	2ml

Mix well and store at 4^0C.

Chloroform/octanol (24:1)

Chloroform	96ml
Octanol	4ml

Mix well and store at 4^0C

1. Extraction of nuclei from whole blood

Briefly, blood from two 5ml EDTA tubes for each patient was transferred into a 50ml Falcon tube, which was filled to 20ml with ice-cold cell lysis buffer and gently inverted 2-3 times. Afterwards the sample was incubated on ice for 10 minutes and then centrifuged at 2500-3000rpm at room temperature in a Beckman model TJ-6 centrifuge (Beckman Coultier, Scotland, UK). After discarding the supernatant, the pellet was resuspended in 20ml ice-cold cell lysis buffer and the incubation and centrifugation step was repeated. Thereafter, the supernatant was discarded, the pellet was resuspended in DNA extraction buffer, and the nuclei extracted were used for DNA extraction.

2. Extraction of DNA from nuclei

Hundred microlitres of proteinase K (10µg/ml) was added to the extracted nuclei and subsequently incubated overnight at 37°C. Following the overnight incubation, 2ml distilled water, 500µl 3M sodium-acetate and 2,5ml phenol/chloroform were added to the nuclei sample. The tubes were inverted and gently shaked for 10

minutes on a Voss rotator (Voss of Maldon, England) at 4°C. Afterwards the mixture was transferred to a 10ml glass Corex tube in order to clearly distinguish the aqueous phase from the organic phase, followed by centrifugation in a Sorvall RC-5B refrigerated super-speed centrifuge (rotor SS 34, Dupont Instruments) at 8000rpm for 10 minutes at 4°C.

The upper aqueous layer, containing the DNA, was transferred to a clean Corex tube and 25ml chloroform/octanol was added, after which the tube was closed with a polypropylene stopper and gently inverted for 10 minutes. This mixture was centrifuged at 4°C, followed by the removal of the upper aqueous phase. The DNA was ethanol precipitated by the addition of two volumes ice-cold 96% ethanol and gently inverted until DNA strands appeared as a white precipitate.

The DNA strands were placed in a clean 1.5ml Eppendorf microfuge tube using a yellow-tipped Gilson pipette and 1µl of 70% ethanol was added to the DNA and subsequently centrifuged in a Beckman microfuge for 3 minutes at 13000rpm. After the ethanol was carefully decanted, the 70% ethanol wash repeated again in order to remove any excess salts. The DNA pellet was air-dried for 30-60 minutes at room temperature by inverting the Eppendorf microfuge tube on Carlton paper. Thereafter, 200µl Tris-EDTA buffer was added and the DNA was resuspended by stationary incubation at 37°C overnight. Next, the DNA mixture were gentle mixed in a Voss rotator at 4°C for a further 3 days, followed by stationary incubation at 4°C until the DNA had been fully resuspended.

After complete resuspension of DNA in the buffer for 1-2 weeks, the optical density (OD) of the DNA was determined at 260nm (OD_{260}) in a Milton Roy series 120i spectrophotometer (USA). The DNA concentration, in µg/µl, was determined and the purity of the DNA was monitored by the OD_{260}//OD_{280} ratio, which should be approximately 1.8 for pure DNA.

APPENDIX IV

LIST OF SUPPLIERS

100bp DNA ladder	Promega
[^{35}S] methionine	PerkinElmer
Acrylamide	Merck
ABI Taqman allelic discrimination assays	Applied Biosystem
Agarose	Whitehead Scientific
Ammonium acetate	B&M Scientific
Ammonium persulphate	Merck
Ampicillin	Roche
Anti-cMyc	Roche
Anti-HA monoclonal antibody	Roche
Aprotinin	Roche
Autoradiography film	Kodak
β-mercapto-ethanol	Sigma
Bacteriophage Lambda gnomic DNA	Promega
Bacto Agar	Merck
Bacto tryptone	Fluka
*BamH*I	Promega
BgIII	Fermentas
Bromophenol blue	Merck
BSA	New England Biolabs
Calcium chloride	Merck
Calf intestinal alkaline phosphatise	Promega
Complete protease inhibitor cocktail tablets	Roche
dATP	Boerhinger Mannheim
dCTP	Boerhinger Mannheim
dGTP	Boerhinger Mannheim
Dimethylformamide	Merck
di-Sodium tetraborate decahydrate	Merck
DMEM	Highveld biological
dNTP mix	TaKaRa
DraII	Fermentas
DsRed-Monomer-C1	BD Bioscience
DTT	Roche
dTTP	Boerhinger Mannheim
*Eco*RI	Promega

EcoRV	Promega
EDTA	Boerhinger Mannheim
Ethanol	Boerhinger Mannheim
Ethidium bromide	Roche
Ex Taq™ polymerase	TaKaRa
Ex Taq™ polymerase Mg^{2+}-containing reaction buffer	TaKaRa
Fetal bovine serum	Delta Bioproducts
GeneJuice®	Novagen
GeneJET ™ Plasmid Miniprep Kit	Fermentas
GenElute ™ PCR Clean-up kit	Sigma
Glass beads (450-600µm)	Sigma
Glucose	Kimix
Glycerol	Promega
Glycine	Kimix
Goat Anti-SNX3 PAb	Santa Cruz
Goat Anti-ANKRD1 MAb	Santa Cruz
H9C2	American Type Culture Collection
HCl	Merck
Herring sperm DNA	Promega
*Hind*III	Promega
HiPerfect transfection agent	Qiagen
Horse serum	Sigma
Hoechst H-33342	Sigma
Hybond N+ nylon membranes	Amersham Pharmacia
Isopropanol	Merck
Isoproterenol	Sigma
K-acetate	Sigma
Kanamycin	Roche
KpnI	Fermentas
L-adenine hemisulphate	Sigma
LiAc	Sigma
Matchmaker™ Two-hybrid system 3	BD Biosciences
MboII	Fermentas
$MgCl_2$	Bioline
$MgSO_4.7H_2O$	Merck
Mouse Anti-RACK1 MAb	Santa Cruz
Mouse Anti-ACTC1 MAb	Santa Cruz
Mouse Anti-cMyBPC K16-MAb	Santa Cruz

Mouse Anti-HA Mab	Santa Cruz
Mouse HRP-conjugated secondary MAb	Santa Cruz
Mouse TxRed-conjugated secondary MAb	Santa Cruz
Mouse Anti-ENO1 PAb	Abnova
Mouse Anti-LGMN PAb	Abnova
Mouse Anti-FBXL10 PAb	Abnova
Mouse Anti-GFP/YFP (JL8) MAb	Clontech
Mouse Anti-DsRed MAb	Clontech
Mouse Anti-β-actin MAb	Cell Signalling
Na_2CO_3	Merck
NaCl	BDH Chemicals
$Na_2HPO_4.7H_2O$	Merck
$Na_2H_2PO_4.H_2O$	Merck
NaOH	Sigma
*Nde*I	Promega
NheI	Fermentas
Oligonucleotide primers	Department of Molecular and Cell Biology, University of Cape Town (UCT), Cape Town, South Africa
ortho-Nitrophenyl-β-galactoside	Roche
pACT2	BD Biosciences
PBS	Sigma
pdEYFP-C1amp	RZPD
PEG4000	Merck
Penicillin/streptomycin	Highveld Biological
Peptone	Difco
pGBKT7	BD Biosciences
$pGFP^2$-C3	PerkinElmer
Phenol/chloroform/isoamyl	Sigma
PIPES	Merck
PMSF	Roche
Wizard® SV Gel and PCR Clean-Up System	Promega
Protein G agarose	Kirkegaard and Perry Laboratories
QDO	BD Biosciences
Quantitect® Primer Assay for *COMMD4*	Qiagen
Quantitect® Primer Assay for *TRFR*	Qiagen
Quantitect® Primer Assay for *GAPDH*	Qiagen
Quantitect® Primer Assay for *HSP1β*	Qiagen
Quantitect® Reverse Transcription Kit	Qiagen

Quantifast™ SYBR® green PCR kit	Qiagen
RNAse wipes	Ambion
RNeasy® Plus Mini Kit	Qiagen
SacI	Fermentas
*Sac*II	Promega
SD^{-Ade}	BD Biosciences
SD^{-Leu}	BD Biosciences
$SD^{-Leu-Trp}$	BD Biosciences
SD^{-Met}	BD Biosciences
SDS	Sigma
SD^{-Trp}	BD Biosciences
SD^{-Ura}	BD Biosciences
siRNA	Qiagen
siRNA Suspension Buffer	Qiagen
Spectra ™ Multicolor Broad Range Protein Ladder	Fermentas
SuperSignal® enhanced chemiluminescence (ECL) kit	Thermo Scientific
T4 DNA Ligase	Promega
Taqman® SNP genotyping assay mix	Applied Biosystems
Taqman® Universal PCR Mastermix	Applied Biosystems
Taq polymerase	Bioline
TDO	BD Biosciences
TEMED	Sigma
TNT® Quick Coupled transcription/translation system	BD Biosciences
Tris	Merck
Tris-HCl	Merck
Triton X-100	Sigma
Trypsin	Highveld Biological
Tween 20	Merck
Whatman 3M paper	Whatman international
Wizard® Purefection plasmid purification kit	BD Bioscences
XhoI	Fermentas
X-α-galactosidase	Southern Cross
Yeast Extract	Difco
Yeast nitrogen base (without amino acids)	BD Biosciences
Zeocin™	Invitrogen
Zyppy™ Plasmid Miniprep kit	Zymo Research

APPENDIX V

BACTERIAL STRAIN PHENOTYPE

E.coli strain DH5α

Φ 80d *lacZΔM15* recA1, *end*A1, *Gry* A96 *thi*-1, *hsd*R17 *sup*E44, *rel*A1, *deoR* Δ(*lacZ*YA *argF*)u169

YEAST STRAIN PHENOTYPES

Yeast strain AH109

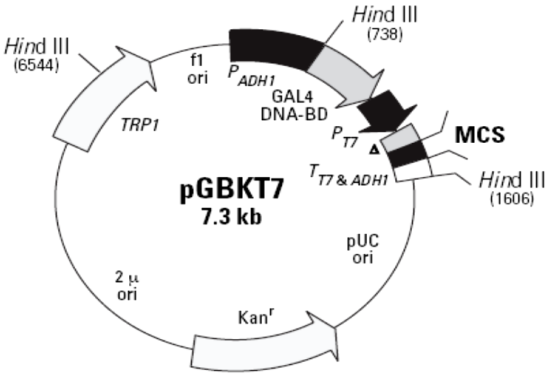
*MAT*a, *trp1-901*, *leu2-3, ura3-5, his3-200, gal4*Δ, *gal80*Δ, *LYS2::GAL1*$_{uas}$*-GAL1*$_{TATA}$*-HIS3, GAL2*$_{UAS}$*-GAL2*$_{TATA}$*-ADE2, URA3::MEL1*$_{UAS}$*-MEL1*$_{TATA}$*-lacZ* (James *et al.*, 1996)

Yeast strain Y187

*MAT*α, *ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4*Δ, *met*$^{-}$, *gal80*Δ, *URA3::GAL1*$_{UAS}$*-GAL1*$_{TATA}$*-lacZ* (Harper *et al.*, 1993)

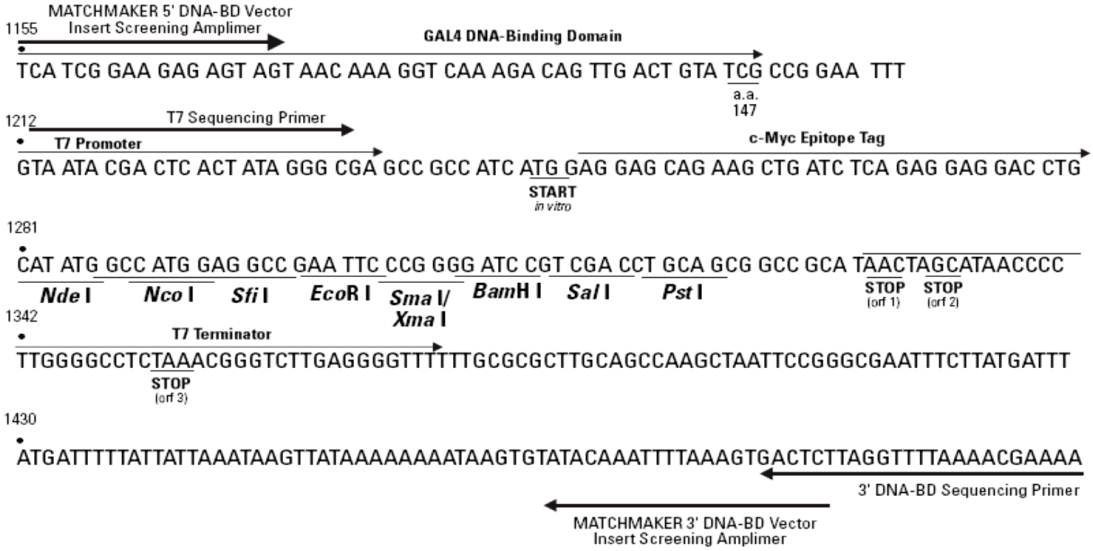
APPENDIX VI

A



B

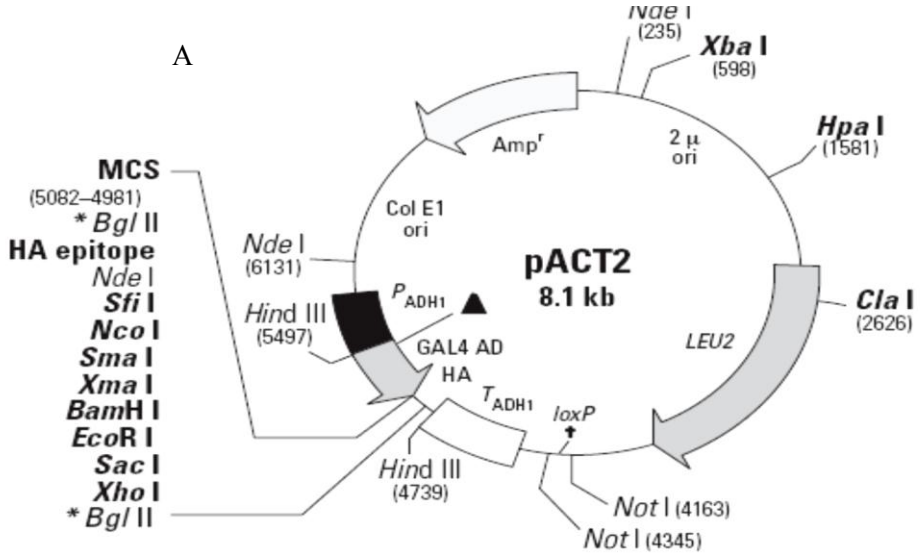
▲ c-Myc epitope tag



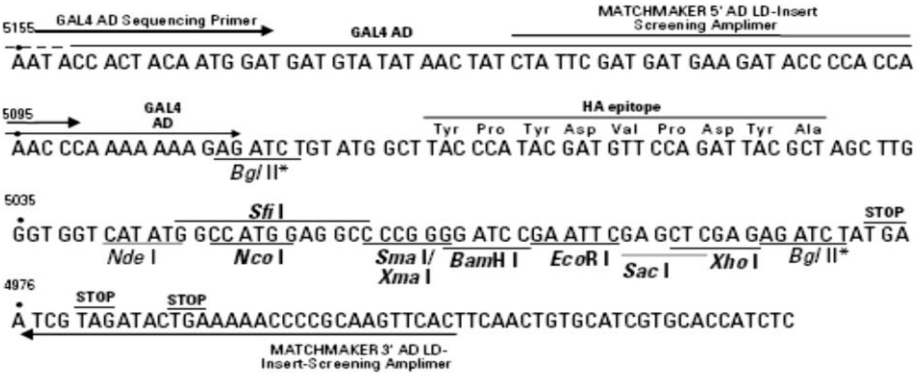
Restriction map and multiple cloning site of pGBKT7 Y2H bait vector.

A) The positions of the kanamycin resistance gene (*Kan*^{r}), the *TRP1* nutritional marker for selection in yeast, *GAL4-BD* coding sequences, f1 bacteriophage and pUC plasmid origins of replication, the truncated *S.cerevisiae ADH1* promoter sequence (P_{ADH1}), the T7 RNA polymerase promoter, c-Myc epitope tag, transcription termination sequences (T7 and *ADH1*) and the MCS are indicated on the map. B) Nucleotide sequence of the pGBKT7 MCS. The position of all unique restriction enzyme recognition sequences, stop codons in the T7 termination sequence, *GAL4-BD* coding sequence, T7 promoter sequence, c-Myc epitope tag, positions of the screening and sequencing primers are indicated on the sequence (taken from Clontech MATCHMAKER vector handbook)

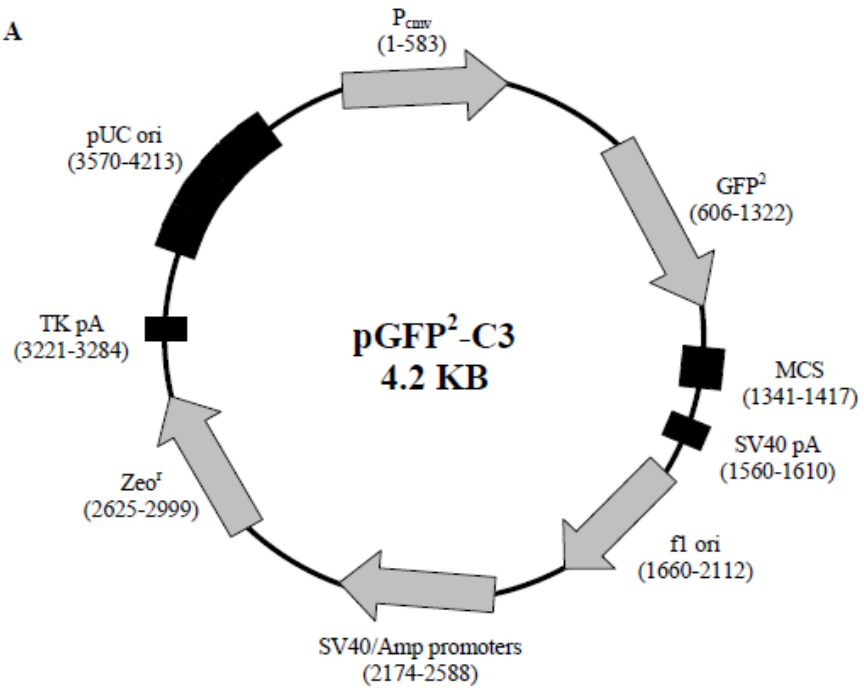
A



B



Restriction map and multiple cloning site of pACT2 Y2H prey vector. A) The positions of unique restriction enzyme sites are indicated in bold font. The position of the ampicillin resistance gene (*Amp*^{r}), *LEU2*, *GAL4-AD* coding sequence, 2µ and pBR322 plasmid origin of replication, the *S.cerevisiae ADH1* promoter, *S.cerevisiae ADH1* termination sequence, Lox sites (Lox 1 and Lox 2), the haemagglutinin (HA) epitope tag and the MCS are indicated on the map. B) Nucleotide sequence of the pACT2 MCS. The position of restriction enzyme sites, stop codons, the *GAL4-AD* coding sequence, HA epitope tag, and the positions of the pACT2-F and pACT2-R primers are all indicated (taken from Clontech MATCHMAKER vector handbook).



