Protein quality control systems in muscle: From disease to mechanism

Doctoral thesis for obtaining the

Academic degree Doctor of Natural Science

(Dr.rer.nat)

Submitted by

Marija (Milan) Marković

at the

Faculty of Science

Department of Biology

Konstanz, 2020

Date of the oral examination: 13.11.2020.

1. Reviewer: Prof. Dr. Olga Mayans

2. Reviewer: Prof. Dr. Kay Diederichs

Science is the rationalization of your gut feeling. Without the gut feeling, there is nothing to rationalize and without rationalization, it is just a gut feeling.

To my beloved mother, who showed me how not to give up.

Acknowledgments

If there is anything I have learned from peeking into the micro world, it is that every single process that occurs is a beautiful network tailored by nature that requires molecular community. As in the cell, biology doesn't really change when it comes to the macro world. This work is my example of that fundamental law of nature. I am grateful to my supervisors, Professor Olga Mayans and Professor Kay Diederich, who both decided to give me an opportunity for my scientific development at the University of Konstanz and Professor Michael Kovermann who supported us on this journey. Special thanks to Olga, for giving me the chance to be a part of the MuRF1 adventure, for giving me the freedom to fold spontaneously and provided me the chaperons when I needed them the most. I wish to thank Professor Hans-Jürgen Apell for being support, for sharing scientific and philosophical discussions when were much needed. My deepest thanks to Dr. Barbara Franke, who has been my lab mentor and my MuRF1 partner. I thank Dr. Jennifer Flaming for being an amazing teacher and sharing her tricks with me. But more than that, thank you both for being great friends, for sharing your fearlessness and energy. I wish also to thank all past and current members of AG Mayans and AG Diederichts for being a part of my journey and also to my students who helped me to learn more. I am grateful for having a Konstanz family during the past four years. Thank you, Chris, Bruno, Eva, Regina, Michael, Chiara, Zhou and sweetest thank to Erik. Writing this thesis during the pandemic and self-isolation was a very challenging task. I wish to thank my life partner, Davide for sharing this challenge with me and keeping me sane. More than anything for showing me that life can be beautiful also after the greatest loss. I am also grateful for my friends who accept and love me as I am and who gave me a support from all the distances around the globe, for that thank you Jelica, Danilo, Vuk, Bojana, Stanislava, Jelena, Silvia, Valentina, Miša, Filip and Djordje. My family created beautiful conditions for a child to dream, to grow and to become whatever it wants. This child chose to become a scientist. To my brother Nikola and to Tamara, for your unconditional love and support. To my father, who made the world better place to live in and who helped me to get up every time. I know you will always be walking with me and I am infinitely grateful for being your daughter. To the endless inspiration, the strongest women and the kindness human I have a privilege to call my mother. *Hvala ti, mama, što si odabrala nas i što nas biraš svakog dana. Zbog tebe sam ponosna I tebi posvećujem ovaj rad. Za naš četverac!*

Table of contents

List of abbreviations

ACD	Assembly competence domain
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
Approx.	Approximately
ATP	Adenosine triphosphate
BAG	Bcl-2 associated athanogene
BHD	B-box coiled-coil
BLI	Biolayer interferometry
CARP	Cardiac ankyrin repeat protein
CC	Coiled-coil
COS	Cos-box domain
CSA	Cross-section area
DCM	Dilated cardiomyopathy
dKO	Double knockout
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E1	E1-ubiquitin activating enzyme
E2	E2-ubiquitin conjugating enzyme
E3	E3-ubiquitin ligase
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
GST	Glutathione S-transferases
HD	Helical domain
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
IL-1	Cytokine interleukin
IMAC	Immobilized metal affinity chromatography
IPTG	isopropyl-b-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KO	Knockdown
LB	Luria-Bertani
LC-MS/MS	Liquid chromatography–mass spectrometry
MA	Muscle atrophy

MABA-ADP	N8-(4-N'- methylanthraniloylaminobutyl)-8-aminoadenosine 5'-Diphosphate
MAFbx	Muscle atrophy F box
MBP	Maltose-binding protein
MD	Molecular dynamics
MES	2-(N-morpholino)ethanesulfonic acid
MFC	MuRF conserved region
MOPS	2-Morpholino-propanesulphonic acid
MR	Molecular replacement
MuRF	Muscle specific RING finger protein
MYH7	Myosin II isoform
MyHC	Myosin heavy chain
MyLC	Myosin light chain
NBD	Nucleotide binding domain
NEF	Nucleotide exchange factor
PBS	Phosphate-buffered saline
PDB	Protein data bank
PEG	Polyethylene glycol
PQCS	Protein quality control system
RBCC	RING B-box coiled-coil
RING	Really interesting new gene
RMB	RING MFC B-box
SBD	Substrate binding domain
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TAMRA	Tetramethylrhodamine
TEV	Tobacco etch virus
TK	Titin kinase
TRIM	Tripartite motif
Tris	Tris-(hydroxymethyl)-aminomethane
Ub	Ubiquitin
UPS	Ubiquitin- proteasome system
WT	Wildtype
YTH	Yeast two-hybrid

List of figures

List of tables

Abstract

The myofibril is among the most organized macromolecular structures in eukaryotes and it is maintained by intricately balanced protein quality control (PQCs) pathways of chaperone mediated protein folding and ubiquitin-proteasome degradation (Kim, Löwe and Hoppe, 2008). Although the main players of the PQCs are well studied, the exact mechanisms by which they recognize clients and control processes remains obscure. The presented research focuses on two distinct muscle disease models - muscle atrophy and dilated cardiomyopathy - in order to reveal the underlying mechanisms that control Muscle RING Finger protein 1 (MuRF1) dependent ubiquitination and Nucleotide Exchange Factor (NEF) BAG3 mediated protein refolding.

MuRF1 is a RING E3 ubiquitin ligase that was found to target sarcomere proteins such as myosin and titin for degradation during the muscle atrophy (Bodine *et al.*, 2007; Clarke *et al.*, 2007). To investigate how MuRF1 recruits sarcomere protein substrates, BioLayer interferometry experiments (BLI) were combined with an *in vitro* ubiquitination assays. Ubiquitination assays uncovered MuRF1 specific sites in myosin heavy chain, titin and CARP, suggesting the absence of sequence or motif specificity in the substrate targeting. The BLI data, supported with co-immunoprecipitation and *in cell* colocalization experiments, surprisingly revealed that the RING domain of MuRF1, known to mediate ubiquitin transfer, is also involved in substrate recruiting. To investigate MuRF1's *in situ* contribution to the filament depletion, *in vitro* MuRF1 ubiquitination assays were adapted to a single muscle fiber experimental setup. Resulting experiments showed a decreased muscle force production as a direct consequence of the MuRF1 activity. Together these results provide the new evidence towards understanding a mechanism by which MuRF1 regulates muscle trophicity.

The molecular co-chaperone BAG3 enhances the nucleotide exchange in HSP70 by interacting with the nucleotide binding domain (NBD) (Rauch, Zuiderweg and Gestwicki, 2016). The inheritable point mutation E455K in the BAG domain of BAG3 causes a fatal dilated cardiomyopathy. To investigate how the E455 residue mediates interaction of BAG3 with HSP70, NBD in complex with the BAG domain of BAG3 in the wild type and mutant (E455K) forms were elucidated to 2.3 and 2.2 Å, respectively. The obtained structures showed that the overall complex structure is not affected by the mutation. Therefore, in order to assess the role of the BAG and its mutation in the conformational freedom of HSP70, these crystal structures were subjected to the molecular dynamic simulations. The results suggest

that BAG binding restricts HSP70-NBD conformational dynamics in a narrow range required for an efficient nucleotide exchange, while the mutation alters this range putting the HSP70 in an arrested state. This hypothesis was further supported by the kinetic experiments showing that the E455K mutation has a drastic influence on the dissociation rate of the BAG3/HSP70 complex in the presence of nucleotides, providing a new insight into the modulatory mechanism exerted by BAG co-chaperones.

Zusammenfassung

Die Myofibrille gehört zu den am besten organisierten makromolekularen Strukturen in Eukaryonten und wird durch kompliziert ausbalancierte Proteinqualitätskontrollpfade (PQCs) der Chaperon-vermittelten Proteinfaltung und Ubiquitin-Proteasom-Degradation aufrechterhalten (Kim, Löwe and Hoppe, 2008). Obwohl die Hauptakteure der PQCs gut erforscht sind, bleiben die genauen Mechanismen unklar, durch die sie zubehandelnde Proteine erkennen und Abläufe steuern. Die vorgelegte Arbeit konzentriert sich auf zwei verschiedene Muskelkrankheitsmodelle, Muskelatrophie und dilatative Kardiomyopathie, um die zugrundeliegenden Mechanismen aufzudecken, die die von Muskel-RING-Finger-Protein 1 (MuRF1) abhängige Ubiquitinierung und die durch den Nukleotid-Austauschfaktor (NEF) BAG3 vermittelte Proteinrückfaltung kontrollieren.

MuRF1 ist eine RING E3-Ubiquitin-Ligase, die auf Sarkomer-Proteine wie Myosin und Titin zum Abbau während der Muskelatrophie abzielt (Bodine *et al.*, 2007; Clarke *et al.*, 2007)Um zu untersuchen, wie MuRF1 Sarkomer-Protein-Substrate rekrutiert, wurden BioLayer-Interferometrie-Experimente (BLI) mit In-vitro-Ubiquitinations-Assays kombiniert. Ubiquitinations-Assays deckten MuRF1-spezifische Stellen in der schweren Kette von Myosin, Titin und CARP auf, was auf das Fehlen einer Sequenz- oder Motivspezifität beim Substrat-Targeting schließen lässt. Die BLI-Daten, unterstützt durch Koimmunopräzipitation und Kolokalisationsexperimenten in Zellen, zeigten überraschenderweise, dass die RING-Domäne von MuRF1, von der bekannt ist, dass sie den Ubiquitin-Transfer vermittelt, auch an der Substratrekrutierung beteiligt ist. Um den in-situ-Beitrag von MuRF1 zur Filamentdepletion zu untersuchen, wurden in-vitro-MuRF1-Ubiquitinations-Assays an einen Versuchsaufbau mit einer einzigen Muskelfaser angepasst. Die die damit durchgeführten Experimente zeigten eine verminderte Muskelkrafterzeugung als direkte Folge der MuRF1-Aktivität. Zusammen liefern diese Ergebnisse neue Hinweise auf einen Mechanismus, bei dem MuRF1 die Muskeltrophizität reguliert.

Das molekulare Co-Chaperon BAG3 verstärkt den Nukleotidaustausch in HSP70, indem es mit der Nukleotidbindungsdomäne (NBD) interagiert (Rauch, Zuiderweg and Gestwicki, 2016). Die vererbbare Punktmutation E455K in der BAG-Domäne von BAG3 verursacht eine tödliche dilatative Kardiomyopathie. Um zu untersuchen, wie der E455-Rest die Interaktion von BAG3 mit HSP70 verändert, wurde die Struktur der NBD im Komplex mit der BAG-Domäne von BAG3 in Wildtyp- und Mutantenform (E455K) mit 2,3 bzw. 2,2 Å

aufgeklärt. Die erhaltenen Strukturen zeigten, dass die komplexe Gesamtstruktur durch die Mutation nicht verändert wird. Um die Rolle von BAG und seiner Mutation für die Konformationsdynamik von HSP70 zu beurteilen, wurden diese Kristallstrukturen daher molekulardynamischen Simulationen unterzogen. Die Ergebnisse deuten darauf hin, dass die BAG-Bindung die Konformationsdynamik von HSP70-NBD auf einen engen Bereich einschränkt, der für einen effizienten Nukleotidaustausch erforderlich ist, während die Mutation diesen Bereich verändert und HSP70 in einem Zustandfesthält. Diese Hypothese wurde durch kinetische Experimente weiter untermauert, die zeigen, dass die E455K-Mutation einen drastischen Einfluss auf die Dissoziationsrate des BAG3/HSP70-Komplexes in Gegenwart von Nukleotiden hat, was einen neuen Einblick in den modulatorischen Mechanismus gibt, der von BAG-Co-Chaperonen ausgeübt wird.

General introduction

In the course of life, efficient force production, sarcomere organization and maintenance of structural and motor proteins results from a fine balance between muscle protein synthesis and muscle protein breakdown. Stress caused adaptation of myofibril leads to an unbalance in protein production and protein degradation processes, culminating as muscle hypertrophy (increase in total protein amount) or muscle atrophy (decrease of a total protein amount).

During the cell cycle, stress conditions, metabolic challenges (cancer), ageing or mutations lead to protein destabilization, misfolding or unfolding (Hartl and Hayer-Hartl, 2002). Misfolded/unfolded proteins lose their primary functions and activities and often show a tendency to aggregate, which is toxic to the cell (Dobson, 2003). Accumulation of an unfolded and aggregated proteins in brain tissue is the main cause of Hamilton disease, Parkinson disease and Alzheimer's, while in the muscle tissue aggregated proteins cause hypertrophies and cardiomyopathies. As a consequence, the cell has developed an elaborate network to monitor and maintain protein homeostasis in the cell- a proteostasis (Frydman, 2001; Hartl and Hayer-Hartl, 2002). The proteostasis is attained through the protein quality control system (PQCS) that involves either refolding, degrading or sequestering misfolded protein.

The PQCS is a network purposed to maintain the proteostasis in the cell via three main pathways: refolding of misfolded proteins is achieved via chaperon activity pathway, degradation of misfolded proteins via the ubiquitin-proteasome pathway or via autophagy (Figure 1). As the interplay between these PQCS pathways is key in the regulation of quality control, any dysfunction in the PQCS could have a catastrophic effect on the cell homeostasis.

The majority of soluble misfolded proteins are removed through the ubiquitin-proteasome system (UPS), the major eukaryotic proteolytic pathway (Ciechanover, 1998). An enzymatic E1/E2/E3 ubiquitination cascade tightly controls degradation by labelling proteins with ubiquitin for degradation by the 26S proteasome. In the case of misfolded proteins being insoluble and/or aggregated, autophagy becomes a major degradation pathway (Kroemer, Mariño and Levine, 2010). Dysfunction of UPS in the form of increased activity leads to fast degradation of not only damaged, but also healthy proteins, turning the protein quality control system into an annihilator. For instance, in the numerous forms of muscle atrophy, a common muscle-specific E3 ubiquitin ligase MuRF1 was found to be the main contributor to enhanced degradation of muscle sarcomere components causing the impaired contractility function (Bodine and Baehr, 2014). In contrast to UPS and autophagy, some misfolded proteins can actually be rescued from degradation and subjected to refolding. A central role in rescue of misfolded proteins have molecular chaperons (Mayer and Bukau, 2005). Among the molecular chaperons, HSP70 has a predominant role, and along with molecular co-chaperons, regulates the majority of protein refolding in the cell (Bukau, Weissman and Horwich, 2006). Malfunction of HSP70 machinery leads to accumulation of misfolded proteins, which are not common targets of UPS or autophagy, hence they remain trapped in the cell. Such a scenario in the heart muscle damages the contractility and efficiency of a heart, resulting in dilated cardiomyopathy (DC).

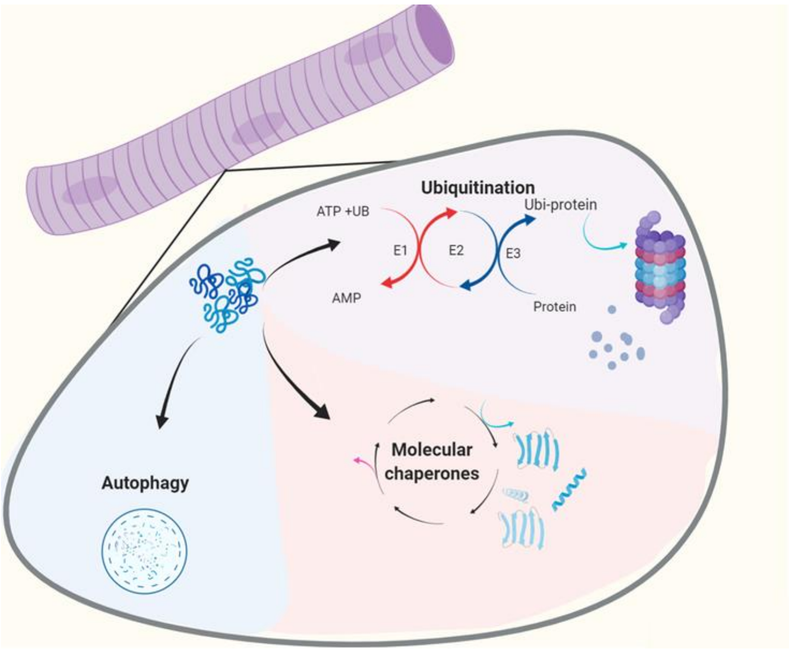


Figure 1. Schematic representation of the protein quality control system in the muscle cell.
Misfolded proteins are removed from the cell via the ubiquitin-dependent proteasome (ubiquitination) or by autophagy. Protein refolding is mediated by molecular chaperones.

Interestingly, there has been evidence in the skeletal muscle of a cross-talk between chaperones, ubiquitination and autophagy, where the ubiquitin-dependent degradation of a chaperone substrate is degraded by the lysosome. Protein quality control systems as an intricate network, rather than simply three distinct pathways, where it is essential to fully understand each of the players and their connections in order to develop methods to modulate their activity in pathological conditions and rescue cellular integrity.

1

Structural and functional characterization of muscle-specific E3 ubiquitin ligase MuRF1

1.1 Introduction

The rapid loss of muscle mass during muscle atrophy (MA) is due to elevated protein degradation mediated by the ubiquitin-proteasome pathway (Solomon and Goldberg, 1996; Lecker, Goldberg and Mitch, 2006). Muscle specific RING finger protein 1 (MuRF1), has been identified as the critical E3 ubiquitin ligase highly upregulated in numerous MA triggering conditions including denervation, immobilization, hind limb suspension, treatment with the cachectic cytokine interleukin-1 (IL-1) and treatment with the glucocorticoid dexamethasone (Bodine *et al.*, 2001). Additional yeast two-hybrid screens using heart cDNA libraries and MuRF1 as bait, revealed two RING finger proteins with high sequence homology to MuRF1- MuRF2 (62%) and MuRF3 (77%). All three MuRFs were found to localize on the M-line of the sarcomere, whilst MuRF1 and MuRF3 were found to localize also on Z-disc (Centner *et al.*, 2001; McElhinny *et al.*, 2002; Franke *et al.*, 2014). In addition, MuRF1 and MuRF2 were found in the nucleus (Lange *et al.*, 2005). Although partially sharing their localization, MuRFs are considered to have different roles in muscle remodeling. MuRF2 and MuRF3 are suggested to be important for microtubule stability (Spencer *et al.*, 2000; McElhinny *et al.*, 2004), whereas MuRF1 is the only family member associated with muscle atrophy and protein degradation.

Knockdown of MuRF1 in C2C12 myotubes results in the attenuation of myofibril protein loss (Clarke *et al.*, 2007) and deletion of MuRF1 in mice has been shown as effective in functional sparing of muscle mass following denervation (Moraes *et al.*, 2015) hindlimb unloading (Labeit *et al.*, 2010), acute lung injury (Files *et al.*, 2012) (Files *et al.*, 2012) and dexamethasone treatment (Castillero *et al.*, 2013) suggesting that selective inhibition of

MuRF1 could be an effective strategy in MA treatment. Comparing the atrophy response in wild type mice and mice expressing the RING-less MuRF1 Cohen et al. (Cohen *et al.*, 2009) concluded that the ubiquitination and degradation of myosin heavy chain (MyHC) is MuRF1 dependent, together with the loss of myosin-binding protein C (MyBP-C) and myosin light chains 1 and 2 (MyLC1 and MyLC2), but not actin. In addition, it was suggested that the RING domain is essential for functional MuRF1 targeting.

In yeast two-hybrid screens and immunoprecipitation, numerous sarcomeric proteins beside MyHC were proposed as MuRF1 interacting partners (Witt *et al.*, 2005), only few however, have been verified *in vitro* techniques. In *in vitro* ubiquitination assays, MuRF1 ubiquitinates MyHC (Marblestone *et al.*, 2012), the titin A band fraction (Bogomolovas *et al.*, 2014), and troponinI (Kedar *et al.*, 2004).

MuRF1 is not only associated with different myofibril compartments, but also with a network of diverse proteins related to energy metabolisms (creatine kinase, mitochondria oxidative phosphorylation, and ATP regeneration) (Witt *et al.*, 2005; Koyama *et al.*, 2008), transcriptional regulation (McElhinny *et al.*, 2002) signaling, particularly through its binding of titin kinase at the sarcomeric M-line and potential role in SUMO mediated pathways (Nowak *et al.*, 2019).

1.1.1 Ubiquitination

The ubiquitin proteasome degradation system (UPS) has a leading role in regulating cellular homeostasis of eukaryotes (Ciechanover, 1998; Hershko and Ciechanover, 1998). UPS is dominantly driven by utilization of ubiquitin tag (Ub). Ubiquitin is a small, highly conserved protein of 76 amino acids heavily abundant in the cell. Covalent attachment of ubiquitin to substrate proteins creates a signal to the cell and has profound effects on protein stability, localization and interaction (Kim, Löwe and Hoppe, 2008). Ubiquitin targeting of a single protein can then globally alter the fate of the cell cycle, proliferation and apoptosis. The complex fate of a targeted protein and the cell cycle by a simple ubiquitin molecule is determined by a unique signature created by an ubiquitination cascade and recognised by the cell regulatory mechanism. Ubiquitination, the most dominant posttranslational modification of proteins, is a highly orchestrated, multi enzyme reaction resulting in the creation of an isopeptide bond between ubiquitin's C-terminal carboxyl group and typically a substrate's Lysine ε-NH2 group, in an ATP dependent manner. Non-canonical ubiquitination has also been observed in mammalian cells as ubiquitin ligation to the N-terminal α-NH2 group, or to a substrate's serine, threonine or cysteine residues forming peptide, ester or thioester linkages, respectively (McDowell and Philpott, 2013).

The ubiquitination enzymatic cascade involves three enzymes with distinct biochemical activities: E1 (Ub-activating), E2 (Ub-conjugating) and E3 (Ub-ligating) (Hershko and Ciechanover, 1998). Firstly, ubiquitin is activated via an ATP dependent mechanism in a reaction with an E1, where ATP is hydrolysed to AMP and pyrophosphate, and ubiquitin is activated through ubiquitin C-terminal glycine residue (G76). Activated ubiquitin is next transferred to a sulfhydryl group on the ubiquitin-conjugating enzymes (E2) via a transthioesterification reaction (Figure 1.1). The final step of the ubi-cascade involves E3 mediated conjugation of ubiquitin to the substrate (Scheffner, Nuber and Huibregtse, 1995).

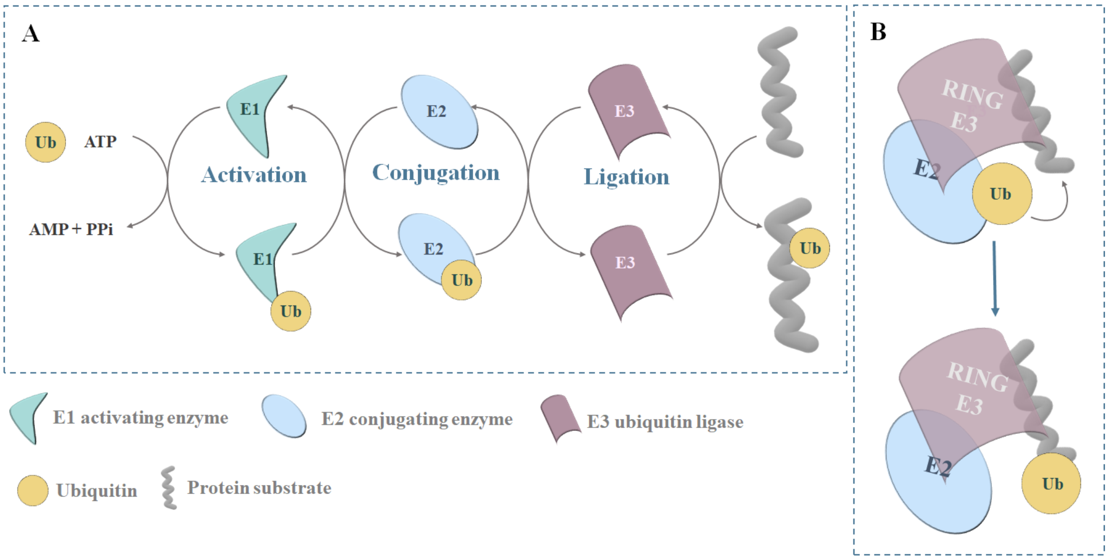


Figure 1.1 Schematic representation of the ubiquitination cascade.
A The initial step in the ubiquitination cascade is the ATP-dependent activation of ubiquitin (Ub) by the ubiquitin-activating enzyme (E1). Ubiquitin is next transferred to the ubiquitin-conjugating enzymes (E2) via a transthioesterification reaction (Conjugation). Ubiquitin ligase (E3) catalyzes the transfer of ubiquitin to the substrate protein (Ligation). **B** RING type E3 ligase brings E2-Ub complex in proximity of protein substrate for Ub transfer to occur, without direct interaction with ubiquitin.

Ubiquitin attachment to another ubiquitin molecule via one of seven Lysine residues present in the molecule (K6, K11, K27, K29, K33, K48 and K63) is named polyubiquitination. Accordingly, ubiquitination substrates can be subjected to monoubiquitination if one lysine is modified by a single ubiquitin; poly-monoubiquitinated if multiple lysines are monoubiquitinated or polyubiquitinated if one or more lysines are modified via polyubiquitin chains (Oh, Akopian and Rape, 2018). The type of ubiquitin chain formed is a unique recognition signature for numerous cell processes and is mostly defined by the E2. Actually, in the human body, there are only a few known E1s and more than 50 E2 genes (Dikic, 2017). Thus, Ub tags can differ in their polyubiquitin chain structure and dictate the fate of modified substrates, as well as recruiting Ub-binding domain receptors, including transport factors that deliver substrates to the proteasome. For instance, polyubiquitination through K^{48} linkage and branched K^{48}-K^{11} chains are the most dominant signals for proteasome degradation (Passmore and Barford, 2004). Monoubiquitination can serve in protein localization alteration and, as recent systematic analyses have revealed, monoubiquitination can be involved in up to 40% of proteins degradation by the proteasome (Braten *et al.*, 2016). This type of diversity in E2 activities illustrates a complexity of signaling pathways in the cell, but is not associated with substrate recognition or specificity.

Hence, cells introduced a third partner in the ubiquitination cascade, ubiquitin-ligases or E3s. E3s specifically recognize and bind to the potential ubiquitination substrates.

Simultaneously, E3s bind the E2-ubiquitin complex, mediating the ubiquitin conjugation to the substrate. There are hundreds of genes in humans coding for E3 ubiquitin ligases, facilitating substrate specificity. E3 ubiquitin ligases are classified based on the mechanism of ubiquitin conjugation in two main types: HECT - and RING-type. In the HECT (Homology to E6-AP Carboxy Terminus) type of E3s ubiquitin is transferred first to the E3 via a transthioesterification reaction and then transferred from the E3 to the substrate (Scheffner, Nuber and Huibregtse, 1995). The other type of mechanism does not involve ubiquitin transfer to E3, but a direct transfer from E2 to the substrate, where E3 has a role in bringing the E2-ubiquitin complex in proximity of the protein substrate (Figure 1.1 B). These enzymes belong to the RING (Really Interesting New Gene) or Ubox group of E3 ligases. The third type, a hybrid mechanism has been discovered, in which E3s interact with the E2-ubiquitin in a similar manner to the RING/Ubox type, but ubiquitin transfer involves cysteinyl residue prior conjugation to the substrate, similar to a HECT-type mechanism (Stone *et al.*, 2005). Regardless of the mechanism, E3 ubiquitin ligases show substrate and tissue specificity.

Within the muscle specific E3 ubiquitin ligases, MuRF1 and MAFbx (atrogin 1) have become known as the critical factors in sarcomere remodeling during the state of muscle atrophy. In a variety of muscle atrophy induced conditions, MuRF1 and MAFbx genes were found to be highly upregulated and knocking out these genes in the mouse models has shown muscle sparing under atrophy models (Labeit *et al.*, 2010; Bodine and Baehr, 2014; Nguyen *et al.*, 2020). MuRF1- and MAFbx-KO mice experiments referred to the MuRF1 deletion being more effective in muscle sparing than the deletion of MAFbx. Moreover, the mass that is preserved in MuRF1-KO's has a functional force output. Together these results suggest that MuRF1 would be a more suitable candidate for drug development in muscle atrophy treatment (Eddins *et al.*, 2011; Bodine and Baehr, 2014). Hereby, this work further focuses on MuRF1 in muscle protein degradation.

1.1.2 MuRFs and TRIM/RBCC fold in substrate targeting and catalysis

MuRFs belong to the TRIpartite Motif (TRIM) superfamily of proteins, which consists of E2-conjugating enzymes and E3-ubiquitin ligases, which are main contributors to the major biological processes, such as transcription, differentiation, cell cycle regulation, muscular physiology and innate immune response (Meroni, 2020) TRIMs (including MuRFs) share N-terminal tripartite fold (RBCC) consisting of a RING finger (R), one or two RING finger-like B-box domains (B), and a helical domain (HD) that forms a coiled-coil-like motif. The individuality among TRIM family members is assigned by their C-terminal fraction (*e.g.* SPRY or B30.2, PHD, PRY-SPRY), which provides targeting functions to these proteins and define family classes (Figure 1.2). Unlike the majority of TRIM classes (nine out of eleven), MuRFs lack the C-terminal domain and consist of a RING domain, B-box2 and coiled-coil domain (Coiled-coil), but also share MuRF family conserved region (MFC), COS-box motif (COS) and acidic tail (Reymond *et al.*, 2001; Mayans and Labeit, 2012) (Figure 1.2).

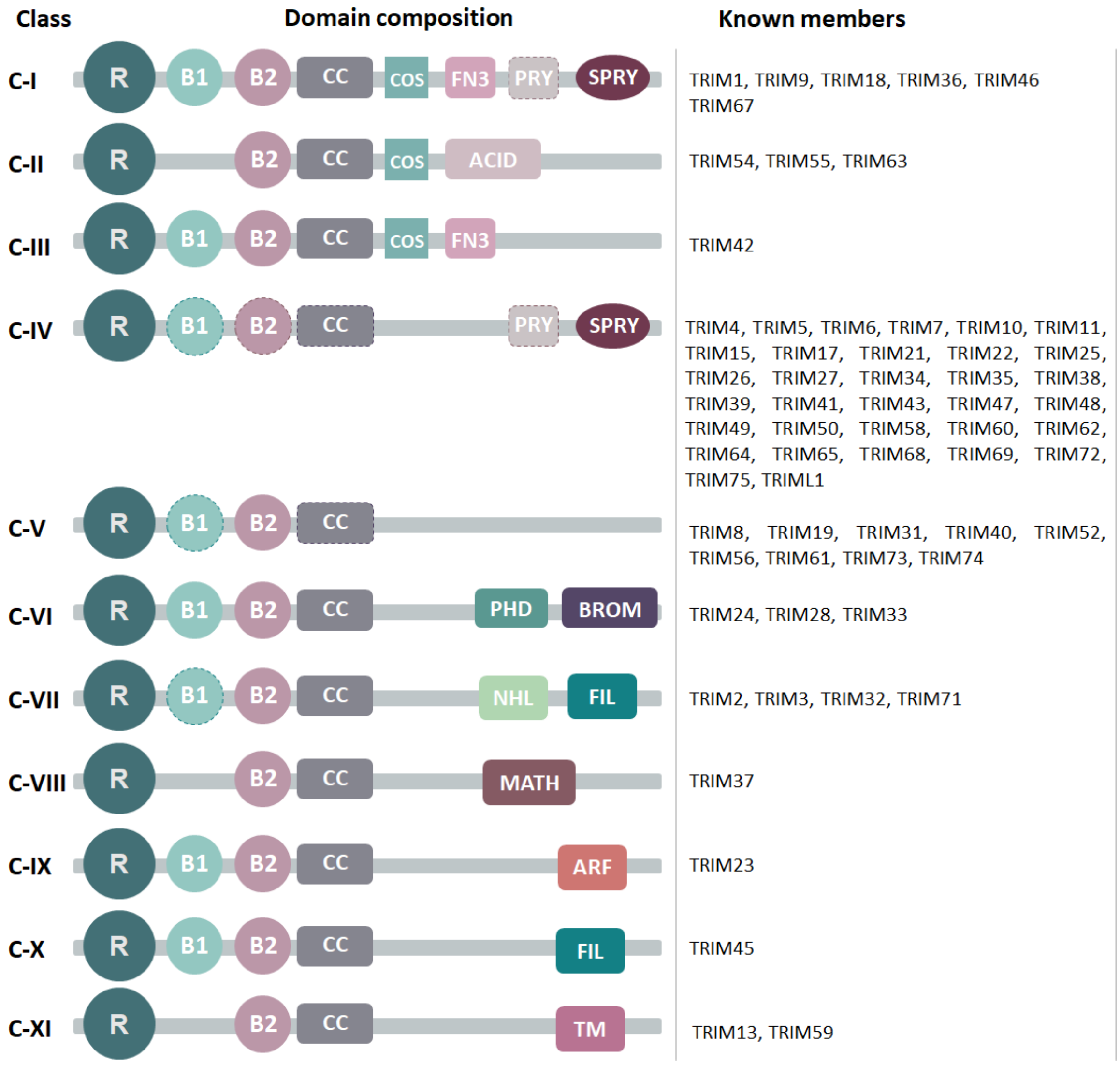


Figure 1.2 Human TRIM/RBCC protein family classification.
The TRIM family has 11 subfamilies (classes C-I to C-XI) sharing RING-finger domain (R), one or two B-box domains (B) and a coiled-coil domain (CC). Additional domains determine the classification, such as: COS-box motif (COS), fibronectin type III repeat (FN3), PRY domain (PRY), SPRY domain (SPRY), acid-rich region (ACID), PHD domain (PHD), bromodomain (BROMO), NCL1, HT2A and LIN41 domain (NHL), filamin-type I G domain (FIL), meprin and TRAF-homology domain (MATH) and transmembrane region (TM). Some members of a certain class are missing one or more domains specific for that class (shown as dash outline). Class II and Class V (MuRFs) lack the C-terminal domain after the RBCC fraction.(Short and Cox, 2006)

Regardless of their C-terminal composition, TRIM proteins derive their E3 ligase activity from the N-terminal RING domain (Koliopoulos *et al.,* 2016). The RING-finger motif was first identified in the protein encoded by the *Really Interesting New Gene 1* (Freemont, Hanson and Trowsdale, 1991) and the canonical RING finger motif is defined as Cys-X2-Cys-X11-16-Cys-X-His-X2-Cys-X2-Cys-X7-74-Cys-X2-Cys (C3HC4 type), where X can be any amino acid. Three-dimensional analyses of RING domains have shown that the RING finger motif is characterized by a conserved pattern of cysteine and histidine residues that coordinate two Zn^{2+} ions in a cross-braced fashion (Deshaies and Joazeiro, 2009). The cross-braced fashion is characterized by the first Zn^{2+} being coordinated by Cys in position 1, 2, 5, 6 and the second Zn^{2+} ion being coordinated by Cys or His in positions 3, 4, 7 and 8, with the fifth coordination site being Cys (C2 type) (Figure 1.3) in the TRIM family or His (H2 type) found in other RING domains (Freemont, 2000). Crystal structures of TRIM25 (PDB: 5FER, (Koliopoulos *et al.*, 2016)), TRIM21 (PDB:6D53; (Kiss *et al.*, 2019)), TRIM5α (PDB: 4TKP; (Yudina *et al.*, 2015) RING domains suggested dimerization via two short helices, which facilitates interaction with E2-cojugating enzyme-Ubiquitin complex (Yudina *et al.*, 2015; Anandapadamanaban *et al.*, 2019; Kiss *et al.*, 2019). In the crystal structure, E2 binds to the proximal RING, while the ubiquitin interacts with both proximal and distal RING in the dimer, promoting "closed" conformation of the E2–Ub conjugate and priming ubiquitin for a transfer (Yudina *et al.*, 2015) (Figure 1.3). As other TRIM E3 ligases, MuRFs are assumed to act as mediators that bring the ubiquitin (Ub)-conjugating E2 enzyme and the protein substrate together via the RING domain. Yet, the structural and functional analysis of MuRF1 interactions and the role of the RING domain are not complete.