B

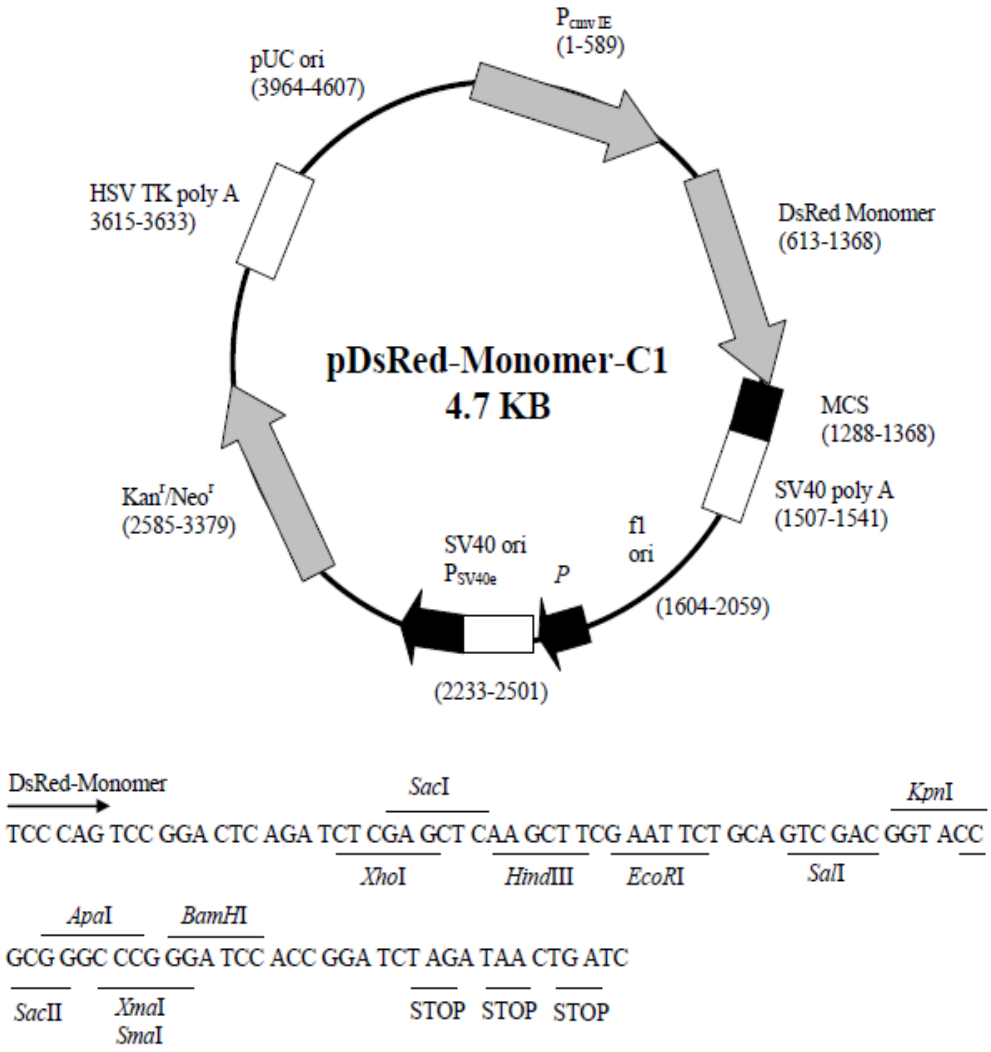
TAC AAG TCC GGA CTC AGA TCT GGC GAG CTC TCG AGA ATT CTC ACG CGT
GFP2 BspEI BglII SacI XhoI EcoRI

CTG CAG GAT ATC AAG CTT GCG GTA CCG CGG GCC CGG GAT CCA CCG
PstI EcoRV HindIII KpnI SacII ApaI BamHI

Stop Stop Stop
GAT CTA GAT AAC TGA TCA
XbaI

* Frame changes characterising pGFP2-C3 occur after this site

Restriction map and multiple cloning site of pGFP^2-C3 vector. A) The position of the P_{CMV}, SV40 early poly (A) signal, *GFP^2* gene, TK poly (A) signal, P_{SV40}/P_{amp} promoters, Zeo^r, f1 origin, pUC sequences and the MCS are indicated on the map. B) Nucleotide sequence of the pGFP^2-C3 MCS. The position of restriction enzyme sites and stop codons are all indicated (taken from BRET^2$_{TM}$ technical data sheet).



Restriction map and multiple cloning site of pDsRed-Monomer-C1 vector. pDsRed-Monomer-C1 is a mammalian expression vector that encodes DsRed-Monomer, a monomeric mutant derived from the tetrameric *Discosoma* sp. red fluorescent protein DsRed. The multiple cloning site (MCS) is positioned between the DsRed-Monomer coding sequence and the SV40 polyadenylation signal (SV40 poly A). The position of the f1 ori, P_{CMV} early promoter, *Kan/Neo resistance* genes, SV40 origin of replication, pUC plasmid replication origin and the Herpes simplex virus (HSV) thymidine kinase (TK) are indicated on the map (Taken from Clontech vector information booklet).

APPENDIX VII

THE FULL SCORING OF THE 520 CLONES IDENTIFIED FROM Y2H ANALYSIS

Colony nr	A Growth on TDO *HIS3* activation	B Growth on QDO *ADE2* activation	C X-α-Galactosidase assay (colour) *MEL1* activation
22	++++	+++	+++ (medium blue)
26	++++	++++	+++ (dark blue)
27	++++	++++	+++ (medium blue)
28	++++	++++	+++ (medium blue)
29	++++	++++	+++ (medium blue)
30	++++	++++	+++ (medium blue)
31	++++	++++	+++ (medium blue)
38	++++	++++	+++ (medium blue)
41	++++	+++	++ (medium blue)
42	++++	+++	++ (medium blue)
45	++++	++++	+++ (medium blue)
46	++++	++++	+++ (medium blue)
47	++++	+++	+++ (medium blue)
48	++++	+++	+++ (medium blue)
51	++++	++++	+++ (medium blue)
52	++++	++++	+++ (medium blue)
53	++++	++++	+++ (medium blue)
54	++++	++++	+++ (medium blue)
57	++++	++++	+++ (medium blue)
58	++++	++++	+++ (medium blue)
64	++++	++++	+++ (medium blue)
65	++++	++++	+++ (medium blue)
71	++++	+++	+++ (medium blue)
77	++++	+++	+++ (medium blue)
78	++++	+++	+++ (medium blue)
79	++++	+++	+++ (medium blue)
80	++++	+++	+++ (medium blue)
81	++++	++++	+++ (medium blue)
83	++++	+++	+++ (medium blue)
86	++++	+++	+++ (medium blue)
87	++++	+++	+++ (medium blue)
88	++++	+++	+++ (medium blue)
89	++++	++++	+++ (medium blue)

Colonies in red font represent the primaries and the black represent the secondary clones. TDO – solid media lacking Leu, Trp and His; QDO-solid media lacking Leu, Trp, His and Ade. Growth of clones on solid media: ++++ = robust, +++ = good, ++ = weak, - = no growth

Colony nr	A Growth on TDO *HIS3* activation	B Growth on QDO *ADE2* activation	C X-α-Galactosidase assay (colour) *MEL1* activation
91	++++	+++	+++ (medium blue)
117	++++	++++	+++ (medium blue)
122	++++	++++	+++ (medium blue)
123	++++	+++	+++ (medium blue)
124	++++	++++	+++ (medium blue)
128	++++	++++	+++ (medium blue)
129	++++	++++	+++ (medium blue)
130	++++	++++	+++ (medium blue)
135	++++	+++	+++ (medium blue)
136	++++	+++	+++(dark blue)
141	++++	++++	++ (medium blue)
142	++++	++++	+++ (medium blue)
143	++++	++++	+++ (medium blue)
147	++++	++++	+++ (medium blue)
148	++++	++++	+++ (medium blue)
149	++++	++++	+++ (dark blue)
150	++++	++++	+++ (medium blue)
158	++++	++++	+++ (medium blue)
164	++++	++++	+++ (medium blue)
165	++++	++++	+++ (medium blue)
166	++++	+++	++ (medium blue)
170	++++	+++	+++ (medium blue)
171	++++	+++	++ (medium blue)
172	++++	+++	+++ (medium blue)
177	++++	++++	+++ (medium blue)
178	++++	++++	+++ (medium blue)
183	++++	++++	+++ (medium blue)
184	++++	++++	+++ (medium blue)
185	++++	++++	+++ (medium blue)
186	++++	++++	+++ (medium blue)
209	++++	++++	+++ (medium blue)
211	++++	++++	++ (dark blue)
217	++++	++++	+++ (medium blue)
227	++++	++++	+++ (medium blue)
232	++++	++++	+++ (medium blue)
234	++++	++++	++ (dark blue)
236	++++	++++	+++ (medium blue)
247	++++	+++	+++ (dark blue)
254	++++	++++	+++ (medium blue)
255	++++	++++	+++ (medium blue)
258	++++	++++	+++ (medium blue)
262	++++	++++	+++ (dark blue)
264	++++	++++	++ (medium blue)

Colonies in red font represent the primaries and the black represent the secondary clones. TDO – solid media lacking Leu, Trp and His; QDO-solid media lacking Leu, Trp, His and Ade. Growth of clones on solid media: ++++ = robust, +++ = good, ++ = weak, - = no growth

Colony nr	A Growth on TDO *HIS3* activation	B Growth on QDO *ADE2* activation	C X-α-Galactosidase assay (colour) *MEL1* activation
268	++++	+++	+++ (medium blue)
281	++++	++++	+++ (medium blue)
287	++++	++++	+++ (medium blue)
288	++++	++++	+++ (dark blue)
289	++++	+++	+++ (medium blue)
291	++++	+++	++ (medium blue)
293	++++	+++	+++ (medium blue)
294	++++	+++	+++ (medium blue)
295	++++	+++	+++ (medium blue)
296	++++	+++	++ (medium blue)
299	++++	+++	+++ (medium blue)
300	++++	+++	+++ (medium blue)
301	++++	+++	++ (medium blue)
334	++++	++++	+++ (medium blue)
335	++++	++++	+++ (dark blue)
339	++++	++++	+++ (dark blue)
341	++++	++++	+++ (medium blue)
342	++++	+++	+++ (medium blue)
344	++++	++++	+++ (medium blue)
345	++++	++++	+++ (medium blue)
346	++++	+++	+++ (medium blue)
351	++++	++++	+++ (medium blue)
353	++++	++++	+++ (medium blue)
356	++++	++++	+++ (medium blue)
373	++++	++++	+++ (medium blue)
377	++++	++++	+++ (medium blue)
381	++++	++++	+++ (medium blue)
383	++++	+++	++ (dark blue)
388	++++	++++	+++ (medium blue)
396	++++	++++	+++ (medium blue)
397	++++	++++	++ (dark blue)
399	++++	++++	+++ (medium blue)
407	++++	++++	+++ (medium blue)
410	++++	+++	++ (dark blue)
411	++++	+++	+++ (medium blue)
412	++++	++++	+++ (medium blue)
416	++++	++++	+++ (medium blue)
429	++++	+++	++ (dark blue)
432	++++	++++	+++ (medium blue)
442	++++	++++	+++ (medium blue)
443	++++	++++	+++ (medium blue)
456	++++	++++	+++ (medium blue)
458	++++	++++	+++ (dark blue)
473	++++	++++	+++ (medium blue)
491	++++	+++	+++ (medium blue)
514	++++	++++	+++ (medium blue)
515	++++	++++	+++ (medium blue)

Colonies in red font represent the primaries and the black represent the secondary clones. TDO – solid media lacking Leu, Trp and His; QDO-solid media lacking Leu, Trp, His and Ade. Growth of clones on solid media: ++++ = robust, +++ = good, ++ = weak, - = no growth

APPENDIX VIII

GRAPHICAL OUTPUTS OF ALLELE FREQUENCIES FOR EACH MARKER GENOTYPED

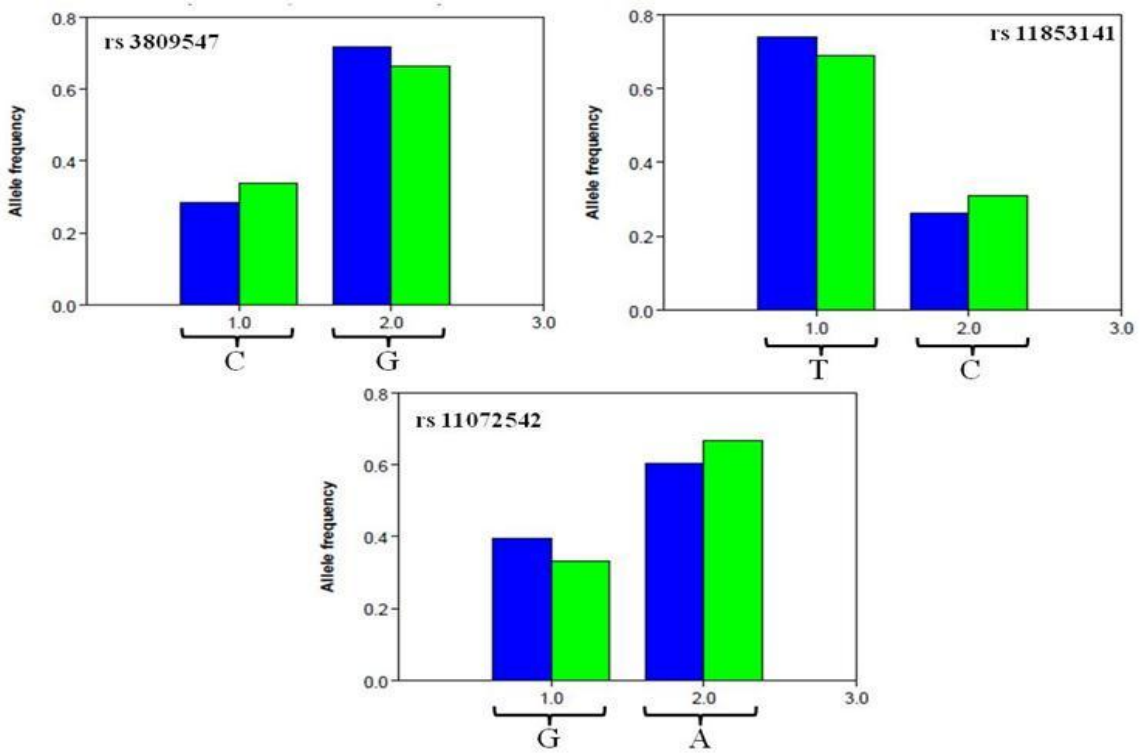


Figure 1. Allele frequencies for SNPs in *COMMD4*.
Allele frequencies for the whole co-hort is indicated in blue while the green bar represents the allele frequencies for unrelated individuals. Allele frequencies are plotted on the Y-axis and allele number is plotted on the X-axis.