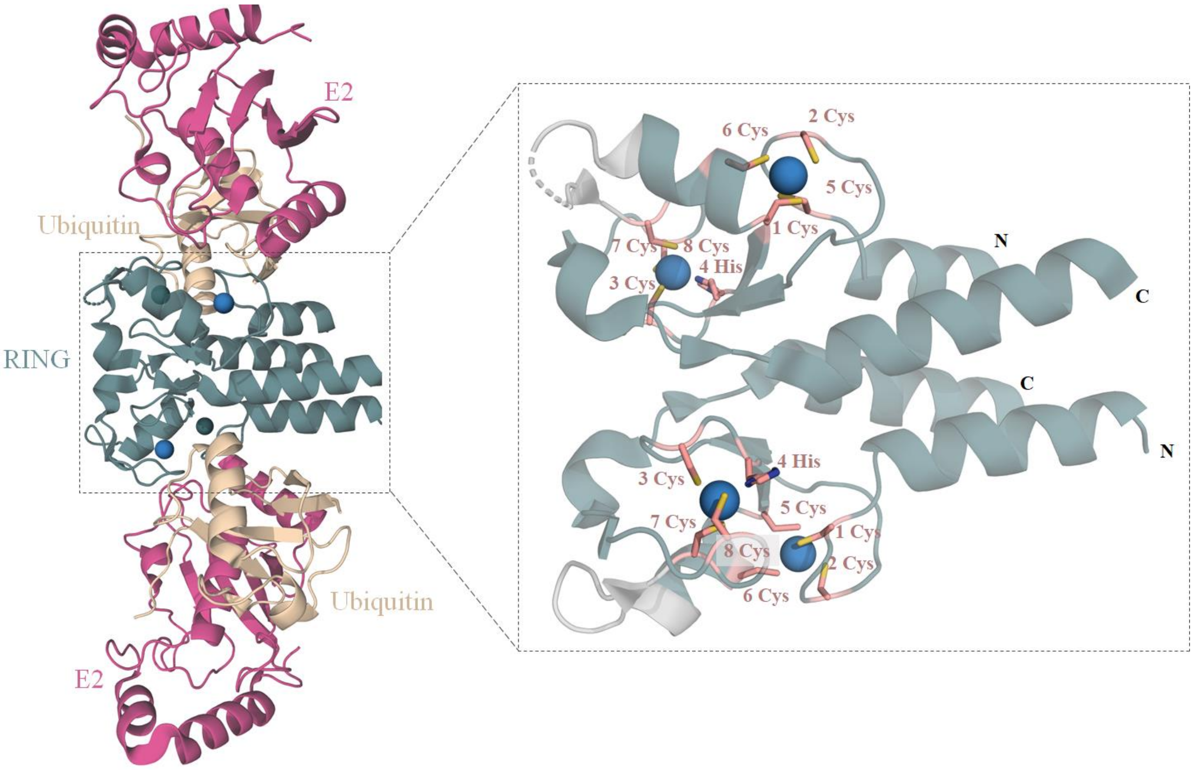


Figure 1.3 RING domain mediates ubiquitin transfer in TRIM E3 ligases.
RING domain (dark green) of the TRIM 21 in complex with an E2-conjugating enzyme (pink) and ubiquitin (golden) (PDB ID 6S53; Kiss *et al.*, 2019) is shown in the cartoon representation. One N-terminal and one C-terminal helix mediate dimerization of the RING domain. The TRIM RING domain belongs to the C2 RING type where conserved patterns of cysteine and histidine residues coordinate two Zn^{2+} ions in a cross-braced fashion. Cysteine (Cys) and histidine (His) residues coordinating Zn^{2+} ions are shown in salmon and numbers indicate position in coordination sites. Structure covering sequence which is non-conserved in the TRIM family of RING domains is shown in grey.

Although all MuRFs are suggested to control the trophic state of the muscle, only MuRF1 has been related to the muscle atrophy. In an early development high protein content of MuRF1was closely associated with the late myofibril atrophy and the severity of weakness (Wollersheim *et al.*, 2014), whilst MuRF2 and MuRF3 were found to be essential for the development and microtubule stability (Perera, Mankoo and Gautel, 2012; Silvestre *et al.*, 2019).

MuRFs share high sequence similarity, with the RING and B-box domains being well conserved (~81% seq.id.) and CC fraction being more diversified (~36% seq.id.) as it is ilustrated in the Figure 1.4. (Spencer *et al.*, 2000) Therefore, it remains unclear how specific roles and substrate specificity are arranged across MuRFs. It has, however, been proposed that MuRFs interact through their coiled-coil fraction with targeting substrates (Mrosek *et al.*, 2007) and the COS box mediates sarcomere binding (Franke *et al.*, 2014).

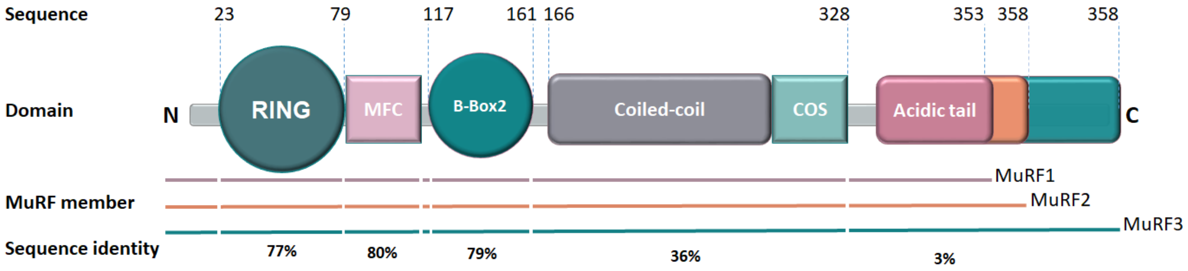


Figure 1.4 MuRF family domain composition and conservation.
MuRFs share TRIM domain composition of a RING finger (RING), a B-box2 domain (B-box2) and a coiled-coil domain (Coiled-coil), but also share MuRF family conserved region (MFC), COS-box motif (COS) and acidic tail. Acidic tail is the most diverse in the MuRF1 family (3% sequence identity), followed by coiled-coil domain, while RING, MFC and B-box domains are highly conserved. Identity is shown under the domain schematic (Spencer *et al*., 2000).

There is no crystal structure available of full length MuRF1, but recent EPR data of HD domain strongly suggested that MuRF1 can be considered to exhibit a canonical fold of the TRIM proteins (Stevens *et al.*, 2019). The canonical TRIM fold has been proposed based on available crystal structures of the helical domain of TRIM5α (PDB ID 4TN3; (Goldstone *et al.*, 2014)), TRIM20 (PDB ID:4CG4; (Weinert *et al.*, 2015)), TRIM25 (PDB ID:4LTB; (Sanchez *et al.*, 2014)) and TRIM69 (PDB ID:(Li *et al.*, 2014)) , which consists of a long helix (H1) followed by two shorter helices (H2 and H3). The HD domain dimerizes through H1 helix forming a long antiparallel coiled-coil (H1-H1) with helices 2 and 3 folding back against it (Figure 1.5). C-terminal functional domains in TRIM25 (PDB ID:4LTB;(Sanchez *et al.*, 2014)) and TRIM20 (PDB ID:4CG4;(Weinert *et al.*, 2015)) are found to be located on the central part of HD domain (Figure 1.4 C and D) while the B-box acts as a capping feature of a coiled-coil in TRIM5α (PDB ID 4TN3;(Goldstone *et al.*, 2014)) (Figure 1.5 B).

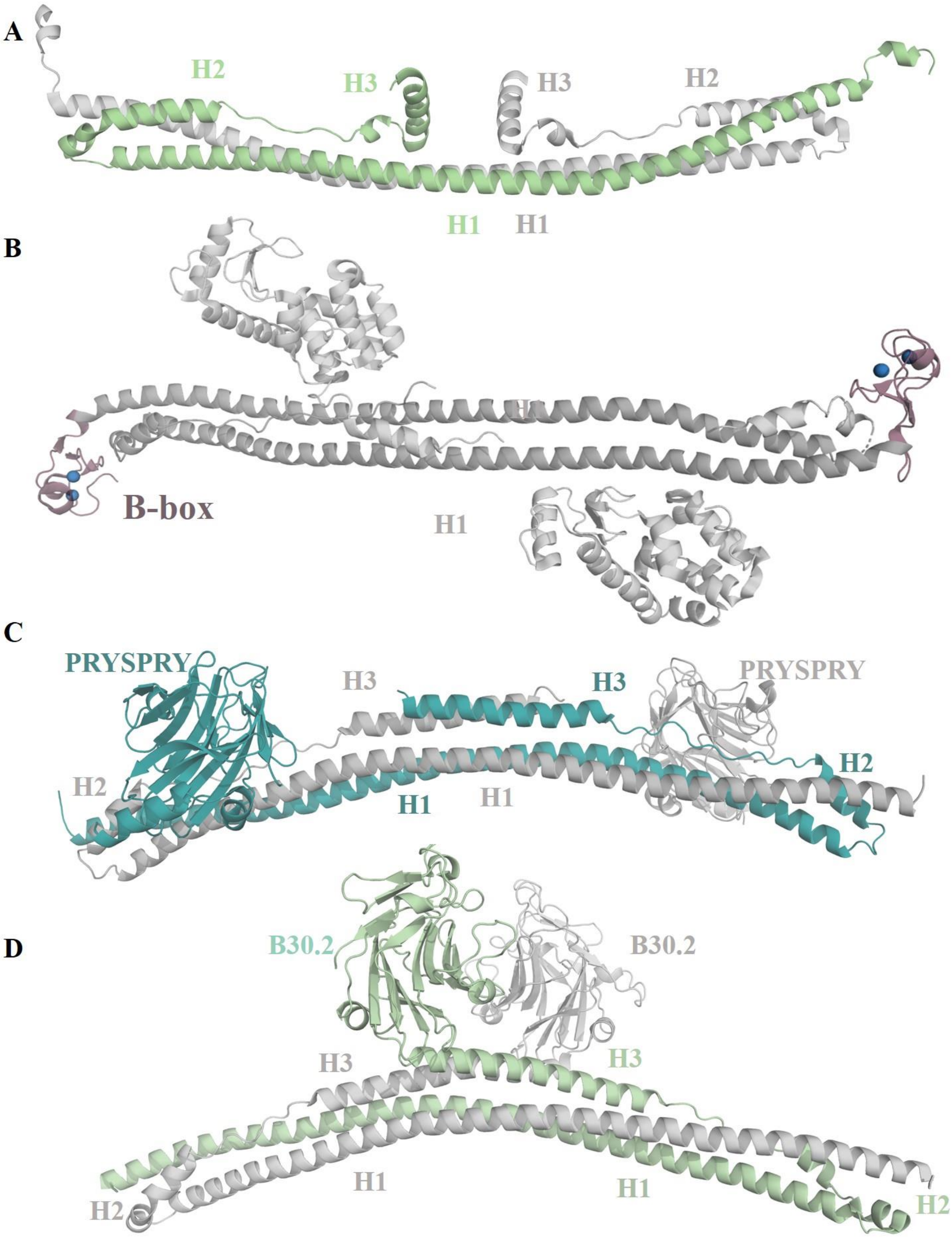


Figure 1.5 Helical domain of TRIM proteins fold in an antiparallel coiled-coil.
A Crystal structure of **A** TRIM 69 HD domain (PDB ID: 4NQJ;(Li *et al.*, 2014)), **B** TRIM 5α BHD (PDB ID 4TN3; (Goldstone *et al.*, 2014)), **C** TRIM 25 HD-PRYSPRY (PDB ID:4LTB; (Sanchez *et al*., 2014)) and **D** TRIM 20 HD- B30.2 (PDB ID:4CG4; (Weinert *et al.*, 2015)) dimers in a cartoon representation. Antiparallel dimerization of the long helix (H1) forms a coiled-coil with short helices (H2 and H3) folding against the CC. Functional C-terminal domains, PRYSPRY (**C**) and B30.2 (**D**) fold onto the central coiled-coil.

1.1.3 B-box mediated MuRF1 high order assembly

The presence of a B-box is the determinant of the TRIM fold, as it has been found only in members of this family. The MuRF1 B-box and the helical fraction (CC-COS) both form dimers in isolated states (Mrosek *et al.*, 2008), but when combined in a BHD construct and full length MuRF1, dimerization is overpowered by the higher order oligomerization states, which can often lead to aggregation.

The B-box and RING domains, are Zinc-finger domains, which coordinates Zn^{2+} ion through highly conserved Cys and His residues ensuring proper folding of a domain. The effect of Zn^{2+} depletion from the B-box on the oligomeric state was observed for the BHD construct via metal chelator Ethylenediaminetetraacetic acid (EDTA) (Mrosek *et al.*, 2008). Mrosek and colleagues showed that EDTA incubation shifts BHD equilibria from forming large oligomers (>500 kDa) to dimers, suggesting that abolishing the ability of one domain to cause high order assembly in a multidomain protein, could be a good strategy in preventing large oligomers and aggregation. Additionally, the crystal structure of the B-box domain revealed a hydrophobic interface clustering onto the F148 residue, mediating B-box dimerization (Figure 1.6), which could be easily disrupted by the alteration of a single residue from highly hydrophobic to highly charged. This type of remodeling interacting interface can be sufficient for an effective disruption of naturally occurring interactions (Schwab *et al.*, 2008; Schneider *et al.*, 2015), as introducing charge in the hydrophobic core is energetically unfavorable. Based on this rationale, Dr. Barbara Franke performed site directed mutagenesis of F148 in a MuRF1-BHD construct in order to disrupt hydrophobic interaction interface and restrain MuRF1-BHD in a dimer form (Franke *et al.*, in preparation).

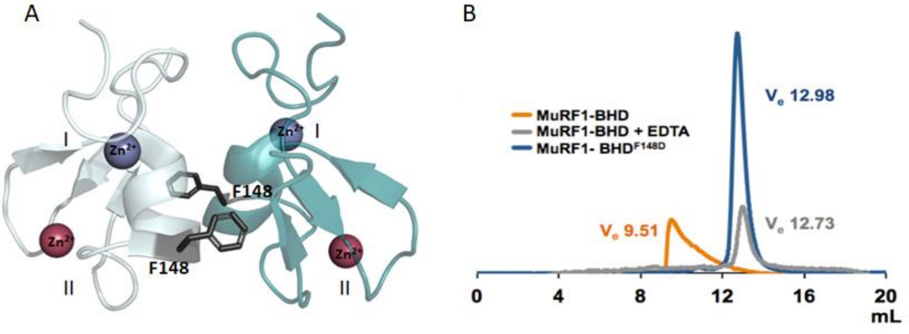


Figure 1.6 Role of the B-box domain in the high-order association of MuRF1-BHD
A. Crystal structure of MuRF1-Bbox dimer (PDB 3DDT). Zinc ions are shown as red and purple spheres and the residue F148 is shown as black in the hydrophobic interacting core; **B.** Size exclusion chromatograms of wild-type MuRF1-BHD and its point mutant MuRF1-BHD^{F148D} collected on a Superdex Increase 200 10/300 GL column (GE Healthcare) in 50mM Tris pH 8.0 100mM NaCl, 10 µM ZnCl$_2$, 1mM DTT (100 mM EDTA). An acute decrease in the MM of the MuRF1-BHD^{F148D} sample is observed, which displays exclusion values similar to those of MuRF1-BHD incubated with EDTA. (Figure adapted from Franke *et al.*, in preparation)

Prior to the beginning of this study, MuRF1-BHD^{F148D} was produced in *E.coli* and successfully purified for the crystallization and biophysical experiments. Size exclusion chromatography confirmed that MuRF1-BHD^{F148D} forms dimers in solution, and adopts the canonical anti-parallel dimer seen in all TRIM BHD structures to date. It remains unclear if the same mutation in the full length MuRF1 could be a simple and effective strategy for disrupting high-order assembly and to study the resulting effects on substrate targeting and catalytic activity.

Aims

Since the initial discovery of MuRF1 expression being upregulated in numerous atrophy-inducing conditions, MuRF1 has been recognized as a key marker of muscle atrophy and as such widely studied. One major gap in our understanding of MuRF1, however, is the lack of reliable data on molecular level. The main obstacle to obtaining biophysical data in *in vitro* experiments is MuRF1 tendency to aggregate. The main aim of this research is, thereby to develop an assay suitable for the MuRF1 interaction studies. Throughout BLI and ubiquitination assays with *in vivo* studies this research aims to reveal the basis of MuRF1 substrate targeting and structural, associative and functional properties that define MuRF1 as a muscle atrophy contributor.

1.2 Material and methods

1.2.1 Protein production and purification

1.2.1.1 MuRF1WT and MuRF1F148D

Clones for murine full-length MuRF1 (UniProtKB Q38HM4) in wild type and mutant form were provided in the pETM41 vector (EMBL collection) and by Dr Alexander Gash (Mannheim Klinikum). The pETM41 vector contains a His_6-MBP tag and a Tobacco Etch Virus (TEV) protease cleavage site N-terminally to the target insert (Appendix A). The constructs were tested for expression in *Escherichia coli* BL21 (DE3) cells (Novagen) and *Escherichia coli (E. Coli)* Rosetta (DE3) cells (Novagen). Due to the higher production in *E. Coli* BL21 (DE3), these cells were used for further production. MuRF1^{F148D} and was purified following the same protocol as the wild type construct. Cells were grown in Luria-Bertani (LB) medium supplemented with 25 µg/^{mL} kanamycin at 37°C. At OD_{600}= 0.6 and protein production was induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). At the moment of induction, LB medium was additionally supplemented with 100 µM $ZnCl_2$. Upon induction, cells were grown for 20 h at 18°C. Harvesting was achieved by centrifugation at 5000 rpm (9760 xg) at 4°C for 20 minutes. Pellets were resuspended in lysis buffer (50 mM trisaminomethane-cloride (Tris-HCl) pH 8.0, 300 mM NaCl, 2 mM dithiothreitol (DTT), 10 µM $ZnCl_2$) supplemented with ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail (Roche) a lysed by sonication on ice. After clarification by centrifugation at 18000 rpm (47850 xg) at 4°C for 45 minutes, the supernatant was applied to a HisTrap HP (GE Healthcare) Immobilized Metal Affinity Chromatography (IMAC) column pre-equilibrated in the lysis buffer with 20 mM imidazole. MuRF1^{WT} and MuRF1^{F148D} were eluted using an increasing gradient of imidazole concentrations from 20 mM to 1M. Protein containing fractions were dialyzed into a low salt buffer (50 mM Tris-HCl pH 8.0, 50 mM NaCl, 2 mM DTT, 10 µM $ZnCl_2$). Dialyzed samples were loaded on an ion exchange chromatography 5 mL MonoQ® 5/50 GL column (GE Healthcare). Protein separation was achieved by using an increasing gradient increase of a salt concentrations (50-1000 mM). A final purification step was carried out by size exclusion chromatography (SEC) by loading the samples on a Superdex S200 26/60 (GE Healthcare) column in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DTT, 10 µM $ZnCl_2$. Purity of the sample was assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Protein samples were flash frozen in liquid nitrogen and stored at -80°C until the further use.

1.2.1.2 MuRF1-BHDWT, MuRF1-BHDF148D and MuRF1-B-box

Human MuRF1-BHD^{WT} (UniProtKB Q969Q1 residues:117-161) and MuRF1-B-box(UniProtKB Q969Q1 residues:117-328) (Figure 1.7) were into pETM-11 vector (EMBL collection) and produced in *E. coli* Rosetta™ (DE3) (Novagen) as described before (Mrosek *et al.*, 2008). MuRF1-BHD^{F148D} was obtained via site directed mutagenesis by Dr. Barbara Franke (Appendix A).

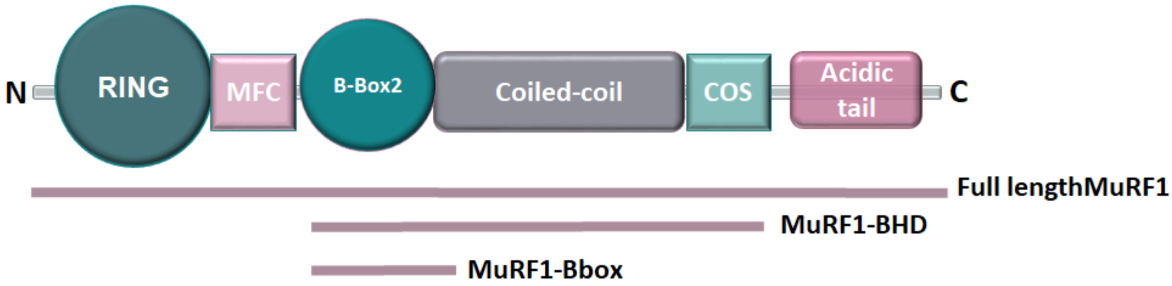


Figure 1.7 MuRF1 constructs composition and descriptions.
MuRF1-BHD is composed of B-box, coiled-coil and cos box;

All the cultures were grown at 37°C to an OD_{600} =0.6 and induced with 250 µM IPTG, followed by addition of 100 µM $ZnCl_2$. and were further grown at 18°C for 18h. Cells were harvested by centrifugation at 5000 rpm (g) at 4°C for 20 minutes and pellets were resuspended in the MuRF1 lysis buffer and sonicated for lysis. Lysed cells were centrifuged on 18 000 rpm for 50 minutes and supernatant was loaded on 5 mL HisTrap HP (GE Healthcare) column for IMAC purification. Protein was eluted with 250 mM imidazole in the lysis buffer during the gradient elution step. His_6-tag was removed by TEV protease in an overnight incubation reaction followed by reverse IMAC purification on a 5 mL HisTrap HP (GE Healthcare) column. Further sample purification followed MuRF1 purification steps where for MuRF1-B-box 2 sample HighLoad Superdex 75 16/60 PG column (GE Healthcare) was used instead of Superdex 200 26/60. All the proteins samples were flash frozen in liquid nitrogen and stored at -80°C until further use.

1.2.1.3 Titin A168-170 Titin A170-M1

Titin A168-A170 from human titin (UniProtKB Q8WZ42 residues:24431–24731) and titin A170-M1 (UniProtKB Q8WZ42 residues: 24622-25170) (Figure 1.8) in pETM-11 vector (Appendix A) was produced in *E. coli* Rosetta™ (DE3)

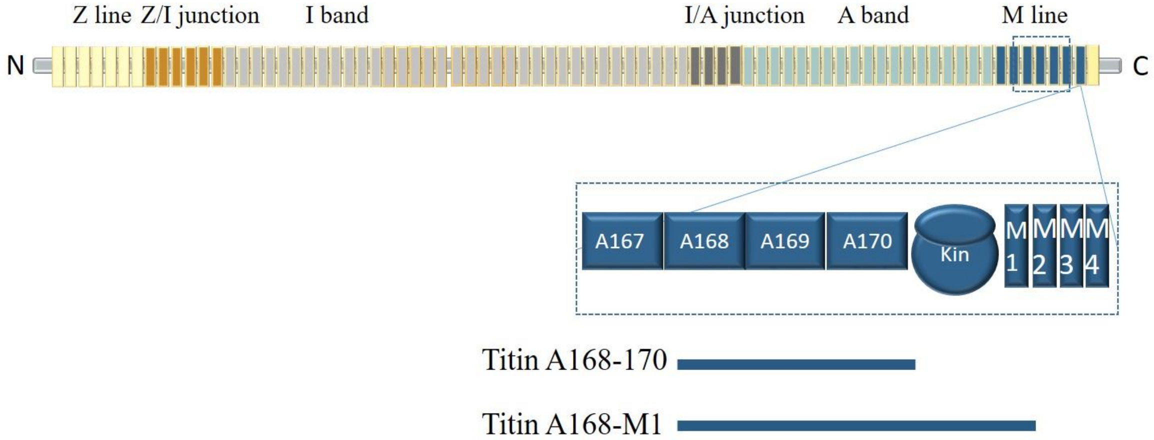


Figure 1.8 Titin domain composition and constructs definition.

For the BLI experiments, constructs were cloned in pOPINJ containing a His$_6$-GST tag (Appendix A) by Dr. Bruno Manso (University of Konstanz). Vectors were transfected in *E. coli* BL21 (DE3) and cells were grown at 37°C to OD$_{600}$ of 0.6. Protein overexpression was induced with addition of IPTG to the total concentration of 500 μM. After overnight incubation at 18°C cells were harvested by centrifugation at 5000 rpm for 20 minutes. Lysis of the cells was obtained in the lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 2 mM DTT) on ice using sonication. Lysates were clarified by centrifugation at 18000 rpm for 50 minutes and the supernatant was loaded on 5 mL HisTrap HP (GE Healthcare) for affinity chromatography. For BLI experiments His$_6$-GST tag was not removed and the purification was improved by size exclusion chromatography on Superdex S200 26/60 (GE Healthcare) column in 50 mM Tris-HCl pH 8.0, 150 mM NaCl buffer. Sample was used fresh or flash frozen and kept at -80°C until further use.

1.2.1.4 Myosin heavy chain MYH7$^{964\text{-}1088}$ fragment

Human MYH7$^{964\text{-}1085}$ (UniProtKB P12883 residues 964-1088) was cloned into pOPIN-J vector (Addgene 26045) by Dr. Thomas Zaharachenko.

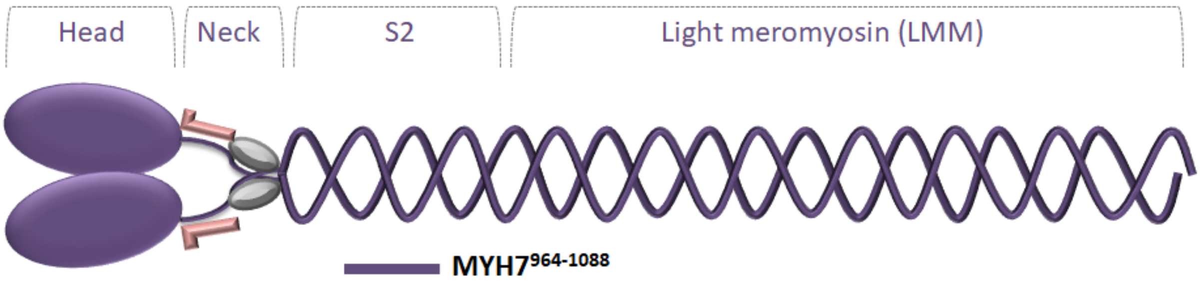


Figure 1.9 Myosin heavy chain fragment (MYH7$^{964\text{-}1088}$) is a part of a S2 domain of myosin heavy chain.

Vector was transformed into *E. coli* Rosetta PlacI (Novagen). pOPIN-J includes N-terminal His$_6$-(glutathione S transferase) GST and 3C protease cleavage site (Appendix A). Overnight cultures were supplemented with 2% glucose. Cells were inoculated in 1:100 ratios into LB supplemented with 50 µg/mL ampicillin and 35 µg/mL chloramphenicol and were grown at 37°C to an OD$_{600}$ = 0.6. Protein expression was induced with 0.5 mM IPTG and cells were further grown for 18h at 18°C. Cells were harvested by centrifugation at 5000 rpm (9760 xg) at 4^{o}C. Bacterial pellets were resuspended in lysis buffer (50mM Tris pH 8.0, 300mM NaCl, 2mM DTT) containing an EDTA-free protease inhibitor cocktail (Roche). Cells were lysed by sonication on ice, followed by clarification of the lysate by centrifugation at 18000 rpm (47850 xg). Cell lysates were applied to a 5mL HisTrap HP (GE Healthcare) equilibrated in a lysis buffer containing 20 mM imidazole. Sample elution was carried out with a continuous imidazole gradient in the concentration range 20-1000mM. Eluted protein was subjected to a buffer exchange via dialysis against 50mM Tris pH 8.0, 50mM NaCl, 2mM DTT. Affinity tags were removed for ubiquitination assays by overnight incubation with 3C protease at 4°C, while samples prepared for Biolayer interferometry (BLI) assays preserved their His$_6$-GST tags. Both samples were further purified by applying them to a 5 mL MonoQ® 5/50 GL column (GE Healthcare). Protein separation was achieved by using an increasing gradient increase of a salt concentrations (50-1000 mM). Eluted protein samples were dialyzed against SEC buffer (50mM Tris pH 8.0, 150mM NaCl, 2mM DTT) and loaded on Superdex S200 26/60 (GE Healthcare). Purified samples were analyzed on 12% SDS-PAGE for impurities and applied for ubiquitination or BLI experiments, or were flash frozen in liquid nitrogen and kept in -80°C until further use.

1.2.1.5 CARP

Human CARP$^{106\text{-}319}$ (UniProtKB Q15327, residues 106-319) was produced and purified following a previously reported protocol (Zhou *et al.*, 2016).

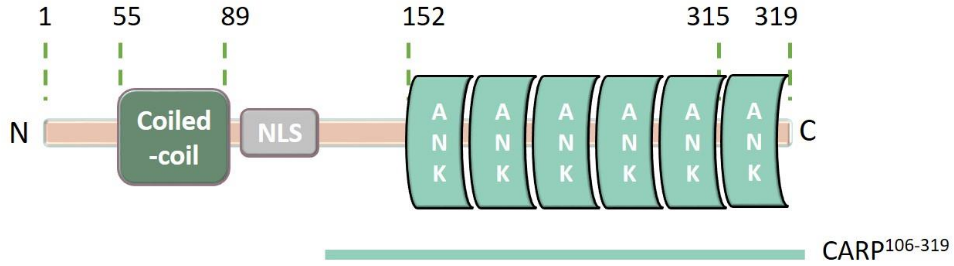


Figure 1.10 CARP composition and construct definition

Shortly, protein was expressed in *E. coli* Rosetta (Novagen). Cells were grown at 37°C to OD_{600} = 0.6 in LB media supplemented with 50 µg/mL ampicillin and 35 µg/mL chloramphenicol. Protein production was induced by 500 µM IPTG and cells were further grown overnight at 18°C. Cells were harvested by centrifugation and lysed by sonication in 25 mM HEPES pH7.5, 300 mM NaCl supplemented with EDTA-free protease inhibitor cocktail (Roche). After clarification of the lysate by centrifugation, the lysate was loaded on the 5mL Ni^+-NTA column (GE HEalthcare) for affinity purification. His_6-tag was removed by overnight incubation with TEV protease and affinity chromatography on 5 mL HisTrap HP (GE Healthcare). Tag free CARP was purified from contaminants using size exclusion chromatography Superdex S75 16/60 (GE Healthcare) column in 25 mM HEPES pH7.5, 100 mM NaCl.

1.2.1.6 Ubiquitin

Ubiquitin plasmids and the protein for ubiquitination experiments were a kind gift provided by professor Martin Scheffner (University of Konstanz). Ubiquitin containing pET3ab plasmids (Appendix A) were transfected into *E. coli* BL21 (DE3) cells and the LB medium containing 50 µg/mL ampicillin was inoculated with bacteria to an OD_{600} = 0.1. Cells were grown at 37°C to the OD_{600} = 0.6-0.8. Protein overexpression was induced with IPTG to the final concentration of 400 µM. Cells were further incubated at 37°C for 5h and harvested by centrifugation at 5000 rpm (9760 xg) for 20 min at 4°C. Pellets were lysed in 5 mL of Phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) supplemented with ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail (Roche) and lysed by sonication on ice. After centrifugation at 10000 rpm (7000 xg) for 15 minutes for clarification, the supernatant was then incubated for 30 minutes at 70°C for heat precipitation of the contaminants. As Ubiquitin is thermostable at 70°C it is resistant to the heat induced denaturation. Unlike ubiquitin, most of the contaminants in the cell lysate are sensitive to the heat denaturation and this advantage of ubiquitin is used in the purification process. Purified protein is flash frozen and stored at -80°C until further use.