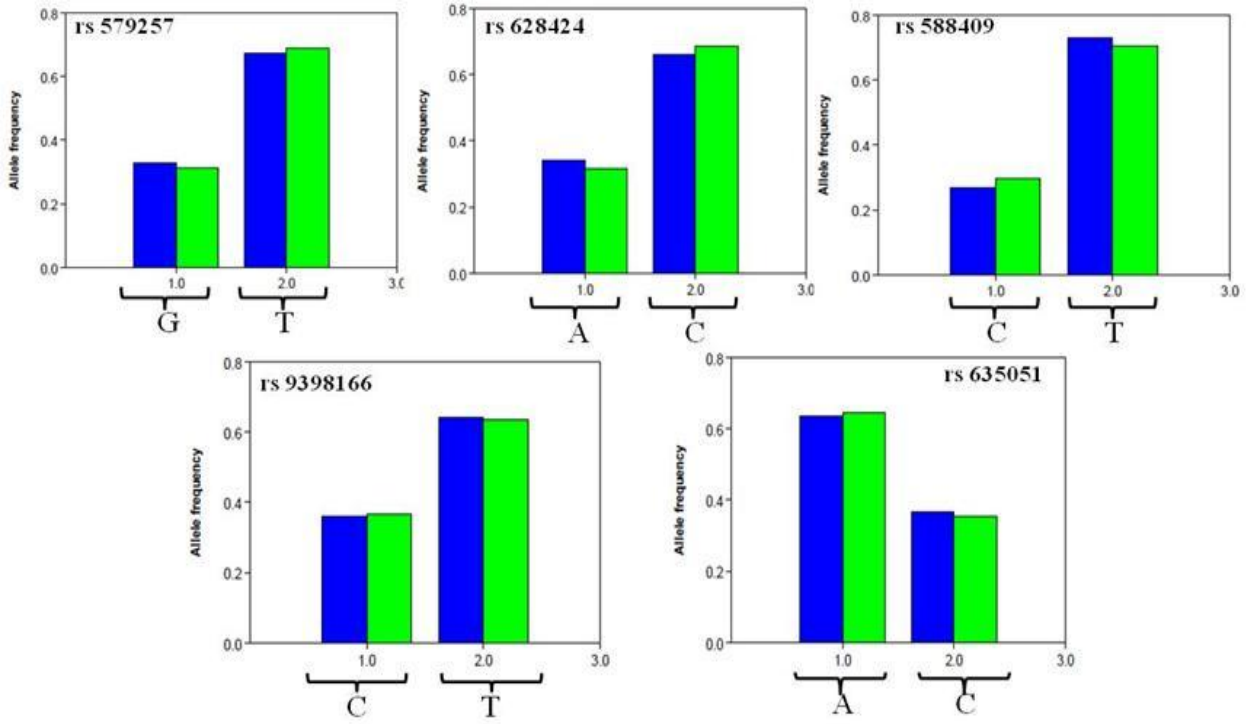


Figure 2. Allele frequencies for SNPs in *SNX3*
Allele frequencies for the whole co-hort is indicated in blue while the green bar represents the allele frequencies for unrelated individuals. Allele frequencies are plotted on the Y-axis and allele number is plotted on the X-axis.

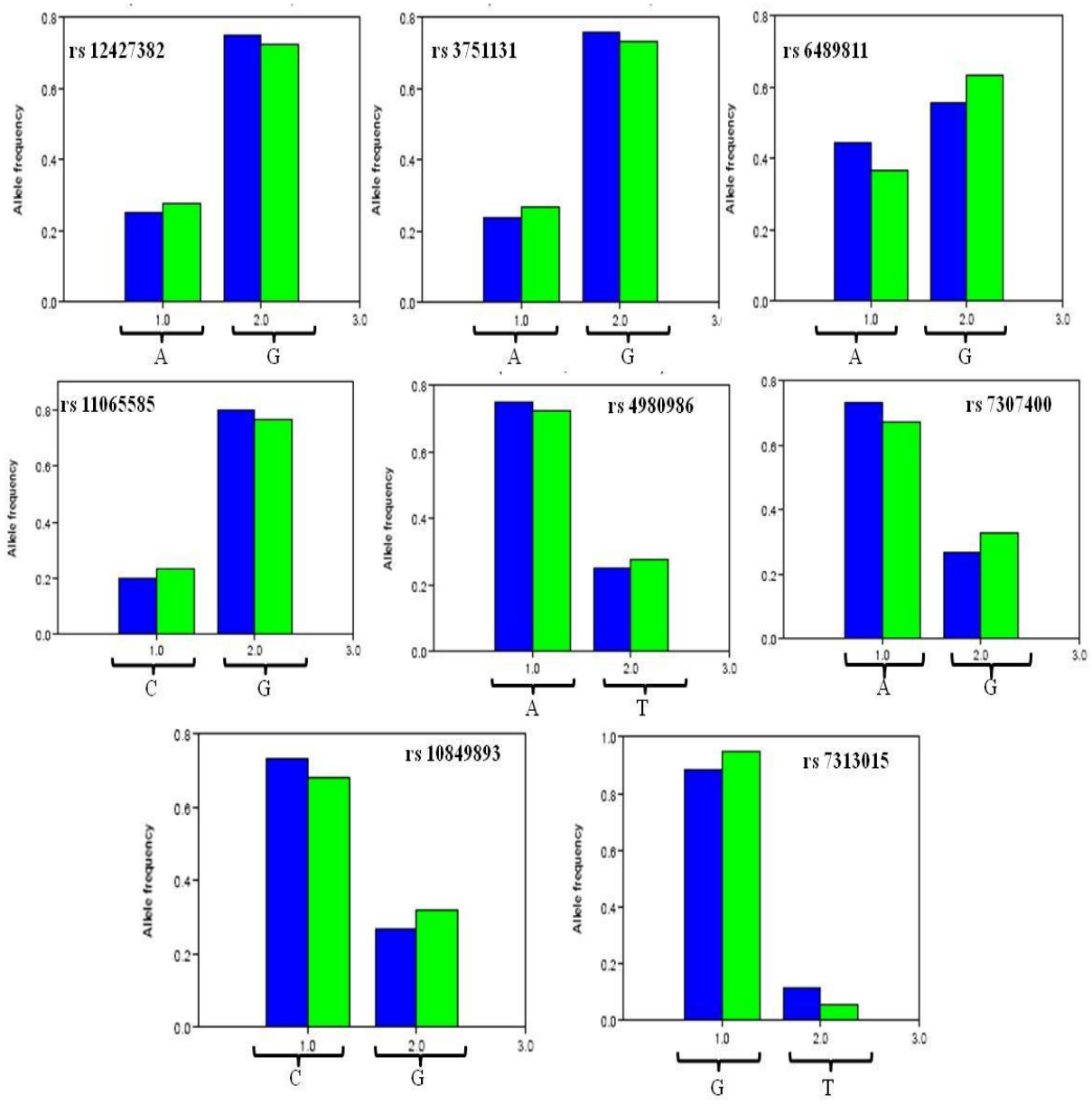


Figure 3. Allele frequencies for SNPs of *FBXL10*

Allele frequencies for the whole co-hort is indicated in blue while the green bar represents the allele frequencies for unrelated individuals. Allele frequencies are plotted on the Y-axis and allele number is plotted on the X-axis.

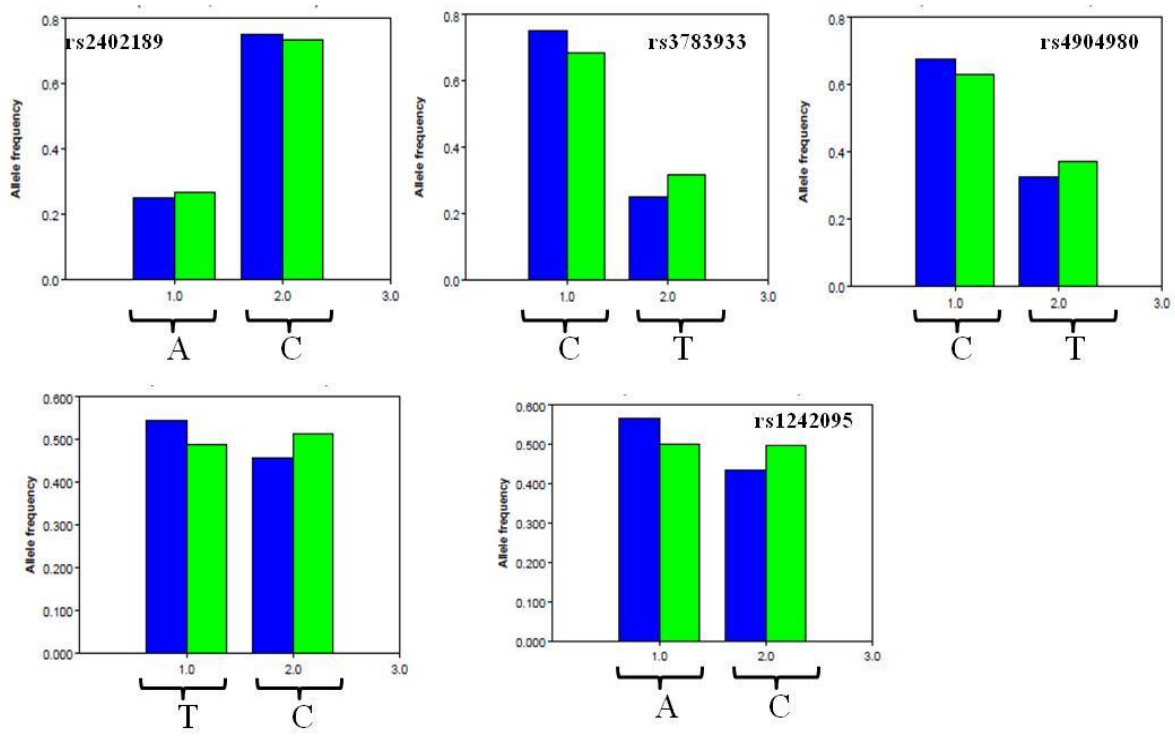


Figure 4. Allele frequencies for SNPs of *LGMN*
Allele frequencies for the whole co-hort is indicated in blue while the green bar represents the allele frequencies for unrelated individuals. Allele frequencies are plotted on the Y-axis and allele number is plotted on the X-axis.

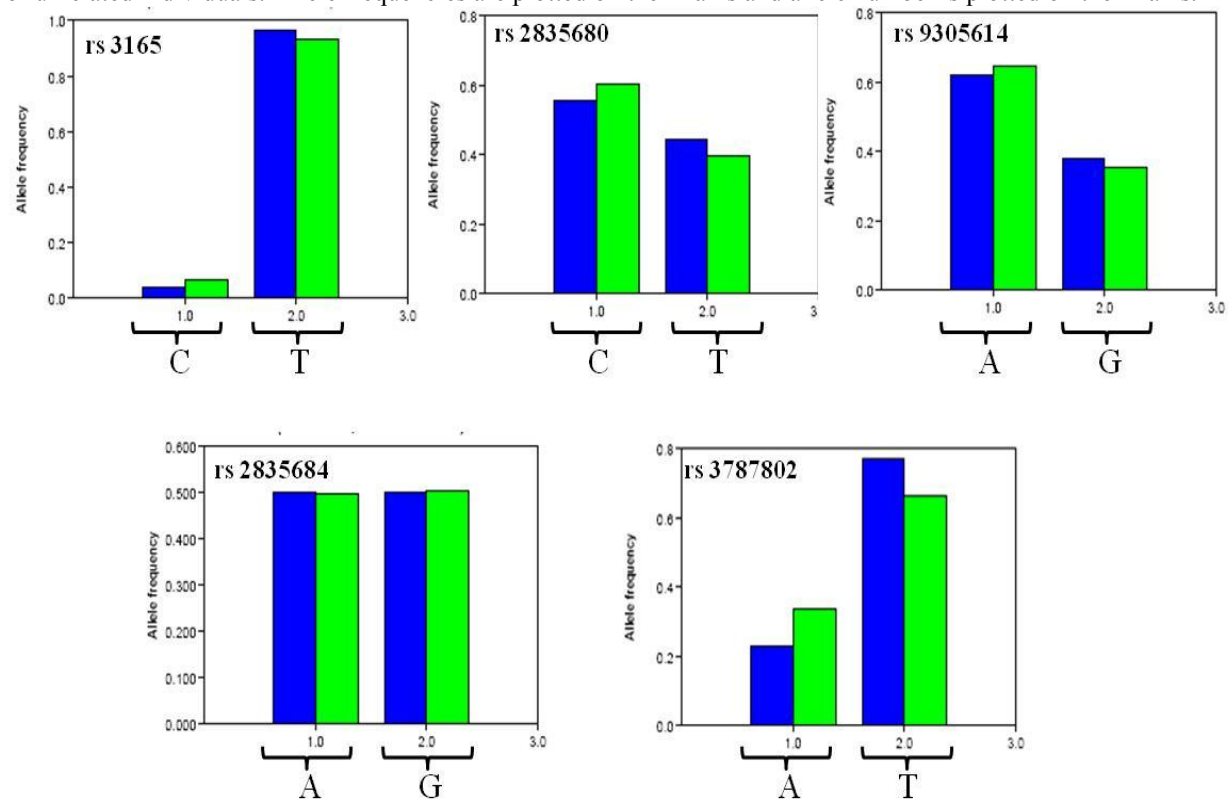


Figure 5. Allele frequencies for SNPs of *DSCR3*
Allele frequencies for the whole co-hort is indicated in blue while the green bar represents the allele frequencies for unrelated individuals. Allele frequencies are plotted on the Y-axis and allele number is plotted on the X-axis.

APPENDIX IX

Attached video clip showcasing trafficking of EYFP-COMMD4 and SNX3

THESIS REFERENCES

American Society of Echocardiography:

http://www.asecho.org/guidelines.php

HCM databases:

http://www.angis.org.au/Databases/Heart,

http://www.hgmd.cf.ac.uk/ac/index.php

http://cardiogenomics.med.harvard.edu/genes/genelisthttp://www.angis.org.au/Databases/Heart/.

http://www.hgmd.cf.ac.uk/ac/gene.php?gene=MYBPC3

NCBI Genbank and Ensemble databases:

http://www.ncbi.nlm.nih.gov

http://www.ensembl.org

Publicly available databases:

http://au.expasy.org/

http://www.dsi.univ-paris5.fr/genatlas/

http://harvester.fzk.de/harvester/

http://www.uniprot.org/uniprot/Q9H0A8

QTDT website

http://www.sph.umich.edu/csg/abecasis/QTDT/)

Subcellular prediction programs:

http://www.cs.ualberta.ca/~bioinfo/PA/Sub

http://www.imtech.res.in/raghava/eslpred/),

Aasland R, Gibson TJ and Stewart AF. The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci* 20: 56-59, 1995.

Abecasis GR, Cookson WO and Cardon LR. Pedigree tests of transmission disequilibrium. *Eur J Hum Genet* 8: 545-551, 2000.

Ackerman MJ, VanDriest SL, Ommen SR, Will ML, Nishimura RA, Tajik AJ and Gersh BJ. Prevalence and age-dependence of malignant mutations in the beta-myosin heavy chain and troponin T genes in hypertrophic cardiomyopathy: a comprehensive outpatient perspective. *J Am Coll Cardiol* 39: 2042-2048, 2002.

Ackermann MA, Hu LY, Bowman AL, Bloch RJ and Kontrogianni-Konstantopoulos A. Obscurin interacts with a novel isoform of MyBP-C slow at the periphery of the sarcomeric M-band and regulates thick filament assembly. *Mol Biol Cell* 20: 2963-2978, 2009.

Adams V, Mangner N, Gasch A, Krohne C, Gielen S, Hirner S, Thierse HJ, Witt CC, Linke A, Schuler G and Labeit S. Induction of MuRF1 is essential for TNF-alpha-induced loss of muscle function in mice. *J Mol Biol* 384: 48-59, 2008.

Adler V, Pincus MR, Minamoto T, Fuchs SY, Bluth MJ, Brandt-Rauf PW, Friedman FK, Robinson RC, Chen JM, Wang XW, Harris CC and Ronai Z. Conformation-dependent phosphorylation of p53. *Proc Natl Acad Sci U S A* 94: 1686-1691, 1997.

Ahmad F, Seidman JG and Seidman CE. The genetic basis for cardiac remodeling. *Annu Rev Genomics Hum Genet* 6: 185-216, 2005.

Al-Khayat HA, Morris EP, Kensler RW and Squire JM. Myosin filament 3D structure in mammalian cardiac muscle. *J Struct Biol* 163: 117-126, 2008.

Alcalai R, Seidman JG and Seidman CE. Genetic basis of hypertrophic cardiomyopathy: from bench to the clinics. *J Cardiovasc Electrophysiol* 19: 104-110, 2008.

Arad M, Benson DW, Perez-Atayde AR, McKenna WJ, Sparks EA, Kanter RJ, McGarry K, Seidman JG and Seidman CE. Constitutively active AMP kinase mutations cause glycogen storage disease mimicking hypertrophic cardiomyopathy. *J Clin Invest* 109: 357-362, 2002.

Arad M, Seidman JG and Seidman CE. Phenotypic diversity in hypertrophic cardiomyopathy. *Hum Mol Genet* 11: 2499-2506, 2002.

Arad M, Maron BJ, Gorham JM, Johnson WH, Jr., Saul JP, Perez-Atayde AR, Spirito P, Wright GB, Kanter RJ, Seidman CE and Seidman JG. Glycogen storage diseases presenting as hypertrophic cardiomyopathy. *N Engl J Med* 352: 362-372, 2005.

Arndt V, Rogon C and Hohfeld J. To be, or not to be--molecular chaperones in protein degradation. *Cell Mol Life Sci* 64: 2525-2541, 2007.

Au Y. The muscle ultrastructure: a structural perspective of the sarcomere. *Cell Mol Life Sci* 61: 3016-3033, 2004.

Bache KG, Brech A, Mehlum A and Stenmark H. Hrs regulates multivesicular body formation via ESCRT recruitment to endosomes. *J Cell Biol* 162: 435-442, 2003.

Bachmair A and Varshavsky A. The degradation signal in a short-lived protein. *Cell* 56: 1019-1032, 1989.

Bahrudin U, Morisaki H, Morisaki T, Ninomiya H, Higaki K, Nanba E, Igawa O, Takashima S, Mizuta E, Miake J, Yamamoto Y, Shirayoshi Y, Kitakaze M, Carrier L and Hisatome I. Ubiquitin-proteasome system impairment caused by a missense cardiac myosin-binding protein C mutation and associated with cardiac dysfunction in hypertrophic cardiomyopathy. *J Mol Biol* 384: 896-907, 2008.

Bai C, Sen P, Hofmann K, Ma L, Goebl M, Harper JW and Elledge SJ. SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* 86: 263-274, 1996.

Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet* 7: 781-791, 2006.

Bechet D, Tassa A, Taillandier D, Combaret L and Attaix D. Lysosomal proteolysis in skeletal muscle. *Int J Biochem Cell Biol* 37: 2098-2114, 2005.

Benjamin EJ and Levy D. Why is left ventricular hypertrophy so predictive of morbidity and mortality? *Am J Med Sci* 317: 168-175, 1999.

Bertola LD, Ott EB, Griepsma S, Vonk FJ and Bagowski CP. Developmental expression of the alpha-skeletal actin gene. *BMC Evol Biol* 8: 166, 2008.

Biasio W, Chang T, McIntosh CJ and McDonald FJ. Identification of Murr1 as a regulator of the human delta epithelial sodium channel. *J Biol Chem* 279: 5429-5434, 2004.

Blair E, Redwood C, Ashrafian H, Oliveira M, Broxholme J, Kerr B, Salmon A, Ostman-Smith I and Watkins H. Mutations in the gamma(2) subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis. *Hum Mol Genet* 10: 1215-1220, 2001.

Blanchard E, Seidman C, Seidman JG, LeWinter M and Maughan D. Altered crossbridge kinetics in the alphaMHC403/+ mouse model of familial hypertrophic cardiomyopathy. *Circ Res* 84: 475-483, 1999.

Boateng SY and Goldspink PH. Assembly and maintenance of the sarcomere night and day. *Cardiovasc Res* 77: 667-675, 2008.

Bolte S and Cordelieres FP. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* 224: 213-232, 2006.

Bonne G, Carrier L, Bercovici J, Cruaud C, Richard P, Hainque B, Gautel M, Labeit S, James M, Beckmann J, Weissenbach J, Vosberg HP, Fiszman M, Komajda M and Schwartz K. Cardiac myosin binding protein-C gene splice acceptor site mutation is associated with familial hypertrophic cardiomyopathy. *Nat Genet* 11: 438-440, 1995.

Bosch-Comas A, Lindsten K, Gonzalez-Duarte R, Masucci MG and Marfany G. The ubiquitin-specific protease USP25 interacts with three sarcomeric proteins. *Cell Mol Life Sci* 63: 723-734, 2006.

Bottinelli R, Coviello DA, Redwood CS, Pellegrino MA, Maron BJ, Spirito P, Watkins H and Reggiani C. A mutant tropomyosin that causes hypertrophic cardiomyopathy is expressed in vivo and associated with an increased calcium sensitivity. *Circ Res* 82: 106-115, 1998.

Boulkroun S, Ruffieux-Daidie D, Vitagliano JJ, Poirot O, Charles RP, Lagnaz D, Firsov D, Kellenberger S and Staub O. Vasopressin-inducible ubiquitin-specific protease 10 increases ENaC cell surface expression by deubiquitylating and stabilizing sorting nexin 3. *Am J Physiol Renal Physiol* 295: F889-F900, 2008.

Brown M, McGuinness M, Wright T, Ren X, Wang Y, Boivin GP, Hahn H, Feldman AM and Jones WK. Cardiac-specific blockade of NF-kappaB in cardiac pathophysiology: differences between acute and chronic stimuli in vivo. *Am J Physiol Heart Circ Physiol* 289: H466-H476, 2005.

Brown WJ, DeWald DB, Emr SD, Plutner H and Balch WE. Role for phosphatidylinositol 3-kinase in the sorting and transport of newly synthesized lysosomal enzymes in mammalian cells. *J Cell Biol* 130: 781-796, 1995.

Brugada R, Kelsey W, Lechin M, Zhao G, Yu QT, Zoghbi W, Quinones M, Elstein E, Omran A, Rakowski H, Wigle D, Liew CC, Sole M, Roberts R and Marian AJ. Role of candidate modifier genes on the phenotypic expression of hypertrophy in patients with hypertrophic cardiomyopathy. *J Investig Med* 45: 542-551, 1997.

Bujny MV, Popoff V, Johannes L and Cullen PJ. The retromer component sorting nexin-1 is required for efficient retrograde transport of Shiga toxin from early endosome to the trans Golgi network. *J Cell Sci* 120: 2010-2021, 2007.

Burkhead JL, Morgan CT, Shinde U, Haddock G and Lutsenko S. COMMD1 forms oligomeric complexes targeted to the endocytic membranes via specific interactions with phosphatidylinositol 4,5-bisphosphate. *J Biol Chem* 284: 696-707, 2009.

Burstein E, Ganesh L, Dick RD, van de SB, Wilkinson JC, Klomp LW, Wijmenga C, Brewer GJ, Nabel GJ and Duckett CS. A novel role for XIAP in copper homeostasis through regulation of MURR1. *EMBO J* 23: 244-254, 2004.

Burstein E, Hoberg JE, Wilkinson AS, Rumble JM, Csomos RA, Komarck CM, Maine GN, Wilkinson JC, Mayo MW and Duckett CS. COMMD proteins, a novel family of structural and functional homologs of MURR1. *J Biol Chem* 280: 22222-22232, 2005.

Cambronero F, Marin F, Roldan V, Hernandez-Romero D, Valdes M and Lip GY. Biomarkers of pathophysiology in hypertrophic cardiomyopathy: implications for clinical management and prognosis. *Eur Heart J* 30: 139-151, 2009.

Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, Altshuler D, Ardlie KG and Hirschhorn JN. Demonstrating stratification in a European American population. *Nat Genet* 37: 868-872, 2005.

Caplan AJ, Mandal AK and Theodoraki MA. Molecular chaperones and protein kinase quality control. *Trends Cell Biol* 17: 87-92, 2007.

Cardozo T and Pagano M. The SCF ubiquitin ligase: insights into a molecular machine. *Nat Rev Mol Cell Biol* 5: 739-751, 2004.

Carlton J, Bujny M, Rutherford A and Cullen P. Sorting nexins--unifying trends and new perspectives. *Traffic* 6: 75-82, 2005.

Carrier L, Bonne G, Bahrend E, Yu B, Richard P, Niel F, Hainque B, Cruaud C, Gary F, Labeit S, Bouhour JB, Dubourg O, Desnos M, Hagege AA, Trent RJ, Komajda M, Fiszman M and Schwartz K. 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: 427-434, 1997.

Carrier L, Schlossarek S, Willis MS and Eschenhagen T. The ubiquitin-proteasome system and nonsense-mediated mRNA decay in hypertrophic cardiomyopathy. *Cardiovasc Res* 85: 330-338, 2010.

Carstens N, The role of rennin-angiotensin-aldosterone system (RAAS) genes in the development of hypertrophy in hypertrophic cardiomyopathy. *MSc Thesis*. 2004; University of Stellenbosch.

Carstens N, van der ML, Revera M, Heradien M, Goosen A, Brink PA and Moolman-Smook JC. Genetic variation in angiotensin II type 2 receptor gene influences extent of left ventricular hypertrophy in hypertrophic cardiomyopathy independent of blood pressure. *J Renin Angiotensin Aldosterone Syst* 2010.

Cenciarelli C, Chiaur DS, Guardavaccaro D, Parks W, Vidal M and Pagano M. Identification of a family of human F-box proteins. *Curr Biol* 9: 1177-1179, 1999.

Chan CB, Abe M, Hashimoto N, Hao C, Williams IR, Liu X, Nakao S, Yamamoto A, Zheng C, Henter JI, Meeths M, Nordenskjold M, Li SY, Hara-Nishimura I, Asano M and Ye K. Mice lacking asparaginyl endopeptidase develop disorders resembling hemophagocytic syndrome. *Proc Natl Acad Sci U S A* 106: 468-473, 2009.

Chandra M, Rundell VL, Tardiff JC, Leinwand LA, De Tombe PP and Solaro RJ. Ca(2+) activation of myofilaments from transgenic mouse hearts expressing R92Q mutant cardiac troponin T. *Am J Physiol Heart Circ Physiol* 280: H705-H713, 2001.

Chen JM, Dando PM, Rawlings ND, Brown MA, Young NE, Stevens RA, Hewitt E, Watts C and Barrett AJ. Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. *J Biol Chem* 272: 8090-8098, 1997.

Chen JM, Fortunato M and Barrett AJ. Activation of human prolegumain by cleavage at a C-terminal asparagine residue. *Biochem J* 352 Pt 2: 327-334, 2000.

Chen Q, Liu J, Horak K, Zheng H, Kumarapeli A, Li J, Li F, Gerdes A, Wawrousek E, Wang X. Intrasarcoplasmic Amyloidosis impairs proteolytic function of proteasomes in cardiomyocytes by compromising substrate uptake. *Circ Res*. 97:1018-1026, 2005

Chew EH and Hagen T. Substrate-mediated regulation of cullin neddylation. *J Biol Chem* 282: 17032-17040, 2007.

Chiu C, Tebo M, Ingles J, Yeates L, Arthur JW, Lind JM and Semsarian C. Genetic screening of calcium regulation genes in familial hypertrophic cardiomyopathy. *J Mol Cell Cardiol* 43: 337-343, 2007.

Clarke BA, Drujan D, Willis MS, Murphy LO, Corpina RA, Burova E, Rakhilin SV, Stitt TN, Patterson C, Latres E and Glass DJ. The E3 Ligase MuRF1 degrades myosin heavy chain protein in dexamethasone-treated skeletal muscle. *Cell Metab* 6: 376-385, 2007.

Clerin V, Shih HH, Deng N, Hebert G, Resmini C, Shields KM, Feldman JL, Winkler A, Albert L, Maganti V, Wong A, Paulsen JE, Keith JC, Jr., Vlasuk GP and Pittman DD. Expression of the cysteine protease legumain in vascular lesions and functional implications in atherogenesis. *Atherosclerosis* 201: 53-66, 2008.

Cohen S, Brault JJ, Gygi SP, Glass DJ, Valenzuela DM, Gartner C, Latres E and Goldberg AL. During muscle atrophy, thick, but not thin, filament components are degraded by MuRF1-dependent ubiquitylation. *J Cell Biol* 185: 1083-1095, 2009.

Coleman PS, Parmacek MS, Lesch M and Samarel AM. Protein synthesis and degradation during regression of thyroxine-induced cardiac hypertrophy. *J Mol Cell Cardiol* 21: 911-925, 1989.

Collins A, Lau W and de l, V. Mapping genes for common diseases: the case for genetic (LD) maps. *Hum Hered* 58: 2-9, 2004.

Collins BM. The structure and function of the retromer protein complex. *Traffic* 9: 1811-1822, 2008.

Colson BA, Bekyarova T, Locher MR, Fitzsimons DP, Irving TC and Moss RL. Protein kinase A-mediated phosphorylation of cMyBP-C increases proximity of myosin heads to actin in resting myocardium. *Circ Res* 103: 244-251, 2008.

Corfield VA, Moolman JC, Martell R and Brink PA. Polymerase chain reaction-based detection of MN blood group-specific sequences in the human genome. *Transfusion* 33: 119-124, 1993.

Craig R and Offer G. Axial arrangement of crossbridges in thick filaments of vertebrate skeletal muscle. *J Mol Biol* 102: 325-332, 1976.

Craig R and Offer G. The location of C-protein in rabbit skeletal muscle. *Proc R Soc Lond B Biol Sci* 192: 451-461, 1976.

Dantuma NP, Groothuis TA, Salomons FA and Neefjes J. A dynamic ubiquitin equilibrium couples proteasomal activity to chromatin remodeling. *J Cell Biol* 173: 19-26, 2006.

Daw EW, Chen SN, Czernuszewicz G, Lombardi R, Lu Y, Ma J, Roberts R, Shete S and Marian AJ. Genome-wide mapping of modifier chromosomal loci for human hypertrophic cardiomyopathy. *Hum Mol Genet* 16: 2463-2471, 2007.

Day RN and Schaufele F. Imaging molecular interactions in living cells. *Mol Endocrinol* 19: 1675-1686, 2005.

de Bie P, van de SB, Klomp L and Wijmenga C. The many faces of the copper metabolism protein MURR1/COMMD1. *J Hered* 96: 803-811, 2005.

de Bie P, van de SB, Burstein E, Duran KJ, Berger R, Duckett CS, Wijmenga C and Klomp LW. Characterization of COMMD protein-protein interactions in NF-kappaB signalling. *Biochem J* 398: 63-71, 2006.

de Bie P, van de SB, Burstein E, van de Berghe PV, Muller P, Berger R, Gitlin JD, Wijmenga C and Klomp LW. Distinct Wilson's disease mutations in ATP7B are associated with enhanced binding to COMMD1 and reduced stability of ATP7B. *Gastroenterology* 133: 1316-1326, 2007.

de la Vega, Isaac HI and Scafe CR. A tool for selecting SNPs for association studies based on observed linkage disequilibrium patterns. *Pac Symp Biocomput* 487-498, 2006.

Debonneville C, Flores SY, Kamynina E, Plant PJ, Tauxe C, Thomas MA, Munster C, Chraibi A, Pratt JH, Horisberger JD, Pearce D, Loffing J and Staub O. Phosphorylation of Nedd4-2 by Sgk1 regulates epithelial Na(+) channel cell surface expression. *EMBO J* 20: 7052-7059, 2001.

Decker RS, Decker ML, Kulikovskaya I, Nakamura S, Lee DC, Harris K, Klocke FJ and Winegrad S. Myosin-binding protein C phosphorylation, myofibril structure, and contractile function during low-flow ischemia. *Circulation* 111: 906-912, 2005.

Devlin AM, Solban N, Tremblay S, Gutkowska J, Schurch W, Orlov SN, Lewanczuk R, Hamet P and Tremblay J. HCaRG is a novel regulator of renal epithelial cell growth and differentiation causing G2M arrest. *Am J Physiol Renal Physiol* 284: F753-F762, 2003.

Dhandapany PS, Sadayappan S, Xue Y, Powell GT, Rani DS, Nallari P, Rai TS, Khullar M, Soares P, Bahl A, Tharkan JM, Vaideeswar P, Rathinavel A, Narasimhan C, Ayapati DR, Ayub Q, Mehdi SQ, Oppenheimer S, Richards MB, Price AL, Patterson N, Reich D, Singh L, Tyler-Smith C and Thangaraj K. A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia. *Nat Genet* 41: 187-191, 2009.

Diao G and Lin DY. Improving the power of association tests for quantitative traits in family studies. *Genet Epidemiol* 30: 301-313, 2006.

Dickinson DP. Cysteine peptidases of mammals: their biological roles and potential effects in the oral cavity and other tissues in health and disease. *Crit Rev Oral Biol Med* 13: 238-275, 2002.

Digilio MC, Conti E, Sarkozy A, Mingarelli R, Dottorini T, Marino B, Pizzuti A and Dallapiccola B. Grouping of multiple-lentigines/LEOPARD and Noonan syndromes on the PTPN11 gene. *Am J Hum Genet* 71: 389-394, 2002.

Donadio S, Alfaidy N, De KB, Micoud J, Feige JJ, Challis JR and Benharouga M. Expression and localization of cellular prion and COMMD1 proteins in human placenta throughout pregnancy. *Placenta* 28: 907-911, 2007.

Doolan G, Nguyen L, Chung J, Ingles J and Semsarian C. Progression of left ventricular hypertrophy and the angiotensin-converting enzyme gene polymorphism in hypertrophic cardiomyopathy. *Int J Cardiol* 96: 157-163, 2004.

Du M, Zhu MJ, Means WJ, Hess BW and Ford SP. Effect of nutrient restriction on calpain and calpastatin content of skeletal muscle from cows and fetuses. *J Anim Sci* 82: 2541-2547, 2004.

Duda DM, Borg LA, Scott DC, Hunt HW, Hammel M and Schulman BA. Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* 134: 995-1006, 2008.

Ehlermann P, Weichenhan D, Zehelein J, Steen H, Pribe R, Zeller R, Lehrke S, Zugck C, Ivandic BT and Katus HA. Adverse events in families with hypertrophic or dilated cardiomyopathy and mutations in the MYBPC3 gene. *BMC Med Genet* 9: 95, 2008.

Elsherif L, Wang L, Saari JT and Kang YJ. Regression of dietary copper restriction-induced cardiomyopathy by copper repletion in mice. *J Nutr* 134: 855-860, 2004.

Epstein ND, Fananapazir L, Lin HJ, Mulvihill J, White R, Lalouel JM, Lifton RP, Nienhuis AW and Leppert M. Evidence of genetic heterogeneity in five kindreds with familial hypertrophic cardiomyopathy. *Circulation* 85: 635-647, 1992.

Erdmann J, Raible J, Maki-Abadi J, Hummel M, Hammann J, Wollnik B, Frantz E, Fleck E, Hetzer R and Regitz-Zagrosek V. Spectrum of clinical phenotypes and gene variants in cardiac myosin-binding protein C mutation carriers with hypertrophic cardiomyopathy. *J Am Coll Cardiol* 38: 322-330, 2001.

Erdmann J, Daehmlow S, Wischke S, Senyuva M, Werner U, Raible J, Tanis N, Dyachenko S, Hummel M, Hetzer R and Regitz-Zagrosek V. Mutation spectrum in a large cohort of unrelated consecutive patients with hypertrophic cardiomyopathy. *Clin Genet* 64: 339-349, 2003.

Evangelou E, Trikalinos TA, Salanti G and Ioannidis JP. Family-based versus unrelated case-control designs for genetic associations. *PLoS Genet* 2: e123, 2006.

Fananapazir L and Epstein ND. 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, 1994.

Flashman E, Redwood C, Moolman-Smook J and Watkins H. Cardiac myosin binding protein C: its role in physiology and disease. *Circ Res* 94: 1279-1289, 2004.

Flashman E, Watkins H and Redwood C. Localization of the binding site of the C-terminal domain of cardiac myosin-binding protein-C on the myosin rod. *Biochem J* 401: 97-102, 2007.

Flashman E, Korkie L, Watkins H, Redwood C and Moolman-Smook JC. Support for a trimeric collar of myosin binding protein C in cardiac and fast skeletal muscle, but not in slow skeletal muscle. *FEBS Lett* 582: 434-438, 2008.

Flavigny J, Souchet M, Sebillon P, Berrebi-Bertrand I, Hainque B, Mallet A, Bril A, Schwartz K and Carrier L. 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: 443-456, 1999.