1.2.1.7 UbcH5B (E2)

UbcH5b plasmids and the protein for optimization of ubiquitination experiments were a kind gift provided by professor Martin Scheffner (University of Konstanz). For the further ubiquitination study during the course of the project protein was purified following the protocol.

Construct plasmid contains His_6 tag followed by TEV protease cleavage site N-terminally to the insert (Appendix A). Transfected *E. coli* BL21 (DE3) cells were grown at 37°C to the OD_{600} = 0.1-0.6 and induced with IPTG to the final concentration of 400 µM. Expression was conducted at 37°C for an additional 5h. After harvesting cells at 5000 rpm (9760 xg) for 20 minutes at 4°C, cell pellets were resuspended in 30 mL PBS supplemented with 20 mM imidazole and 1 mM DTT and lysed by sonication. Clarified lysate was applied on 5 mL HisTrap HP (GE Healthcare) for affinity purification using gradient increase of imidazole from 20 mM to 1 M. Protein was eluted at the concentration of imidazole of 250-300 mM. Imidazole was removed from the solution by dialysis against 50 mM Tris pH 7.5, 300 mM NaCl, 0.5 mM DTT in the presence of TEV protease. Further purification was obtained by reverse affinity chromatography on 5mL HisTrap HP (GE Healthcare) and size exclusion chromatography on Superdex S750 16/60 (GE Healthcare) column in 50 mM Tris pH 7.5, 150 mM NaCl, 0.5mM DTT. Purified protein was flash frozen in liquid nitrogen and stored at -80°C until the use.

1.2.1.8 GST and MBP

His_6-GST and His_6-MBP were used in the BLI experiments as a control to investigate if the observed binding is affected by the presence of a protein tag. Hence, for the simulation of the control experiments, protein tags were produced from the same plasmids as the experimental proteins.

For His_6-MBP sample preparation *E. coli* BL21 (DE3) cells transfected with the empty pETM-41 vector (EMBL collection) were grown at 37°C to the OD_{600} = 0.6, followed by induction with IPTG to final 500 µM and further expression overnight at 18°C. Cells were harvested and lysed as described for His_6-MBP-MuRF1. Purification steps were consisted of IMAC using 5 mL HisTrap HP (GE Healthcare), dialysis against size exclusion buffer (50 mM Tris pH 8.0, 150 mM NaCl, 2 mM DTT) and size exclusion chromatography on Superdex S200 26/60 (GE Healthcare).

For His_6-GST sample preparation *Escherichia coli* placI (DE3) cells were transfected with empty pOPIN-J plasmid (Addgene 26045) with addition of 0.5% glucose in the overnight medium. After inoculation, cells were grown and harvested as described for $MYH7^{964-1085}$. Protein sample purification was achieved by two step chromatography purification: IMAC on 5 mL HisTrap HP (GE Healthcare) in 50 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT and elution using gradient increase imidazole concentration and size exclusion chromatography on Superdex S200 26/60 (GE Healthcare) in 50 mM Tris pH 8.0, 150 mM NaCl, 2 mM DTT. Proteins were flash frozen in liquid nitrogen and stored at -80°C until further use in BLI experiments.

1.2.2 Muscle protein substrates *in vitro* ubiquitination

In order to identify and quantitatively analyze three different known targets of MuRF1, $MYH7^{964-1085}$, titin A168-170 and $CARP^{106-319}$ were subjected to the ubiquitination assay reaction for 5h at 37°C in the same reaction as the autoubiquitination assay. UBA1 was a kind gift from professor Martin Scheffner (University of Konstanz). Set of five reactions was set per target. MuRF1, as other E3 ubiquitin ligases, ubiquitinates itself. Autoubiquitination of MuRF1 can serve as a positive control for MuRF1 activity. Reaction mixture lacked MurF1 and substrate was one of three controls. Second control contained MuRF1, but was lacking ubiquitin and third control was containing ubiquitin and lacking MuRF1. Fifth reaction was composed only of a substrate and reaction buffer. The reactions were stopped by adding 7 µL of SDS-PAGE sample loading buffer and boiling for 5 min at 90°C. Reaction products (30 µg) were loaded on gradient commercially available Mini-PROTEAN® TGX Stain-Free™ Precast Gels (BioRAD) and separated for 40 min at 40 mA. Gels were stained by Coomasie brilliant blue staining (0.01% Coomassie brilliant blue R250, 45% (v/v) methanol and 10% (v/v) glacial acetic acid) overnight min at room temperature and destained in destaining solution (50% ethanol, 40 % glacial acetic acid) for 30 minutes. Ubiquitination was monitored by identification of a higher molecular mass species (corresponding to mono- or poly- ubiquitination) on the stained gels existing only in the reaction mixture, but in none of the controls. Destained gels were imaged using Gel Doc™ EZ Imager (BioRad) and analyzed with Image Lab™ Software 5.2.1. Corresponding bands were cut out from the SDS-PAGE gel and submitted to Proteomics facility at University of Konstanz for identification of di-glycines.

1.2.3 Identification of ubiquitination sites by mass spectrometry

Mass spectrometry analysis were done by the Mass spectrometry analysis facility at the University of Konstanz. All samples were analysed by reversed phase liquid chromatography nanospray tandem mass spectrometry (LC-MS/MS) using an LTQ-Orbitrap mass spectrometer (Thermo Fisher) and an Eksigent nano-HPLC. The dimensions of the reversed-phase LC column were 5 µm, 100 Å pore size C18 resin in a 75 µm i.d. × 10 cm long piece of fused silica capillary (Acclaim PepMap100, Thermo Scientific). After sample injection, the column was washed for 5 min with 95% mobile phase A (0.1% formic acid) and 5% mobile phase B (0.1% formic acid in acetonitrile), and peptides were eluted using a linear gradient of 5% mobile phase B to 40% mobile phase B in 65 min, then to 80% B in an additional 5 min, at 300 nL/min. The LTQ-Orbitrap mass spectrometer was operated in a data dependent mode in which each full MS scan (30 000 resolving power) was followed by five MS/MS scans where the five most abundant molecular ions were dynamically selected and fragmented by collision-induced dissociation (CID) using a normalized collision energy of 35% in the LTQ ion trap. Dynamic exclusion was allowed. Tandem mass spectra were searched against the Swissprot human protein database using Mascot (Matrix Science) with "Trypsin/P" enzyme cleavage, static cysteine alkylation by iodoacetamide and variable methionine oxidation.

1.2.4 Biolayer interferometry (BLI)

For protein-protein interaction studies BLI was used with an Octet K2 instrument (Pall Forte Bio). All the assays were performed at 30 °C in solid black 96-well plates (Geiger Bio-One) with agitation set to 1000 rpm. Octet K2 is equipped with 2 biosensor-channels, which enabled simultaneous binding assay with one sensor loaded with a ligand (His$_6$-GST-MYH$^{964-1085}$ or His$_6$-GST-A168-170) and the second sensor loaded with a control (His$_6$-GST).

In order to optimize the immobilization onto dip and read Anti-GST biosensors (ForteBio) varying concentration of His$_6$-GST-MYH$^{964-1085}$ (3.0 µmol/mL, 7.5 µmol/mL, 15 µmol/mL and 30 µmol/mL) in 50 mM Tris pH 7.5, 300 mM NaCl were tested for 400 seconds at 30°C incubation. A 120 seconds biosensor washing step was applied post-loading. After the optimal concentration was determined for His$_6$-GST-MYH$^{964-1085}$, the same optimization protocol was implemented for optimization of loading His$_6$-GST-A168-170 and His$_6$-GST.

His_6-GST-MYH$^{964-1085}$ or His_6-GST-A168-170 were immobilized onto anti-GST biosensors (Pall Forte Bio) by incubating the sensor in 30 μg/mL protein sample in PBS buffer for 300 s at 30°C. Sensor loading was followed by a 120 s washing step buffer to remove excess of unbound sample. The analyte samples in the test wells, His_6-MBP-MuRF1 or MuRF1-BHD, were assayed in MuRF1 buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 μM $ZnCl_2$, 2 mM DTT) for 600 in the concentration range 0.5-10 μM. Dissociation was monitored for 600s by MuRF1 buffer. In separate control experiments, the possible unspecific binding of MBP and GST tags with each other, with the empty sensor, and with the samples assayed was examined. The control data were used to correct any systematic baseline drift or unspecific binding. Data analysis and curve fitting were done using Octet Forte Bio Data Analysis software v10.0 and the 2:1 heterogeneous ligand (HL) binding mode. K_D values were calculated from steady-state kinetic analyses calculated using the estimated response at equilibrium for each analyte concentration. Additionally, steady-state kinetic analyses were done for every data set to calculate the K_D using the estimated response at equilibrium for each analyte concentration rather than the k_{on} and k_{off} values.

1.2.5 EDTA Zn^{2+} depletion from the BHD and the MuRF1

RING fingers and B-box 2 are both Zn-finger domains, which coordinate Zn^{2+} ion through Cys and His residues and this interaction is essential for a proper folding of the domains. To observe the effect of Zn depletion from the Zn-finger domains on the MuRF1 stability, samples were incubated with the metal chelator Ethylenediaminetetraacetic acid EDTA. Purified MuRF1^{WT} sample concentrated to a final concentration of 1 mg/mL. MuRF1^{WT}-EDTA sample contained 100 μL of protein and 10 μL of 100 mM EDTA. Native sample was prepared by mixing 100 μL of protein and 10 μL of buffer. Both samples were incubated at 4°C overnight and loaded on the analytical size exclusion chromatography column Tricorn S200 30/10 (GE Healthcare) equilibrated in the 50 mM Tris pH 8, 100 mM NaCl, 10 μM $ZnCl_2$, 1 mM DTT (10 mM EDTA). Samples were eluted in the same equilibration buffers with a flow rate of 0.75 mL min^{-1}. Elution peak shift was observed as indication of the oligomerization state change. Same experiment was conducted for the full-length MuRF1 and MuRF1^{F148D}.

1.3 Results

1.3.1 Sample preparation /Protein purification

In order to investigate and characterize MuRF1 interactions with its substrates, myosin, titin and CARP, protein samples were recombinantly produced in E. coli and purified using a combination of chromatographic techniques previously described in section 1.2.1. The final purity and homogeneity of samples prepared for Biolayer interferometry analysis was confirmed through size exclusion chromatography (SEC) and visual inspection of proteins by SDS-PAGE stained with Coomassie-Brilliant blue (Figure 1.11). His$_6$-GST-MYH$^{964-1088}$, His$_6$-GST-Titin A168-170, BHD^{WT} and BHD^{F148D} were successfully purified with the production yield of average 14 mg/L, 18 mg/L, 20 mg/L and 22mg/L, respectively. Proteins were concentrated in SEC buffer and stored at -80°C for up to two months.

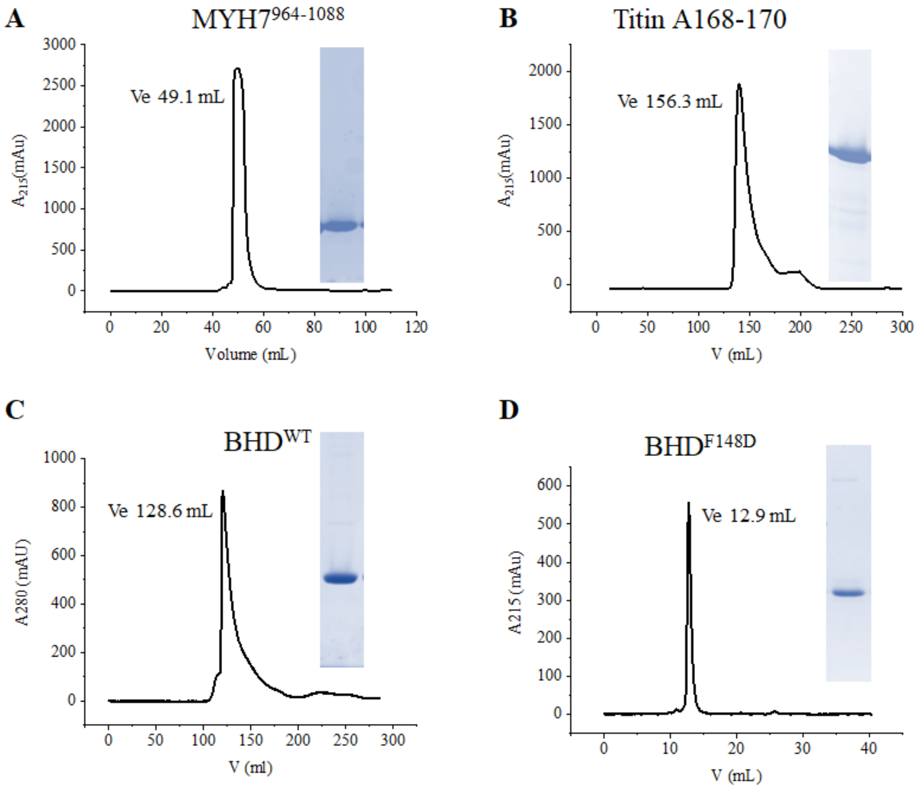


Figure 1.11 Protein quality assessment for BLI experiments
Size exclusion chromatography profiles with the SDS-PAGE insets showing fractions corresponding to the main peaks of proteins used for the biophysical characterization by BLI. **A** Myosin heavy chain fragment aa 964-1088 SEC on Superdex S75 16/60 column (GE Healthcare); **B** Titin A168-170 sample purified on SEC on Superdex S200 26/60 (GE Healthcare); BHD^{WT} SEC elution profile on Superdex S200 26/60; BHD^{F148D} sample elation profile collected on a Superdex Increase 200 10/300 GL column (GE Healthcare); All samples were in 50mM Tris pH 8.0 100mM NaCl, 1mM DTT and C and D buffer contained 10 µM $ZnCl_2$.

Previous research showed that unfolding of the B-box domain prevents high order oligomerization in BHD construct (Mrosek *et al.*, 2008). Dr. Barbara Franke designed the F148D mutation in the B-box in order to disrupt high order assembly, while preserving structural integrity of the B-box domain in the BHD construct (Franke *et al.*, in preparation). Therefore, the F148D mutation was inserted in the full length MuRF1 to investigate if disrupting the hydrophobic interface in the B-box would have an effect on full length MuRF1. His_6-MBP-$MuRF1^{WT}$ and His_6-MBP-$MuRF1^{F148D}$ were overexpressed in *E. coli* with the majority of sample being in soluble form and were purified to a high homogeneity with the final yield of 2-5mg/L of bacterial culture. The attempt of His_6-MBP tag removal led to increased sample instability and precipitation, hence the tag was preserved during the purification process and protein as such was used for all the experiments. His_6-MBP-$MuRF1^{F148D}$ showed higher sample quality and stability than the wild type and was concentrated to 18 mg/mL without precipitation. Wild type samples were concentrated up to 2 mg/mL with any further concentrating attempt resulting in precipitation.

Size exclusion chromatography on Superdex S200 26/60 (GE Healthcare) for wild type and mutant MuRF1 show that the mutation F148D in the B-box domain causes no acute shift in the elution profiles, with both samples being eluted close to the void volume (V_0=108 mL) of the column (Figure 1.12). These results show that the mutation in the B-box domain of MuRF1 does not fully prevent formation of high oligomeric states, but increases sample quality and protein stability.

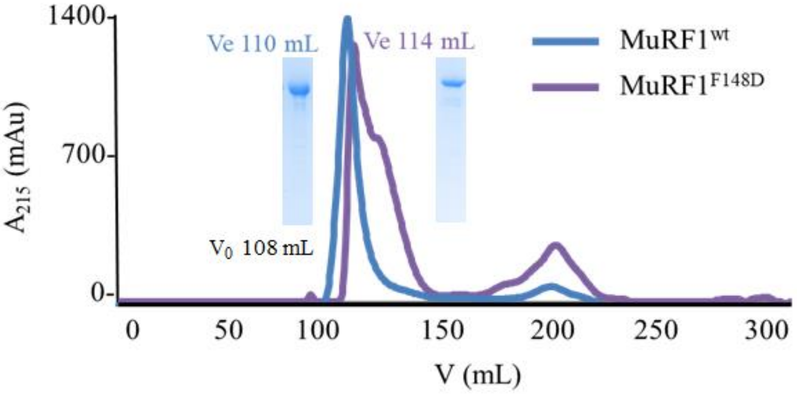


Figure 1.12 Size exclusion chromatography of $MuRF1^{WT}$ and $MuRF1^{F148D}$ elution profile comparison.
Proteins were eluted on a Superdex S200 26/60 column (GE Healthcare) in 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 2 mM DTT. The mutation F148D in the B-box domain causes no acute shift in the elution profile of MuRF1. Elution volumes (V_e) indicate that samples elute as high molecular mass species close to the column void (V_0=108 mL).

1.3.2. Zn^{2+} depletion has no acute effect on the MuRF1 high order oligomerization

The F148D mutation in the B-box domain of MuRF1 effect on the high order assembly of MuRF1 was not detected by using size exclusion chromatography elution profiles comparison. To further investigate if the interruption of structural integrity of two Zn^{2+} binding domains, RING finger and B-box would affect oligomerization properties of the MuRF1, EDTA chelation of the endogenous Zn^{2+} was obtained by overnight incubation of the full length MuRF1 samples with 100 mM EDTA and corresponding elution profiles on SEC were compared.

In the SEC profiles of the MuRF1 in native (Ve= 8.2 mL) form and MuRF1 incubated with EDTA (Ve= 8.62 mL) no significant change in the elution volume was detected and both samples were found to elute as a high molecular mass species (>600 kDa). Compared SEC elution profiles revealed that Zn^{2+} depletion, although stabilizes dimerization of BHD construct, does not have the same effect on the full length MuRF1 (Figure 1.13). The potential differences in the size of formed oligomeric states in the native MuRF1, MuRF1^{F148D} and the MuRF1 in the presence of EDTA should be further investigated by utilizing methods suitable for analyzing high molecular mass particles and aggregates, such as Asymmetrical flow field-flow fractionation (AF4).

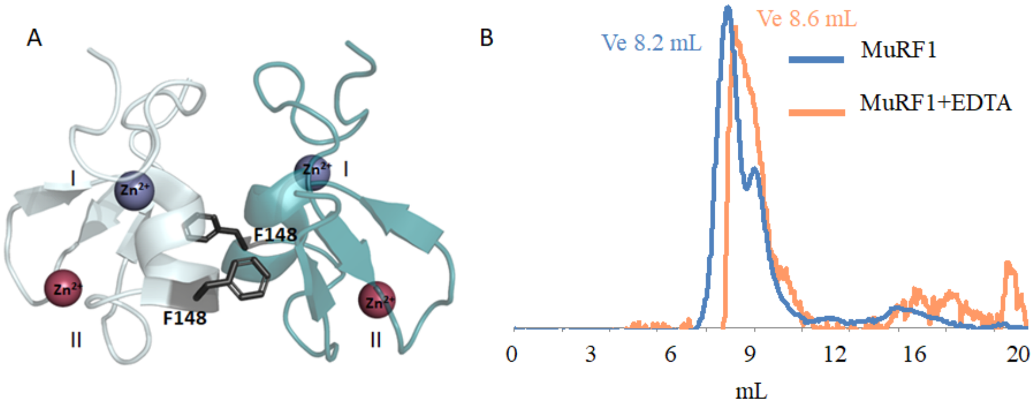


Figure 1.13 Role of the B-box domain in the high-order association of MuRF1
A. Crystal structure of MuRF1-Bbox dimer (PDB 3DDT). Zinc ions are shown as red and purple spheres and the residue F148 is shown as black in the hydrophobic interacting core; **B.** Size exclusion chromatograms of full length MuRF1 collected on a Superdex Increase 200 10/300 GL column (GE Healthcare) in 50mM Tris pH 8.0 100mM NaCl, 10 µM ZnCl$_2$, 1mM DTT (100 mM EDTA). No acute change in the elution profile of MuRF1 incubated with EDTA and native form of the samples are observed.

1.3.3 MuRF1 catalytic activity is affected by the B-box F148D mutation

The mutation F148D in the B-box domain of MuRF1 was shown to be effective in preventing higher order oligomerization in the BHD construct, but no severe effect was observed for the full length MuRF1 assembly. In order to assess if the mutation affects catalytic activity of MuRF1, *in vitro* ubiquitination assays were developed.

MuRF1 E3 ubiquitin ligase activity has only been observed using silver staining or Western-Blot for detection of low abundant ubiquitinated species in the SDS-PAGE analysis of muscle tissue samples and in vitro ubiquitination assays (Marblestone *et al.*, 2012). Here, MuRF1 ubiquitination assays were optimized to increase the reaction yield and enable simple detection via a standard Coomassie Brilliant blue staining protocol. The optimized ubiquitination assay allowed for a fast monitoring of MuRF1 catalytic activity using three selected targets: titin A168-170 as a well-established target suggested to have a role in signaling and as degradation target (Centner *et al.*, 2001; Mrosek *et al.*, 2007; Bogomolovas *et al.*, 2014); myosin heavy chain (MYH7) suggested to be a degradation target (Clarke *et al.*, 2007; Cohen *et al.*, 2009; Marblestone *et al.*, 2012) and the C-terminal ankyrin-repeat fraction of CARP (Witt *et al.*, 2008) where the role of ubiquitination is still obscure. Additionally, MuRF1^{wt} and MuRF1^{F148D} activities were directly compared for autoubiquitination and all the selected targets using the same experimental conditions.

The ubiquitination of A168-170, MYH and CARP substrates and autoubiquitination of MuRF1^{WT} and MuRF1^{F148D} was detected as a molecular mass band shift on a gradient (4-20%) SDS-TGX electrophoresis gel (BioRAD) using Coomassie Brilliant blue staining (Figure 1.14). The autoubiquitination activity revealed a decrease in catalytic activity of MuRF1 due to the F148D mutation. In the sample of wild type MuRF1 the majority of a sample is observed as ubiquitinated (Figure 1.14 A, MuRF1), while in the sample of MuRF1^{F148D} strong band corresponding to non-ubiquitinated protein was observed with addition of low intensity higher molecular mass bands corresponding to autoubiquitination (Figure 1.14 B, MuRF1).

The qualitative comparative study of three independent targets of MuRF1 showed that titin A168-170 is the most ubiquitinated by MuRF1, where the ubiquitination is indicated with one clear band approximately 8 kD upstream from the titin sample, corresponding to mono-ubiquitination (Figure 1.14 A and B, Titin). Myosin mono-ubiquitination is observed with a lower intensity band compared to the titin (Figure 1.14 A and B, MYH7). Unlike titin and myosin samples where only monoubiquitination is observed, in CARP ubiquitination samples both mono- and di-ubiquitination corresponding bands were detected (Figure 1.14

A, CARP). Despite the lower catalytic activity, MuRF1^{F148D} mediated ubiquitination was detected for all three substrates, resembling target specific mono- and di-ubiquitination observed for the wild type. All the ubiquitination bands were confirmed to be ubiquitinated substrates by mass spectrometry.

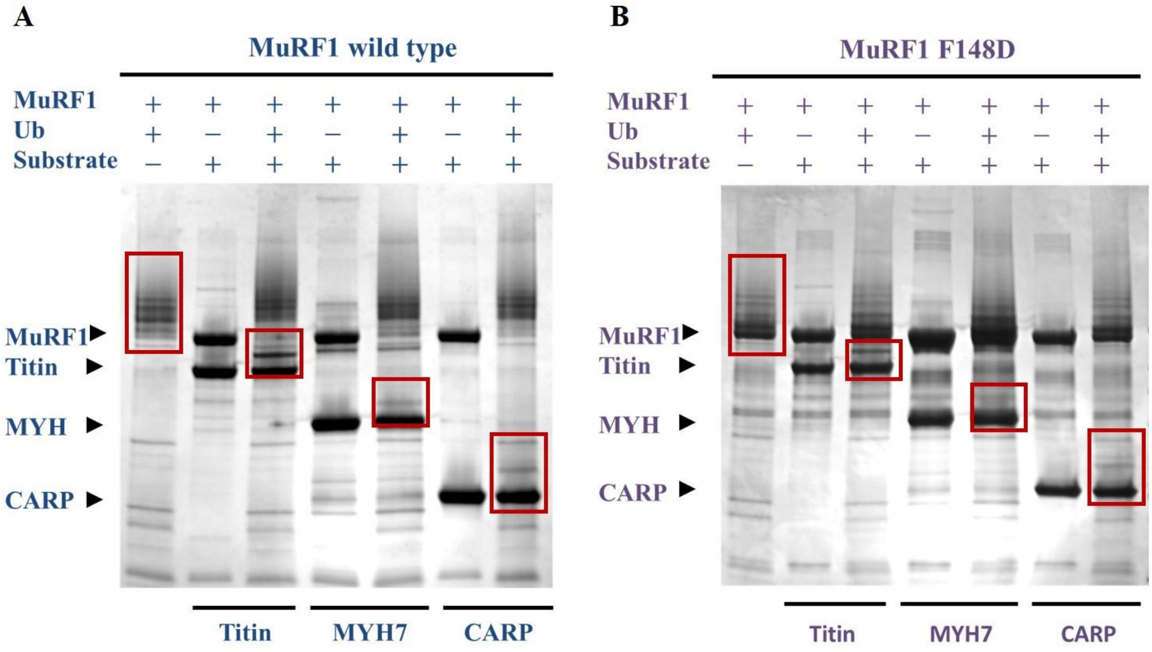


Figure 1.14 MuRF1 ubiquitination of muscle proteins.
Coomassie-stained gradient (4-20%) TGX precast gels (BioRAD) of MuRF^{WT} **A** and MuRF^{F148D} **B** ubiquitination reactions with titin A168-170 (Titin), Myosin heavy chain aa 964-1088 (MYH7) and CARP 166-306 (CARP). Red boxes indicate areas of the detected ubiquitinated species.

Titin was previously ubiquitinated *in vitro* and ubiquitination sites were identified (Bogomolovas *et al*., in preparation), but for the first time in this study ubiquitination sites for myosin heavy chain fragments and CARP were observed. For MYH7, a total of 19 ubiquitination sites were found and their potential role in filament formation is discussed in detail in Chapter 2. Ten ubiquitination sites were identified for CARP: K108, K132, K138, K143, K153, K203, K208, K211, K217, K281. To investigate whether a certain pattern of ubiquitination exists, detected sites were mapped on the predicted 3D model of C-terminal domains of CARP including amino acid residues 106-319. As there is no crystal structure of CARP available, a 3D-model of the CARP$^{106–319}$ fold was calculated using i-TASSER (Roy, Kucukural and Zhang, 2010) (estimated TM-score = 0.82 ± 0.08; confidence score *C*-score = 0.79) and used for mapping of MuRF1 specific ubiquitination sites (Figure 1.15). The C-terminal domain of CARP was predicted to fold into six ankyrin repeat (AR) motifs (Zhou *et al*., 2016). AR motifs (~33 residues) fold into two short α-helices linked by a tight turn, forming a helix-loop-helix α-hairpin structure. Based on this prediction, MuRF1 ubiquitination seems to target lysine residues located in the loops (8), rather than lysines located in the α-helices (2) (Figure 1.15 B).

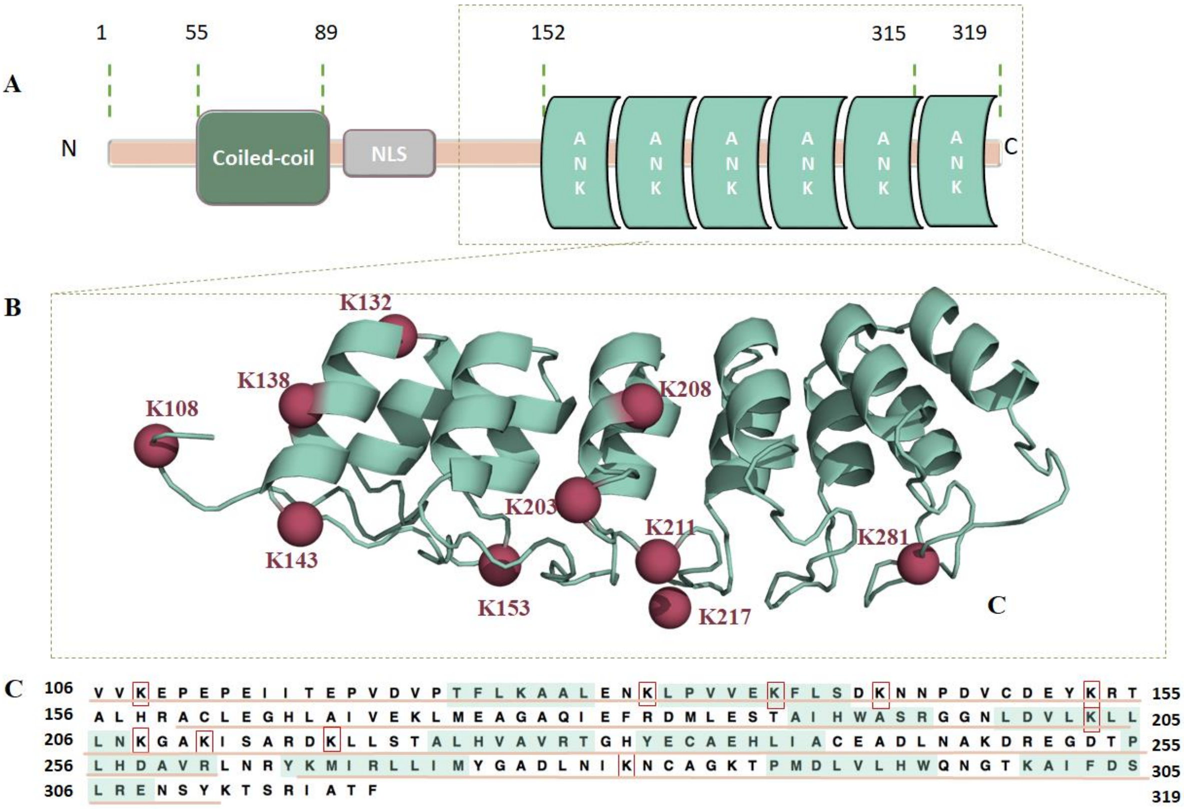


Figure 1.15 CARP ubiquitination by MuRF1.
A Schematic representation of domain composition of CARP with boxed construct used for ubiquitination. **B** Sequence based 3D model of CARP$^{106–319}$ predicted by i-TASSER (Roy, Kucukural and Zhang, 2010) showing CARP folding in six ankyrin repeat motifs in cartoon and MuRF1 ubiquitinated lysine C-α as red balls. **C** Sequence based secondary structure prediction by PCI-SS server with α-helices highlighted in green and ubiquitinated lysines (K) are shown in red boxes. (Underlined sequence corresponds to mass spectrometry identified peptides. (90.72% coverage) (Green, Korenberg and Aboul-Magd, 2009).

1.3.4 MuRF1 substrate binding requires the RING domain

MuRF1 acts as an E3 ubiquitin ligase, mediating ubiquitination of muscle sarcomeric proteins through a RING finger-dependent mechanism (Bodine and Baehr, 2014). To achieve its role, MuRF1 interacts with an E2 conjugating enzyme via its RING domain and with numerous substrates through, yet unknown mechanism. It was originally suspected that the coiled-coil fraction of MuRF1 interacts with the substrates ((Mrosek *et al.*, 2007), but no kinetic data has been provided up to date. To investigate the substrate recognition mechanism of MuRF1, BLI experiments for *in vitro* binding were designed and performed for MYH7^{964-}

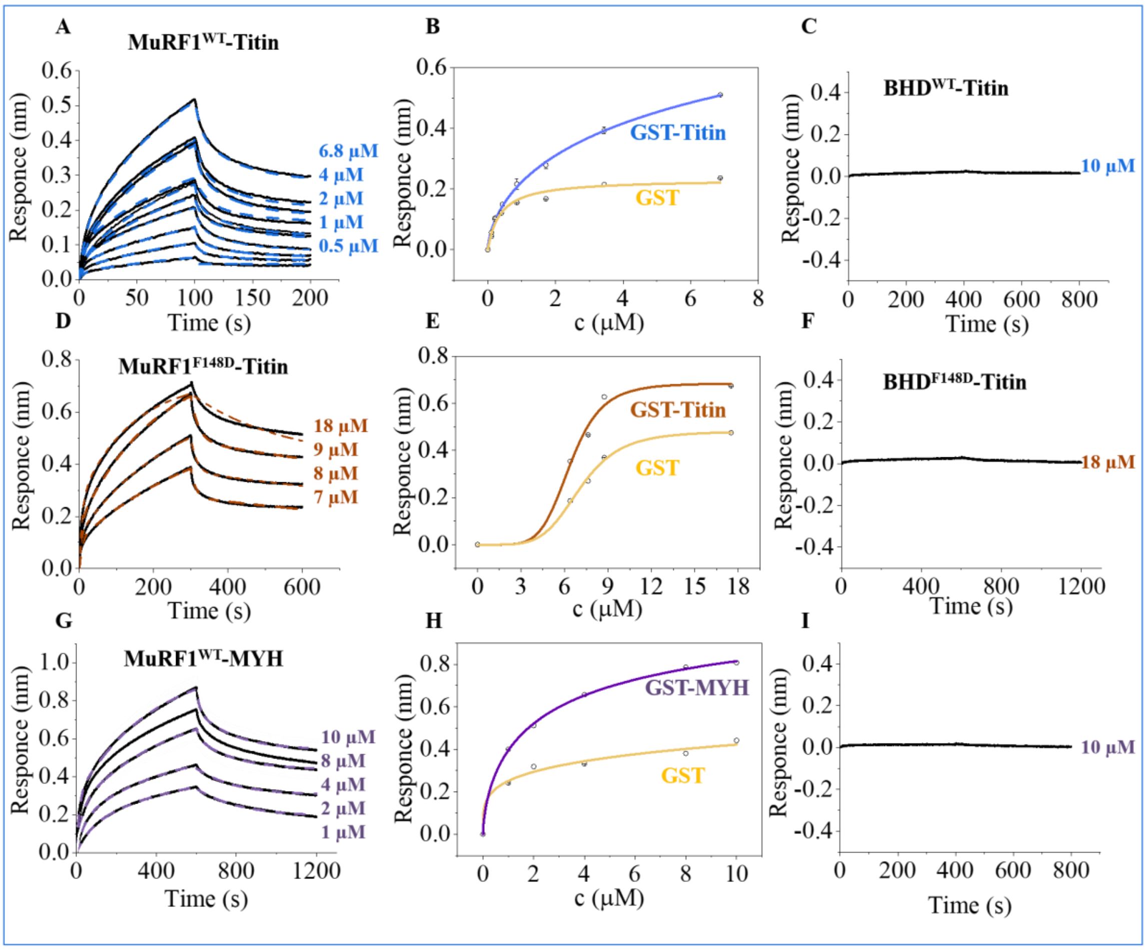
^{1088} and titin A168-170 with the wild type and F148D mutant full length MuRF1 and its C-terminal fraction BHD (B-box-coiled-coil) and N-terminal fraction RMB (RING-MFC-B-box).

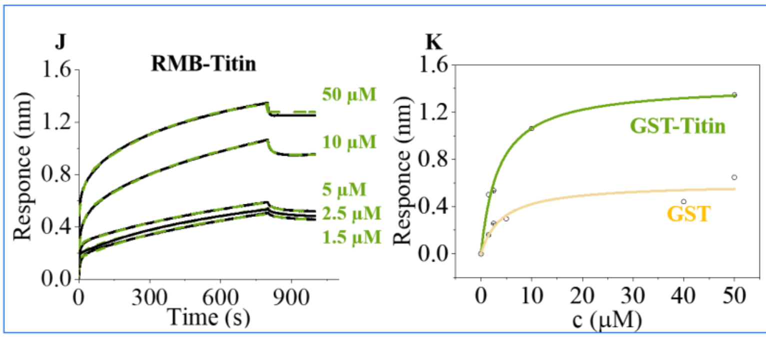
Biolayer interferometry is a relatively new label free ligand-capture method in biophysics, where light interferometry enables time resolved monitoring of protein interaction (Concepcion *et al.*, 2009). For experimental measurements, one of the interaction partners is loaded onto the optical fiber biosensor's surface (His$_6$-GST-titin A168-170 or His$_6$-GST-MYH7$^{964-1088}$), while the other binding partner remains in solution (MuRF1, BHD or RMB). Sensor surface thickness increases due to the protein interaction and directly affects BLI signal. Monitoring the signal change in the association and the dissociation steps in a concentration dependent manner allows for calculation of the kinetic parameters.

In order to optimize the loading step, an anti-GST sensors (ForteBio) were loaded with varying concentrations of His$_6$-GST-titin A168-170 or His$_6$-GST-MYH7$^{964-1088}$ and the optimal concentration was found to be 30 μg/mL during 150s loading time for both proteins. Loaded biosensors were transferred to the well containing the assay buffer to obtain baselines, prior to the association step. For the association step, biosensors were transferred to wells containing varying concentrations (1.0, 2.0, 4.0, 8.0 and 10.0 μM for MYH7$^{964-1088}$ and 0.5, 1.0, 2.0, 4.0, 8.0 μM for titin A168-170) of MuRF1^{WT}, followed by a dissociation step in the well containing the assay buffer. To observe the effect of F148D mutation on MuRF1 substrate binding affinity, biosensors loaded with titin A168-170 were transferred to a well containing varying concentration of MuRF1^{F148D} (7.0, 8.0, 9.0 and 18 μM) for the association steps, followed by a dissociation step in assay buffer.

Results show a significant shift of BLI signal (0.1-0.9 nm) recorded for both substrates, titin A168-170 (Figure 16 A) and MYH7$^{964-1088}$ (Figure 1.16 G). In addition, MuRF1^{F148D} binding curves show BLI signal shift similar (0.3-0.7 nm) to one observed for the wild type (Figure 1.16 D).

Strikingly, the removal of the RING domain reduces significantly MuRF1 interaction with both substrates, titin A168-170 (Figure 1.16 C) and myosin (Figure 1.16 F). No significant shift of the BLI signal (0.001-0.05 nm) was detected in the association step of titin A168-170 and MYH7$^{964-1088}$ with BHD^{WT}, nor BHD^{F148D}, indicating that the C-terminal domains of MuRF1 are not essential for the effective interaction. On the contrary, F148D mutation in the B-box of full length MuRF1 has little to no effect on MuRF1-titin interaction (Figure 1.16 D and E). When, however, the interaction between N-terminal RMB fraction of MuRF1 was studied with titin as a substrate (Figure 1.16 J) the strong signal was observed.





Interaction	Kd (µM)	R^2
MuRF1^{WT}-Titin	9.68±5.224	0.9998
MuRF1^{F148D}-Titin	6.78±0.091	0.9864
MuRF1^{WT}-MYH	1.148±0.067	0.9896
RMB-Titin	3.49±0.252	0.9917

Figure 1.16 The RING finger domain is sufficient for an effective MuRF1 substrate binding.
Biolayer interferometry real-time binding data of the full length MuRF1, RING-less MuRF1-BHD and CC-less MuRF1-RMB to its substrates obtained from an anti-GST sensors tip covered with: **A, C, D, F** and **J** Titin A168-170 (Titin); **G** and **I** MYH7$^{964\text{-}1088}$ (MYH). Raw data of the MuRF1 concentration dependent binding association and dissociation steps is shown as a black line and fitting results colored dash lines (**A, D** and **G**). No binding was detected for the MuRF1-BHD construct with either of the substrates (**C, F** and **I**), while strong signal was detected for the MuRF1-RMB (**J**). **B, E, H** and **K** show steady-state analysis of real-time binding data obtained in **A**, **D, G** and **J**, respectively by plotting the response derived from the raw binding data *versus* protein concentrations. **L** Affinity constraints (K_d) and quality of the fit (R^2) are shown in the table.

These results confirm the ability of MuRF1 to efficiently bind its substrates. The real time binding curves recorded in the BLI measurements were fit locally to the heterogeneous binding model (Figure 1.16 A, D and G dashed colored lines) and the dissociation constants were calculated by plotting the BLI signal values observed at the end of dissociation step against the concentration and corresponding fitting (Figure 1.16 B, E, H). Calculated Kd values show that the binding affinity for both substrates is in the micro molar range, with myosin (1.148 ± 0.067 μM) having slightly higher affinity than titin (9.844 ± 5.224 μM) and that the F148D mutation has no effect of binding affinity of MuRF1 for titin (6.780 ± 0.091 μM). More, obtained data shows that upon the deletion of RING domain binding affinity for titin and myosin heavy chains are lost. Moreover, the RMB binding was detected as a high signal (Figure 1.16 J) and the calculated affinity constant was in the same range of affinity as the full length MuRF1 (3.49 ± 0.252 μM), indicating that the RING domain is a key mediator for MuRF1 effective substrate binding.

1.3.5 The role of the RING in MuRF1

The BLI data showed that C-terminal fraction of MuRF1 is not sufficient for substrate recognition while the N-terminal fraction RMB resembles the binding of the full length MuRF1, indicated the potential new role of the RING domain. RING domain of TRIM E3-ubiquitin ligases is known to stabilize closed conformation of E2-ubiquitin complex and prime the thioester bond for ubiquitin transfer to occur (Dou *et al.*, 2012; Plechanovová *et al.*, 2012). The E2-ubiquitin complex is a dynamic structure that exists between two extreme conformations, open and close, and the RING type of E3-ubiquitin ligases acts as an adaptor that brings E2-ubiquitin complex in proximity of substrate and promotes ubiquitin transfer (Berndsen and Wolberger, 2014). To investigate the RING domain role in the substrate binding, in cell colocalization experiments and immunoprecipitation were performed with Dr. Bruno Manso. As the BLI data indicated that the RMB fraction is sufficient for the substrate binding *in vitro*, it was further investigated if the RMB fraction is also sufficient for the ubiquitination reaction *in vitro*. Ubiquitination assays of the RMB construct with titin as substrate were performed with Dr. Barbara Franke and Francesca Ward.

For colocalization of titin A168-TK with MuRF1 and its RING-less (MuRF1$^{\Delta R}$) RING-MFC-less (MuRF1$^{\Delta R\text{-}MFC}$) constructs were transfected into H9C2 cells and visualized using confocal microscopy. H9C2 cells (ATCC CRL-1446) were transfected with C-terminally tagged MuRF1-pmCherry, MuRF1$^{\Delta R}$-pmCherry, MuRF1$^{\Delta R\text{-}MFC}$-pmCherry or N-terminally tagged pEGFP-A168-TK. Samples were stained with DAPI, fixed with 4% paraformaldehyde and mounted on microscope slides using Aqua-poly/mount media (Polysciences). Sample imaging was done by laser scanning confocal Microscope (LSM700, Carl Zeiss) using a 63x/1.4 oil immersion objective. Deletion of MuRF1-RING and MuRF1-

RING-MFC fraction seems to have an acute effect on MuRF1 localization and MuRF1 and titin A168-TK colocalization pattern *in vivo.* (Figure 1.17 A Merged). Diffused fluorescence signal for MuRF1$^{\Delta R}$ (Figure 1.17 B) and MuRF1$^{\Delta R\text{-}MFC}$ (Figure 1.17 C), however, indicate that presence of the RING domain may be required for a high affinity substrate targeting of MuRF1. In addition, co-immunoprecipitation experiments confirmed that the deletion of the RING/MFC domain abolishes MuRF1 interactions with titin A168-TK (Figure 1.18), whilst deletion of coiled-coil fraction has no effect on the binding (Figure 1.18). Together these data show that loss of binding affinity upon deletion of RING/MFC domain observed in vitro BLI experiments is in consistency with the *in vivo* data.

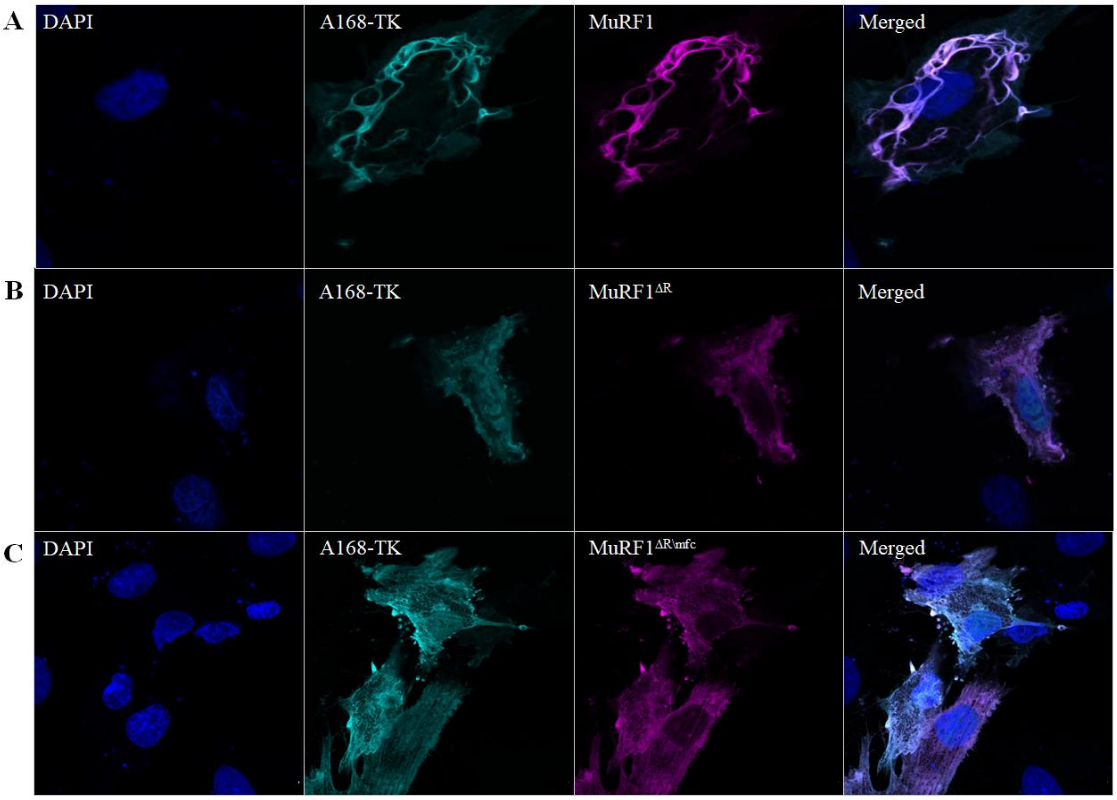


Figure 1.17 The RING domain regulates cellular targeting of MuRF1
Co-localization of MuRF1 and titin A168-TK coexpressed in H9C2 cells. MuRF1-mCherry and its truncated variants MuRF1^{DR} -mCherry, MuRF1$^{\Delta R\text{-}MFC}$-mCherry are shown in pink, titin EGFP-A168-TK localization is shown in green and DAPI staining allows the visualization of the nucleus. **A** Coexpression of MuRF1 (pink) with EGFP-A168-TK (green) shows colocalization (Merged) while truncated variants of MuRF1 lacking their N-terminal fraction (spanning the RING domain or the RING-MFC fractions) result in an uneven mCherry diffuse background (**B** and **C**). pEGFP and mCherry fluorescence was excited using a 488 nm and 561 nm diode laser, respectively (Merged). Emitted fluorescence for GFP (up to 550 nm), and mCherry (from 578 nm), as well as transmitted light images was recorded at Nyquist sampling intervals (85 nm pixel size). Single channel and composite images were prepared using Fiji (Schindelin *et al.*, 2012). (Figure provided by Dr. Bruno Manso).

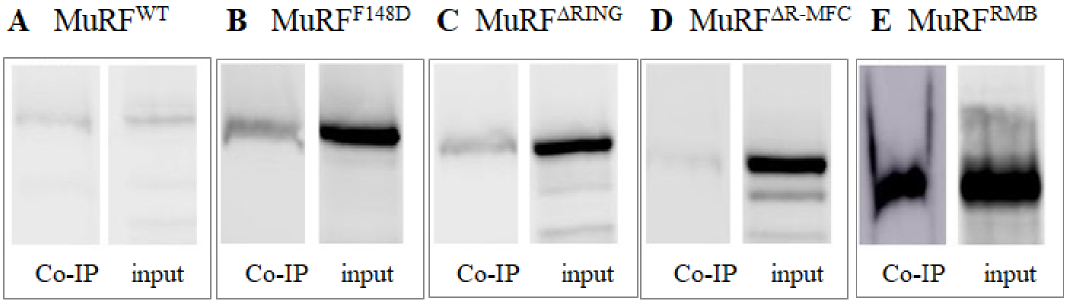


Figure 1.18 Co-immunoprecipitation of MuRF1 and titin A168-TK.
Cell lysates from H9C2 cells transfected with GFP-A168-TK and each of the MuRF1-mCherry constructs were incubated with α-GFP nanobodies immobilized on sepharose beads. Binding of titin kinase (Titin A168-TK) was investigated by western blot using an anti-mCherry antibody with full length MuRF1^{WT} (**A**), MuRF1^{F148D}. Deletion of the RING (**C**) and RING/MFC (**D**) domains from MuRF1 reduces its ability to bind to A168-TK, while a construct composed of only the RING, MFC and B-box domains shows strong binding (**E**).

For determining which domains in MuRF1 are sufficient for the catalytic activity RING/MFC/B-box (MuRF1-RMB) and B-box/CC (BHD) constructs were subjected to the ubiquitination assays described in Section 1.2.2. except reaction mixtures contained 3 µg of MuRF1-RMB and 3 µg of substrate. Surprisingly, SDS-PAGE gel of MuRF1-RMB ubiquitination reaction showed a high level of both autoubiquitination and substrate ubiquitination activity (Figure 1.19) and no activity for the BHD construct. Obtained results confirm that the RING/MFC/Bbox is sufficient fraction of MuRF1 for the catalytic activity.

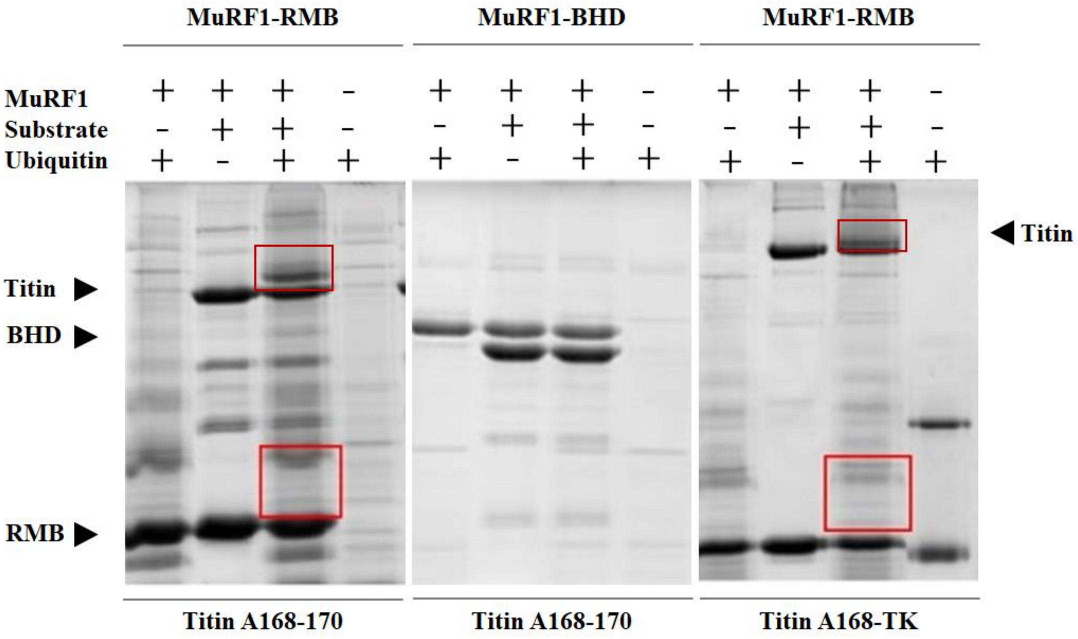


Figure 1.19 Ubiquitination activity of the N-terminal domains of MuRF1-RMB.
RMB construct shows autoubiquitination activity, as well as substrate ubiquitination for Titin A168-170 and titin A168-TK. RING-less construct-BHD has no catalytic activity. Red boxes indicate ubiquitinated species that appear upon addition of ubiquitin in the reaction mixture (+). (Figure provided by Francesca Ward)

1.4 Discussion

The Biolayer interferometry (BI) study was successfully designed and optimized for the real time monitoring of the MuRF1 interactions with myosin heavy chain (MYH7) and an A-band fraction of titin (titin A168-170). Surprisingly, BLI results provided the strong evidence for the RING domain being necessary for an effective binding of MuRF1 with substrate proteins. One of the key steps in understanding all biological processes are the underlying interactions between biomolecules. Despite the high interest, MuRF1 interactions with its substrates have only been observed in the pull-down assays and co-immunoprecipitation (Koyama *et al.*, 2008; Witt *et al.*, 2008). Although powerful in detecting strong interaction partners, these methods give no significant information about the nature, mechanism and kinetics of the interaction. MuRF1, together with other TRIM/RBCC proteins demonstrated a tendency to aggregate, which can be a challenge for biophysical characterization experiments. Standard methods used for protein-protein interaction studies, such as ITC (isothermal calorimetry), require high molecular concentrations, significant amounts of samples and do not differentiate between protein interaction and aggregation in the sample, and as such are not suitable for the TRIM proteins. Hereby, through experimental design and optimization, it was shown that the relatively novel sensor based method, BLI can be used for highly sensitive investigation of the difficult MuRF1 sample in (i) obtaining equilibrium dissociation constants (K_D) for the interactions between MuRF1 and its substrates (MYH7 and titin), (ii) the qualitative analysis of substrate recruiting region of MuRF1 required for effective binding of myosin and titin. These applications of the BLI to study MuRF1 highlight the advantage of this technique in characterization of difficult protein samples and allow for dissecting the target preferences and targeting mechanisms that MuRF1 employs for the substrate recognition. Furthermore, establishing the operative method for analytical studies of samples with tendency to oligomerize and aggregate, opens up new avenues not just for the MuRF1, but also for other TRIM family members.

Surprisingly, the BLI experimental setups led to the discovery of the RING domain in MuRF1 being essential for the effective substrate recognition. There has been no evidence of the RING domain having a role in the substrate recognition, on contrary previous studies suggested the coiled-coil (CC) domain of MuRF1 to mediate protein interactions (Mrosek *et al.*, 2007; Witt *et al.*, 2008). Mrosek *et al.*, used the ITC for binding experiments of MuRF1-CC and titin A168-170 reporting the high affinity binding (Kd=35 nM). However, the ITC results also reported an unusual molar binding ratio of 10 (MuRF1):1 (titin) and potential issues in determination of protein concentration (Mrosek *et al.*, 2007). Together with previously discussed high tendency of the TRIM proteins to aggregate, the lack of possibility in the ITC experiments to differentiate between an aggregation and the specific interaction, and in the light of the new evidence, it might be assumed that the interactions observed in ITC experiments were affected by the sample aggregation. In the BLI experiments, no technical issues have been detected and all the data were reproducible, however there was no significant binding detected for neither CC, nor B-box-CC (BHD) area of MuRF1, while the strong binding was detected for the RING-MFC-B-box (RMB) suggesting that the RING domain is the key contributor for the substrate binding.

Additional experiments, including *in vivo* colocalization data and *in vitro* ubiquitination data, strongly supported the hypothesis of the N-terminal fraction of MuRF1 being sufficient to bind substrates and also to catalyze ubiquitination of the titin substrate and autoubiquitination, further suggesting that BHD is not required for neither, substrate targeting nor the catalytic activity of MuRF1. Similar phenomena was observed for TRIM25, where the catalytic activity exists independent of the C-terminal fraction of the protein, whilst could be allosterically enhanced by CC and B-box via oligomerization (Li *et al.*, 2014). Together these results indicate that the RING domain is a key fragment of MuRF1 for both substrate targeting and catalytic activity. Unlike the majority of TRIM E3 ubiquitin ligases, MuRF1 does not have an additional C-terminal domain specified for substrate targeting, the RING domain might adopt an additional role playing as an interaction interface for the protein substrates. This study revealed a novel role of a well-studied and established protein fold RING finger, while opening up a new question of a BHD role in MuRF1.

TRIM ligases dimerize through their coiled-coil region and the B-box domain, and while self-association has been considered as a difficulty of sample handling and preparation, it has also been proposed to be an enhancer for catalytic activity in TRIM proteins (Streich *et al.*, 2013). A systematic analysis of the relationship between MuRF1 oligomerization and catalytic activity is currently missing, thereby disruption of the B-box interacting interface was an attempt for a simple and elegant alteration of MuRF1 high order assembly. Introducing the F148D mutation in the B-box showed a strong effect on catalytic power by decreasing MuRF1 autoubiquitination and substrate ubiquitination activity. Although full length MuRF1^{F148D} showed lower tendency to precipitate and higher stability in the purification process, it remained unclear what is the F148D mutation effect on the partial disruption of MuRF1 self-assembly. The size exclusion chromatography as a method is not sensible for detecting differences in such high molecular weight oligomers, therefore the future work is needed to determine in which extend oligomerization is affected and how does it impair catalytic activity of MuRF1. It might be foreseen that MuRF1 forms a substrate recruiting interface by quaternary domain rearrangement with the RING domain being a main contributor. It is essential for future studies to address this hypothesis and its potential as a foundation for the therapeutic strategies.

2

MuRF1 ubiquitination of myosin heavy chain - a RING for disassembly

2.1 Introduction

Skeletal muscle tissue is a core of the human mobility. It accounts for over 40% of a body mass and as such, it also represents the amino acids reservoir, muscle proteins can be quickly degraded and resulting amino acids can be redirected for the further use (Frontera and Ochala, 2015).

To maintain its function, the muscle tissue undergoes remodeling, replacing damaged proteins from the sarcomere with the newly synthesized. Some misfolded proteins can be rescued from the degradation via chaperon mediated refolding, but most of the damaged proteins are removed from the sarcomere by the ubiquitin proteasome degradation pathways (Haas, Woodruff and Broadie, 2007). The fine balance between rates of protein degradation pathways and protein synthesis keeps the sarcomere in shape (Hnia, Clausen and Moog-Lutz, 2019). However, under stress and disease conditions, protein degradation becomes the dominant process. If this shift of balance remains impaired, due to enhanced degradation, protein mass in the muscle decreases continuously, until the muscle becomes dysfunctional. This state is known as muscle atrophy and has been described in Chapter 1. Shortly, muscle atrophy occurs due to numerous stressed conditions such as cancer, immobilization, disuse, cachexia, dexamethasone treatment and become one of the most occurring life threatening conditions of the modern age. Yet, all available treatments for muscle atrophy have shown low efficiency (Ding *et al.*, 2018). In the first chapter of this thesis, the main focus was on the key actor of muscle atrophy and potential drug target E3 ubiquitin ligase MuRF1. Here, the main focus will alternate towards the best known target of MuRF1, myosin heavy chain. This research focuses on understanding what features define myosin heavy chain for being targeted by MuRF1 and what is the mechanism that enables non diffusible molecules, like myosin heavy chain, hidden and well packed in the sarcomere, to be ubiquitinated and degraded. Understanding these central characteristics will help to establish novel strategies for the drug development. Therefore, it is essential to understand myosin and its role in the sarcomere.

The muscle sarcomere is the basic structural and functional unit of a striated muscle that through its specific protein arrangement enables every contraction in the human body (Gordon, Huxley and Julian, 1966). The architecture of a muscle sarcomere is based on the strict arrangement of its main components: I band, A band, Z disk and M line. The bearing pillars which define the borders of each sarcomere are the Z disk (Figure 2.1.). The main components of the Z-disc are titin and alpha-actin, with additional regulatory proteins, establishing the connection between A band and I band components. Titin is a giant sarcomeric protein that has its C-terminal end based in the Z disc and from there it crosses along the whole sarcomere length, acting as a molecular spring (Granzier and Labeit, 2006).The A- band and I-band alternate along the sarcomere. The I-band is an actin based filament (also labeled as thin filament) and the A-band is a myosin based filament (thick filament). The M-line or middle zone is a place where the myosin filament meets other proteins such as titin and meromyosin, which further assist in filament cross-linking, sustaining the structure of a thick filament (Craig and Megerman, 1977).

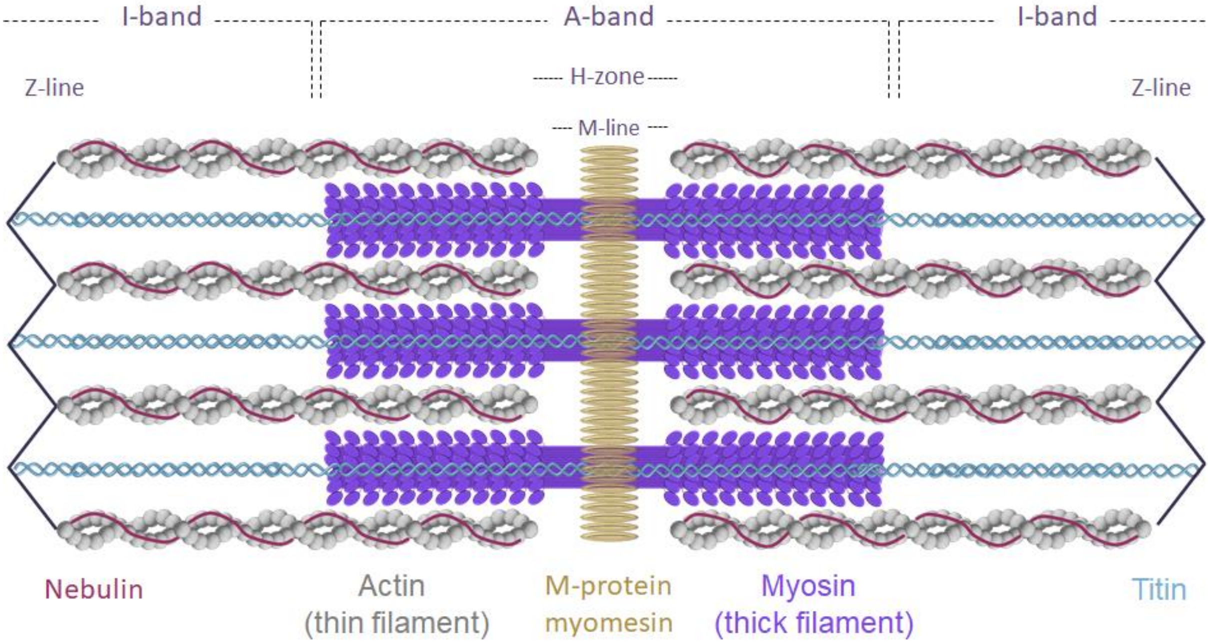


Figure 2.1 The striated muscle sarcomere organization.
Z-lines border each sarcomere and assist titin communication between two sarcomeres. In the middle of each sarcomere is a M-line. Myosin molecules form a bipolar structure called the thick filament in which myosin heads are protruding from the line of thick filament to interact with actin molecules embedded in the thin filaments. Actin-myosin interaction is the key to muscle contraction.

Thick filaments in striated muscles of vertebrates are about 1600 nm long with a diameter of 30 nm and contain hundreds of muscle myosin II isoform (MYH7) hexamers together with a number of accessory proteins (Al-Khayat HA. 2013; Squire JM. 1973). The thick filament is a bipolar formation, consisting of myosin tails meeting in the M- line and myosin head domains protruding on the surface in a helix lattice manner in the A-band region. Myosin is in the core of the thick filament formation and function. Hence, in order to understand the thick filament, it is essential to understand myosin.

2.1.1 Myosin is a power force forming the nano-machinery in muscle

Myosin is a protein abundantly found in the muscle, which through its dual role, as a molecular building block and as an energy production unit for contraction and intracellular motility, was brought to attention decades ago (Pollard and Korn, 1973). Yet, today myosin is still one of the top interests in diverse fields. Muscle myosin belongs to a family consisting of over 15 classes and they are found to be highly conserved within the species (Hartman and Spudich, 2012). Tight regulation of the isoform expression is essential for the variability of the muscle force required by different muscles (Fili and Toseland, 2019). Until now eleven muscle myosin isoforms have been identified in vertebrates' skeletal muscle, each encoded by separate genes (Lee *et al.*, 2019) and the all belong to a fraction called conventional myosins, as they have a propensity for building filaments.

Conventional myosin is a hexamer composed of two heavy chains (MyHC) forming a parallel dimer complexed with two essential and two regulatory light chains (MLC). Myosin light chains are small Ca^{2+} binding proteins that have regulatory roles in the myosin catalytic activity (Bagchi, Kemp and Means, 1989). MyHC N-terminal head domain is catalytically active and interacts with actin molecules from the thin filament enabling the muscle contraction, in an ATP dependent manner (Ruppel and Spudich, 1996). When the myosin head domain binds an ATP molecule, it is in the relaxed state in the thick filament, but upon hydrolysis of the ATP, the head domain changes its conformation by extending through the thick filament into the thin filament and binds to the myosin-binding sites in actin molecules, this process is known as the power stroke (Cooke and Holmes, 1986). During the power stroke, myosin heads swing and pull the actin molecule to slide towards the M-line and ADP is released. For myosin to detach from actin, a new ATP molecule has to bind to the ATP binding site in the head domain. ATP dependent interaction between myosin head domain and actin mediates force and contraction production through the repetitive cycles (Spudich, 2001). Each myosin-actin interaction complex is called a cross-bridge and the total force produced by a sarcomere is linearly dependent on the number of cross-bridges formed (Colombini *et al.*, 2010). The myosin head domain continues to the neck domain, which assists in the head movements and is followed by a long tail domain. As the head domain is directly involved in the muscle contraction, the neck domain is assisting in head domain positioning, it remains obscure what is the role of the long rod domain and whether it is essential for ensuring complete regulation of myosin function.

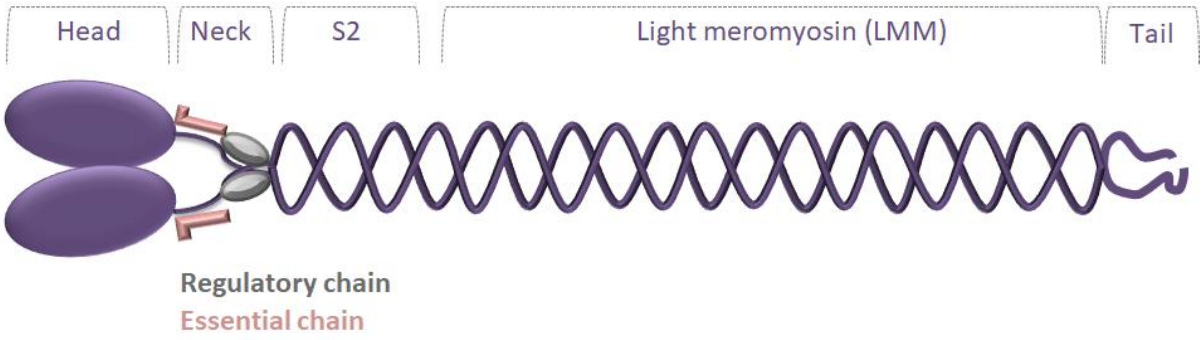


Figure 2.2 Myosin domain composition.
Conventional myosin is a hexamer composed of a myosin heavy chain dimer, two regulatory and two essential chains, called myosin light chains. Myosin heavy chains fold into a catalytically active head domain, flexible neck domain and long rod domain. The rod domain consists of S2 or hinge domain, light meromyosin and the disordered tail domain.