Foucault G, Vacher M, Merkulova T, Keller A and rrio-Dupont M. Presence of enolase in the M-band of skeletal muscle and possible indirect interaction with the cytosolic muscle isoform of creatine kinase. *Biochem J* 338 (Pt 1): 115-121, 1999.

Fougerousse F, Edom-Vovard F, Merkulova T, Ott MO, Durand M, Butler-Browne G and Keller A. The muscle-specific enolase is an early marker of human myogenesis. *J Muscle Res Cell Motil* 22: 535-544, 2001.

Freiburg A and Gautel M. 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: 317-323, 1996.

Freisinger P, Horvath R, Macmillan C, Peters J and Jaksch M. Reversion of hypertrophic cardiomyopathy in a patient with deficiency of the mitochondrial copper binding protein Sco2: is there a potential effect of copper? *J Inherit Metab Dis* 27: 67-79, 2004.

Frescas D, Guardavaccaro D, Bassermann F, Koyama-Nasu R and Pagano M. JHDM1B/FBXL10 is a nucleolar protein that represses transcription of ribosomal RNA genes. *Nature* 450: 309-313, 2007.

Frey N, Katus HA, Olson EN and Hill JA. Hypertrophy of the heart: a new therapeutic target? *Circulation* 109: 1580-1589, 2004.

Friedrich FW, Bausero P, Sun Y, Treszl A, Kramer E, Juhr D, Richard P, Wegscheider K, Schwartz K, Brito D, Arbustini E, Waldenstrom A, Isnard R, Komajda M, Eschenhagen T and

Carrier L. A new polymorphism in human calmodulin III gene promoter is a potential modifier gene for familial hypertrophic cardiomyopathy. *Eur Heart J* 30: 1648-1655, 2009.

Ganesh L, Burstein E, Guha-Niyogi A, Louder MK, Mascola JR, Klomp LW, Wijmenga C, Duckett CS and Nabel GJ. The gene product Murr1 restricts HIV-1 replication in resting CD4+ lymphocytes. *Nature* 426: 853-857, 2003.

Garcia-Castro M, Coto E, Reguero JR, Berrazueta JR, Alvarez V, Alonso B, Sainz R, Martin M and Moris C. [Mutations in sarcomeric genes MYH7, MYBPC3, TNNT2, TNNI3, and TPM1 in patients with hypertrophic cardiomyopathy]. *Rev Esp Cardiol* 62: 48-56, 2009.

Gautel M, Zuffardi O, Freiburg A and Labeit S. Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? *EMBO J* 14: 1952-1960, 1995.

Geier C, Gehmlich K, Ehler E, Hassfeld S, Perrot A, Hayess K, Cardim N, Wenzel K, Erdmann B, Krackhardt F, Posch MG, Osterziel KJ, Bublak A, Nagele H, Scheffold T, Dietz R, Chien KR, Spuler S, Furst DO, Nurnberg P and Ozcelik C. Beyond the sarcomere: CSRP3 mutations cause hypertrophic cardiomyopathy. *Hum Mol Genet* 17: 2753-2765, 2008.

Geisterfer-Lowrance AA, Christe M, Conner DA, Ingwall JS, Schoen FJ, Seidman CE and Seidman JG. A mouse model of familial hypertrophic cardiomyopathy. *Science* 272: 731-734, 1996.

Gietz RD, Triggs-Raine B, Robbins A, Graham KC and Woods RA. Identification of proteins that interact with a protein of interest: applications of the yeast two-hybrid system. *Mol Cell Biochem* 172: 67-79, 1997.

Glickman MH and Ciechanover A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev* 82: 373-428, 2002.

Golemis EA, Serebriiskii I and Law SF. The yeast two-hybrid system: criteria for detecting physiologically significant protein-protein interactions. *Curr Issues Mol Biol* 1: 31-45, 1999.

Goll DE, Dayton WR, Singh I and Robson RM. Studies of the alpha-actinin/actin interaction in the Z-disk by using calpain. *J Biol Chem* 266: 8501-8510, 1991.

Goll DE, Thompson VF, Taylor RG and Christiansen JA. Role of the calpain system in muscle growth. *Biochimie* 74: 225-237, 1992.

Goll DE, Thompson VF, Li H, Wei W and Cong J. The calpain system. *Physiol Rev* 83: 731-801, 2003.

Goll DE, Neti G, Mares SW and Thompson VF. Myofibrillar protein turnover: the proteasome and the calpains. *J Anim Sci* 86: E19-E35, 2008.

Gordon D and Finch SJ. Factors affecting statistical power in the detection of genetic association. *J Clin Invest* 115: 1408-1418, 2005.

Gordon D, Haynes C, Blumenfeld J and Finch SJ. PAWE-3D: visualizing power for association with error in case-control genetic studies of complex traits. *Bioinformatics* 21: 3935-3937, 2005.

Greene WC. How resting T cells deMURR HIV infection. *Nat Immunol* 5: 18-19, 2004.

Gruen M, Prinz H and Gautel M. 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: 254-259, 1999.

Gruen M and Gautel M. 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: 933-949, 1999.

Guinto PJ, Haim TE, Dowell-Martino CC, Sibinga N and Tardiff JC. Temporal and mutation-specific alterations in Ca2+ homeostasis differentially determine the progression of cTnT-related cardiomyopathies in murine models. *Am J Physiol Heart Circ Physiol* 297: H614-H626, 2009.

Gurevich EV and Gurevich VV. Arrestins: ubiquitous regulators of cellular signaling pathways. *Genome Biol* 7: 236, 2006.

Gustafsson AB and Gottlieb RA. Recycle or die: the role of autophagy in cardioprotection. *J Mol Cell Cardiol* 44: 654-661, 2008.

Haft CR, de la Luz SM, Barr VA, Haft DH and Taylor SI. Identification of a family of sorting nexin molecules and characterization of their association with receptors. *Mol Cell Biol* 18: 7278-7287, 1998.

Haft CR, de la Luz SM, Bafford R, Lesniak MA, Barr VA and Taylor SI. Human orthologs of yeast vacuolar protein sorting proteins Vps26, 29, and 35: assembly into multimeric complexes. *Mol Biol Cell* 11: 4105-4116, 2000.

Haglund K, Di Fiore PP and Dikic I. Distinct monoubiquitin signals in receptor endocytosis. *Trends Biochem Sci* 28: 598-603, 2003.

Harris SP, Bartley CR, Hacker TA, McDonald KS, Douglas PS, Greaser ML, Powers PA and Moss RL. Hypertrophic cardiomyopathy in cardiac myosin binding protein-C knockout mice. *Circ Res* 90: 594-601, 2002.

Harris SP, Rostkova E, Gautel M and Moss RL. Binding of myosin binding protein-C to myosin subfragment S2 affects contractility independent of a tether mechanism. *Circ Res* 95: 930-936, 2004.

Hartmann-Petersen R, Seeger M and Gordon C. Transferring substrates to the 26S proteasome. *Trends Biochem Sci* 28: 26-31, 2003.

Hartzell HC and Sale WS. Structure of C protein purified from cardiac muscle. *J Cell Biol* 100: 208-215, 1985.

Hattersley AT and McCarthy MI. What makes a good genetic association study? *Lancet* 366: 1315-1323, 2005.

Hayashi T, Arimura T, Itoh-Satoh M, Ueda K, Hohda S, Inagaki N, Takahashi M, Hori H, Yasunami M, Nishi H, Koga Y, Nakamura H, Matsuzaki M, Choi BY, Bae SW, You CW, Han KH, Park JE, Knoll R, Hoshijima M, Chien KR and Kimura A. Tcap gene mutations in hypertrophic cardiomyopathy and dilated cardiomyopathy. *J Am Coll Cardiol* 44: 2192-2201, 2004.

Hayden MS and Ghosh S. Signaling to NF-kappaB. *Genes Dev* 18: 2195-2224, 2004.

He J, Kallin EM, Tsukada Y and Zhang Y. The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15(Ink4b). *Nat Struct Mol Biol* 15: 1169-1175, 2008.

Herron TJ, Rostkova E, Kunst G, Chaturvedi R, Gautel M and Kentish JC. Activation of myocardial contraction by the N-terminal domains of myosin binding protein-C. *Circ Res* 98: 1290-1298, 2006.

Hlubocka Z, Marecek Z, Linhart A, Kejkova E, Pospisilova L, Martasek P and Aschermann M. Cardiac involvement in Wilson disease. *J Inherit Metab Dis* 25: 269-277, 2002.

Hoeller D, Hecker CM, Wagner S, Rogov V, Dotsch V and Dikic I. E3-independent monoubiquitination of ubiquitin-binding proteins. *Mol Cell* 26: 891-898, 2007.

Hu YH, Warnatz HJ, Vanhecke D, Wagner F, Fiebitz A, Thamm S, Kahlem P, Lehrach H, Yaspo ML and Janitz M. Cell array-based intracellular localization screening reveals novel functional features of human chromosome 21 proteins. *BMC Genomics* 7: 155, 2006.

Hua S and Sun Z. Support vector machine approach for protein subcellular localization prediction. *Bioinformatics* 17: 721-728, 2001.

Huang J and Forsberg NE. Role of calpain in skeletal-muscle protein degradation. *Proc Natl Acad Sci U S A* 95: 12100-12105, 1998.

Huang Y, Wu M and Li HY. Tumor suppressor ARF promotes non-classic proteasome-independent polyubiquitination of COMMD1. *J Biol Chem* 283: 11453-11460, 2008.

Hung IH, Suzuki M, Yamaguchi Y, Yuan DS, Klausner RD and Gitlin JD. Biochemical characterization of the Wilson disease protein and functional expression in the yeast Saccharomyces cerevisiae. *J Biol Chem* 272: 21461-21466, 1997.

Hurley JH. ESCRT complexes and the biogenesis of multivesicular bodies. *Curr Opin Cell Biol* 20: 4-11, 2008.

Huxford T, Huang DB, Malek S and Ghosh G. The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. *Cell* 95: 759-770, 1998.

Ishii S. Legumain: asparaginyl endopeptidase. *Methods Enzymol* 244: 604-615, 1994.

Jackson PK, Eldridge AG, Freed E, Furstenthal L, Hsu JY, Kaiser BK and Reimann JD. The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases. *Trends Cell Biol* 10: 429-439, 2000.

Jacques AM, Copeland O, Messer AE, Gallon CE, King K, McKenna WJ, Tsang VT and Marston SB. Myosin binding protein C phosphorylation in normal, hypertrophic and failing human heart muscle. *J Mol Cell Cardiol* 45: 209-216, 2008.

Jeffries CM, Whitten AE, Harris SP and Trewhella J. Small-angle X-ray scattering reveals the N-terminal domain organization of cardiac myosin binding protein C. *J Mol Biol* 377: 1186-1199, 2008.

Jin J, Cardozo T, Lovering RC, Elledge SJ, Pagano M and Harper JW. Systematic analysis and nomenclature of mammalian F-box proteins. *Genes Dev* 18: 2573-2580, 2004.

Jin X, Wang LS, Xia L, Zheng Y, Meng C, Yu Y, Chen GQ and Fang NY. Hyper-phosphorylation of alpha-enolase in hypertrophied left ventricle of spontaneously hypertensive rat. *Biochem Biophys Res Commun* 371: 804-809, 2008.

Jones WK, Brown M, Ren X, He S and McGuinness M. NF-kappaB as an integrator of diverse signaling pathways: the heart of myocardial signaling? *Cardiovasc Toxicol* 3: 229-254, 2003.

Jovic M, Sharma M, Rahajeng J and Caplan S. The early endosome: a busy sorting station for proteins at the crossroads. *Histol Histopathol* 25: 99-112, 2010.

Kachur TM and Pilgrim DB. Myosin assembly, maintenance and degradation in muscle: Role of the chaperone UNC-45 in myosin thick filament dynamics. *Int J Mol Sci* 9: 1863-1875, 2008.

Kai H, Muraishi A, Sugiu Y, Nishi H, Seki Y, Kuwahara F, Kimura A, Kato H and Imaizumi T. Expression of proto-oncogenes and gene mutation of sarcomeric proteins in patients with hypertrophic cardiomyopathy. *Circ Res* 83: 594-601, 1998.

Kaake R, Wang X and Huang L. Profiling of Protein Interaction Networks of Protein Complexes Using Affinity Purification. *Mol & Cell Proteomics 9.8: 1650-1665*

and Quantitative Mass Spectrometry*

Kaler SG. Diagnosis and therapy of Menkes syndrome, a genetic form of copper deficiency. *Am J Clin Nutr* 67: 1029S-1034S, 1998.

Ke Y, Butt AG, Swart M, Liu YF and McDonald FJ. COMMD1 downregulates the epithelial sodium channel through Nedd4-2. *Am J Physiol Renal Physiol* 298: F1445-F1456, 2010.

Kedar V, McDonough H, Arya R, Li HH, Rockman HA and Patterson C. Muscle-specific RING finger 1 is a bona fide ubiquitin ligase that degrades cardiac troponin I. *Proc Natl Acad Sci U S A* 101: 18135-18140, 2004.

Keller A, Rouzeau JD, Farhadian F, Wisnewsky C, Marotte F, Lamande N, Samuel JL, Schwartz K, Lazar M and Lucas M. Differential expression of alpha- and beta-enolase genes during rat heart development and hypertrophy. *Am J Physiol* 269: H1843-H1851, 1995.

Keller A, Peltzer J, Carpentier G, Horvath I, Olah J, Duchesnay A, Orosz F and Ovadi J. Interactions of enolase isoforms with tubulin and microtubules during myogenesis. *Biochim Biophys Acta* 1770: 919-926, 2007.

Keller DI, Coirault C, Rau T, Cheav T, Weyand M, Amann K, Lecarpentier Y, Richard P, Eschenhagen T and Carrier L. Human homozygous R403W mutant cardiac myosin presents disproportionate enhancement of mechanical and enzymatic properties. *J Mol Cell Cardiol* 36: 355-362, 2004.

Kensler RW and Harris SP. The structure of isolated cardiac Myosin thick filaments from cardiac Myosin binding protein-C knockout mice. *Biophys J* 94: 1707-1718, 2008.

Klarlund JK, Guilherme A, Holik JJ, Virbasius JV, Chawla A and Czech MP. Signaling by phosphoinositide-3,4,5-trisphosphate through proteins containing pleckstrin and Sec7 homology domains. *Science* 275: 1927-1930, 1997.

Klomp AE, van de SB, Klomp LW and Wijmenga C. The ubiquitously expressed MURR1 protein is absent in canine copper toxicosis. *J Hepatol* 39: 703-709, 2003.

Klues HG, Schiffers A and Maron BJ. Phenotypic spectrum and patterns of left ventricular hypertrophy in hypertrophic cardiomyopathy: morphologic observations and significance as assessed by two-dimensional echocardiography in 600 patients. *J Am Coll Cardiol* 26: 1699-1708, 1995.

Koegl M, Hoppe T, Schlenker S, Ulrich HD, Mayer TU and Jentsch S. A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* 96: 635-644, 1999.

Kong X, varez-Castelao B, Lin Z, Castano JG and Caro J. Constitutive/hypoxic degradation of HIF-alpha proteins by the proteasome is independent of von Hippel Lindau protein ubiquitylation and the transactivation activity of the protein. *J Biol Chem* 282: 15498-15505, 2007.

Konno T, Shimizu M, Ino H, Matsuyama T, Yamaguchi M, Terai H, Hayashi K, Mabuchi T, Kiyama M, Sakata K, Hayashi T, Inoue M, Kaneda T and Mabuchi H. 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: 781-786, 2003.

Koretz JF. Effects of C-protein on synthetic myosin filament structure. *Biophys J* 27: 433-446, 1979.

Korolchuk VI, Menzies FM and Rubinsztein DC. Mechanisms of cross-talk between the ubiquitin-proteasome and autophagy-lysosome systems. *FEBS Lett* 584: 1393-1398, 2010.

Koyama-Nasu R, David G and Tanese N. The F-box protein Fbl10 is a novel transcriptional repressor of c-Jun. *Nat Cell Biol* 9: 1074-1080, 2007.

Kulikovskaya I, McClellan G, Flavigny J, Carrier L and Winegrad S. Effect of MyBP-C binding to actin on contractility in heart muscle. *J Gen Physiol* 122: 761-774, 2003.

Kulikovskaya I, McClellan GB, Levine R and Winegrad S. Multiple forms of cardiac myosin-binding protein C exist and can regulate thick filament stability. *J Gen Physiol* 129: 419-428, 2007.

Kunst G, Kress KR, Gruen M, Uttenweiler D, Gautel M and Fink RH. 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: 51-58, 2000.

Kuo YM, Zhou B, Cosco D and Gitschier J. The copper transporter CTR1 provides an essential function in mammalian embryonic development. *Proc Natl Acad Sci U S A* 98: 6836-6841, 2001.

Labeit S, Gautel M, Lakey A and Trinick J. Towards a molecular understanding of titin. *EMBO J* 11: 1711-1716, 1992.

Laing NG and Nowak KJ. When contractile proteins go bad: the sarcomere and skeletal muscle disease. *Bioessays* 27: 809-822, 2005.

Laird NM and Lange C. Family-based designs in the age of large-scale gene-association studies. *Nat Rev Genet* 7: 385-394, 2006.

Lang F, Bohmer C, Palmada M, Seebohm G, Strutz-Seebohm N and Vallon V. (Patho)physiological significance of the serum- and glucocorticoid-inducible kinase isoforms. *Physiol Rev* 86: 1151-1178, 2006.

Lange EM, Sun J, Lange LA, Zheng SL, Duggan D, Carpten JD, Gronberg H, Isaacs WB, Xu J and Chang BL. Family-based samples can play an important role in genetic association studies. *Cancer Epidemiol Biomarkers Prev* 17: 2208-2214, 2008.

Lankford EB, Epstein ND, Fananapazir L and Sweeney HL. Abnormal contractile properties of muscle fibers expressing beta-myosin heavy chain gene mutations in patients with hypertrophic cardiomyopathy. *J Clin Invest* 95: 1409-1414, 1995.

Lecarpentier Y, Vignier N, Oliviero P, Guellich A, Carrier L and Coirault C. Cardiac Myosin-binding protein C modulates the tuning of the molecular motor in the heart. *Biophys J* 95: 720-728, 2008.

Lechin M, Quinones MA, Omran A, Hill R, Yu QT, Rakowski H, Wigle D, Liew CC, Sole M, Roberts R and . Angiotensin-I converting enzyme genotypes and left ventricular hypertrophy in patients with hypertrophic cardiomyopathy. *Circulation* 92: 1808-1812, 1995.

Lecker SH, Solomon V, Mitch WE and Goldberg AL. Muscle protein breakdown and the critical role of the ubiquitin-proteasome pathway in normal and disease states. *J Nutr* 129: 227S-237S, 1999.

Lee IH, Dinudom A, Sanchez-Perez A, Kumar S and Cook DI. Akt mediates the effect of insulin on epithelial sodium channels by inhibiting Nedd4-2. *J Biol Chem* 282: 29866-29873, 2007.

Lekanne Deprez RH, Muurling-Vlietman JJ, Hruda J, Baars MJ, Wijnaendts LC, Stolte-Dijkstra I, Alders M and van Hagen JM. Two cases of severe neonatal hypertrophic cardiomyopathy caused by compound heterozygous mutations in the MYBPC3 gene. *J Med Genet* 43: 829-832, 2006.

Levine R, Weisberg A, Kulikovskaya I, McClellan G and Winegrad S. Multiple structures of thick filaments in resting cardiac muscle and their influence on cross-bridge interactions. *Biophys J* 81: 1070-1082, 2001.

Levy D, Anderson KM, Savage DD, Kannel WB, Christiansen JC and Castelli WP. Echocardiographically detected left ventricular hypertrophy: prevalence and risk factors. The Framingham Heart Study. *Ann Intern Med* 108: 7-13, 1988.

Li C, Kelley J and Ha T. Activation of nuclear factor-kappaB. *Methods Mol Biol* 315: 141-150, 2006.

Li DN, Matthews SP, Antoniou AN, Mazzeo D and Watts C. Multistep autoactivation of asparaginyl endopeptidase in vitro and in vivo. *J Biol Chem* 278: 38980-38990, 2003.

Li HH, Kedar V, Zhang C, McDonough H, Arya R, Wang DZ and Patterson C. Atrogin-1/muscle atrophy F-box inhibits calcineurin-dependent cardiac hypertrophy by participating in an SCF ubiquitin ligase complex. *J Clin Invest* 114: 1058-1071, 2004.

Li HH, Willis MS, Lockyer P, Miller N, McDonough H, Glass DJ and Patterson C. Atrogin-1 inhibits Akt-dependent cardiac hypertrophy in mice via ubiquitin-dependent coactivation of Forkhead proteins. *J Clin Invest* 117: 3211-3223, 2007.

Li P, Zhang S and Fan C. COMMD6 from amphioxus Branchiostoma belcheri (BbCOMMD6) interacts with creatine kinase and inhibits its activity. *Int J Biochem Cell Biol* 41: 2459-2465, 2009.

Lian M and Zheng X. HSCARG regulates NF-kappaB activation by promoting the ubiquitination of RelA or COMMD1. *J Biol Chem* 284: 17998-18006, 2009.

Lindmo K and Stenmark H. Regulation of membrane traffic by phosphoinositide 3-kinases. *J Cell Sci* 119: 605-614, 2006.

Liu YV and Semenza GL. RACK1 vs. HSP90: competition for HIF-1 alpha degradation vs. stabilization. *Cell Cycle* 6: 656-659, 2007.

Liu YV, Baek JH, Zhang H, Diez R, Cole RN and Semenza GL . RACK1 competes with HSP90 for binding to HIF-1alpha and is required for O(2)-independent and HSP90 inhibitor-induced degradation of HIF-1alpha. *Mol Cell* 25: 207-217, 2007.

Liu YV, Hubbi ME, Pan F, McDonald KR, Mansharamani M, Cole RN, Liu JO and Semenza GL. Calcineurin promotes hypoxia-inducible factor 1alpha expression by dephosphorylating RACK1 and blocking RACK1 dimerization. *J Biol Chem* 282: 37064-37073, 2007.

Lowey S, Lesko LM, Rovner AS, Hodges AR, White SL, Low RB, Rincon M, Gulick J and Robbins J. Functional effects of the hypertrophic cardiomyopathy R403Q mutation are different in an alpha- or beta-myosin heavy chain backbone. *J Biol Chem* 283: 20579-20589, 2008.

Luther PK, Bennett PM, Knupp C, Craig R, Padron R, Harris SP, Patel J and Moss RL. 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, 2008.

Luzio JP, Piper SC, Bowers K, Parkinson MD, Lehner PJ and Bright NA. ESCRT proteins and the regulation of endocytic delivery to lysosomes. *Biochem Soc Trans* 37: 178-180, 2009.

Maine GN, Mao X, Komarck CM and Burstein E. COMMD1 promotes the ubiquitination of NF-kappaB subunits through a cullin-containing ubiquitin ligase. *EMBO J* 26: 436-447, 2007.

Maine GN and Burstein E. COMMD proteins and the control of the NF kappa B pathway. *Cell Cycle* 6: 672-676, 2007.

Maine GN and Burstein E. COMMD proteins: COMMing to the scene. *Cell Mol Life Sci* 64: 1997-2005, 2007.

Maine GN, Mao X, Muller PA, Komarck CM, Klomp LW and Burstein E. COMMD1 expression is controlled by critical residues that determine XIAP binding. *Biochem J* 417: 601-609, 2009.

Malik B, Price SR, Mitch WE, Yue Q and Eaton DC. Regulation of epithelial sodium channels by the ubiquitin-proteasome proteolytic pathway. *Am J Physiol Renal Physiol* 290: F1285-F1294, 2006.

Maniatis N, Collins A, Xu CF, McCarthy LC, Hewett DR, Tapper W, Ennis S, Ke X and Morton NE. The first linkage disequilibrium (LD) maps: delineation of hot and cold blocks by diplotype analysis. *Proc Natl Acad Sci U S A* 99: 2228-2233, 2002.

Marian AJ, Yu QT, Workman R, Greve G and Roberts R. Angiotensin-converting enzyme polymorphism in hypertrophic cardiomyopathy and sudden cardiac death. *Lancet* 342: 1085-1086, 1993.

Marian AJ, Wu Y, Lim DS, McCluggage M, Youker K, Yu QT, Brugada R, DeMayo F, Quinones M and Roberts R. A transgenic rabbit model for human hypertrophic cardiomyopathy. *J Clin Invest* 104: 1683-1692, 1999.

Marian AJ, Salek L and Lutucuta S. Molecular genetics and pathogenesis of hypertrophic cardiomyopathy. *Minerva Med* 92: 435-451, 2001.

Marian AJ. Modifier genes for hypertrophic cardiomyopathy. *Curr Opin Cardiol* 17: 242-252, 2002.

Marian AJ. Update on hypertrophic cardiomyopathy. *Tex Heart Inst J* 37: 322-323, 2010.

Marston S, Copeland O, Jacques A, Livesey K, Tsang V, McKenna WJ, Jalilzadeh S, Carballo S, Redwood C and Watkins H. Evidence from human myectomy samples that MYBPC3 mutations cause hypertrophic cardiomyopathy through haploinsufficiency. *Circ Res* 105: 219-222, 2009.

Martin AF, Rabinowitz M, Blough R, Prior G and Zak R. Measurements of half-life of rat cardiac myosin heavy chain with leucyl-tRNA used as precursor pool. *J Biol Chem* 252: 3422-3429, 1977.

Mayosi BM, Keavney B, Watkins H and Farrall M. Measured haplotype analysis of the aldosterone synthase gene and heart size. *Eur J Hum Genet* 11: 395-401, 2003.

Mcalister-Henn L, Gibson N and Panisko E. Applications of the yeast two-hybrid system. *Methods* 19: 330-337, 1999.

McClellan G, Kulikovskaya I and Winegrad S. Changes in cardiac contractility related to calcium-mediated changes in phosphorylation of myosin-binding protein C. *Biophys J* 81: 1083-1092, 2001.

McConnell BK, Jones KA, Fatkin D, Arroyo LH, Lee RT, Aristizabal O, Turnbull DH, Georgakopoulos D, Kass D, Bond M, Niimura H, Schoen FJ, Conner D, Fischman DA, Seidman CE and Seidman JG. Dilated cardiomyopathy in homozygous myosin-binding protein-C mutant mice. *J Clin Invest* 104: 1235-1244, 1999.

McGrath MJ, Cottle DL, Nguyen MA, Dyson JM, Coghill ID, Robinson PA, Holdsworth M, Cowling BS, Hardeman EC, Mitchell CA and Brown S. Four and a half LIM protein 1 binds myosin-binding protein C and regulates myosin filament formation and sarcomere assembly. *J Biol Chem* 281: 7666-7683, 2006.

McIntyre LM, Martin ER, Simonsen KL and Kaplan NL. Circumventing multiple testing: a multilocus Monte Carlo approach to testing for association. *Genet Epidemiol* 19: 18-29, 2000.

Mearini G, Gedicke C, Schlossarek S, Witt CC, Kramer E, Cao P, Gomes MD, Lecker SH, Labeit S, Willis MS, Eschenhagen T and Carrier L. Atrogin-1 and MuRF1 regulate cardiac MyBP-C levels via different mechanisms. *Cardiovasc Res* 2009.

Medin M, Hermida-Prieto M, Monserrat L, Laredo R, Rodriguez-Rey JC, Fernandez X and Castro-Beiras A. Mutational screening of phospholamban gene in hypertrophic and idiopathic dilated cardiomyopathy and functional study of the PLN -42 C>G mutation. *Eur J Heart Fail* 9: 37-43, 2007.

Meiners S, Dreger H, Fechner M, Bieler S, Rother W, Gunther C, Baumann G, Stangl V and Stangl K. Suppression of cardiomyocyte hypertrophy by inhibition of the ubiquitin-proteasome system. *Hypertension* 51: 302-308, 2008.

Mercer JF. Menkes syndrome and animal models. *Am J Clin Nutr* 67: 1022S-1028S, 1998.

Merkulova T, Lucas M, Jabet C, Lamande N, Rouzeau JD, Gros F, Lazar M and Keller A. Biochemical characterization of the mouse muscle-specific enolase: developmental changes in electrophoretic variants and selective binding to other proteins. *Biochem J* 323 (Pt 3): 791-800, 1997.

Merkulova T, Keller A, Oliviero P, Marotte F, Samuel JL, Rappaport L, Lamande N and Lucas M. Thyroid hormones differentially modulate enolase isozymes during rat skeletal and cardiac muscle development. *Am J Physiol Endocrinol Metab* 278: E330-E339, 2000.

Messer AE, Jacques AM and Marston SB. Troponin phosphorylation and regulatory function in human heart muscle: dephosphorylation of Ser23/24 on troponin I could account for the contractile defect in end-stage heart failure. *J Mol Cell Cardiol* 42: 247-259, 2007.

Michele DE, Gomez CA, Hong KE, Westfall MV and Metzger JM. Cardiac dysfunction in hypertrophic cardiomyopathy mutant tropomyosin mice is transgene-dependent, hypertrophy-independent, and improved by beta-blockade. *Circ Res* 91: 255-262, 2002.

Minamisawa S, Sato Y, Tatsuguchi Y, Fujino T, Imamura S, Uetsuka Y, Nakazawa M and Matsuoka R. Mutation of the phospholamban promoter associated with hypertrophic cardiomyopathy. *Biochem Biophys Res Commun* 304: 1-4, 2003.

Mizukami Y, Li J, Zhang X, Zimmer MA, Iliopoulos O and Chung DC. Hypoxia-inducible factor-1-independent regulation of vascular endothelial growth factor by hypoxia in colon cancer. *Cancer Res* 64: 1765-1772, 2004.

Mizutani R, Yamauchi J, Kusakawa S, Nakamura K, Sanbe A, Torii T, Miyamoto Y and Tanoue A. Sorting nexin 3, a protein upregulated by lithium, contains a novel phosphatidylinositol-binding sequence and mediates neurite outgrowth in N1E-115 cells. *Cell Signal* 21: 1586-1594, 2009.

Mogensen J, Klausen IC, Pedersen AK, Egeblad H, Bross P, Kruse TA, Gregersen N, Hansen PS, Baandrup U and Borglum AD. Alpha-cardiac actin is a novel disease gene in familial hypertrophic cardiomyopathy. *J Clin Invest* 103: R39-R43, 1999.

Mohiddin SA, Begley DA, McLam E, Cardoso JP, Winkler JB, Sellers JR and Fananapazir L. Utility of genetic screening in hypertrophic cardiomyopathy: prevalence and significance of novel and double (homozygous and heterozygous) beta-myosin mutations. *Genet Test* 7: 21-27, 2003.

Molkentin JD, Lu JR, Antos CL, Markham B, Richardson J, Robbins J, Grant SR and Olson EN. A calcineurin-dependent transcriptional pathway for cardiac hypertrophy. *Cell* 93: 215-228, 1998.

Moolman-Smook J, Flashman E, de LW, Li Z, Corfield V, Redwood C and Watkins H. Identification of novel interactions between domains of Myosin binding protein-C that are modulated by hypertrophic cardiomyopathy missense mutations. *Circ Res* 91: 704-711, 2002.

Moolman-Smook JC, De Lange J, Brink A and Corfield A. Hypertrophic cardiomyopathy repealing tenets in South Africa. *Cardiovasc J S Afr* 11: 202-209, 2000.

Moolman JC, Corfield VA, Posen B, Ngumbela K, Seidman C, Brink PA and Watkins H. Sudden death due to troponin T mutations. *J Am Coll Cardiol* 29: 549-555, 1997.

Moos C, Offer G, Starr R and Bennett P. Interaction of C-protein with myosin, myosin rod and light meromyosin. *J Mol Biol* 97: 1-9, 1975.

Moos C, Mason CM, Besterman JM, Feng IN and Dubin JH. 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: 571-586, 1978.

Morita Y, Araki H, Sugimoto T, Takeuchi K, Yamane T, Maeda T, Yamamoto Y, Nishi K, Asano M, Shirahama-Noda K, Nishimura M, Uzu T, Hara-Nishimura I, Koya D, Kashiwagi A and Ohkubo I. Legumain/asparaginyl endopeptidase controls extracellular matrix remodeling through the degradation of fibronectin in mouse renal proximal tubular cells. *FEBS Lett* 581: 1417-1424, 2007.

Morner S, Richard P, Kazzam E, Hellman U, Hainque B, Schwartz K and Waldenstrom A. Identification of the genotypes causing hypertrophic cardiomyopathy in northern Sweden. *J Mol Cell Cardiol* 35: 841-849, 2003.

Mufti AR, Burstein E, Csomos RA, Graf PC, Wilkinson JC, Dick RD, Challa M, Son JK, Bratton SB, Su GL, Brewer GJ, Jakob U and Duckett CS. XIAP Is a copper binding protein deregulated in Wilson's disease and other copper toxicosis disorders. *Mol Cell* 21: 775-785, 2006.

Mufti AR, Burstein E and Duckett CS. XIAP: cell death regulation meets copper homeostasis. *Arch Biochem Biophys* 463: 168-174, 2007.

Mukhopadhyay D and Riezman H. Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science* 315: 201-205, 2007.

Muller PA, van de SB, Groot AJ, Verbeek D, Vonk WI, Maine GN, Burstein E, Wijmenga C, Vooijs M, Reits E and Klomp LW. Nuclear-cytosolic transport of COMMD1 regulates NF-kappaB and HIF-1 activity. *Traffic* 10: 514-527, 2009.

Murthy RV, Arbman G, Gao J, Roodman GD and Sun XF. Legumain expression in relation to clinicopathologic and biological variables in colorectal cancer. *Clin Cancer Res* 11: 2293-2299, 2005.

Muthuchamy M, Pieples K, Rethinasamy P, Hoit B, Grupp IL, Boivin GP, Wolska B, Evans C, Solaro RJ and Wieczorek DF. Mouse model of a familial hypertrophic cardiomyopathy mutation in alpha-tropomyosin manifests cardiac dysfunction. *Circ Res* 85: 47-56, 1999.

Nabetani A, Hatada I, Morisaki H, Oshimura M and Mukai T . Mouse U2af1-rs1 is a neomorphic imprinted gene. *Mol Cell Biol* 17: 789-798, 1997.

Nagayama T, Takimoto E, Sadayappan S, Mudd JO, Seidman JG, Robbins J and Kass DA. Control of in vivo left ventricular [correction] contraction/relaxation kinetics by myosin binding protein C: protein kinase A phosphorylation dependent and independent regulation. *Circulation* 116: 2399-2408, 2007.

Nair R and Rost B. Inferring sub-cellular localization through automated lexical analysis. *Bioinformatics* 18 Suppl 1: S78-S86, 2002.

Nakai A, Yamaguchi O, Takeda T, Higuchi Y, Hikoso S, Taniike M, Omiya S, Mizote I, Matsumura Y, Asahi M, Nishida K, Hori M, Mizushima N and Otsu K. The role of autophagy in cardiomyocytes in the basal state and in response to hemodynamic stress. *Nat Med* 13: 619-624, 2007.

Nakai K and Kanehisa M. A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14: 897-911, 1992.

Nakamura A, Hattori M and Sakaki Y. Isolation of a novel human gene from the Down syndrome critical region of chromosome 21q22.2. *J Biochem* 122: 872-877, 1997.

Narindrasorasak S, Kulkarni P, Deschamps P, She YM and Sarkar B. Characterization and copper binding properties of human COMMD1 (MURR1). *Biochemistry* 46: 3116-3128, 2007.