In the assembly of a sarcomere, myosin molecules first join in the bipolar thick filament, which by further assembly forms the A-band of the sarcomere. It is suggested that the thick filament formation is directed mainly by the rod domain of myosin (Taylor *et al.*, 2015). As the rod domain folds in a long coiled-coil, for understanding its role in the filament assembly and contraction, it is essential to understand the principle in the nano mechanics of a coiled-coil fold (Root *et al.*, 2006).

2.1.2 Coiled-coils

In 1952, Francis Crick described a new type of quaternary protein structure found in fibrous proteins, which he called the coiled-coil (CC). Muscle myosin, together with alpha keratin and tropomyosin, are among the first proteins in which CC's was identified and it have served as classical examples in the CC studies. By supercoiling an a-helix, one could obtain a repeating structure in which two or more adjacent a-helices pack together in a knobs-into-holes arrangement (CRICK, 1952). The canonical coiled-coil is characterized by a pattern of alternating hydrophobic and hydrophilic residues within a seven-amino acid (heptad) repeat.

The position of an amino acid residues in the heptad repeat are labelled ***a***, ***b***, ***c***, ***d***, ***e***, ***f*** and ***g***, where hydrophobic amino acids occupy positions ***a*** and ***d*** (Figure 2.3). The conservation of hydrophobic residues alternates between three and four residues (average 3.5) and is close to the 3.6 amino acids per turn periodicity of a regular a-helix. Consequently, helices deriving from such repeating sequences exhibit a distinct amphipathic character, with both hydrophobic and polar faces. The association of two helices via their hydrophobic faces drives the coiled-coil formation. However, in order to pack two helices together and maintain hydrophobic contacts, the knobs-into-holes packing of side chains requires that these residues occupy equivalent positions, turn after turn. The distance required for the superhelix to complete a full turn is called the pitch. This allows the positions of side chains to repeat after two turns, or seven residues, instead of drifting continuously on the helical surface (Lupas and Gruber, 2005).

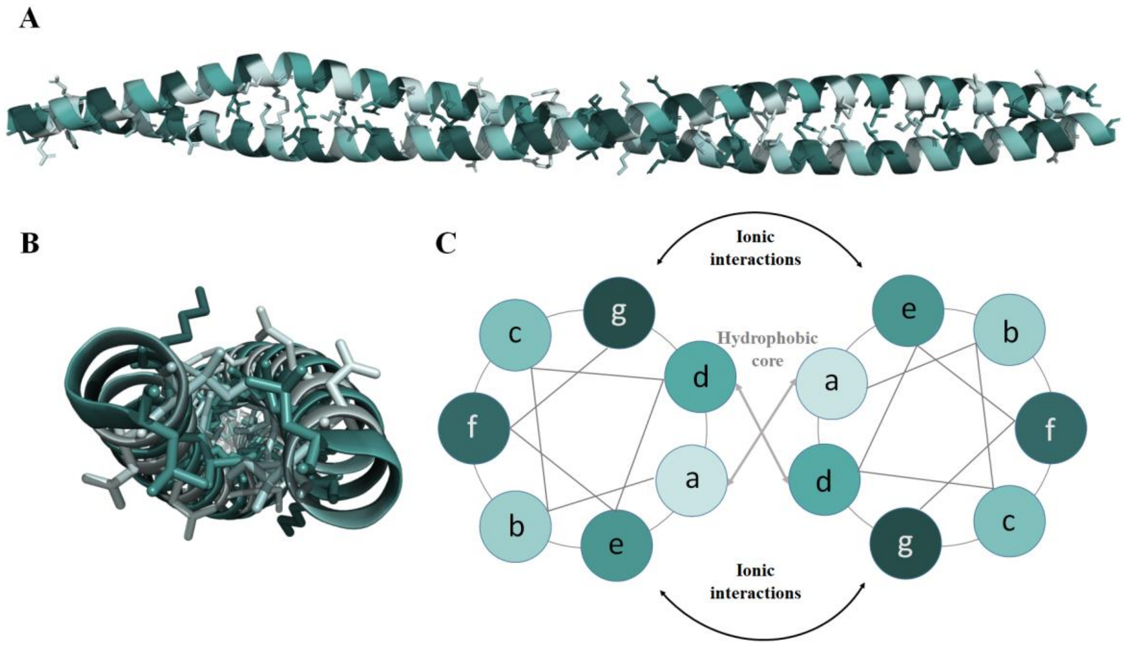


Figure 2.3 The coiled-coil structure of myosin II.
A Crystal structure of Myosin 838-964 (PDB ID: 2fxm). **B** C-terminal side view of Myosin 838-964 coiled-coil. **C** The helical wheel as a representation of the parallel coiled-coil heptad repeat showing how the positions **a**–**g** appears when viewed from the C-terminal end of the parallel helices, indicating hydrophobic interactions (**a-d**) establishing the core supported by ionic interactions between **e** and **g** positions, and how **b**, **c** and **f** positions are solvent exposed on the opposite side of a hydrophobic core. All panels are colored as in C, showing the corresponding heptad position of each residue.

Discontinuities in a CC, such as insertions of one or more residues into the heptad pattern, can cause disturbance in the higher order packing of the CCs (Strelkov and Burkhard, 2002). Insertions consisting of the one residue are called *skips*, the three-residue insertions are *stammers*, while the four-residue insertions are *stutters*. All discontinuities can result in the backbone continuity perturbations or in the side-chain packing disturbance. However, in some cases discontinuities are essential for introducing flexibility which is further required for the high order structure assembly (Seo and Cohen, 1993) and one well studied example of this phenomena is the myosin heavy chain (Taylor *et al.*, 2015).

Initial observations of myosin led to the assumption that it forms a simple coiled-coil motif, however, soon after the myosin sequence became available and was analyzed, it was clear that the structure must be more complex than a conventional CC (McLachlan and Karn, 1982). Based on a coiled-coil prediction, myosin showed areas with the low CC propensity, based on an uncanonical distribution of hydrophobic and charged residues (McLachlan and Karn, 1982; Lupas, Van Dyke and Stock, 1991). Four skip residues (T1188, E1385, E1582 and G1807) have been identified and assumed to introduce a local flexible hinge into the cardiac myosin rod required for the filament formation (Taylor *et al.*, 2015) (Figure 2.4). Furthermore, mutations in skip residues areas were closely related with a development of cardiomyopathies and numerous muscle myopathies (Wolny *et al.*, 2013; Peckham, 2016).

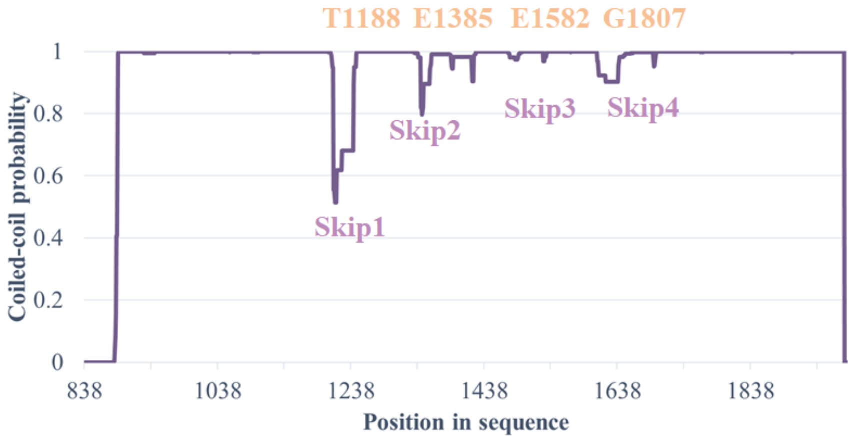


Figure 2.4 Marcoil based myosin rod domain sequence prediction of coiled-coil propensity.
The Marcoil algorithm is based on a Hidden Markov Model (Delorenzi and Speed, 2002) and the results show four skip areas.

The folding of a coiled-coils is a highly cooperative process, triggered by a higher CC propensity sequence within the protein, termed a trigger sequence. It was assumed that the myosin rod domain consists of the trigger sequence between 1301 and 1330 (Wolny *et al.*, 2013) that would assist its coiling. It was further shown that smaller myosin fragments, not containing the trigger sequence, could have a lower propensity to form a CC (Korkmaz *et al.*, 2016). Thereby, the proper folding of recombinantly produced myosin fragments became one of the main obstacles in completing myosin rod domain structures. Moreover, myosin fragment production exhibits a strong salt-dependent aggregation in the form of paracrystals, mainly due to the periodic clusters of positively and negatively charged residues along the sequence. Combining a highly controlled ionic strength of solutions and introduction of a folding domain such as leucine zipper from GCN4 transcription factor that can be fused to a CC protein (O'Shea *et al.*, 1991), enabled high resolution structural information to be obtained for more than 30% of the myosin rod domain. Completing the myosin rod domain structure would be of great asset to fully understand thick filament assembly and disassembly and the mechanism behind related pathologies.

2.1.3 Myosin assembly and disassembly

Myosin filament formation is assumed to occur in a three step process, which includes myosin dimerization, antiparallel tail-tail assembly and a high order assembly (Srikakulam and Winkelmann, 2004). Myosin rod domain has the leading role, while the head domain proper folding is not required for a filament formation (Chow *et al.*, 2002). Myosin dimerization is led by CC formation. Further assembly of myosin CCs is mediated by the assembly competence domain (ACD), a highly conserved C-terminal 29 amino acid sequence including 4 consecutive negatively charged residues surrounded by positive charges (Sohn *et al.*, 1997). An ACD mediated antiparallel tail-tail interaction represents a nucleation site for the filament assembly. Parallel assembly of myosin molecules at the nucleation site creates a bipolar filament, with myosin heads facing outward and the long rod regions facing inward.

Although four different hypotheses have been proposed and structural information has become more available (Squire, Paul and Morris, 2017), the detailed mechanism of the thick filament assembly into the sarcomere remains unclear (Sanger *et al.*, 2005). However, myosin filament formation can occur spontaneously *in vitro*, but incorporation in the sarcomere seems to require association of other sarcomere proteins, such as nebulin and titin (Weinert *et al.*, 2006). Additionally, a proper folding of the myosin head is mediated by the molecular chaperones HSP70, Hsp90 and UNC-45 (Barral *et al.*, 2002; Hutagalung *et al.*, 2002), and although it is not involved in filament formation, myosin head formation is necessary for a proper function.

In the muscle cell protein degradation is predominantly driven by ubiquitin proteasome systems (UPS), both in physiological and pathological conditions, such as muscle atrophy. The half-life of myosin in skeletal muscle is estimated to be 3h and it is assumed that over 90% of myosin molecules from a single thick filament will be replaced within 12h (Ojima *et al.*, 2015). Non-muscle myosin filaments have an even faster turnover than the muscle myosins indicating that the rate limiting step of myosin degradation could be an actual release from the sarcomere (Hu *et al.*, 2016). It was originally proposed that cytosolic proteases may mediate myosin release from the sarcomere (Lecker *et al.*, 1999), enhancing the rate of ubiquitination, but recent studies showed that ubiquitination can effectively occur in the absence or inhibition of the proteases (Plant *et al.*, 2009). In muscle atrophy the E3 ubiquitin ligase MuRF1 was found to be the main contributor to myosin degradation (Bodine and Baehr, 2014), yet the mechanism behind myosin ubiquitination and filament depletion remains obscure.

Therefore, it seems to be essential to resolve the fundamental mechanisms by which myofibril formation is controlled at the molecular level, in order to develop an effective approach for therapeutic purposes.

Aims

Myosin filament formation is dependent myosin heavy chain coiled-coil. Specific positions in the heptad repeat of the coiled-coil are more prone to be exposed to the surface and participate in inter molecular interactions, whilst amino acids occupying certain positions are prone to participate in the formation of salt bridges that stabilize the coiled-coil. Despite myosin filament being well packed in the pseudocrystaline structure, myosin heavy chain was found to be heavily ubiquitinated during the muscle atrophy. It remained unclear how MuRF1 targets myosin heavy chain. This chapter aims to investigate: (i) if MuRF1 has a specific pattern of recognition in the myosin heavy chain; (ii) Can MuRF1 attack lysine residues in the heptad position of myosin coiled-coil that are engaged in the salt bridge formation in; (iii) How does MuRF1 activity affects the sarcomere architecture in the absence of activated degradation pathways;

2.2 Material and methods

2.2.1 Protein production and purification

Full length myosin heavy chain (MYH7) extracted from rabbit muscle was commercially available (Sigma). Myosin heavy chain fragments ($MYH7^{x-y}$) were recombinantly produced in *E. coli* and purified for further use (figure 2.1.1). All myosin fragments were cloned in the pOPINJ vectors by Dr Thomas Zacharchenko prior to this work (Addgene plasmid 26045). Fragment definition and used abbreviations of myosin fragments are schematically described in Figure 2.5.

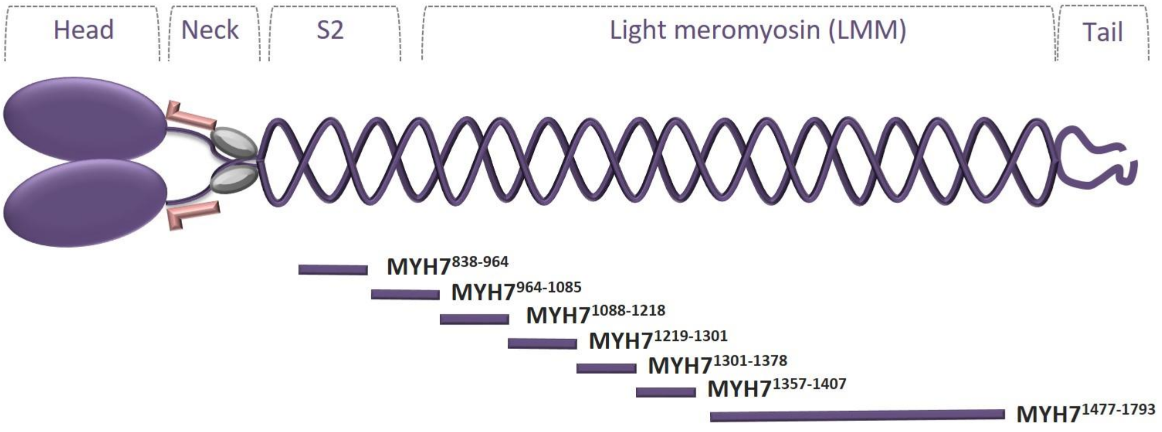


Figure 2.5 Human Myosin heavy chain fragments (MYH7).
Schematic representation of MYH7 fragments, location, and definition.

2.2.1.1 Myosin heavy chain fragments purification

Vectors containing the DNA sequences of the desired constructs were transfected into *Escherichia coli* Rosetta (DE3) PlacI cells (Novagen). The pOPIN-J contains N-terminal His_6-(glutathione S transferase) GST tag and 3C protease cleavage site prior to the protein of interest (Appendix A). Transformations used Luria-Bertani (LB) medium supplemented with 2% glucose. Transformed cells were inoculated in a 1:100 ratio into the LB supplemented with 50 µg/mL ampicillin and 35 µg/mL chloramphenicol and were grown at 37°C to an OD_{600} = 0.6. Protein expression was induced with 0.5 mM IPTG and cells were further grown for 18h at 18°C. The harvest was obtained by centrifugation at 5000 rpm (9760 xg) at 4°C. Bacterial pellets were resuspended in a lysis buffer (50mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT) containing an EDTA-free protease inhibitor cocktail (Roche). Cells were lysed by sonication on ice, followed by clarification of the lysate by centrifugation at 18000 rpm (47850 xg). Cell lysates were applied to a 5mL HisTrap HP (GE Healthcare) equilibrated in a lysis buffer containing 20 mM imidazole. Sample elution was carried out with a continuous imidazole gradient in the concentration range 20-1000 mM. Eluted protein was subjected to a buffer exchange via dialysis against 50mM Tris pH 8.0, 50 mM NaCl, 2 mM DTT. Affinity tags were removed for ubiquitination assays by overnight incubation with 3C protease at 4°C, while samples prepared for the Biolayer interferometry (BLI) assays preserved their His_6-GST tag. Both samples were further purified by ion exchange chromatography on a 5 mL MonoQ® 5/50 GL column (GE Healthcare) and an elution step was achieved through a gradient increase of NaCl concentration in the range of 50-1000 mM. Eluted protein samples were dialyzed against the SEC buffer (50mM Tris pH 8.0, 150mM NaCl, 2mM DTT) and loaded on Superdex S200 26/60 HiLoad (GE Healthcare) for samples $MYH7^{1477-1793}$ samples and Superdex S75 26/60 HiLoad (GE Healthcare) for $MYH7^{838-964}$, $MYH7^{964-1088}$, $MYH7^{1088-1218}$, $MYH7^{1219-1301}$, $MYH7^{1301-1378}$ and $MYH7^{1357-1407}$. Purified samples were analyzed on 12% SDS-PAGE for impurities and used for ubiquitination or BLI experiments or were flash frozen in the liquid nitrogen and stored at -80°C until the further use.

2.2.1.2 TAMRA-ubiquitin purification

Fluorescent labelling of ubiquitin with TAMRA was achieved following the protocol developed in the AG Scheffner (University of Konstanz). Recombinantly purified ubiquitin (described in Section 1.2.1) was concentrated in 50 mM $NaCH_3COO$ pH 7.2 to the final concentration of 5.3 µg/µL. A 368 µg of TAMRA-NHS (Sigma-Aldrich) was dissolved in 20 µL of dimethyl-sulfoxide (DMSO) and mixed with ubiquitin. Coupling reaction was incubated at 4°C overnight, followed by quenching via adding 21 µL of Tris-HCl pH 8.0 and 30 minutes' incubation at 30°C. The reaction mixture was diluted in 30 mL 25 mM $NaCH_3COO$ pH 4.0 and purified using cation exchange chromatography on 1 mL HiTrap SP HP column (GE Healthcare). Samples were eluted through a gradient increase of a NaCl concentration in the range 0-1 M. Eluted samples were collected and loaded onto the Amicon®Ultra 4 mL centrifugal filters with the cut off 3 kDa (Amicon) and centrifuged at 4°C and 4500 xg until the final volume of 500 µL. Sample was washed three times with 3 mL of 25 mM Tris-HCl pH 7.5, 50 mM NaCl and kept at -20°C until the further use.

2.2.2 MuRF1 mediated ubiquitination of MYH7 *in vitro*

Ubiquitination reactions were carried out in 30 µL reaction volumes. Each reaction contained 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM DTT, 1 mM ATP, 1mM $MgCl_2$, 0.2 µg UBA1, 0.2 µg UBcH5b, 15 µg ubiquitin, 6.3 µg His_6-MBP-MuRF1 and 5 µg of substrate and was incubated for 5h at 37°C. Reactions were stopped by addition of 7 µL of 4x SDS-PAGE loading buffer (200 mM Tris-Cl (pH 6.8), 400 mM DTT, 8% SDS, 0.4% bromophenol blue, 40% glycerol) and incubating for 5 min at 90°C. Full reactions were loaded onto 8% SDS-PAGE gels. Ubiquitination was detected by staining with Coomassie brilliant blue (CBB) and ubiquitination sites were determined by mass spectrometry analysis from gel extraction samples.

2.2.3 Crystallization trials of MYH7$^{964\text{-}1085}$

The protein was finally purified and its homogeneity was confirmed by loading onto a Superdex S75 26/60 HiLoad (GE Healthcare) size exclusion column. One peak was observed and analyzed by SDS-PAGE confirming MYH7$^{964\text{-}1085}$. The protein was concentrated in 50 mM Tris-HCl pH 8, 50 mM NaCl, 1mM DTT to the final concentration of 10 mg/mL. MYH7$^{964\text{-}1085}$ fragment was screened for crystallization using an automated nano volume-dispensing crystallization robot (Crystal Gryphon, Art Robbins Instruments) in 96-well Intelli-plates (Art Robbins Instruments) and commercial crystallization screens (Table 2.1).

Table 2.1 Commercial crystallization screens used for MYH7$^{964\text{-}1085}$ crystallization study

Company	**Screens**
Molecular Dimensions	JCSG plus Eco, Structure Screen 1 and 2, Morpheus, Proplex 1 and 2
Hampton Research	PEGRx HT, Natrix 1 and 2, SaltRx 1 and 2
Rigaku Reagents	Wizard I and II, Cryo Wizard

MYH7$^{964\text{-}1088}$ crystals were obtained at 19°C from a sitting drops containing equal volumes (200 nL) of protein and reservoir solutions (0.1 M MES, 0.1 M imidazole pH 6.5; 0.09 M $NaNO_3$, 0.09 M Na_2HPO_4, 0.09M $(NH_4)_2SO_4$; 20% w/v PEG 500 MME; 10% [w/v] PEG). Crystals were cryoprotected with mother liquor containing 30% [v/v] glycerol prior to cryocooling in liquid nitrogen. X-ray diffraction data was collected on beamline PXIII (Swiss Light Source, Switzerland) and indexed using XDS.

2.2.4 Generation of a myosin ubiquitination sites database for sequence studies

To study the potential effect of ubiquitination on myosin heavy chain folding and fiber assembly, a unique and comprehensive database with all known ubiquitination sites of myosin heavy chain was created for sequence analysis. A database MYH7 Ubi-sites was generated collecting and combining available entries in the mUbiSiDa (Chen *et al.*, 2014) and PhosphositePlus v.6.5.8 web databases (Hornbeck *et al.*, 2012) together with sites extracted from the extensive literature search, as not all the identified sites are deposited on the servers. All entries were filtered for the ub-sites in the heavy chains of human, mouse or rat MYH7. Collected information consisted of location, source (tissue and organism) and method of ubiquitination (*in vitro* or *in vivo*). To simplify mapping of ubiquitination sites human myosin sequence (UniProtKB P12883) was used, considering that most of the ubiquitination sites were found in human samples and that sequence identity among species is very high (99%). Collected data was further used in mapping ubiquitination sites in the known structures, studying the potential role of ubiquitinated lysine in inter- and intra-chain interactions and position in heptad repeat of myosin coiled-coil.

2.2.5 Muscle fiber isometric tension measurements

To study the effect of MuRF1 and its ubiquitination activity of myosin heavy chain on the mechanical properties of muscle, muscle fibers were isolated from the rabbit psoas muscle, treated with ubiquitination mixture and tested for isometric force production.

For extraction of muscle fibers, the New Zealand white rabbit was heavily sedated and euthanized according to the national directives for the care and use of laboratory animals (Directive 2010/63/EU) approved by the ethics committee of the University of Thessaly (decision 2-1/10-10-2012) and the scientific committee of the University Hospital of Larissa, Greece (decision 1/4-1-2012). Psoas muscle samples were dissected and fixed to the sterile holders using surgical thread and immersed into the buffer solution for chemical skinning (120 mM potassium acetate (KAc), 50 mM 3-(N-Morpholino) propanesulfonic acid (MOPS), 5 mM magnesium acetate ($MgAc_2$), 4 mM ethylene glycol tetraacetic acid (EGTA), 50 %

glycerol volume per volume; pH 7 at 0°C) supplemented with protease inhibitor (SIGMA-P8340). The skinning solution was replaced with a fresh solution and muscle samples were stored at -20°C until further use.

To study the effect of MuRF1 mediated ubiquitination on muscle fibers, MuRF1 ubiquitination reaction was adopted from *in vitro* studies described in section 2.2 and upscaled to the total volume of 500 µL. The main deviations from the *in vitro* reactions were introduced as adaptations to a muscle fiber as a substrate instead of soluble myosin. Specifically, *in vitro* reactions were incubated at 37°C, but because muscle fibers are temperature sensitive, ubiquitination reactions *ex vivo* were performed at 4°C instead. To facilitate handling and preserve fibers 10% v/v of glycerol was added to the ubiquitination reaction. To confirm that MuRF1 efficiency is not affected, additional *in vitro* experiments were carried out.

Psoas muscle fibre bundles were dissected from the harvested and chemically skinned psoas muscle, described above, for ubiquitination reaction. Dissection was performed under a stereomicroscope on a cold stage with the fiber bundles still immersed in the skinning solution. Two muscle fiber bundles per condition were then incubated in control or reaction solutions described in Table 2.8.1. Three control groups were designed (Table 2.2 Control, MuRF1 Control and Ubiquitin Control) to assess the potential effect of buffer conditions, endogenous E3 ubiquitin ligases and MuRF1 without triggered reaction, and two reaction groups (Table 2.2 Reaction and Pre-incubated reaction). All incubations were performed at 4°C.

Table 2.2 Incubation conditions for the muscle fiber force production measurement

No.	Label	Incubation conditions
1	Control	10% glycerol, 1 mM DTT, 5 mM ATP, 2 mM $MgCl_2$ (24h).
2	Ubiquitin Control	10% glycerol, 1 mM DTT, 5mM ATP, 2 mM $MgCl_2$, 46 nM UBA1 (E1), 330 nM UBcH5b (E2), 56 µM ubiquitin (24h).
3	MuRF1 Control	10% glycerol, 1 mM DTT, 5 mM ATP, 2 mM $MgCl_2$, 46 nM UBA1 (E1), 330 nM UBcH5b (E2) and 1 µM His_6-MBP-MuRF1 (24h).
4	Reaction	10% glycerol, 1 mM DTT, 5 mM ATP, 2mM $MgCl_2$, 46 nM UBA1 (E1), 330 nM UBcH5b (E2), 56 µM ubiquitin and 1 µM His_6-MBP-MuRF1 (24h).
5	Pre-incubated reaction	10% glycerol, 1 mM DTT, 5 mM ATP, 2 mM $MgCl_2$, 46 nM UBA1 (E1), 330 nM UBcH5b (E2), 1 µM His_6-MBP-MuRF1 (24h); addition of 56 µM ubiquitin (further 24h).

After the incubation, fiber bundles were washed in 50% glycerol and a single fibers were dissected and the fiber ends were attached between two tissue mounts of a muscle micro-dynamometer (SI Heidelberg/WPI, Germany), connected to a force transducer (0.4 μN-4 N) and a motor arm (used as a fixed end for isometric assessments) (Figure 2.6). The fiber was then immersed in wells containing experimental solutions. The micro-dynamometer consisted of five independent temperature controlled wells and a He-Ne laser for sarcomere length measurements. The temperature of the wells was controlled using a cooling/heating water circulator (Thermo Electron Haake WKL26, Thermo Fisher Scientific, USA) and rapid switching between solutions allowed the change in force upon temperature-jumps (t-jumps) to be recorded.

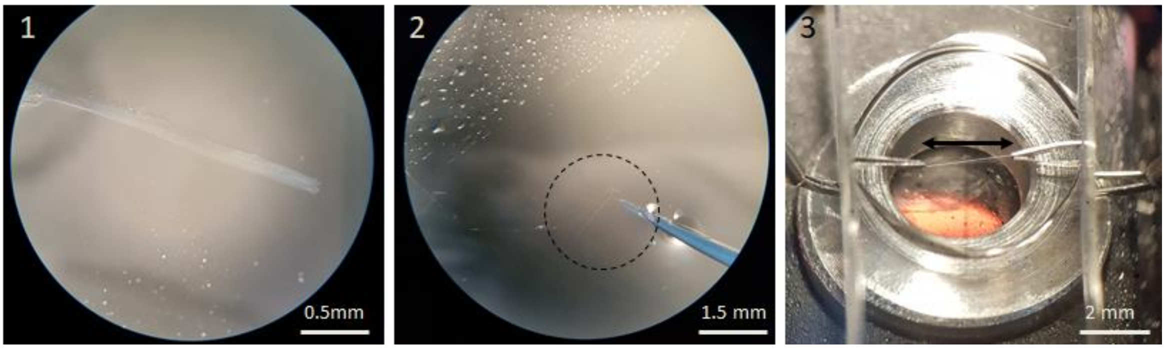


Figure 2.6 Single muscle fiber sample preparation.
Muscle fiber bundle (1) was washed from the incubation solution and dissection was completed in the 50% glycerol solution (2). The single fiber of an average length of 2.5 mm was loaded between tissue clamps connected to a force transducer and motor arm (3).

Force generation capacity was assessed in a total of 63 single muscle fibers at 10°C and 30°C and pH 7. Each fiber was first immersed for one minute in a well containing rigor solution to wash out excess incubation solution. The experimental fiber was then immersed in a relaxing solution (120 mM KAc, 50 mM MOPS pH 7, 5 mM $MgAc_2$ 1 mM EGTA, 5 mM adenosine triphosphate (ATP)) at 10°C until signal stabilization and then switched to an activating solution (120 mM KAc, 50 mM MOPS pH 7, 5 mM $MgAc_2$ 1 mM EGTA, 5mM ATP, 1.1 mM $CaCl_2$) in well 2 at 10°C to measure the isometric tension. After the signal stabilization (allowing the sarcomere arrangement stabilization), fiber was briefly exposed to the near physiological temperature of 30°C in activation solution (well 3) to obtain maximal isometric force and returned to the relaxing solution in well 1 (Figure 2.7).

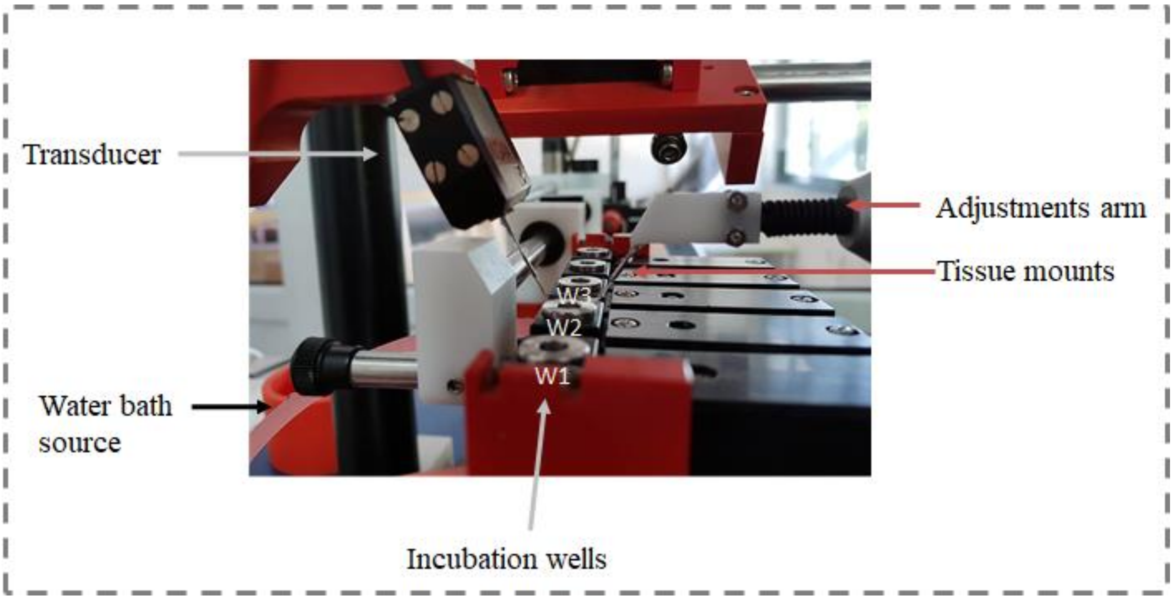


Figure 2.7 SI Heidelberg (WPI) micro dynamometer setup for the muscle fiber isometric tension measurement.
Five wells are temperature controlled through the water bath circulating system. Wells 1 and 2 were set to 10°C and wall 3 to 30°C. Muscle fiber is loaded between two tissue mounts, one connected to the force transducer and one connected to a motor arm and transferred between the wells for measurements.

Force data were recorded in both absolute (µN) and specific values (mN/mm) *i.e.*, absolute force corrected for the calculated cross-sectional area (CSA). For further comparison of the fibers only specific values were used. Data are reported as mean ± standard error of the mean.

2.2.6 Muscle fibers imaging

To monitor if MuRF1 has an effect on sarcomeric arrangement in muscle fibers and what is the location of ubiquitination sites, muscle fibers were subjected to ubiquitination reactions with tetramethylrhodamine (TAMRA)-labeled ubiquitin. TAMRA is fluorescent label covalently introduced to Ubiquitin-Lysine residues as described in section 2.1.2. TAMRA-ubiquitin was tested *in vitro* conditions and it was confirmed that the tag does not affect MuRF1 activity, thereby is suitable for the experiment.

First incubations were conducted in the same manner as for the mechanic experiments described in Section 2.8, but as the production of TAMRA-ubiquitin is more expensive and time consuming than production of unlabeled ubiquitin, the experimental setup for fiber incubations had to be adopted in a more cost effective way. For that purpose, a special flow cell was designed (Figure 2.8). The flow cell was specifically designed with the aim to reduce incubation volume from 500 µL to 25 µL and to enable washing steps in situ. In this way, handling of a muscle fiber was reduced and simplified. All the glass parts were fixed using glass specific glue and the glass slide was flexible, attached and detached using silicon. Single muscle fibers were dissected as described in section 2.8. and each end of a single fiber was carefully fixed using nail polish, in the middle of a glass plate. After fixing the glass slide on top, a flow cell was created and using a pipet the fiber was easily washed under the flow. Control and reaction conditions were the same as described in Table 2.2. and only ubiquitin was replaced with TAMRA-ubiquitin. After incubation, fibers were washed with 2 mL of rigor solution and 2 mL of 50% glycerol and remained in 50% glycerol for the imaging.