Nau PN, Van NT, Ralphe JC, Teneyck CJ, Bedell KA, Caldarone CA, Segar JL and Scholz TD. Metabolic adaptation of the fetal and postnatal ovine heart: regulatory role of hypoxia-inducible factors and nuclear respiratory factor-1. *Pediatr Res* 52: 269-278, 2002.

Neitzel H. A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum Genet* 73: 320-326, 1986.

Niimura H, Bachinski LL, Sangwatanaroj S, Watkins H, Chudley AE, McKenna W, Kristinsson A, Roberts R, Sole M, Maron BJ, Seidman JG and Seidman CE. Mutations in the gene for cardiac

myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy. *N Engl J Med* 338: 1248-1257, 1998.

Nyamsuren O, Faggionato D, Loch W, Schulze E and Baumeister R. A mutation in CHN-1/CHIP suppresses muscle degeneration in Caenorhabditis elegans. *Dev Biol* 312: 193-202, 2007.

Offer G, Moos C and Starr R. A new protein of the thick filaments of vertebrate skeletal myofibrils. Extractions, purification and characterization. *J Mol Biol* 74: 653-676, 1973.

Okagaki T, Weber FE, Fischman DA, Vaughan KT, Mikawa T and Reinach FC. The major myosin-binding domain of skeletal muscle MyBP-C (C protein) resides in the COOH-terminal, immunoglobulin C2 motif. *J Cell Biol* 123: 619-626, 1993.

Palmer BM, McConnell BK, Li GH, Seidman CE, Seidman JG, Irving TC, Alpert NR and Maughan DW. Reduced cross-bridge dependent stiffness of skinned myocardium from mice lacking cardiac myosin binding protein-C. *Mol Cell Biochem* 263: 73-80, 2004.

Pancholi V. Multifunctional alpha-enolase: its role in diseases. *Cell Mol Life Sci* 58: 902-920, 2001.

Parrish JR, Gulyas KD and Finley RL, Jr. Yeast two-hybrid contributions to interactome mapping. *Curr Opin Biotechnol* 17: 387-393, 2006.

Patel R, Lim DS, Reddy D, Nagueh SF, Lutucuta S, Sole MJ, Zoghbi WA, Quinones MA, Roberts R and Marian AJ. Variants of trophic factors and expression of cardiac hypertrophy in patients with hypertrophic cardiomyopathy. *J Mol Cell Cardiol* 32: 2369-2377, 2000.

Patton EE, Willems AR and Tyers M. Combinatorial control in ubiquitin-dependent proteolysis: don't Skp the F-box hypothesis. *Trends Genet* 14: 236-243, 1998.

Petris MJ, Mercer JF, Culvenor JG, Lockhart P, Gleeson PA and Camakaris J. Ligand-regulated transport of the Menkes copper P-type ATPase efflux pump from the Golgi apparatus to the plasma membrane: a novel mechanism of regulated trafficking. *EMBO J* 15: 6084-6095, 1996.

Pickart CM and Cohen RE. Proteasomes and their kin: proteases in the machine age. *Nat Rev Mol Cell Biol* 5: 177-187, 2004.

Pilia G, Chen WM, Scuteri A, Orru M, Albai G, Dei M, Lai S, Usala G, Lai M, Loi P, Mameli C, Vacca L, Deiana M, Olla N, Masala M, Cao A, Najjar SS, Terracciano A, Nedorezov T, Sharov A, Zonderman AB, Abecasis GR, Costa P, Lakatta E and Schlessinger D. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* 2: e132, 2006.

Piper RC and Katzmann DJ. Biogenesis and function of multivesicular bodies. *Annu Rev Cell Dev Biol* 23: 519-547, 2007.

Pohlmann L, Kroger I, Vignier N, Schlossarek S, Kramer E, Coirault C, Sultan KR, El-Armouche A, Winegrad S, Eschenhagen T and Carrier L. Cardiac myosin-binding protein C is required for complete relaxation in intact myocytes. *Circ Res* 101: 928-938, 2007.

Pons V, Luyet PP, Morel E, Abrami L, van der Goot FG, Parton RG and Gruenberg J. Hrs and SNX3 functions in sorting and membrane invagination within multivesicular bodies. *PLoS Biol* 6: e214, 2008.

Portbury AL, Willis MS and Patterson C. Tearin' up my heart: proteolysis in the cardiac sarcomere. *J Biol Chem* 286: 9929-9934, 2011.

Powell SR. The cardiac 26S proteasome: regulating the regulator. *Circ Res* 99: 342-345, 2006.

Powell SR. The ubiquitin-proteasome system in cardiac physiology and pathology. *Am J Physiol Heart Circ Physiol* 291: H1-H19, 2006.

Pritchard JK and Donnelly P. Case-control studies of association in structured or admixed populations. *Theor Popul Biol* 60: 227-237, 2001.

Pullan L, Mullapudi S, Huang Z, Baldwin PR, Chin C, Sun W, Tsujimoto S, Kolodziej SJ, Stoops JK, Lee JC, Waxham MN, Bean AJ and Penczek PA. The endosome-associated protein Hrs is hexameric and controls cargo sorting as a "master molecule". *Structure* 14: 661-671, 2006.

Purcell NH, Tang G, Yu C, Mercurio F, DiDonato JA and Lin A. Activation of NF-kappa B is required for hypertrophic growth of primary rat neonatal ventricular cardiomyocytes. *Proc Natl Acad Sci U S A* 98: 6668-6673, 2001.

Purintrapiban J, Wang MC and Forsberg NE. Degradation of sarcomeric and cytoskeletal proteins in cultured skeletal muscle cells. *Comp Biochem Physiol B Biochem Mol Biol* 136: 393-401, 2003.

Pyle GG, Kamunde CN, McDonald DG and Wood CM. Dietary sodium inhibits aqueous copper uptake in rainbow trout (Oncorhynchus mykiss). *J Exp Biol* 206: 609-618, 2003.

Raiborg C and Stenmark H. The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. *Nature* 458: 445-452, 2009.

Ramburan A. Investigation of the N-terminal interactions of cardiac Myosin-Binding Protein C (cMyBPC) under defined phosphorylation states. *Phd Thesis*, 2008, University of Stellenbosch

Razeghi P, Baskin KK, Sharma S, Young ME, Stepkowski S, Essop MF and Taegtmeyer H. Atrophy, hypertrophy, and hypoxemia induce transcriptional regulators of the ubiquitin proteasome system in the rat heart. *Biochem Biophys Res Commun* 342: 361-364, 2006.

Razumova MV, Bezold KL, Tu AY, Regnier M and Harris SP. Contribution of the myosin binding protein C motif to functional effects in permeabilized rat trabeculae. *J Gen Physiol* 132: 575-585, 2008.

Redwood CS, Moolman-Smook JC and Watkins H. Properties of mutant contractile proteins that cause hypertrophic cardiomyopathy. *Cardiovasc Res* 44: 20-36, 1999.

Reid MB. Response of the ubiquitin-proteasome pathway to changes in muscle activity. *Am J Physiol Regul Integr Comp Physiol* 288: R1423-R1431, 2005.

Reinhardt A and Hubbard T. Using neural networks for prediction of the subcellular location of proteins. *Nucleic Acids Res* 26: 2230-2236, 1998.

Revera M, van der ML, Heradien M, Goosen A, Corfield VA, Brink PA and Moolman-Smook JC. Troponin T and beta-myosin mutations have distinct cardiac functional effects in hypertrophic cardiomyopathy patients without hypertrophy. *Cardiovasc Res* 77: 687-694, 2008.

Richard P, Charron P, Carrier L, Ledeuil C, Cheav T, Pichereau C, Benaiche A, Isnard R, Dubourg O, Burban M, Gueffet JP, Millaire A, Desnos M, Schwartz K, Hainque B and Komajda M. Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. *Circulation* 107: 2227-2232, 2003.

Rodriguez-Garcia MI, Monserrat L, Ortiz M, Fernandez X, Cazon L, Nunez L, Barriales-Villa R, Maneiro E, Veira E, Castro-Beiras A and Hermida-Prieto M. Screening mutations in myosin binding protein C3 gene in a cohort of patients with Hypertrophic Cardiomyopathy. *BMC Med Genet* 11: 67, 2010.

Rome E, Offer G and Pepe FA. X-ray diffraction of muscle labelled with antibody to C-protein. *Nat New Biol* 244: 152-154, 1973.

Rotin D, Kanelis V and Schild L. Trafficking and cell surface stability of ENaC. *Am J Physiol Renal Physiol* 281: F391-F399, 2001.

Rottbauer W, Gautel M, Zehelein J, Labeit S, Franz WM, Fischer C, Vollrath B, Mall G, Dietz R, Kubler W and Katus HA. 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: 475-482, 1997.

Ryan KM, Ernst MK, Rice NR and Vousden KH. Role of NF-kappaB in p53-mediated programmed cell death. *Nature* 404: 892-897, 2000.

Ryo A, Suizu F, Yoshida Y, Perrem K, Liou YC, Wulf G, Rottapel R, Yamaoka S and Lu KP. Regulation of NF-kappaB signaling by Pin1-dependent prolyl isomerization and ubiquitin-mediated proteolysis of p65/RelA. *Mol Cell* 12: 1413-1426, 2003.

Sachdev B and Elliott PM. Isolated cardiac manifestations in Fabry disease: the UK experience. *Acta Paediatr Suppl* 91: 28-30, 2002.

Sadayappan S, Gulick J, Osinska H, Martin LA, Hahn HS, Dorn GW, Klevitsky R, Seidman CE, Seidman JG and Robbins J. Cardiac myosin-binding protein-C phosphorylation and cardiac function. *Circ Res* 97: 1156-1163, 2005.

Sadayappan S, Osinska H, Klevitsky R, Lorenz JN, Sargent M, Molkentin JD, Seidman CE, Seidman JG and Robbins J. Cardiac myosin binding protein C phosphorylation is cardioprotective. *Proc Natl Acad Sci U S A* 103: 16918-16923, 2006.

Sadayappan S, Gulick J, Klevitsky R, Lorenz JN, Sargent M, Molkentin JD and Robbins J. Cardiac myosin binding protein-C phosphorylation in a {beta}-myosin heavy chain background. *Circulation* 119: 1253-1262, 2009.

Saha A and Deshaies RJ. Multimodal activation of the ubiquitin ligase SCF by Nedd8 conjugation. *Mol Cell* 32: 21-31, 2008.

Saksena S, Sun J, Chu T and Emr SD. ESCRTing proteins in the endocytic pathway. *Trends Biochem Sci* 32: 561-573, 2007.

Salomons FA, Menendez-Benito V, Bottcher C, McCray BA, Taylor JP and Dantuma NP. Selective accumulation of aggregation-prone proteasome substrates in response to proteotoxic stress. *Mol Cell Biol* 29: 1774-1785, 2009.

Sanbe A, Osinska H, Saffitz J, Glabe C, Kayed R and Maloyan A. Desmin-related cardiomyopathy in transgenic mice: A cardiac amyloidosis. *PNAS* 101: 27: 10132-10136, 2004

Sarikas A, Carrier L, Schenke C, Doll D, Flavigny J, Lindenberg KS, Eschenhagen T and Zolk O. Impairment of the ubiquitin-proteasome system by truncated cardiac myosin binding protein C mutants. *Cardiovasc Res* 66: 33-44, 2005.

Sato N, Kawakami T, Nakayama A, Suzuki H, Kasahara H and Obinata T. A novel variant of cardiac myosin-binding protein-C that is unable to assemble into sarcomeres is expressed in the aged mouse atrium. *Mol Biol Cell* 14: 3180-3191, 2003.

Sato TK, Overduin M and Emr SD. Location, location, location: membrane targeting directed by PX domains. *Science* 294: 1881-1885, 2001.

Scheinberg IH and Sternlieb I. Wilson disease and idiopathic copper toxicosis. *Am J Clin Nutr* 63: 842S-845S, 1996.

Schlossarek S and Carrier L. The ubiquitin-proteasome system in cardiomyopathies. *Curr Opin Cardiol* 26: 190-195, 2011

Schwartz K. Familial hypertrophic cardiomyopathy. Nonsense versus missense mutations. *Circulation* 91: 2865-2867, 1995.

Schwartz K, Carrier L, Guicheney P and Komajda M. Molecular basis of familial cardiomyopathies. *Circulation* 91: 532-540, 1995.

Seaman MN. Recycle your receptors with retromer. *Trends Cell Biol* 15: 68-75, 2005.

Seet LF and Hong W. The Phox (PX) domain proteins and membrane traffic. *Biochim Biophys Acta* 1761: 878-896, 2006.

Semsarian C, Healey MJ, Fatkin D, Giewat M, Duffy C, Seidman CE and Seidman JG. A polymorphic modifier gene alters the hypertrophic response in a murine model of familial hypertrophic cardiomyopathy. *J Mol Cell Cardiol* 33: 2055-2060, 2001.

Shaffer JF, Razumova MV, Tu AY, Regnier M and Harris SP. Myosin S2 is not required for effects of myosin binding protein-C on motility. *FEBS Lett* 581: 1501-1504, 2007.

Shaffer JF, Kensler RW and Harris SP. The myosin-binding protein C motif binds to F-actin in a phosphorylation-sensitive manner. *J Biol Chem* 284: 12318-12327, 2009.

Shaffer JF and Harris SP. Species-specific differences in the Pro-Ala rich region of cardiac myosin binding protein-C. *J Muscle Res Cell Motil* 30: 303-306, 2009.

Shaffer JF, Kensler RW and Harris SP. The myosin-binding protein C motif binds to F-actin in a phosphorylation-sensitive manner. *J Biol Chem* 284: 12318-12327, 2009.

Shaffer JF and Gillis TE. Evolution of the regulatory control of vertebrate striated muscle: the roles of troponin I and myosin binding protein-C. *Physiol Genomics* 42: 406-419, 2010.

Shaffer JF, Wong P, Bezold KL and Harris SP. Functional differences between the N-terminal domains of mouse and human myosin binding protein-C. *J Biomed Biotechnol* 2010: 789798, 2010.

Shi PP, Cao XR, Sweezer EM, Kinney TS, Williams NR, Husted RF, Nair R, Weiss RM, Williamson RA, Sigmund CD, Snyder PM, Staub O, Stokes JB and Yang B. Salt-sensitive hypertension and cardiac hypertrophy in mice deficient in the ubiquitin ligase Nedd4-2. *Am J Physiol Renal Physiol* 295: F462-F470, 2008.

Shirahama-Noda K, Yamamoto A, Sugihara K, Hashimoto N, Asano M, Nishimura M and Hara-Nishimura I. Biosynthetic processing of cathepsins and lysosomal degradation are abolished in asparaginyl endopeptidase-deficient mice. *J Biol Chem* 278: 33194-33199, 2003.

Silverstein RA and Ekwall K. Sin3: a flexible regulator of global gene expression and genome stability. *Curr Genet* 47: 1-17, 2005.

Skaar JR and Pagano M. Control of cell growth by the SCF and APC/C ubiquitin ligases. *Curr Opin Cell Biol* 21: 816-824, 2009.

Slagsvold T, Aasland R, Hirano S, Bache KG, Raiborg C, Trambaiolo D, Wakatsuki S and Stenmark H. Eap45 in mammalian ESCRT-II binds ubiquitin via a phosphoinositide-interacting GLUE domain. *J Biol Chem* 280: 19600-19606, 2005.

Snyder PM. Minireview: regulation of epithelial Na+ channel trafficking. *Endocrinology* 146: 5079-5085, 2005.

Solban N, Jia HP, Richard S, Tremblay S, Devlin AM, Peng J, Gossard F, Guo DF, Morel G, Hamet P, Lewanczuk R and Tremblay J. HCaRG, a novel calcium-regulated gene coding for a nuclear protein, is potentially involved in the regulation of cell proliferation. *J Biol Chem* 275: 32234-32243, 2000.

Solban N, Dumas P, Gossard F, Sun Y, Pravenec M, Kren V, Lewanczuk R, Hamet P and Tremblay J. Chromosomal mapping of HCaRG, a novel hypertension-related, calcium-regulated gene. *Folia Biol (Praha)* 48: 9-14, 2002.

Solomon V, Lecker SH and Goldberg AL. The N-end rule pathway catalyzes a major fraction of the protein degradation in skeletal muscle. *J Biol Chem* 273: 25216-25222, 1998.

Sommerhalter M, Zhang Y and Rosenzweig AC. Solution structure of the COMMD1 N-terminal domain. *J Mol Biol* 365: 715-721, 2007.

Sottile J and Hocking DC. Fibronectin polymerization regulates the composition and stability of extracellular matrix fibrils and cell-matrix adhesions. *Mol Biol Cell* 13: 3546-3559, 2002.

Spee B, Arends B, van Wees AM, Bode P, Penning LC and Rothuizen J. Functional consequences of RNA interference targeting COMMD1 in a canine hepatic cell line in relation to copper toxicosis. *Anim Genet* 38: 168-170, 2007.

Spirito P and Maron BJ. Relation between extent of left ventricular hypertrophy and occurrence of sudden cardiac death in hypertrophic cardiomyopathy. *J Am Coll Cardiol* 15: 1521-1526, 1990.

Spirito P, Seidman CE, McKenna WJ and Maron BJ. The management of hypertrophic cardiomyopathy. *N Engl J Med* 336: 775-785, 1997.

Spirito P, Bellone P, Harris KM, Bernabo P, Bruzzi P and Maron BJ. Magnitude of left ventricular hypertrophy and risk of sudden death in hypertrophic cardiomyopathy. *N Engl J Med* 342: 1778-1785, 2000.

Squire JM, Luther PK and Knupp C. 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: 713-724, 2003.

Starr R and Offer G. The interaction of C-protein with heavy meromyosin and subfragment-2. *Biochem J* 171: 813-816, 1978.