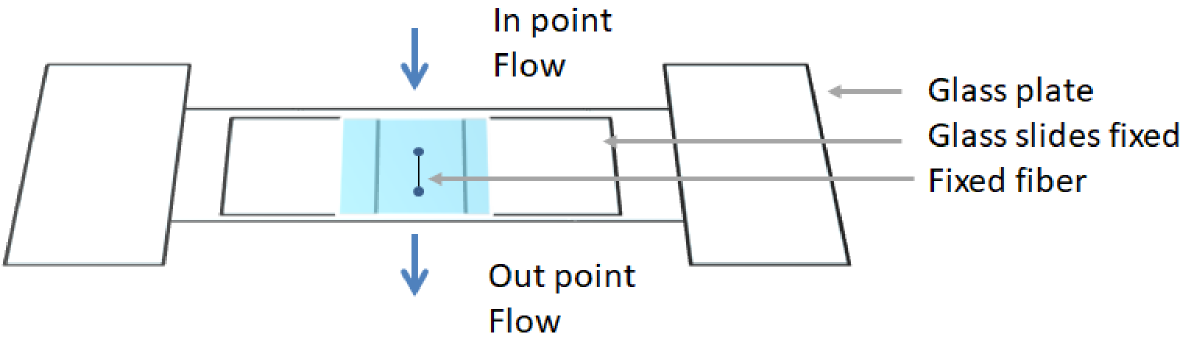


Figure 2.8 Schematic representation of a designed flow cell.
Cell was used for muscle fiber incubation, washing and imaging. Glass slides fixed in the middle create a chamber of approx. Volume of 20 uL and 3D space for muscle fiber fixation. Using a standard microscopy glass plate and cover slides would crash the fiber.

Fibers were imaged under a light microscope (Leica, DM 2000) equipped with an excitation filter of 515–560 nm and a camera (Leica, DFC-425). Collected images were analyzed using ImageJ software (Schindelin *et al.*, 2012). Average sarcomere length was measured for Control sample, Ubiquitin Control, MuRF1 Control, Reaction and Pre-incubated reaction and ubiquitin fluorescence was monitored for the Ubiquitin Control, Reaction and Pre-incubation reaction samples.

2.3 Results

2.3.1 Myosin heavy chain is ubiquitinated by MuRF1 *in vitro*

For studying the effect of MuRF1 on myosin heavy chain, ubiquitination assays have been established to identify the pattern of ubiquitination. Full-length myosin heavy chain, extracted from rabbit muscle was ubiquitinated by MuRF1 *in vitro* and resultant ubiquitination was detected as a higher molecular mass smear, not as a single band, which indicates multiple ubiquitination sites. These results are in agreement with previously shown MuRF1 *in vitro* ubiquitination assisted by different E2-ubiquitin-conjugating enzymes (Marblestone *et al.*, 2012). Despite this experiment confirming that MuRF1 targets myosin heavy chain *in vitro* conditions, the pattern of activity remained unclear.

Full-length MYH7 was shown to be a target by MuRF1 in experimental conditions designed for this study. However, as the sample was isolated from a rabbits' muscle, myosin could contain posttranslational modifications, including ubiquitination, introduced *in vivo*, prior to extraction, that could potentially interfere with the MuRF1 activity. To eliminate sample modifications, seven partially overlapping MYH7 fragments ($MYH7^{838\text{-}963}$, $MYH7^{964\text{-}1085}$, $MYH7^{1088\text{-}1218}$, $MYH7^{1219\text{-}1301}$, $MYH7^{1301\text{-}1378}$, $MYH7^{1357\text{-}1407}$, $MYH7^{1477\text{-}1793}$) were recombinantly produced, purified and subjected to ubiquitination assay. Five fragments were found to be ubiquitinated ($MYH7^{964\text{-}1085}$, $MYH7^{1219\text{-}1301}$, $MYH7^{1301\text{-}1378}$, $MYH7^{1357\text{-}1407}$, $MYH7^{1477\text{-}1793}$) and for two fragments ubiquitination was not detected under used conditions ($MYH7^{838\text{-}963}$, $MYH7^{1088\text{-}1218}$). Lack of detected ubiquitination of two fragments indicates that MuRF1 is selective, rather than promiscuous. Protein bands corresponding to ubiquitination were extracted from the SDS-PAGE gels and analyzed for identification of ubiquitination sites applying mass spectrometry analysis.

Mass spectrometry analysis revealed nineteen lysines modified by ubiquitination in myosin heavy chain (K1022, K1026, K1043, K1279, K1305, K1316, K1326, K1354, K1363, K1374, K1483, K1503, K1528, K1651, K1700, K1757, K1771, K1784, K1791). Human MYH7 sequence (UniProtKB - P12883) including residues 838-1793 contains 110 lysine residues and 17.2% were ubiquitinated by MuRF1. It is, nonetheless, crucial to investigate how presented data corresponds to MuRF1 activity *in vivo*.

MuRF1 *in vivo* ubiquitination was recently studied by Lang et al. in mice model during muscle denervation induced atrophy (Lang *et al.*, 2017). Stable Isotope Labeling by/with Amino acids in Cell culture (**SILAC**) is a mass spectrometry based technique that detects differences in protein abundance among samples using non-radioactive isotopic labeling (Lang *et al.*, 2017) and has been implemented by Lang et al. in identification of MuRF1 related ubiquitination sites. From 10 identified ubiquitination sites in MYH7, only K1022 was suggested as MuRF1 dependent while K936, K1326, K1503, K1644, K1651 as potential MuRF1 ubiquitination sites. The remaining three sites, K1107, K1279 and K1641 were assigned as no-MuRF1 (Figure 2.9). Comparing data presented by Lang et al., and data obtained in this work, four ubiquitination sites were found in both studies, including K1022. More ubiquitination sites were identified *in vitro,* which can be explained by the lower sensitivity of the *in vivo* method. One ubiquitination site K1279 was detected in this study, but was assigned as non-MuRF1 ubiquitination site by Lang et al. This finding suggests that K1279 can be targeted by MuRF1 but could also be a target of other E3 ubiquitin ligases present *in vivo*.

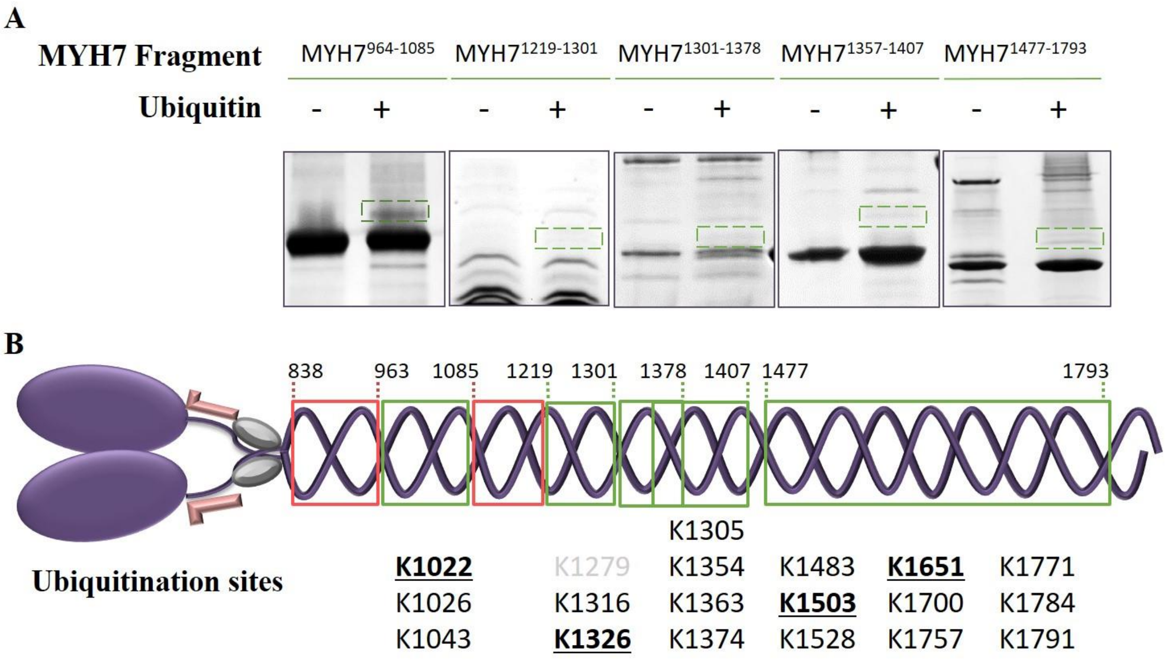


Figure 2.9 MuRF1 mediates ubiquitination of five myosin fragments.
A SDS-PAGE of ubiquitination reactions. Upon addition of ubiquitin (+), higher molecular mass bands appear, corresponding to ubiquitinated myosin fragments. **B** Schematic representation of myosin fragments tested for ubiquitination. Two fragments were not found to be ubiquitinated (red squares), while five were (green squares). Ubiquitination sites indicate ubiquitinated lysine (K) residues found in corresponding fragments. Ubiquitination sites K1022, K1326, K1503 and K1651 were previously reported in literature as potential MuRF1 ubiquitination sites *in vivo* (Lang *et al.*, 2017).

Overall, ubiquitination experiments showed that MuRF1 ubiquitination activity on myosin heavy chain *in vitro* resembles its *in vivo* pattern. The number of ubiquitination sites suggests that MuRF1 targets specific lysine residues, but it is not clear how those sites are selected for. To understand the underlying pattern of substrate recognition and MuRF1's targeting mechanism, the MYH7 sequence and ubiquitination sites were analyzed computationally.

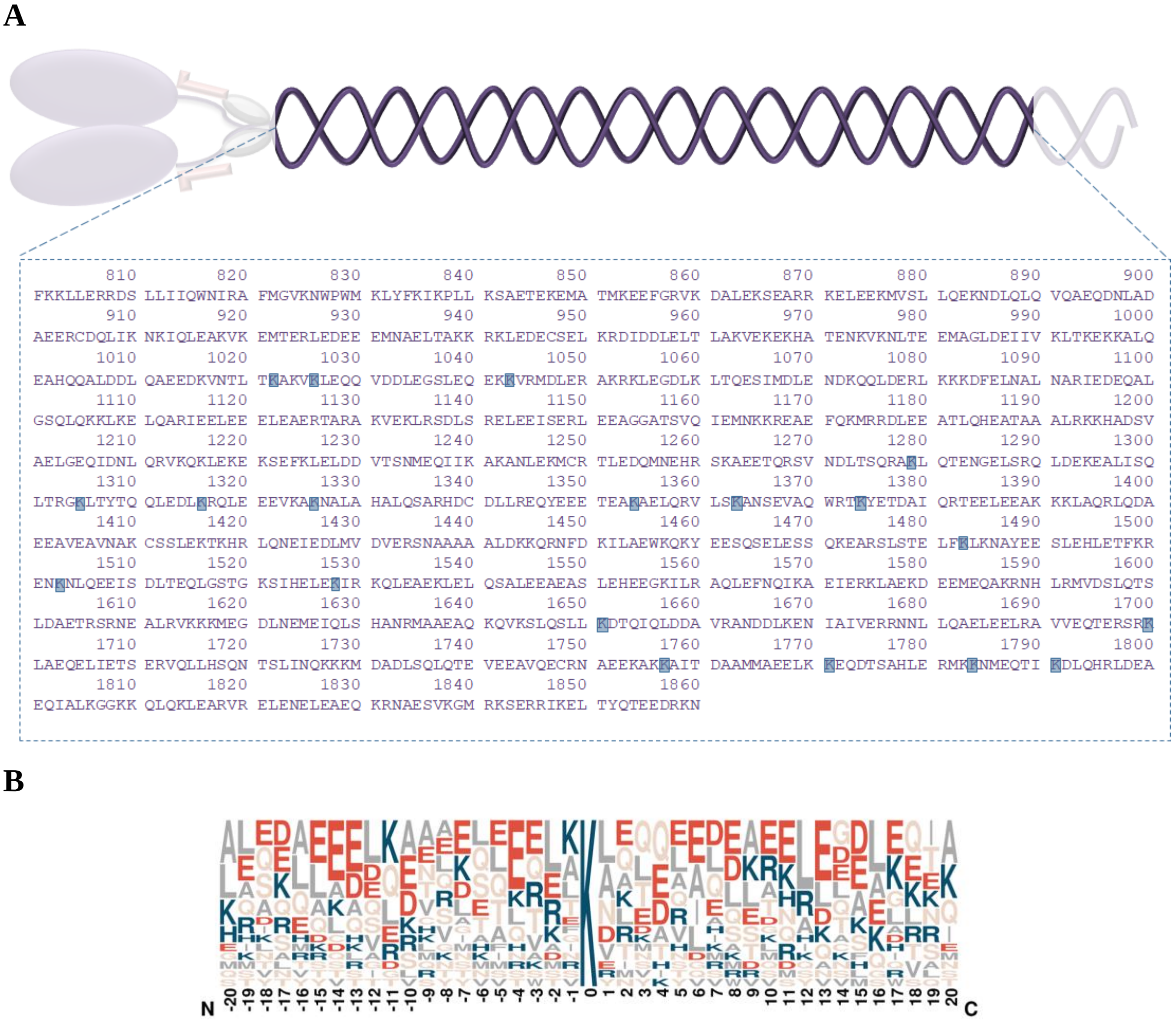
2.3.2 Ubiquitination site analysis

2.3.2.1 Amino acid composition (ACC) of regions surrounding ubiquitination sites does not reveal a conserved recognition pattern in MYH7

Analysis of *in vitro* MuRF1 mediated ubiquitination of MYH7 revealed 19 targeted lysines. MYH7 consists of 1935 residues forming a myosin head (1-838) and a coiled-coil domain (838-1935) called the rod domain (Figure 2.10). The rod domain is involved in myosin oligomerization during the formation of thick filaments in muscle development, representing a key of myosin stability. Thereby, the ubiquitination of this region could lead to myosin unfolding and thick filament depletion.

To explore potential underlying recognition patterns that determine certain lysines for ubiquitination by MuRF1, sequences surrounding ubiquitination sites were analyzed. When searching for an ubiquitination recognition pattern, the amino acid composition (AAC) was used to determine the frequencies of the twenty amino acids surrounding the ubiquitination sites in a specified window length.

The sequence window of 41 amino acids with ubiquitination site centered were collected for 19 MuRF1 in vitro ubiquitination sites. Frequency of each amino acid residue was calculated per position in the analyzed window and compared. As a result, there was no conserved sequence-specific pattern found in the MuRF1 ubiquitination sites of myosin heavy chain. The highest frequency was calculated for the glutamic acid residue in 13. position up and downstream around the central ubiquitinated lysine. Results are graphically represented in the Figure 2.10 by Sequence logo (Schneider and Stephens, 1990).



```
       810        820        830        840        850        860        870        880        890        900
FKKLLERRDS LLIIQWNIRA FMGVKNWPWM KLYFKIKPLL KSAETEKEMA TMKEEFGRVK DALEKSEARR KELEEKMVSL LQEKNDLQLQ VQAEQDNLAD
       910        920        930        940        950        960        970        980        990       1000
AEERCDQLIK NKIQLEAKVK EMTERLEDEE EMNAELTAKK RKLEDECSEL KRDIDDLELT LAKVEKEKHA TENKVKNLTE EMAGLDEIIV KLTKEKKALQ
      1010       1020       1030       1040       1050       1060       1070       1080       1090       1100
EAHQQALDDL QAEEDKVNTL TKAKVKLEQQ VDDLEGSLEQ EKKVRMDLER AKRKLEGDLK LTQESIMDLE NDKQQLDERL KKKDFELNAL NARIEDEQAL
      1110       1120       1130       1140       1150       1160       1170       1180       1190       1200
GSQLQKKLKE LQARIEELEE ELEAERTARA KVEKLRSDLS RELEEISERL EEAGGATSVQ IEMNKKREAE FQKMRRDLEE ATLQHEATAA ALRKKHADSV
      1210       1220       1230       1240       1250       1260       1270       1280       1290       1300
AELGEQIDNL QRVKQKLEKE KSEFKLELDD VTSNMEQIIK AKANLEKMCR TLEDQMNEHR SKAEETQRSV NDLTSQRAKL QTENGELSRQ LDEKEALISQ
      1310       1320       1330       1340       1350       1360       1370       1380       1390       1400
LTRGKLTYTQ QLEDLKRQLE EEVKAKNALA HALQSARHDC DLLREQYEEE TEAKAELQRV LSKANSEVAQ WRTKYETDAI QRTEELEEAK KKLAQRLQDA
      1410       1420       1430       1440       1450       1460       1470       1480       1490       1500
EEAVEAVNAK CSSLEKTKHR LQNEIEDLMV DVERSNAAAA ALDKKQRNFD KILAEWKQKY EESQSELESS QKEARSLSTE LFKLKNAYEE SLEHLETFKR
      1510       1520       1530       1540       1550       1560       1570       1580       1590       1600
ENKNLQEEIS DLTEQLGSTG KSIHELEKIR KQLEAEKLEL QSALEEAEAS LEHEEGKILR AQLEFNQIKA EIERKLAEKD EEMEQAKRNH LRMVDSLQTS
      1610       1620       1630       1640       1650       1660       1670       1680       1690       1700
LDAETRSRNE ALRVKKKMEG DLNEMEIQLS HANRMAAEAQ KQVKSLQSLL KDTQIQLDDA VRANDDLKEN IAIVERRNNL LQAELEELRA VVEQTERSRK
      1710       1720       1730       1740       1750       1760       1770       1780       1790       1800
LAEQELIETS ERVQLLHSQN TSLINQKKKM DADLSQLQTE VEEAVQECRN AEEKAKKAIT DAAMMAEELK KEQDTSAHLE RMKKNMEQTI KDLQHRLDEA
      1810       1820       1830       1840       1850       1860
EQIALKGGKK QLQKLEARVR ELENELEAEQ KRNAESVKGM RKSERRIKEL TYQTEEDRKN
```

Figure 2.10 Sequence analysis of ubiquitination sites.
A Myosin heavy chain rod domain sequence with highlighted ubiquitination sites. **B** Sequence logo of amino acids frequency for position surrounding ubiquitination sites shows no obvious pattern.

In the rod domain there is a high abundance of charged residues (41.85% of sequence). Charged residue abundance is related to the tendency to interact in salt bridge formation (arginine-aspartate and lysine-glutamate pairs). Arginine and Aspartate have similar occupancy in the myosin sequence, 6.7% and 5.6% respectively while Lysine and Glutamate have much higher frequency, 11.7% and 17.85% respectively. Such a high frequency of charged residues is unusual (SwissProt reports average frequency of Glu is 6.72). For example, head domain composition contains 8.7% Lysine, 4% Glutamate, 4.9% Arginine and 5.1% Aspartate. An abundance of high charges in the myosin rod domain is considered to be important for myosin packing in the thick filament, but the exact mechanism is not well understood (Rosenberg *et al.*, 2008).

Because the myosin rod domain is a coiled-coil, for structural stability and protein interaction heptad position plays more important role than a sequence itself, hence pattern of ubiquitination could be hidden there.

2.3.2.2 Prediction of CC repeat and location of Ubi sites

For coiled-coil proteins, an abundance of a certain residue does not give a lot of information itself, but it is rather combined with the position in the heptad repeat for analysis. Prediction of coiled-coils is challenging and rather hard, mostly because a number of coiled-coils are not regular heptad repeats but interrupted by skips and stammers. However, as more structures became available, the algorithms for structure prediction have improved and have become more reliable. For analysis of myosin rod domain sequence MARCOIL was used for calculating heptad repeat position of amino acids. MARCOIL is a hidden MARkov model-based program that predicts the existence and location of potential coiled-coil domains in protein sequences (Delorenzi and Speed, 2002) and its assessment of heptad repeat in myosin heavy chain is shown in the Figure 2.11..

Abundance and frequency of lysine residue were calculated per position (**a**, **b**, **c**, **d**, **e**, **f**, **g**) and ubiquitinated and non-ubiquitinated lysine residue abundances were compared. Comparison of the position abundance between Ubi-sites and those of non-Ubi-sites provides significant information for the identification of a targeting pattern.

Summary of the results for all amino acids is shown in the Table 2.2. While **b**, **c**, **d**, **f** and **g** positions are found 147-148 times, position **a** and **e** are found slightly more frequently (154 and 151, respectively) indicating that MARCOIL predicted the existence of skips and stammers. The most abundant amino acid in myosin rod domain is glutamate (195), followed by leucine (141) and lysine (128).

```
 840       850        860        870        880        890
PL LKSAEREKEM ASMKEEFTRL KEALEKSEAR RKELEEKMVS LLQEKNDLQL QVQAEQD
de fgabcdefga bcdefgabcd efgabcdefg abcdefgabc defgabcdef gabcdef
 900       910        920        930        940        950
LA DAEERCDQLI KNKIQLEAKV KEMNERLEDE EEMNAELTAK KRKLEDECSE LKRDIDD
ab cdefgabcde fgabcdefga bcdefgabcd efgabcdefg abcdefgabc defgabc
 960       970        980        990        1000       1010
EL TLAKVEKEKH ATENKVKNLT EEMAGLDEII AKLTKEKKAL QEAHQQALDD LQAEEDK
ef gabcdefgab cdefgabcde fgabcdefga bcdefgabcd efgabcdefg abcdefg
 1020      1030       1040       1050       1060       1070
NT LTKAKVKLEQ QVDDLEGSLE QEKKVRMDLE RAKRKLEGDL KLTQESIMDL ENDKQQL
bc defgabcdef gabcdefgab cdefgabcde fgabcdefga bcdefgabcd efgabcd
 1080      1090       1100       1110       1120       1130
ER LKKKDFELNA LNARIEDEQA LGSQLQKKLK ELQARIEELE EELEAERTAR AKVEKLR
fg abcdefgabc defgabcdef gabcdefgab cdefgabcde fgabcdefga bcdefga
 1140      1150       1160       1170       1180       1190
DL SRELEEISER LEEAGGATSV QIEMNKKREA EFQKMRRDLE EATLQHEATA AALRKKH
cd efgabcdefg abcdefgabc defgabcdef gabcdefgab cdefgabcde bfgabcd
 1200      1210       1220       1230       1240       1250
DS VAELGEQIDN LQRVKQKLEK EKSEFKLELD DVTSNMEQII KAKANLEKMC RTLEDQM
fg abcdefgabc defgabcdef gabcdefgab cdefgabcde fgabcdefga bcdefga
 1260      1270       1280       1290       1300       1310
EH RSKAEETQRS VNDLTSQRAK LQTENGELSR QLDEKEALIS QLTRGKLTYT QQLEDLK
cd efgabcdefg abcdefgabc defgabcdef gabcdefgab cdefgabcde fgabcde
 1320      1330       1340       1350       1360       1370
QL EEEVKAKNAL AHALQSARHD CDLLREQYEE ETEAKAELQR VLSKANSEVA QWRTKYE
ga bcdefgabcd efgabcdefg abcdefgabc defgabcdef gabcdefgab cdeffga
 1380      1390       1400       1410       1420       1430
DA IQRTEELEEA KKKLAQRLQE AEEAVEAVNA KCSSLEKTKH RLQNEIEDLM VDVERSN
cd efgabcdefg abcdefgabc defgabcdef gabcdefgab cdefgabcde fgabcde
 1440      1450       1460       1470       1480       1490
AA AALDKKQRNF DKILAEWKQK YEESQSELES SQKEARSLST ELFKLKNAYE ESLEHLE
ga bcdefgabcd efgabcdefg abcdefgabc defgabcdef gabcdefgab cdefgab
 1500      1510       1520       1530       1540       1550
FK RENKNLQEEI SDLTEQLGSS GKTIHELEKV RKQLEAEKME LQSALEEAEA SLEHEEG
de fgabcdefga bcdefgabcd efgabcdefg abcdefgabc defgabcdef gabcdef
 1560      1570       1580       1590       1600       1610
IL RAQLEFNQIK AEIERKLAEK DEEMEQAKRN HLRVVDSLQT SLDAETRSRN EALRVKK
ab cdefgabcde fgabcgabcd efgabcdefg abcdefgabf gabcdefgab cdefgab
 1620      1630       1640       1650       1660       1670
ME GDLNEMEIQL SHANRMAAEA QKQVKSLQSL LKDTQIQLDD AVRANDDLKE NIAIVER
de fgabcdefga bcdefgabcd efgabcdefg abcdefgabc defgabcdef gabcdef
 1680      1690       1700       1710       1720       1730
NN LLQAELEELR AVVEQTERSR KLAEQELIET SERVQLLHSQ NTSLINQKKK MDADLSQ
ab cdefgabcde fgabcdefga bcdefgabcd efgabcdefg abcdefgabc defgabc
 1740      1750       1760       1770       1780       1790
QT EVEEAVQECR NAEEKAKKAI TDAAMMAEEL KKEQDTSAHL ERMKKNMEQT IKDL
ef gabcdefgab cdefgabcde fgabcdefga bcdefgabcd efgabcdefg abcd
```

Figure 2.11 MARCOIL predicted heptad repeat position in the myosin rod domain coiled-coil.

Table 2.3 Myosin rod domain amino acid composition and heptad repeat position occupation predicted by MARCOIL. One letter amino acid code annotates 19 amino acids present in myosin heavy chain. Positions **a** and **e** are found more frequently than other positions, indicating that calculation predicted discontinuous in the myosin rod domain.

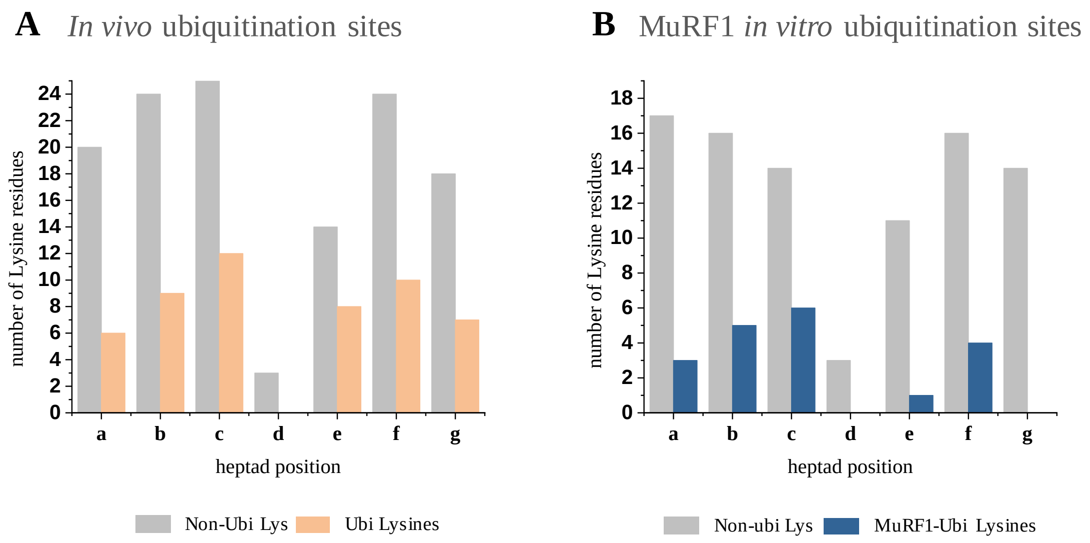
AA	**Heptad position MARCOIL assigned**							
	a	**b**	**c**	**d**	**e**	**f**	**g**	**total**
A	9	22	14	22	5	19	17	108
R	9	10	7	2	9	21	16	74
N	8	12	8	1	6	6	6	47
D	0	14	20	1	7	9	10	61
C	6	0	0	0	0	0	0	6
E	1	34	40	14	45	29	32	195
Q	1	11	14	1	29	9	21	86
G	2	2	2	1	4	5	2	18
H	3	3	3	2	1	4	1	17
I	15	2	2	4	6	3	2	34
L	45	5	5	67	6	7	6	141
K	**20**	**24**	**25**	**3**	**14**	**24**	**18**	**128**
M	7	3	1	8	1	0	5	25
F	4	1	0	3	0	1	0	9
P	0	0	0	1	0	0	0	1
T	3	3	5	7	13	8	5	44
W	0	0	0	2	0	0	0	2
Y	3	0	0	2	0	1	1	7
V	18	2	1	7	5	1	6	40
total	154	148	147	148	151	147	148	

However, amino acids are not equally distributed along the heptad repeat. Heptad positions assessment revealed that positions **a** and **d** are dominantly occupied by hydrophobic residues with the frequency of 0.648 and 0.777, respectively. Position **c** is most frequently occupied by charged residues (0.643), followed by positions **f** (0.592), **b** (0.5743), **g** (0.520) and **e** (0.503) (Table 2.3). From the charged occupied position, only position **f** is prevalently occupied by positively charged residues (Lysine, Arginine and Histidine), while all the other positions are predominantly occupied by negative residues (Glutamate and Aspartate). Lysine most frequently occupies positions **c** (0.170), **f** (0.163), **b** (0.162) and slightly less positions **a** (0.123) and **g** (0.126), while is infrequent in positions **e** (0.093) and **d** (0.020) (Table 2.4).

Table 2.4 Amino acid frequency of a heptad position in myosin rod domain based on a MARCOIL CC prediction. The frequency was calculated as a total number of certain residue (AA) in a defined position normalized by the total number of defined position (**a, b, c, d, e, f, g**).

	AA frequency in heptad position MARCOIL based						
AA	**a**	**b**	**c**	**d**	**e**	**f**	**g**
A	0.0584	0.1486	0.0952	0.1486	0.0331	0.1293	0.1149
R	0.0584	0.0676	0.0476	0.0135	0.0596	0.1429	0.1081
N	0.0519	0.0811	0.0544	0.0068	0.0397	0.0408	0.0405
D	0.0000	0.0946	0.1361	0.0068	0.0464	0.0612	0.0676
C	0.0390	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
E	0.0065	0.2297	0.2721	0.0946	0.2980	0.1973	0.2162
Q	0.0065	0.0743	0.0952	0.0068	0.1921	0.0612	0.1419
G	0.0130	0.0135	0.0136	0.0068	0.0265	0.0340	0.0135
H	0.0195	0.0203	0.0204	0.0135	0.0066	0.0272	0.0068
I	0.0974	0.0135	0.0136	0.0270	0.0397	0.0204	0.0135
L	0.2922	0.0338	0.0340	0.4527	0.0397	0.0476	0.0405
K	**0.1299**	**0.1622**	**0.1701**	**0.0203**	**0.0927**	**0.1633**	**0.1216**
M	0.0455	0.0203	0.0068	0.0541	0.0066	0.0000	0.0338
F	0.0260	0.0068	0.0000	0.0203	0.0000	0.0068	0.0000
P	0.0000	0.0000	0.0000	0.0068	0.0000	0.0000	0.0000
T	0.0195	0.0203	0.0340	0.0473	0.0861	0.0544	0.0338
W	0.0000	0.0000	0.0000	0.0135	0.0000	0.0000	0.0000
Y	0.0195	0.0000	0.0000	0.0135	0.0000	0.0068	0.0068
V	0.1169	0.0135	0.0068	0.0473	0.0331	0.0068	0.0405

Lysine ubiquitination by MuRF1 does not mirror heptad position based frequency (Figure 2.12). There is no obvious preference of MuRF1 for targeting a Lysine is a specific heptad position, as all positions except **d** and **g** are found to be targeted. However, it is interesting that position **g** is less likely to be targeted than position a, although those two positions have the same frequency of Lysine. Similarly, positions **b**, **c** and **f** share a similar frequency of Lysine occupancy, while the Lysine in position **c** seems to be the most frequently ubiquitinated, followed by position **b** and then position **f**.


A *In vivo* ubiquitination sites

B MuRF1 *in vitro* ubiquitination sites

Figure 2.12 Ubiquitination pattern based on heptad repeat of myosin rod domain. Lysine distribution in myosin rod domain is assigned by MARCOIL prediction and ubiquitinated and non-ubiquitinated Lysines per heptad position is presented graphically. **A** Distribution of non-ubiquitinated and ubiquitinated lysine residues obtained Wagner *et al.*, 2012, Jeon *et al.*, 2007 and Lang *et al.*, 2007 without assigned E3-ubiquitin ligase; **B** Distribution of non-ubiquitinated and ubiquitinated lysine residues by MuRF1;

To study if MuRF1 resembles the ubiquitination pattern of other E3 ubiquitin ligases targeting myosin rod domain, the same heptad position frequency occupancy analysis was conducted for the all known ubiquitination sites found in myosin rod domain (Wagner *et al.*, 2012, Jeon *et al.*, 2007 and Lang *et al.*, 2007) and compared to non-ubiquitinated Lysine residues. As a result, similarly to MuRF1 ubiquitination sites, the frequency of no-ubiquitinated lysine in positions **a**, **b** and **c** is highly similar, but the ubiquitination frequency shows a clear preference for position **c**, followed by position **b** and finally position **a**. Non-ubiquitination site was found in heptad position **d**, but unlikely the MuRF1 ubiquitination, other ubiquitination sites were found in position g. These results imply that MuRF1 could share a ubiquitination pattern for positions **a**, **b** and **c** with other ubiquitin ligases, while it has a distinct propensity for positions **e** and **g** of heptad repeat found in myosin rod domain. However, to understand whether Lysines targeted for ubiquitination have a specific role in structural stability and myosin assembly, structural analyses were carried out on available myosin rod domain structures.

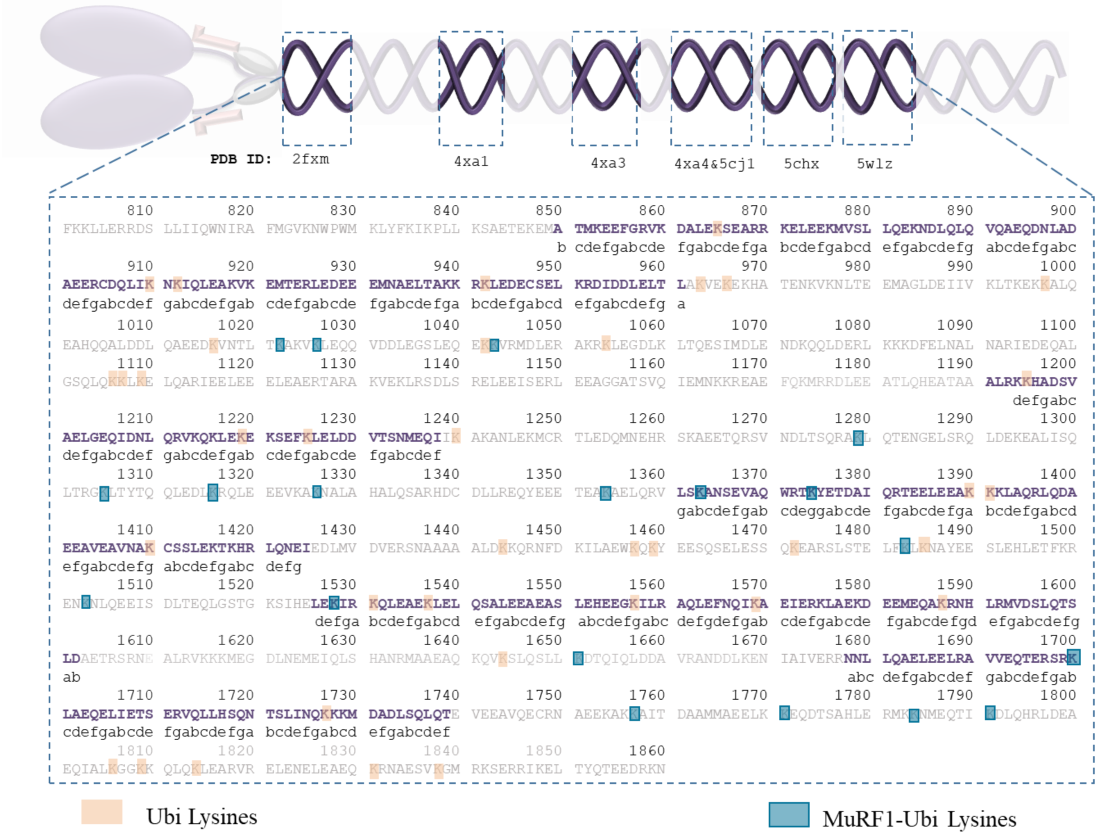
2.3.3 Structural analysis of the ubiquitination targeted lysines in myosin rod domain

2.3.3.1 Structure based analysis of Lysine heptad positions agree with MARCOIL predicted model

MARCOIL prediction is a sequence based method. By applying MARCOIL analysis, it is possible to create a reliable data set for Lysine position and ubiquitination model analysis. However, to verify how MARCOIL prediction coincides with experimental data, a set of available myosin rod domain crystal structures was analyzed for Lysine residue heptad positions using Twister software (Strelkov and Burkhard, 2002). The total of six myosin crystal structures (PDB ID: 2fxm, 4xa1, 4xa3, 5cj1, 5wlz, 5chz) cover 41 Lysine residue, of which 16 are known ubiquitination sites and 4 were discovered in this study. Twister analysis of heptad position highly coincides with the MARCOIL prediction. Direct comparison of 41 Lysine residues between MARCOIL prediction and Twister analysis shows the same position for 39 residues and a disagreement for two (K1569 and K1195).

The fraction that can be analyzed by Twister software, however is not sufficient for the detailed analysis of whether Lysine ubiquitination is affected by its position in the coiled-coil heptad repeat in the whole myosin rod but verifies MARCOIL prediction model (Figure 2.13).

A Myosin rod domain available crystal structures and corresponding sequence with highlighted Ubi sites



B Lysine and ubiquitination sites distribution in myosin rod heptad repeat assigned by Twister

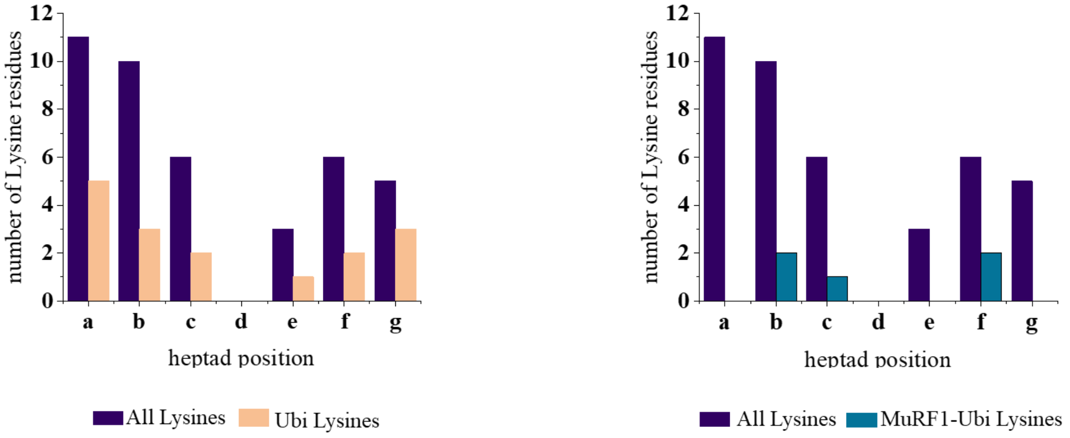


Figure 2.13 Heptad repeat position assigned by Twister analysis of Lysine residue in available myosin structures.

A Schematic representation of available myosin structures and their location in the rod domain with the corresponding sequence coverage and heptad position. **B** Heptad position distribution of 41 analyzed Lysine residues and ubiquitination sites. Ubiquitination sites obtained from Wagner *et al.*, 2012, Jeon *et al.*, 2007 and Lang *et al.*, 2007 without assigned E3-ubiquitin ligase (orange) are plotted separately from the MuRF1- ubiquitination sites (blue).

2.3.3.2 PISA and HBPlus analysis of intermolecular interactions

It is essential to understand what the role of Lysine residues in myosin heavy chain is to further understand how ubiquitination leads sarcomere destabilization. Structure analysis of five available myosin rod domain structures (PDB ID: 2fxm, 4xa1, 4xa3, 5cj1, 5wlz) included the assessment of intramolecular interactions by HBPlus and intermolecular analysis by PISA software.

PDBePISA is an interactive tool for the study of macromolecular interfaces searching for interactions and analyzing surface accessible and buried areas (Krissinel and Henrick, 2007). PISA results indicate that the majority of all ubiquitinated Lysine residues are in the surface accessible area (77%) and that only two Lysines are involved in salt bridge formation with the residues from a neighboring chain in a dimer. Current data shows no evidence of MuRF1 targeting above mentioned or other Lysines in buried areas. Additionally, HBPlus analysis of ubiquitination sites revealed that none of the ubiquitinated Lysine side chains are involved in interchain interactions. Taken together, these structural analysis of the available myosin rod domain indicate that free and accessible lysines are more likely to be targeted for ubiquitination by E3 ubiquitin ligases, yet lysine residues engaged in salt bridge formation are not entirely excluded, but rarely targeted.

All the conclusions in this chapter are made cautiously, as the number of available myosin structures is limited and describes only a third of a rod domain. Thereby, completing the myosin structure seems to be a necessity for comprehension of myosin and its ubiquitination.

2.3.4 Towards completing the myosin rod domain structure

A myosin heavy chain fragment containing amino acid residues 964-1088 was screened for the crystallization against 864 different conditions. Crystals grew from multiple conditions but the only diffracting crystals were obtained using a sitting drop vapor diffusion methods in drops containing 200 nL of protein and 200 nL of reservoir solution (0.1 M MES, 0.1 M imidazole pH 6.5; 0.09 M $NaNO_3$, 0.09 M Na_2HPO_4, 0.09 M $(NH_4)_2SO_4$; 20% w/v PEG 500 MME; 10% w/v PEG). MYH$^{964-1088}$ crystals grew as thin plates in clusters and size of were 480x360x240 µm (Figure 2.13A). From the crystal cluster, a single plate (Figure 2.13 A, red box) was fragmented in multiple single parts, mounted on cryo-loops and flash frozen for the X-ray diffraction experiments. Data diffracted to 2.56 Å resolution (Figure 2.13, B).

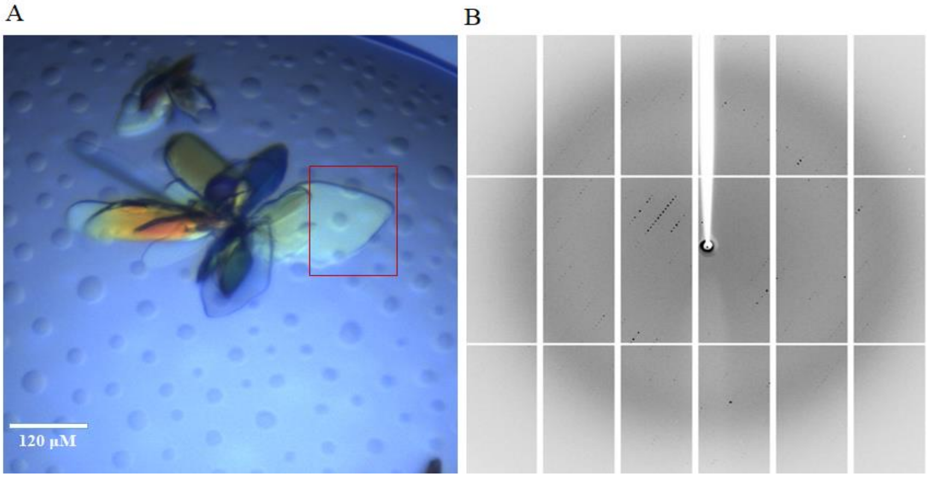


Figure 2.14 MYH$^{964-1088}$ crystallization and X-ray diffraction.
A MYH$^{964-1088}$ crystals in a cluster of thin plates morphology were obtained in 200 nL sitting drops conditions containing 0.1 M MES, 0.1 M imidazole pH 6.5; 0.09M $NaNO_3$, 0.09M Na_2HPO_4, 0.09M $(NH_4)_2SO_4$; 20% w/v PEG 500 MME; 10% w/v PEG. Single plate (red box) was separated from the cluster and mounted on a cryo-loop for X-ray diffraction experiments. **B** Diffraction image for MYH$^{964-1088}$ corresponding to an overlay of 10°. Data was collected using PILATUS 2MF to a maximum of 2.56 Å resolution.

Two datasets obtained from two different fragments of the same crystal were merged. The crystal lattice was identified as monoclinic and the space group was suggested to be $P2_1$ (Table 2.5). Matthews coefficient (Matthews, 1968) calculations suggested the presence of 4, 5 or 6 molecules per asymmetric unit, corresponding to a solvent content of 46% 55% or 64%, respectively. Phasing of the obtained data and attempts to elucidate the crystal structure of MYH$^{964\text{-}1088}$ are ongoing.

Table 2.5 Diffraction data statistics for MYH7$^{964\text{-}1088}$ merged dataset

	MYH7$^{964\text{-}1088}$
X-ray source	SLS-PXIII
Detector	PILATUS 2M-F
Wavelength (Å)	1.00
Space Group	$P2_1$
Unit Cell Dimensions	
a, b, c (Å)	33.6, 76.9, 154.2
α, β, γ (°)	90, 92.5, 90
Solvent Content (%)	46 / 55 / 64
Resolution (Å)	50-2.56 (2.63-2.56)
Unique Reflections	25 091 (1086)
Multiplicity	6.89 (6.70)
Completeness, %	98.6 (97.1)
CC (1/2)	99.9 (71.5)
<I/σ(l)>	8.95 (1.34)
R_{sym} (I)	17.8 (154.4)

2.3.5 MuRF1 dependent ubiquitination leads to physiological disturbances of the force production in isolated muscle fibers

Muscle fibers experiments were performed at University of Thessaly under the supervision of Dr. Christina Karatzaferi and assistance of Dr. Georgia Mitrou.

Force production and contractile properties of single fibers are conducted by the molecular-level interactions of the sarcomeric proteins. In order to assess the effect of MuRF1 activity on muscle function, the pilot experiments included experimental design and initial optimization on a total of 63 single muscle fibers isolated from a rabbit psoas muscle. This study evaluated the potential of single fiber analysis in studying the effect of ubiquitination on muscle properties, analysis of the contractile properties of permeabilized skeletal muscle fibers (further referred as fiber) were adopted for the *ex vivo* ubiquitination reactions as a novel approach.

The results of preliminary study strongly suggest that the single fiber model could be successfully implemented for the ubiquitination research. MuRF1 ubiquitination reactions developed for *in vitro* studies (section 1.2.2) were modified and adopted for non-diffusible substrates of muscle sarcomere. A pre-incubation step of muscle fibers with MuRF1 was introduced to overcome the problem of diffusibility and allow MuRF1 to perfuse into the muscle fiber and access its targets in muscle sarcomere. After a 24h incubation with MuRF1, the ubiquitination reaction is then triggered by addition of E1, E2 and ubiquitin.

Contractile properties of single muscle fibers were assigned by measuring isometric force production upon Ca^{2+} induced activation at 10°C and 30°C. Preliminary data indicates the decrease in maximal isometric force formation upon MuRF1 ubiquitination of muscle fibers (Figure 2.15, A). To examine if the decrease in force production is due to MuRF1 ubiquitination activity, isometric force production was assessed for control fibers incubated with MuRF1 for 48h and fibers incubated with reaction mixture containing E1, E2 and ubiquitin, but not MuRF1. All the control fibers performed at least 75% of the maximal force production of non-treated fiber, while the fibers subjected to MuRF1 ubiquitination performed less than 50% (Figure 2.15, B).

These results strongly suggest that MuRF1 ubiquitination activity can be monitored in single fiber as an experimental model and that it has a distinct effect on contractile properties. For a more detailed analysis of these effects, further studies on a larger scale are required.

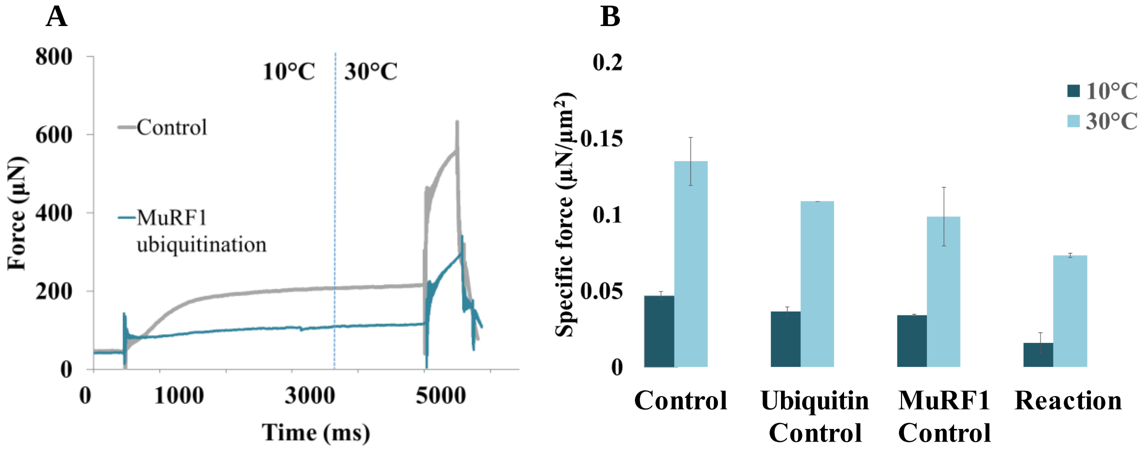


Figure 2.15 Muscle fiber isometric force production.
Muscle fiber isometric force is reduced in the presence of MuRF1 mediated ubiquitination. **A** Isometric force measurement for a control and MuRF1 ubiquitinated fiber. Passive tension developed in a relaxing solution (0-500 ms); Calcium administration induces activation. Maximum isometric force was allowed to develop before the temperature jump (500-5000 ms); **B** Total isometric force was normalized to a fiber cross section area for a direct comparison between fibers showing the decrease in the maximal produced force for MuRF1 ubiquitinated fibers (reaction) at both 10 and 30°C (n=3 for each experiment).

2.3.6 MuRF1 ubiquitination directs muscle sarcomere collapse

Sarcomere length must be maintained due to its key role in muscle function and direct correlation with the maximal isometric force production. To identify ubiquitination localization within the sarcomere and to assess the effect on the sarcomere architecture, single muscle fibers were subjected to the *ex vivo* MuRF1 ubiquitination in the presence of fluorescently labelled ubiquitin (TAMRA-ubiquitin). From the white light images, average relaxed sarcomere length was estimated by measuring sarcomere length in 10 non overlapping regions of an images from a single fiber, whereby each region contained 10-15 sarcomeres (Figure 2.16 B). Sarcomere length was measured from A-band to A-band. The average sarcomere length of a control fiber was 2.47±0.03 μm. Muscle fibers incubated in the reaction mixture lacking MuRF1 (Figure 2.15 Ubiquitin control) showed less than 5% enlargement of sarcomere length (2.58±0.04 μm) suggesting that ubiquitin, E1 and E2 have minimal or no effect on the sarcomere length. However, average sarcomere lengths of a fiber subjected to the ubiquitination reaction varied substantially from the control fibers with the enlargement to 2.84±0.02 μm. The observed enlargement is MuRF1 dependent, but to identify if resultant ubiquitination has a specific pattern on the sarcomere, TAMRA-labeled ubiquitin was detected using the fluorescent imaging.

Fluorescence signal was detected in both samples containing TAMRA-ubiquitin, Ubiquitin control and reaction fibers. Fluorescence detected in ubiquitin control fibers resembles a striated pattern with a uniquely dispersed signal, whilst in reaction fiber along with the striated pattern distinct puncta were observed. The presence of fluorescence signal in ubiquitin control fiber could suggest that more extensive washing steps after incubation are needed to reduce non-specific binding and background noise but could also be due to ubiquitin interacting with sarcomeric proteins in a MuRF1 independent manner.

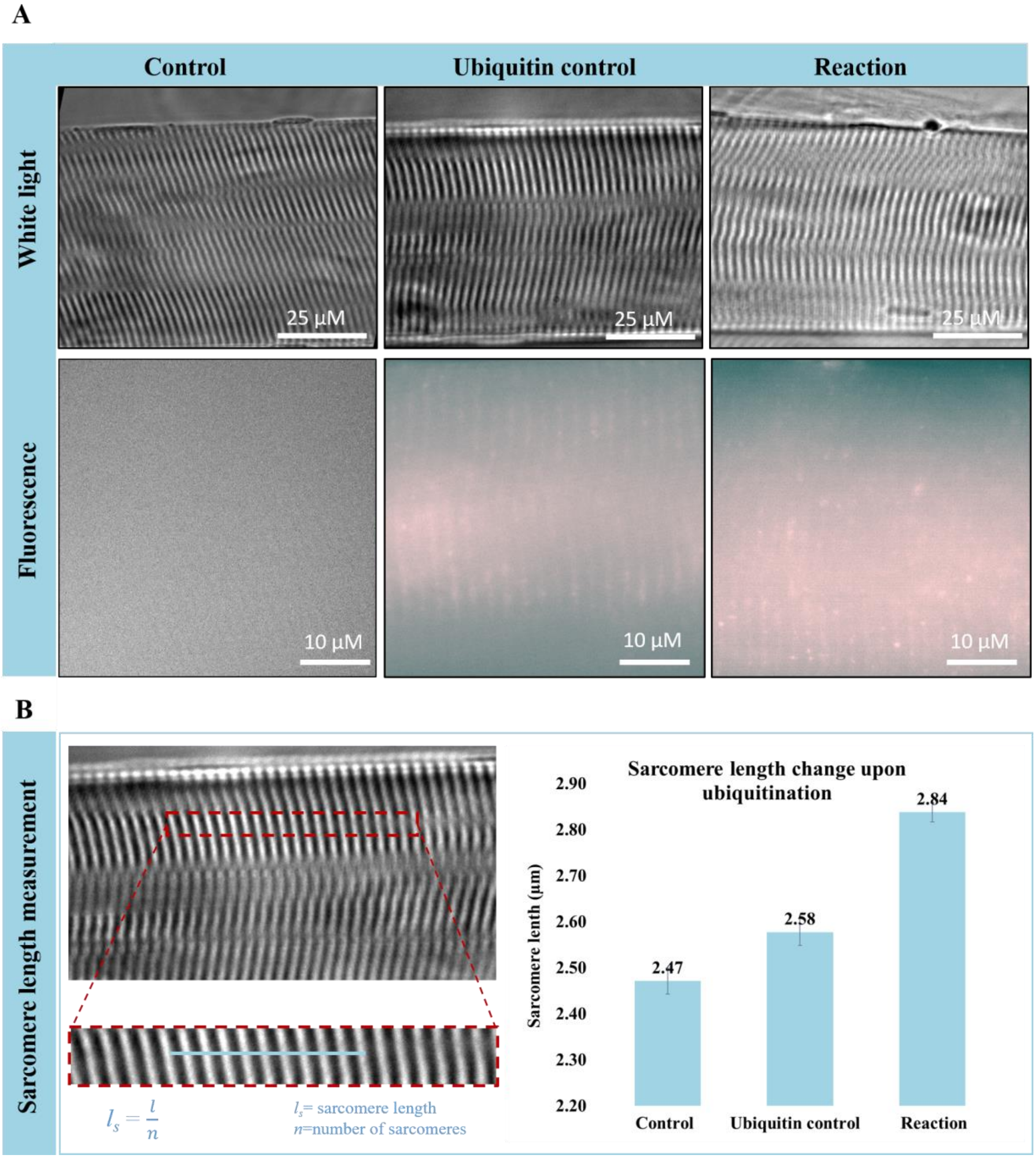


Figure 2.16 Muscle fibers change architecture due to the activity of MuRF1.

A Muscle fibers were treated for the MuRF1-mediated ubiquitination with fluorescently labelled (TAMRA-labeled) ubiquitin. Fibers incubated in control conditions (ubiquitination reaction lacking ubiquitin and MuRF1), ubiquitin control (ubiquitination reaction lacking MurF1) condition as reaction (ubiquitination reaction containing MuRF1 and ubiquitin) were imaged under the optical microscope using white light or fluorescence. Images recorded under the white light show a typical sarcomere striation pattern with alternating dark (I band) and light (A band) zones. TAMRA-ubiquitin fluorescence showed strong signal in both ubiquitin control and reaction samples (pink). Differences were observed in fluorescence patterns, but in the absence of other sarcomeric markers no further assessment can be made.

B Image analysis by ImageJ (Schindelin *et al.*, 2012) were used to determine the average sarcomere length (l_s) by dividing length of selected area (l) by number of containing sarcomeres (n) measured in 10 different areas of a fiber image.

2.4 Discussion

Myosin heavy chain (MYH7) is a dominant player in the muscle contraction because of its role in both creating motion and providing structural support. What enables its role is a dual nature, through ATPase active head domain, myosin interacts with actin molecules creating a motion, while the rod domain provides stiffness and structural support for contraction (Craig and Woodhead, 2006). Maintenance of myosin molecules, along with other sarcomeric proteins must be tightly regulated. However, during muscle atrophy, myosin heavy chain was found to be heavily degraded (Clarke *et al.*, 2007). MYH7 degradation is closely related to the upregulation of a key muscle atrophy marker, the E3 ubiquitin ligase MuRF1 (Bodine *et al.*, 2001). MuRF1 targeting of MYH7 for ubiquitination is well established in both, *in vivo* and *in vitro* experimental studies (Clarke *et al.*, 2007; Fielitz *et al.*, 2007), yet none of the existing evidence can explain the molecular mechanism and the individual effect of MuRF1 on myosin filament degradation. The research field of muscle atrophy has been active for decades studying diverse pathways and looking for possible targets for drug development. To date, however, no effective treatment for muscle atrophy has been found, mostly because the exact mechanism behind protein solubilization and degradation from the sarcomere is not well understood.

This study focused on determining the MuRF1 ubiquitination pattern of MYH7, revealing the molecular mechanism behind the ubiquitination events on myosin depletion, and finally, developing a novel approach in monitoring MuRF1 isolated effects on muscle sarcomere properties.

The *in vitro* ubiquitination assays performed for this study allowed identification of nineteen ubiquitination sites and for the first time, myosin Ub-sites were unambiguously addressed to MuRF1. Our data confirms all six previously detected Ub-sites using *in vivo* techniques (Lang *et al.*, 2017) and adds 13 novel MuRF-specific Ubi-sites. Taken together, all sites identified in literature, combined with our novel 13 sites bring the total of 63 ubiquitination sites found in myosin, regardless of tissue and metabolic state and as such represent a valid dataset for studying the ubiquitination mechanism of myosin. One major question that remains to be answered is what determines whether or not a lysine is targeted for ubiquitination.

It has been proposed that some proteins contain an “embedded code” which serves as a recognition signal for E3 ubiquitin ligases (Jadhav, Wooten and Wooten, 2011). However, no universal recognition code for ubiquitination has been found. For instance, UBR box containing E3 ubiquitin ligases target proteins containing N-degrons (Zenker *et al.*, 2005), Ubc3/Cdc34, Skp1, cullin/Cdc53 and F-box proteins target proteins for fast degradation via proline (P), glutamate (E), serine (S), and threonine (T) (PEST) sequences (Rogers, Wells and Rechsteiner, 1986) and ubiquitination of mitotic cyclins is shown to be mediated by a small motif known as a D-box (Glotzer, Murray and Kirschner, 1991)and KEN-box (Pfleger and Kirschner, 2000). We sought to find such a site existed in the MYH7 by analyzing all identified ubiquitination sites for sequence and structural patterns, however no consensus pattern could be found. Because the myosin rod domain is a coiled-coil, it is more likely that MuRF1 will attack structural patterns rather than sequence specific ones. To look for a structural specific targeting by MuRF1, all known MYH7 ubiquitination sites were mapped on the MARCOIL predicted heptad repeat of MYH7. In comparison to other E3 ubiquitin ligases, MuRF1 shares the highest preference for lysine in position c, but no lysines seems to be targeted in heptad position e and g based on the current data for lysine. Positions e and g in heptad repeats are generally involved in forming salt bridges that maintain the structural integrity of coiled-coils (Lupas and Gruber, 2005). Structural analysis of available myosin rod domain structures strongly supports MARCOIL prediction and further show that the majority of ubiquitination targeted Lysine residues in MYH7 are found in surface available areas. These findings indicate that MuRF1 may not act on signal recognition patterns, but rather attack the most accessible lysine residues within the myosin coiled-coil, interrupting filament packing and increasing myosin diffusibility (Figure 2.16). These results contribute not only to better understanding of the mode of action of MuRF1 but could also serve in future research oriented towards revealing recognition patterns in ubiquitin proteomics.

To investigate if MuRF1 ubiquitination would cause the disturbance in myosin filament packing, which could further affect the sarcomere architecture and performance, prior to degradation, the special *ex vivo* ubiquitination assay was designed in which single muscle fibers were subjected to MuRF1 ubiquitination reaction and the effects were measured by isometric force assessment. This experimental set up allows for monitoring the direct effect of MuRF1 on the muscle sarcomere, abstracted from the numerous parallel pathways, normally activated upon muscle trophy signal *in vivo*. As a surprising result, MuRF1 ubiquitination reactions caused treated muscle fibers to decrease their isometric force production by 50%. Isometric force production is directly related to the number of cross-bridges formed between myosin heads and actin molecules in the sarcomere array (Colombini *et al.*, 2010). Ubiquitin, although a small protein, is a large posttranslational modification which can disrupt intermolecular interactions between neighboring proteins, important for fiber packing in the sarcomere (Figure 2.17). As a matter of fact, average sarcomere length estimation experiments indicate elongation of a sarcomere by almost 20% upon MuRF1 ubiquitination and could possibly explain the decrease in isometric force production. Hereby, it is speculated that the attachment of ubiquitin could interrupt the packing between sarcomeric proteins disturbing the regular array and enhance diffusibility of myosin (Figure 2.16). The role of ubiquitination in stabilization and destabilization of tight packing is a familiar concept proposed for E3 ubiquitin ligase Ozz active in myosin filament formation in early development, where through ubiquitination facilitates myosin isoform replacement, (Campos *et al.*, 2010) and for the CRL complex, important in neurodegenerative disease, where through its ubiquitination activity CRL prevents protein aggregate formation (Chen *et al.*, 2019).

Finally, here is proposed mechanism by which posttranslational modification-ubiquitination mediated by MuRF1 does not only label sarcomeric proteins for degradation, but also affects structural integrity of a sarcomere packing, prior to degradation (Figure 2.17). Myosin rod domain forms a core of the myofibril via tight interactions between hundreds of myosin molecules (Figure 2.17 tight filament), decreasing accessibility for interaction with cytoplasmic proteins. MuRF1 might attack myosin filament via interacting with available surface areas (Figure 2.17 Step1) and mediate ubiquitination of accessible lysine residues (Figure 2.17 Step2). Introduced ubiquitin modification may cause disturbance of a tight packing of the filament and induce loosening of the overall sarcomere packing (Figure 2.17 loose filament). This model of MuRF1 mediated MYH7 ubiquitination could be used to explain a decrease in isometric force and increase in sarcomere length upon MuRF1 ubiquitination detected in *ex vivo* experiments (Figure 2.15; Figure 2.16). It could be speculated that as a result, loose packing in myosin filament allows for single molecules to retreat and become a target of ubiquitin-degradation machinery in sarcoplasm.

Further research, however, should investigate myosin diffusibility from the muscle sarcomere upon ubiquitination and explore applicability of *ex vivo* ubiquitination experiments in drug development. It would also be interesting to monitor the ubiquitination pattern on the sarcomere components and identify if it is the tick filament being primary targeted or all components. To improve detection, TAMRA-ubiquitin experiments require further optimization and addition of sarcomere markers. Nevertheless, here developed approach and presented possible mode of action could provide useful insight into the individual impact of MuRF1 activity on muscle function and may be further used to guide the understanding of muscular pathologies and potential treatments.

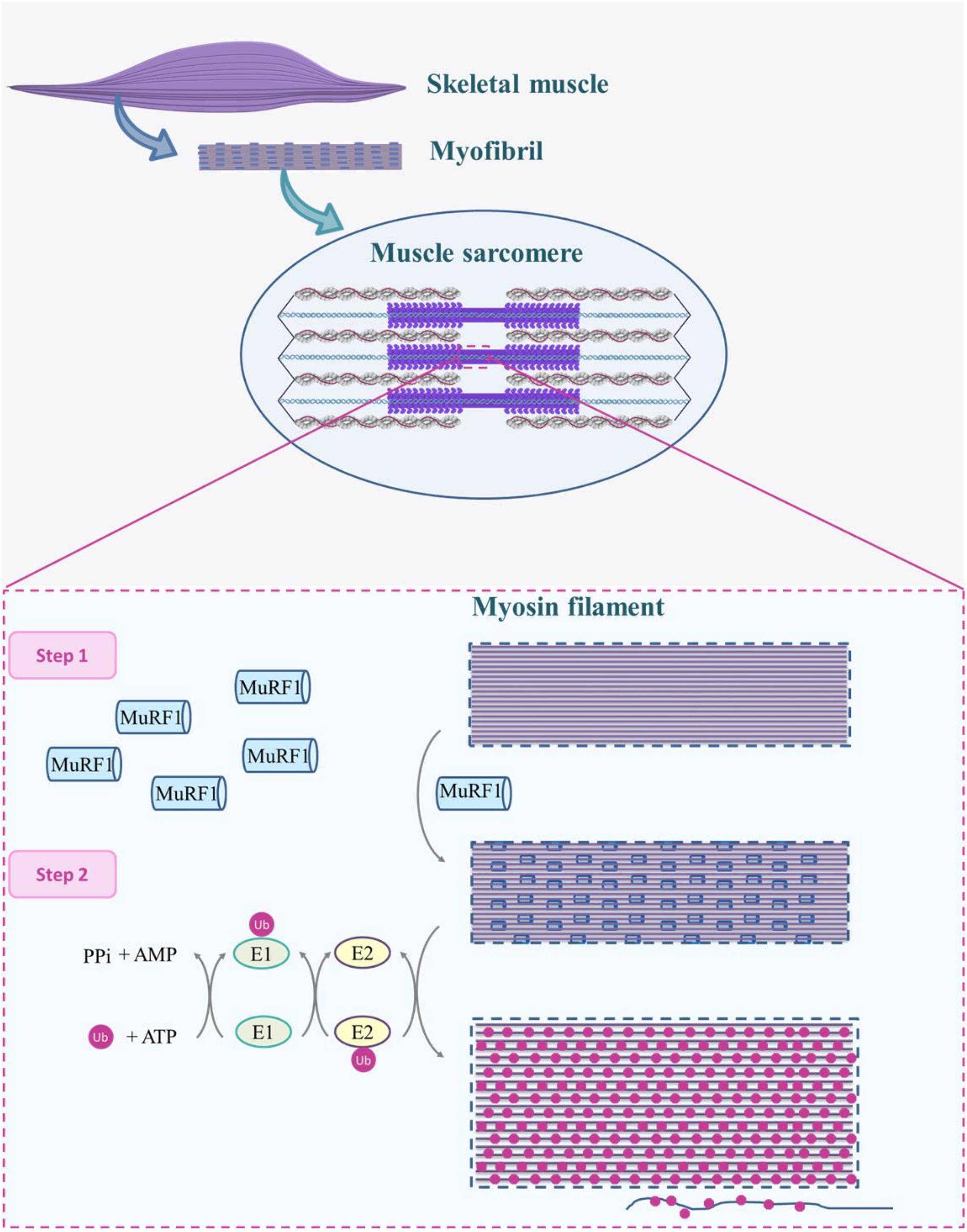


Figure 2.17 Muscle sarcomere ubiquitination.
Skeletal muscles are composed of myofibrils in which proteins are strictly organized in a pseudo crystalline structure of a sarcomere. In the *ex vivo* ubiquitination performed in this study Step 1 was pre-incubation of muscle fibers with MuRF1, allowing protein molecules to settle and specifically bind. In the Step 2 remaining components of the ubiquitin reaction, including E1, E2 and ubiquitin were added. As a result, introduced ubiquitin interrupts packing between sarcomeric proteins directly affecting its performance.