Stelzer JE, Dunning SB and Moss RL. Ablation of cardiac myosin-binding protein-C accelerates stretch activation in murine skinned myocardium. *Circ Res* 98: 1212-1218, 2006.

Stelzer JE, Patel JR, Walker JW and Moss RL. Differential roles of cardiac myosin-binding protein C and cardiac troponin I in the myofibrillar force responses to protein kinase A phosphorylation. *Circ Res* 101: 503-511, 2007.

Strochlic TI, Setty TG, Sitaram A and Burd CG. Grd19/Snx3p functions as a cargo-specific adapter for retromer-dependent endocytic recycling. *J Cell Biol* 177: 115-125, 2007.

Strochlic TI, Schmiedekamp BC, Lee J, Katzmann DJ and Burd CG. Opposing activities of the Snx3-retromer complex and ESCRT proteins mediate regulated cargo sorting at a common endosome. *Mol Biol Cell* 19: 4694-4706, 2008.

Strom CC, Kruhoffer M, Knudsen S, Stensgaard-Hansen F, Jonassen TE, Orntoft TF, Haunso S and Sheikh SP. Identification of a core set of genes that signifies pathways underlying cardiac hypertrophy. *Comp Funct Genomics* 5: 459-470, 2004.

Takeuchi T, Watanabe Y, Takano-Shimizu T and Kondo S. Roles of jumonji and jumonji family genes in chromatin regulation and development. *Dev Dyn* 235: 2449-2459, 2006.

Tan P, Fuchs SY, Chen A, Wu K, Gomez C, Ronai Z and Pan ZQ. Recruitment of a ROC1-CUL1 ubiquitin ligase by Skp1 and HOS to catalyze the ubiquitination of I kappa B alpha. *Mol Cell* 3: 527-533, 1999.

Tanaka T, Inazawa J and Nakamura Y. Molecular cloning of a human cDNA encoding putative cysteine protease (PRSC1) and its chromosome assignment to 14q32.1. *Cytogenet Cell Genet* 74: 120-123, 1996.

Tanaka Y, Guhde G, Suter A, Eskelinen EL, Hartmann D, Lullmann-Rauch R, Janssen PM, Blanz J, von FK and Saftig P. Accumulation of autophagic vacuoles and cardiomyopathy in LAMP-2-deficient mice. *Nature* 406: 902-906, 2000.

Tannous P, Zhu H, Nemchenko A, Berry JM, Johnstone JL, Shelton JM, Miller FJ, Jr., Rothermel BA and Hill JA. Intracellular protein aggregation is a proximal trigger of cardiomyocyte autophagy. *Circulation* 117: 3070-3078, 2008.

Tao TY, Liu F, Klomp L, Wijmenga C and Gitlin JD. The copper toxicosis gene product Murr1 directly interacts with the Wilson disease protein. *J Biol Chem* 278: 41593-41596, 2003.

Tardiff JC, Hewett TE, Palmer BM, Olsson C, Factor SM, Moore RL, Robbins J and Leinwand LA. Cardiac troponin T mutations result in allele-specific phenotypes in a mouse model for hypertrophic cardiomyopathy. *J Clin Invest* 104: 469-481, 1999.

Tasaki T and Kwon YT. The mammalian N-end rule pathway: new insights into its components and physiological roles. *Trends Biochem Sci* 32: 520-528, 2007.

Thierfelder L, Watkins H, MacRae C, Lamas R, McKenna W, Vosberg HP, Seidman JG and Seidman CE. Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell* 77: 701-712, 1994.

Tidball JG and Spencer MJ. Expression of a calpastatin transgene slows muscle wasting and obviates changes in myosin isoform expression during murine muscle disuse. *J Physiol* 545: 819-828, 2002.

Tondeleir D, Vandamme D, Vandekerckhove J, Ampe C and Lambrechts A. Actin isoform expression patterns during mammalian development and in pathology: insights from mouse models. *Cell Motil Cytoskeleton* 66: 798-815, 2009.

Tong CW, Stelzer JE, Greaser ML, Powers PA and Moss RL. Acceleration of crossbridge kinetics by protein kinase A phosphorylation of cardiac myosin binding protein C modulates cardiac function. *Circ Res* 103: 974-982, 2008.

Trinick J. Interaction of titin/connectin with the thick filament. *Adv Biophys* 33: 81-90, 1996.

Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P and Zhang Y. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439: 811-816, 2006.

Tyska MJ, Hayes E, Giewat M, Seidman CE, Seidman JG and Warshaw DM. Single-molecule mechanics of R403Q cardiac myosin isolated from the mouse model of familial hypertrophic cardiomyopathy. *Circ Res* 86: 737-744, 2000.

Uys G, Investigations of the role of myomegalin in the phosphorylation of cardiac myosin binding protein C. *Phd Thesis*, 2010, University of Stellenbosch

van de Sluis BJ, Breen M, Nanji M, van WM, de JP, Binns MM, Pearson PL, Kuipers J, Rothuizen J, Cox DW, Wijmenga C and van Oost BA. Genetic mapping of the copper toxicosis locus in Bedlington terriers to dog chromosome 10, in a region syntenic to human chromosome region 2p13-p16. *Hum Mol Genet* 8: 501-507, 1999.

van de Sluis B, Rothuizen J, Pearson PL, van Oost BA and Wijmenga C. Identification of a new copper metabolism gene by positional cloning in a purebred dog population. *Hum Mol Genet* 11: 165-173, 2002.

van de Sluis B, Muller P, Duran K, Chen A, Groot AJ, Klomp LW, Liu PP and Wijmenga C. Increased activity of hypoxia-inducible factor 1 is associated with early embryonic lethality in Commd1 null mice. *Mol Cell Biol* 27: 4142-4156, 2007.

van de Sluis B, Groot AJ, Wijmenga C, Vooijs M and Klomp LW. COMMD1: a novel protein involved in the proteolysis of proteins. *Cell Cycle* 6: 2091-2098, 2007.

van de Sluis B, Groot AJ, Vermeulen J, van der WE, van Diest PJ, Wijmenga C, Klomp LW and Vooijs M. COMMD1 Promotes pVHL and O2-Independent Proteolysis of HIF-1alpha via HSP90/70. *PLoS One* 4: e7332, 2009.

van der Merwe L, Cloete R, Revera M, Heradien M, Goosen A, Corfield VA, Brink PA and Moolman-Smook JC. Genetic variation in angiotensin-converting enzyme 2 gene is associated with extent of left ventricular hypertrophy in hypertrophic cardiomyopathy. *Hum Genet* 124: 57-61, 2008.

van Dijk SJ, Dooijes D, dos RC, Michels M, Lamers JM, Winegrad S, Schlossarek S, Carrier L, ten Cate FJ, Stienen GJ and van d, V. Cardiac myosin-binding protein C mutations and hypertrophic cardiomyopathy: haploinsufficiency, deranged phosphorylation, and cardiomyocyte dysfunction. *Circulation* 119: 1473-1483, 2009.

Van Driest SL, Vasile VC, Ommen SR, Will ML, Tajik AJ, Gersh BJ and Ackerman MJ. Myosin binding protein C mutations and compound heterozygosity in hypertrophic cardiomyopathy. *J Am Coll Cardiol* 44: 1903-1910, 2004.

Van Driest SL, Gakh O, Ommen SR, Isaya G and Ackerman MJ . Molecular and functional characterization of a human frataxin mutation found in hypertrophic cardiomyopathy. *Mol Genet Metab* 85: 280-285, 2005.

Van Driest SL, Ommen SR, Tajik AJ, Gersh BJ and Ackerman MJ. Sarcomeric genotyping in hypertrophic cardiomyopathy. *Mayo Clin Proc* 80: 463-469, 2005.

Van Criekinge W and Beyaert R. Yeast Two-Hybrid: State of the Art. *Biol Proced Online* 2: 1-38, 1999.

Vandesompele J, De PK, Pattyn F, Poppe B, Van RN, De PA and Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3: RESEARCH0034, 2002.

Verges M. Retromer and sorting nexins in development. *Front Biosci* 12: 3825-3851, 2007.

Verhoef LG, Heinen C, Selivanova A, Halff EF, Salomons FA and Dantuma NP. Minimal length requirement for proteasomal degradation of ubiquitin-dependent substrates. *FASEB J* 23: 123-133, 2009.

Vervoort VS, Viljoen D, Smart R, Suthers G, DuPont BR, Abbott A and Schwartz CE. Sorting nexin 3 (SNX3) is disrupted in a patient with a translocation t(6;13)(q21;q12) and microcephaly, microphthalmia, ectrodactyly, prognathism (MMEP) phenotype. *J Med Genet* 39: 893-899, 2002.

Vidalain PO, Boxem M, Ge H, Li S and Vidal M. Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods* 32: 363-370, 2004.

Vignier N, Schlossarek S, Fraysse B, Mearini G, Kramer E, Pointu H, Mougenot N, Guiard J, Reimer R, Hohenberg H, Schwartz K, Vernet M, Eschenhagen T and Carrier L. Nonsense-mediated mRNA decay and ubiquitin-proteasome system regulate cardiac myosin-binding protein C mutant levels in cardiomyopathic mice. *Circ Res* 105: 239-248, 2009.

Voss TC, Demarco IA and Day RN. Quantitative imaging of protein interactions in the cell nucleus. *Biotechniques* 38: 413-424, 2005.

Wang GL, Jiang BH, Rue EA and Semenza GL. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc Natl Acad Sci U S A* 92: 5510-5514, 1995.

Wang KH, Roman-Hernandez G, Grant RA, Sauer RT and Baker TA. The molecular basis of N-end rule recognition. *Mol Cell* 32: 406-414, 2008.

Watkins H, Seidman CE, MacRae C, Seidman JG and McKenna W. Progress in familial hypertrophic cardiomyopathy: molecular genetic analyses in the original family studied by Teare. *Br Heart J* 67: 34-38, 1992.

Watkins H. Multiple disease genes cause hypertrophic cardiomyopathy. *Br Heart J* 72: S4-S9, 1994.

Watkins H, Conner D, Thierfelder L, Jarcho JA, MacRae C, McKenna WJ, Maron BJ, Seidman JG and Seidman CE. Mutations in the cardiac myosin binding protein-C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. *Nat Genet* 11: 434-437, 1995.

Weber FE, Vaughan KT, Reinach FC and Fischman DA. 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: 661-669, 1993.

Weekes J, Morrison K, Mullen A, Wait R, Barton P and Dunn MJ. Hyperubiquitination of proteins in dilated cardiomyopathy. *Proteomics* 3: 208-216, 2003.

Weisberg A and Winegrad S. 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, 1996.

Weisberg A and Winegrad S. Relation between crossbridge structure and actomyosin ATPase activity in rat heart. *Circ Res* 83: 60-72, 1998.

Welikson RE and Fischman DA. The C-terminal IgI domains of myosin-binding proteins C and H (MyBP-C and MyBP-H) are both necessary and sufficient for the intracellular crosslinking of sarcomeric myosin in transfected non-muscle cells. *J Cell Sci* 115: 3517-3526, 2002.

Westhoff B, Chapple JP, van der SJ, Hohfeld J and Cheetham ME. HSJ1 is a neuronal shuttling factor for the sorting of chaperone clients to the proteasome. *Curr Biol* 15: 1058-1064, 2005.

Whitten AE, Jeffries CM, Harris SP and Trewhella J. Cardiac myosin-binding protein C decorates F-actin: implications for cardiac function. *Proc Natl Acad Sci U S A* 105: 18360-18365, 2008.

Willems AR, Schwab M and Tyers M. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim Biophys Acta* 1695: 133-170, 2004.

Willis MS, Ike C, Li L, Wang DZ, Glass DJ and Patterson C. Muscle ring finger 1, but not muscle ring finger 2, regulates cardiac hypertrophy in vivo. *Circ Res* 100: 456-459, 2007.

Willis MS, Schisler JC and Patterson C. Appetite for destruction: E3 ubiquitin-ligase protection in cardiac disease. *Future Cardiol* 4: 65-75, 2008.

Willis MS, Schisler JC, Portbury AL and Patterson C. Build it up-Tear it down: protein quality control in the cardiac sarcomere. *Cardiovasc Res* 81: 439-448, 2009.

Willis MS, Townley-Tilson WH, Kang EY, Homeister JW and Patterson C. Sent to destroy: the ubiquitin proteasome system regulates cell signaling and protein quality control in cardiovascular development and disease. *Circ Res* 106: 463-478, 2010.

Winegrad S. Cardiac myosin binding protein C. *Circ Res* 84: 1117-1126, 1999.

Witt CC, Gerull B, Davies MJ, Centner T, Linke WA and Thierfelder L. Hypercontractile properties of cardiac muscle fibers in a knock-in mouse model of cardiac myosin-binding protein-C. *J Biol Chem* 276: 5353-5359, 2001.

Witt SH, Granzier H, Witt CC and Labeit S. MURF-1 and MURF-2 target a specific subset of myofibrillar proteins redundantly: towards understanding MURF-dependent muscle ubiquitination. *J Mol Biol* 350: 713-722, 2005.

Woelk T, Sigismund S, Penengo L and Polo S. The ubiquitination code: a signalling problem. *Cell Div* 2: 11, 2007.

Worby CA and Dixon JE. Sorting out the cellular functions of sorting nexins. *Nat Rev Mol Cell Biol* 3: 919-931, 2002.

Xu Q, Dewey S, Nguyen S and Gomes AV. Malignant and benign mutations in familial cardiomyopathies: insights into mutations linked to complex cardiovascular phenotypes. *J Mol Cell Cardiol* 48: 899-909, 2010.

Xu Y, Hortsman H, Seet L, Wong SH and Hong W. SNX3 regulates endosomal function through its PX-domain-mediated interaction with PtdIns(3)P. *Nat Cell Biol* 3: 658-666, 2001.

Yamamoto K and Moos C. The C-proteins of rabbit red, white, and cardiac muscles. *J Biol Chem* 258: 8395-8401, 1983.

Yamane T, Takeuchi K, Yamamoto Y, Li YH, Fujiwara M, Nishi K, Takahashi S and Ohkubo I. Legumain from bovine kidney: its purification, molecular cloning, immunohistochemical localization and degradation of annexin II and vitamin D-binding protein. *Biochim Biophys Acta* 1596: 108-120, 2002.

Yamasaki R, Berri M, Wu Y, Trombitas K, McNabb M, Kellermayer MS, Witt C, Labeit D, Labeit S, Greaser M and Granzier H. Titin-actin interaction in mouse myocardium: passive tension modulation and its regulation by calcium/S100A1. *Biophys J* 81: 2297-2313, 2001.

Yamasaki R, Wu Y, McNabb M, Greaser M, Labeit S and Granzier H. Protein kinase A phosphorylates titin's cardiac-specific N2B domain and reduces passive tension in rat cardiac myocytes. *Circ Res* 90: 1181-1188, 2002.

Yang Q, Sanbe A, Osinska H, Hewett TE, Klevitsky R and Robbins J. In vivo modeling of myosin binding protein C familial hypertrophic cardiomyopathy. *Circ Res* 85: 841-847, 1999.

Yang Q, Hewett TE, Klevitsky R, Sanbe A, Wang X and Robbins J. PKA-dependent phosphorylation of cardiac myosin binding protein C in transgenic mice. *Cardiovasc Res* 51: 80-88, 2001.

Yang Z, McMahon CJ, Smith LR, Bersola J, Adesina AM, Breinholt JP, Kearney DL, Dreyer WJ, Denfield SW, Price JF, Grenier M, Kertesz NJ, Clunie SK, Fernbach SD, Southern JF, Berger S, Towbin JA, Bowles KR and Bowles NE. Danon disease as an underrecognized cause of hypertrophic cardiomyopathy in children. *Circulation* 112: 1612-1617, 2005.

Zhang J, Ping P, Vondriska TM, Tang XL, Wang GW, Cardwell EM and Bolli R. Cardioprotection involves activation of NF-kappa B via PKC-dependent tyrosine and serine phosphorylation of I kappa B-alpha. *Am J Physiol Heart Circ Physiol* 285: H1753-H1758, 2003.

Zheng J, Yang X, Harrell JM, Ryzhikov S, Shim EH, Lykke-Andersen K, Wei N, Sun H, Kobayashi R and Zhang H. CAND1 binds to unneddylated CUL1 and regulates the formation of SCF ubiquitin E3 ligase complex. *Mol Cell* 10: 1519-1526, 2002.

Zheng Q, Li J and Wang X. Interplay between the ubiquitin-proteasome system and autophagy in proteinopathies. *Int J Physiol Pathophysiol Pharmacol* 1: 127-142, 2009.

Zhou M and Veenstra TD. Proteomic analysis of protein complexes. *Proteomics* 7: 2688-2697, 2007.

Zhu LA, Fang NY, Gao PJ, Jin X and Wang HY. Differential expression of alpha-enolase in the normal and pathological cardiac growth. *Exp Mol Pathol* 87: 27-31, 2009.

Zhu X, Bouzekri N, Southam L, Cooper RS, Adeyemo A, McKenzie CA, Luke A, Chen G, Elston RC and Ward R. Linkage and association analysis of angiotensin I-converting enzyme

(ACE)-gene polymorphisms with ACE concentration and blood pressure. *Am J Hum Genet* 68: 1139-1148, 2001.

Zoghbi ME, Woodhead JL, Moss RL and Craig R. Three-dimensional structure of vertebrate cardiac muscle myosin filaments. *Proc Natl Acad Sci U S A* 105: 2386-2390, 2008.

Zolk O, Schenke C and Sarikas A. The ubiquitin-proteasome system: focus on the heart. *Cardiovasc Res* 70: 410-421, 2006.

Zou Y, Evans S, Chen J, Kuo HC, Harvey RP and Chien KR. CARP, a cardiac ankyrin repeat protein, is downstream in the Nkx2-5 homeobox gene pathway. *Development* 124: 793-804, 1997.