3

The mutational alteration of BAG3 causes loss-of function in HSP70 and leads to cardiomyopathy

3.1 Introduction

Misfolded or damaged proteins are involved in the pathogenesis of neurodegenerative disorders, skeletal myopathies and heart failure (Mayer and Bukau, 2005; Kampinga and Bergink, 2016) The 70 kDa Heat Shock Proteins (HSP70s) are a conserved family of molecular chaperones that contribute to the PQCS by preventing newly synthesized proteins from misfolding and misfolded proteins from aggregating (Ganassi *et al.*, 2016) If the client protein cannot be refolded due to a severe damage, HSP70s can recruit protein degradation by facilitating client transfer to the proteasome (Fernández-Fernández and Valpuesta, 2018). Diversity of cellular functions exerted by HSP70 molecular machinery rely on the ATP dependent dynamic cycle.

3.1.1 Dynamic cycle of HSP70

HSP70 proteins consist of an N-terminal nucleotide-binding domain (NBD) of 43 kDa and a C-terminal substrate-binding domain (SBD) of 27 kDa connected by a highly conserved linker, which mediates allosteric regulation of conformational change between NBD and SBD (Figure 3.1 A). The NBD is composed of four subdomains (IA, IB, IIA, and IIB) arranged in two lobes that form a deep crevice between them; ATP binds at the bottom of this crevice and is coordinated by all four subdomains (Figure 3.1 A) (Sondermann *et al.*, 2001). SBD is divided in two subdomains, protein client binding site as two-layered twisted β-sandwiches (SBDβ) and alpha-helical subdomain (SBDα), also called “lid” that ends in 30 residues unstructured stretch (Figure 3.1 A) (Pellecchia *et al.*, 2000). Mechanistically, the functional cycle of HSP70 starts when an ATP molecule binds to the nucleotide binding site of NBD (Figure 3.1 B). In the ATP-bound state, the SBD lid is in "open" configuration and affinity for client binding is relatively low, whilst in ADP-bound state, lid is in "closed" configuration and the affinity for client binding is high (Palleros *et al.*, 1993). The hydrolysis of ATP, thereby is essential for an effective substrate capture in the SBD. As the basal ATPase activity of NBD is low and rate limiting step, HSP70s require assistance of a co-chaperone. Binding of a co-chaperone protein J stimulates the nucleotide hydrolysis and enables efficient binding of the protein client to the SBD (Figure 3.1 B) (Walsh *et al.*, 2004; Behl, 2016; Kampinga and Bergink, 2016). Upon the protein refolding, client release is promoted by SBD lid opening, triggered by a release of ADP from a NBD and binding of a new ATP molecule. As the nucleotide exchange in NBD is a very slow process, specialized co-chaperones-Nucleotide exchange factors (NEFs) are required (Hartl and Hayer-Hartl, 2002; Bukau, Weissman and Horwich, 2006). NEFs by interacting with the NBD domain, facilitate ADP release, enable binding of the new ATP molecule and as a consequence, foster protein client dissociation from SBD (Mayer and Bukau, 2005).

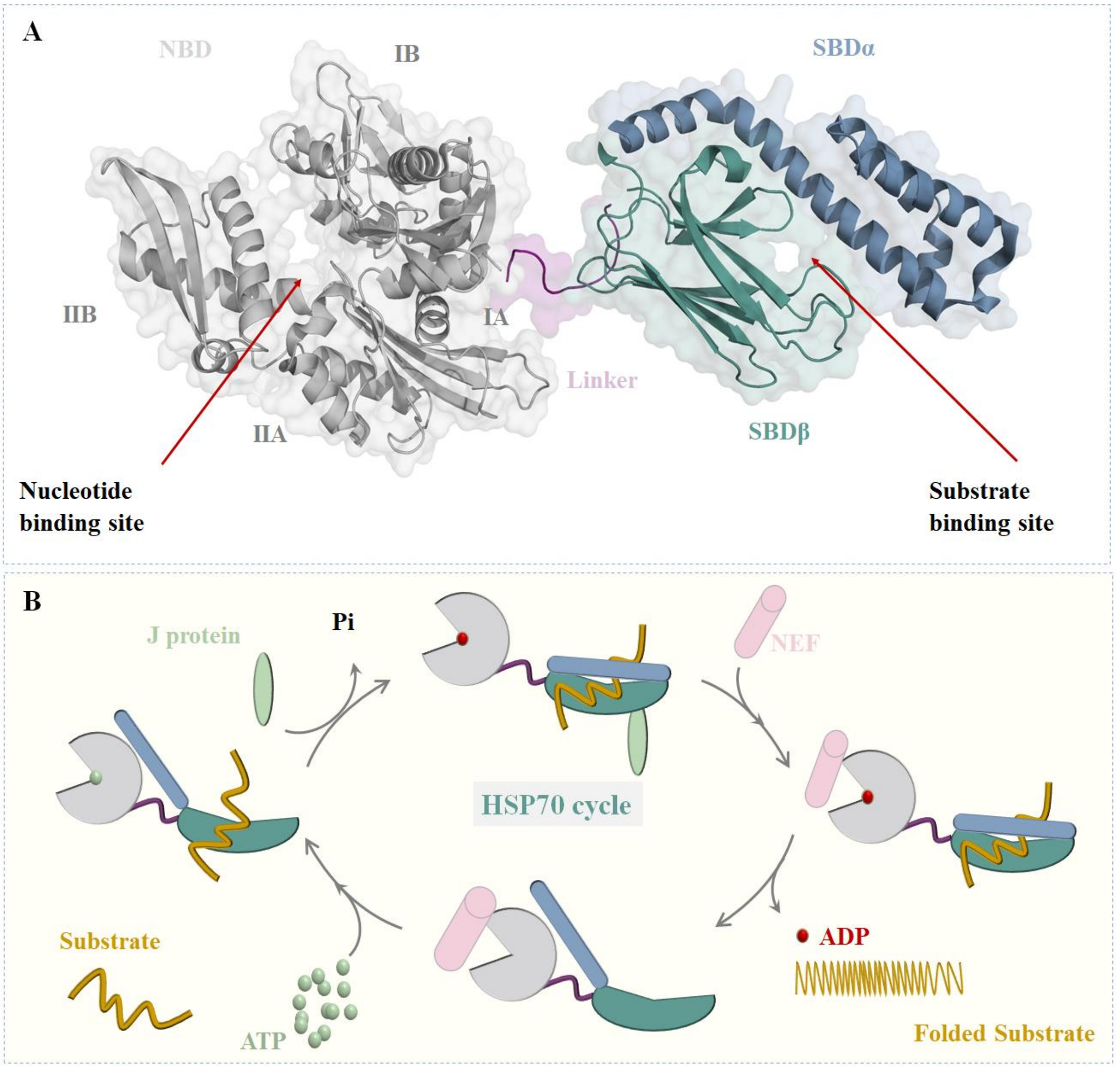


Figure 3.1 HSP70 cycle mediates protein folding.
A Crystal structure of the *E. coli* HSP70 analogue DNaK (PDB ID: 2KHO) showing an N-terminal nucleotide binding domain (NBD) in grey, organized in four subdomains (IA, IIA, IB and IIB) and a substrate binding domain (SBD) connected by a flexible linker (Linker). The SBD is composed of two subdomains, an SBDβ shown in green and an SBDα (lid) shown in blue; **B** The HSP70 cycle starts with binding of an ATP molecule to the NBD and a protein client (Substrate) to the SBD. Nucleotide hydrolysis is enhanced by protein J and causes the lid domain to close. An ADP release is a slow process and requires NEFs assistance. The release of an ADP and binding of a new ATP molecule starts the new cycle and stimulates the opening of the SBD lid domain and substrate release.

3.1.2 BAG3 is a nucleotide exchange factor of HSP70

The Bcl-2 associated athanogene (BAG) family is a multifunctional evolutionarily conserved group of proteins that mediates multiple cellular processes (Takayama, Xie and Reed, 1999) The family currently consists of six members that share the BAG protein domain BAG1, BAG2, BAG3 (BIS/CAIR), BAG4 (SODD), BAG5 and BAG6 (Scythe/BAT3) (Takayama and Reed, 2001). BAG domain folds into three alpha helices arranged in an anti-parallel manner with helices 2 and 3 mediating interaction with HSP70-NBD. Among the BAG proteins, BAG3 has been found to have the highest binding affinity for NBD measured *in vitro* (Rauch and Gestwicki, 2014). Interestingly, BAG proteins are constitutively expressed at the low levels in the cardiac and skeletal muscle, but when triggered by diverse stressors, BAG3 becomes heavily upregulated, whilst the other BAG proteins expression remain at the basal level (Franceschelli *et al.*, 2008). In addition, high expression levels of the BAG3 were also detected in cells of different cancer types, such as neuroblastoma, glioblastomas, pancreatic cancers and ovarian and breast carcinomas (Liao *et al.*, 2001; Franceschelli *et al.*, 2008; Gentilella *et al.*, 2008; Felzen *et al.*, 2015). Multiple functions exerted by the BAG3 in the cell are closely related to its diverse function domain composition (Stürner and Behl, 2017).

The human BAG3 is composed of four characteristic amino acid domains: a WW (tryptophan-tryptophan) domain, two IPV (isoleucine–proline–valine) motifs, a PxxP (proline-rich) region and a BAG domain. Two conserved IPV motifs mediate the binding of BAG3 to the small heat shock proteins HSPB8 (HSP22) as well as HSPB6 (HSP20) (Carra, Seguin and Landry, 2008; Fuchs *et al.*, 2009; Rauch *et al.*, 2017) and the PxxP repeat is known to interact with the motor protein dynein adaptor for targeting aggregated proteins for the autophagy via the microtubule network (Stürner and Behl, 2017). Mutations in the functional domains of BAG3 protein are found as a common cause in cardiac diseases (Domínguez *et al.*, 2018). The point mutation P209L mutation in IPV motif found in patients suffering from myofibrillar myopathy, indicated a stalling effect in HSP70 machinery, which promotes accumulation of aggregated proteins in heart muscle (Meister-Broekema *et al.*, 2018). The E455K in the BAG domain of BAG3 was found in a number of patients with dilated cardiomyopathies (Villard *et al.*, 2011). Binding studies *in vitro* showed a mild decrease in the binding affinity of BAG3 for HSP70 (Rauch, Zuiderweg and Gestwicki, 2016; Fang *et al.*, 2017) that was in contrary to *in vivo* collected evidence of increased aggregation of misfolded proteins in the heart of mice carrying the E455K mutation in BAG3 (Fang *et al.*, 2017). Thus, a direct correlation between the loss of function mutation E455K and the pathology of dilated cardiomyopathy has been proven *in vivo*, yet little is understood about the underlying molecular mechanism.

Although the crystal structure of HSP70 nucleotide binding domain is well known and abundantly studied (Mayer, 2013), certain aspects of its structure-function features are yet to be fully established. The structure of BAG3 complexed to HSP70 is unknown and the molecular mechanism of nucleotide exchange enhancement by BAG3 is currently not well understood. Obtaining high-resolution structural information of the BAG3 wild type and

E455K mutation complexes with NBD will contribute to understanding the molecular mechanisms of the co-chaperone and modulatory mechanism involved in the development of pathological condition

Aims

The main objective of this chapter is to obtain better understanding of the molecular mechanism underlying mutation E455K modified BAG3 binding to HSP70 and inhibition of function of the HSP70 cycle, found to be a cause of dilated cardiomyopathies. Therefore, a combination of X-ray crystallography, molecular dynamic simulations and Bio-Layer interferometry (BLI) is used to study structural changes in complexation and monitor real-time binding of HSP70 to BAG3 in the nucleotide free or nucleotide bound states.

3.2 Material and methods

3.2.1 Recombinant protein production and purification

The NBD domain of murine HSP70 (UniProtKB - Q61696, residues 1–388) and the BAG domain of BAG3 (UniProtKB - O95817, residues 419–509) in its wild type and E455K were overexpressed in *E. coli* Rosetta (Novagen). Constructs were cloned into PETM-11 vectors (EMBL collection) which contain an N-terminal His_6-tag followed by a tobacco etch virus (TEV) protease cleavage site by Dr. Julius Bogomolovas. Cell cultures were grown at 37°C in LB (Luria-Bertani) liquid medium to an OD_{600} of 0.6 and induced with 0.5 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) and cultured further at 18°C for 16h. Cells were harvested by centrifugation at 5000 rpm (9760 xg) at 4°C for 20 minutes and the pellets were resuspended in 50 mM Tris-HCl buffer pH 7.5, 500 mM NaCl (lysis buffer). Lysis of the cells was achieved by sonication on ice. After centrifugation of the lysed cells at 18000 rpm (47850 xg) at 4°C for 45 minutes, the supernatant was applied to a HisTrap HP (GE Healthcare) pre-equilibrated in lysis buffer with 20 mM imidazole. The protein was eluted using an increasing gradient of imidazole concentrations from 20 mM to 1 M. The His_6-tag of HSP70-NBD, BAG3-BAG^{WT} and BAG3-BAG^{E455K} were removed by overnight incubation with TEV protease at 4°C followed by reverse HisTrap purification. Proteins were further purified using size exclusion chromatography, Superdex S200 26/60 prep grade high load column (GE Healthcare) in 50 mM Tris-HCl pH 7.5, 150 mM NaCl. For the BLI experiments, after affinity purification His_6-tag was not removed from BAG3-BAG^{WT} and BAG3-BAG^{E455K} and samples were purified by loading onto the Superdex S75 26/600 column (GE Healthcare) in 50 mM Tris-HCl pH 7.5, 150 mM NaCl. Purified samples were stored at 4°C until further use.

3.2.2. Crystallization

HSP70-NBD and BAG3-BAG or BAG3-BAG^{E455K} were mixed in 1:10 molar ratio and incubated overnight at 4°C. The complexes were purified using a Sephadex G-200 26/600 (GE Healthcare) size exclusion column. Two peaks were observed and analyzed by SDS-PAGE in both cases. The protein complexes were concentrated in 50 mM Tris-HCl pH 7.5 to 19 mg/mL and were screened for crystallization using an automated nano volume-dispensing crystallization robot (Crystal Gryphon LCP, Art Robbins Instruments) in sitting drop 96-well Intelli-plates (Art Robbins Instruments) with the drop size of 200 nL and commercial crystallization screens (Table 3.1).

Table 3.1 Commercial crystallization screens used for BAG-NBD crystallization study

Company	**Screens**
Molecular Dimensions	JCSG plus Eco, Structure Screen 1 and 2, Morpheus, Proplex 1 and 2
Hampton Research	PEGRx HT, Natrix 1 and 2, SaltRx 1 and 2
Rigaku Reagents	Wizard I and II, Cryo Wizard

Crystals grew in multiple conditions were tested for X-ray diffraction and high resolution diffraction data were obtained only for one condition for the wild type complex and one condition for the mutant complex. HSP70-NBD/BAG3-BAG^{E455K} was crystallized at 20°C from sitting drops containing equal volumes (100 nL) of protein and reservoir solutions (0.1 M Na-HEPES, 0.1 M MOPS pH 7.5, 0.1 M L-Na-Glutamate; 0.1 M Alanine (racemic); 0.1 M Glycine; 0.1 M Lysine HCl (racemic); 0.1 M Serine (racemic), 12.5% [v/v] MPD; 12.5% PEG 1000; 12.5% [w/v] PEG 3350) and HSP70-NBD/BAG3-BAG was crystallized in the same conditions and reservoir solutions (0.1 M Na-HEPES, 0.1 M MOPS pH 7.5, 0.1M L-Na-Glutamate; 0.1 M Alanine (racemic); 0.1 M Glycine; 0.1 M Lysine HCl (racemic); 0.1 M Serine (racemic), 40% [v/v] PEG 500 MME; 20 % [w/v] PEG 20000).

3.2.3 Data collection, model building and refinement

Crystals were cryoprotected with mother liquor containing 30% [v/v] glycerol prior to cryocooling in liquid nitrogen. X-ray diffraction data were collected on beamline PXI (Swiss Light Source, Switzerland) and indexed using XDS (Kabsch, 2010). Molecular replacement was conducted with Phaser (McCoy *et al.*, 2007) using HSC70-NBD and the BAG5-BAG domain as individual search models (PDB: 1ATR and 3A8Y) and the resulting solution was refined in PHENIX[41]. Manual building was performed on WinCoot (Emsley *et al.*, 2010). All the structural alignments and figures were prepared with Pymol (http://www.pymol.org).

3.2.4 Biolayer interferometry

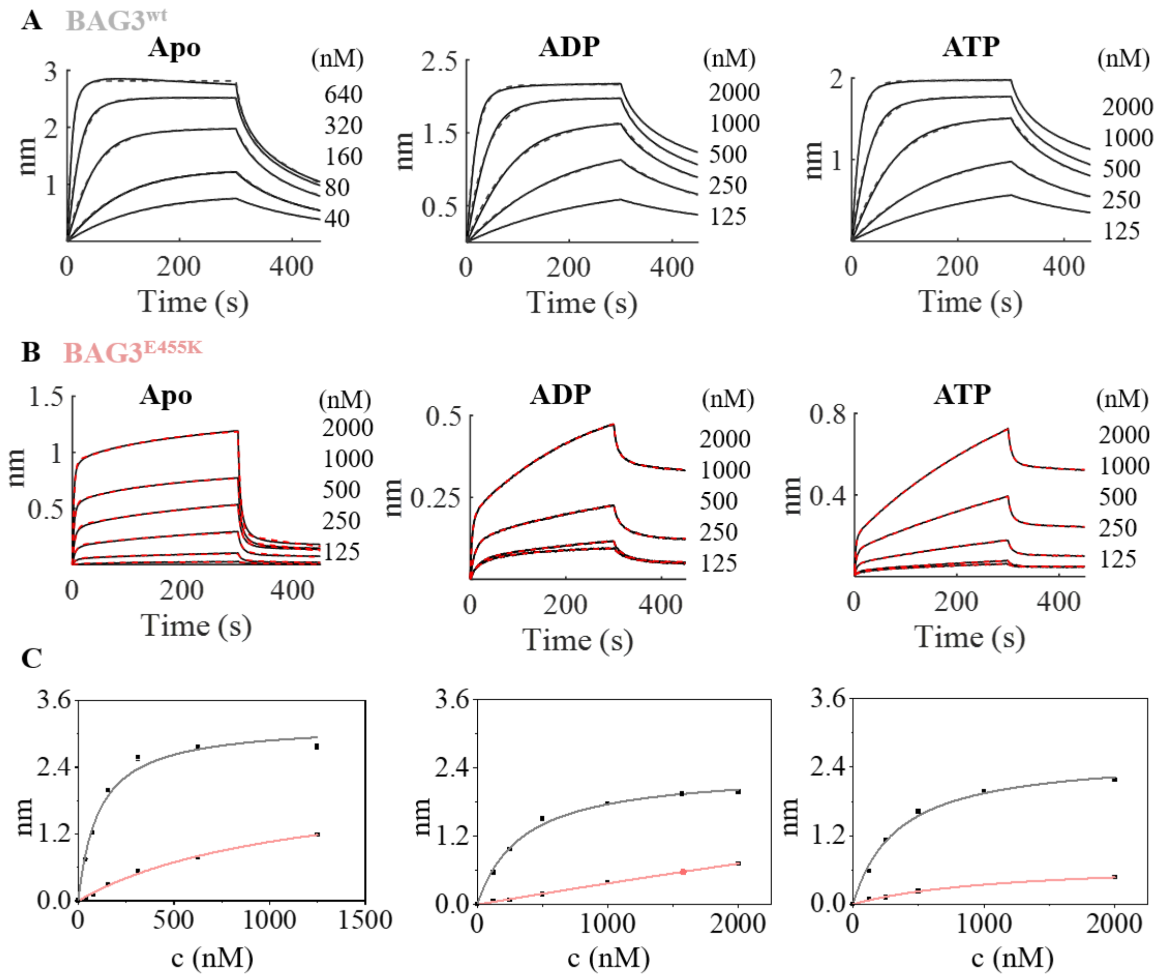
The protein-protein interaction study was performed using a biolayer interferometry Octet K2 instrument (Pall Forte Bio). All assays were performed at 30 °C in solid black 96-well plates (Geiger Bio-One) with the agitation set to 1000 rpm using Ni^{2+}-NTA biosensors (ForteBio). Sensors were pre-hydrated by 10 min incubation in the running buffer (50 mM Tris buffer pH 7.5,180 mM NaCl). Baseline step of 60s was followed by 150s of a loading step with 35 µg/mL His_6-BAG3-BAG or His_6-BAG3-BAG^{E455K}. Prior to the association step, washing and a second baseline were obtained for 60s in the running buffer. Binding of HSP70 to BAG3 was monitored in the association step for 300s in the concentration range 0.1-2 µM in running buffer, followed by 150 s dissociation step in running buffer. For experiments including the nucleotides effect on binding affinities, the running buffer contained 1 mM $MgCl_2$ and 1:2 molar ratio of protein: nucleotide (ATP or ADP). Concentrations used for the wild type APO form were: 40, 80, 160, 320, 600, 1250 nM, while the concentrations used for the other conditions were :125, 250, 500, 1000, 2000 nM. Non-specific binding of HSP70 to the sensor was not detected during the experiments. The control signal was obtained from His_6-BAG^{wt} in the running buffer and was subtracted from the interaction signal for data analysis. Data analysis and curve fitting were performed using Data analysis software version 10.0.1.6 (Pall Forte Bio) and experimental data were fitted with the binding equation available for 2:1 heterogeneous ligand (HL) for all measurements. By plotting the signal recorded at the end of the association step against the HSP70 concentration, Steady-state parameters were calculated using the Michaelis-Menten equation. Data was plotted using Origin 2020 software (OriginLab Corporation).

3.3 Results

3.3.1 The E455K mutation of BAG domain impairs the HSP70-BAG3 binding

Nucleotide exchange in the HSP70 cycle, as a rate limiting step, is enhanced by the direct binding of BAGs to the nucleotide binding domain of HSP70 (Sondermann *et al.*, 2001). In order to monitor NBD interaction with the BAG domain of BAG3 and assess the effect of the E455K point mutation on binding affinity, Biolayer interferometry (BLI) binding experiments were conducted in absence or presence of nucleotides (ATP and ADP). The increase of a BLI signal shift in the association step, upon immersing BAG-coated Ni^{2+}-NTA (Forte Bio) biosensor in varying concentration solutions of NBD in assay buffer (or buffer containing nucleotides) was recorded for 300s, followed by a dissociation step in the assay buffer (or buffer containing nucleotides) for 150s (Figure 3.2 A and B). The recorded data were analyzed using the global 2:1 heterogeneous ligand model fit and the steady-state analysis (Figure 3.2 C) were used for calculating the equilibrium dissociation contracts (Kd) (Figure 3.2 D). Obtained results shows approx. 10-fold decrease in binding affinity of NBD for BAG due to the E455K mutation (Figure 3.2 A and B Apo), but more interestingly, that binding of the BAG^{E455K} domain is mostly disturbed in the presence of ATP (Figure 3.2 A and B ATP and Figure 3.3 B).

The binding affinity of $BAG3^{WT}$ for the NBD-APO is 0.1 ± 0.015 µM and for NBD-ADP and NBD-ATP is 0.34 ± 0.04 µM and 0.35 ± 0.04 µM, respectively. Upon the mutation, the binding affinity for the NBD-APO form decreases to 1.1 ± 0.2 µM and to the NBD-ADP 1.4 ± 0.1 µM. Strikingly, a drastic decrease in the binding affinity was observed for the NBD-ATP with the calculated binding affinity of 26 ± 23 µM. Moreover, increasing the concentration of ATP to 1mM completely abolished detection of NBD-ATP binding to the $BAG3^{E455K}$ (Figure 3.3). Together these results indicate that the residue E455 in BAG3 is not critical for the complex formation itself, but it contributes to the affinity of complexation, specifically in the presence of ATP.



D

Apo		ADP		ATP	
Kd (µM)	R^2	Kd (µM)	R^2	Kd (µM)	R^2
0.13±0.015	0.9875	0.35±0.04	0.9864	0.34±0.04	0.9864
1.1±0.2	0.9919	1.4±0.1	0.9986	26±23	0.9968

Figure 3.2 Biolayer interferometry measurements of HSP70-NBD binding affinity to BAG3-BAG.

Ni^{2+}-NTA sensors were loaded with **A** His_6-BAG^{wt} and **B** His_6-BAG^{E455K} and the BLI signal of nm shift was recorded for association step (0-300s) in absence (Apo) or presence of nucleotides (ATP and ADP) in 1:2 molar ration with HSP70-NBD in varying concentrations, followed by a dissociation step (300-450s) in buffer solution containing nucleotides only. The experimental data is shown in black lines and corresponding fitting curves to the 2:1 heterogeneous ligand binding model are shown as dash lines in grey for the wild type and in salmon for mutant. **C** Steady-state analysis of obtained data was used to determine equilibrium binding affinity (Kd) by plotting the BLI response recorded at the end of association step against the HSP70-NBD concentration and calculated values are reported **D**.

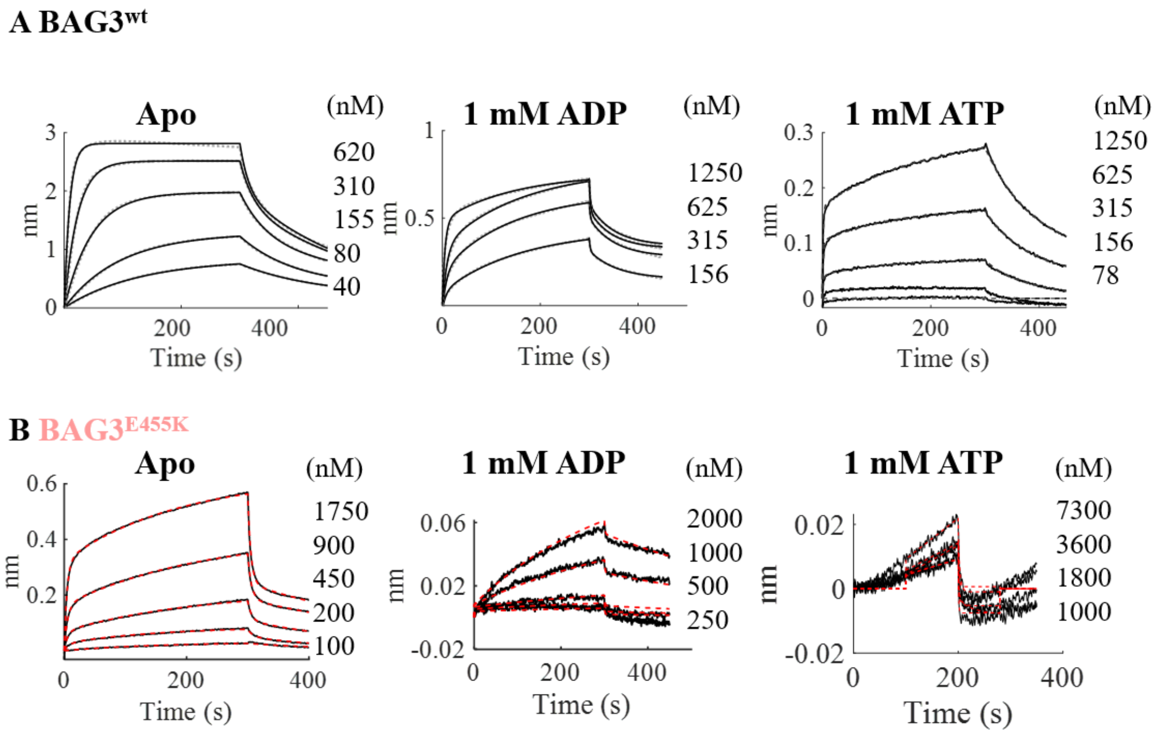


Figure 3.3 Increased concentration of nucleotides to 1mM decreases the BLI signal of HSP70-NBD binding to BAG3-BAG.
Ni^{2+}-NTA sensors were loaded with **A** His_6-BAG^{wt} and (**B**) His_6-BAG^{E455K} and the BLI signal of nm shift was recorded for association step (0-300s) in absence (Apo) or presence of nucleotides (ATP and ADP) in 1mM molar concentration, followed by a dissociation step (300-450s) in buffer solution containing nucleotides only. **B** Low and unstable signal obtained in the presence of 1mM nucleotides unbaled calculation of the kinetic parameters.

3.3.2 E455 in BAG3 is essential for its NEF activity

The BAG domain acts as a nucleotide exchange factor for the Hsp70 family of chaperones, through accelerating the nucleotide exchange rate (Jennifer N. Rauch and Gestwicki, 2014). To measure the effect of the E455K mutation on the ADP exchange rate for NBD of HSP70, Dr. Julius Bogomolovas implemented fluorescent MABA-ADP in a stopped-flow experimental setup. Previous studies reported the suitability of this approach for studying NEF activity (Theyssen *et al.*, 1996). The decay of the MABA-ADP fluorescence signal, resulting from the exchange of analogue to unlabeled ADP upon mixing of MABA-ADP/NBD Hsp70 complex with excess of unlabeled ATP was recorded. As seen from the concentration dependent acceleration of signal decay, both $BAG3^{WT}$and $BAG3^{E455K}$ stimulated the off rate of MABA-ADP for NBD of Hsp70 (Figure 3.4 A). However, $BAG3^{WT}$

acted as a more potent NEF, than BAG3^{E455K}. Both studied BAG3 constructs enhanced spontaneous ADP off-rate (k_{off}^{ADP}= 0.0786 s^{-1}) in a concentration dependent manner. BAG3^{WT} accelerated this process up to 289 times in the studied concentration range, whereas BAG3^{E455K} only up to 4.8. Overall, BAG3^{WT} was up to 60 times more efficient NEF than BAG3^{E455K} in the studied concentration range. Taken together, obtained data indicate that BAG3^{E455K} has severely impaired NEF activity towards HSP70.

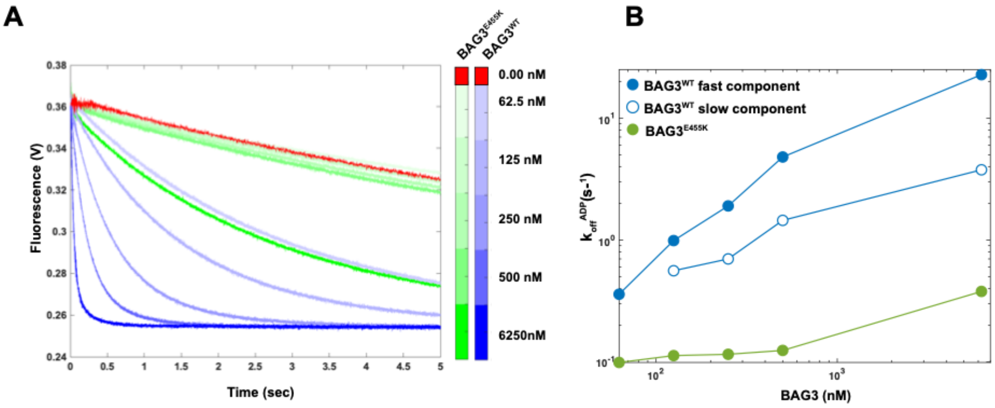


Figure 3.4 Stimulation of ADP release from HSP70 NBD by BAG3.
A Dissociation curves of MABA-ADP·HSP70-NBD complex upon mixing with excess of unlabeled ADP in the absence (red line) and presence of varying concentrations of BAG3^{WT} (blue) or BAG3^{E455K}(green). Color gradient corresponds to BAG3 concentrations.
B The k_{off}^{ADP} values of MABA-ADP for Hsp70 NBD were determined at various BAG3 concentrations, and the rates were plotted against the BAG3 concentration. For higher concentrations of BAG3^{WT} second slow component could be identified by fitting a double exponential curve (hollow circles). (Figure provided by Dr. Julius Bogomolovas)

3.3.3 BAG-NBD complex structure is not affected by E455K mutation

The effect of E455K mutation on functional properties of BAG3 effective binding to NBD and NEF activity have been observed in BLI and nucleotide exchange experiments, yet in order to understand the mechanical properties of the BAG-NBD interaction and how E455K mutation impairs the function of HSP70, structural studies were conducted. As the affinity of NBD-BAG^{E455K} binding is significantly weakened in the presence of nucleotides, in the co-crystallization attempt no crystals were found for the ATP and ADP bound complexes. Nevertheless, the BAG and BAG^{E455K} complexed to the APO NBD were crystallized and a 2.3 and 2.2 Å X-ray reflection data sets were collected, respectively (Table 3.2). The X-ray structures were determined by molecular replacement, using HSC70 (PDB ID: 1ATR) and the BAG (PDB ID: 13A8Y) as individual search models. Structures were refined to an R/R$_{free}$ of 0.18/0.23 for the mutant and 22.5/30 for the wild type complex (Table 3.2). Crystal of NBD-BAG contains one copy per asymmetric unit, while the asymmetric unit of NBD-BAG^{E455K} contains two copies of the complex with NBD in a different conformational arrangement (Figure 3.5 A). By convention, the two E455K complex copies are labelled as Copy A and Copy B (Figure 3.5 A). The overall folds of Copy A and Copy B in the asymmetric unit differ only by the RMSD 1.4 Å (for all atoms) with a main difference in the relative position of the helix 2 in IIB subdomain of NBD (Figure 3.5 B). The overall fold of the BAG domain remained intact upon the mutation (Figure 3.5 C). The structures show a typical BAG domain consisting of three alpha helices arranged in an antiparallel manner with helix2 and helix3 forming an interaction interface with the NBD and the E455K mutation located in helix2.

Table 3.2 X-ray data collection and refinement statistics

	HSP70- BAG3	**HSP70-BAG3^{E455K}**
X-ray source	SLS-PXIII	SLS-PXI
Detector	PILATUS 2M-F	EIGER 16M X
Wavelength (Å)	1.000	1.000
Data collection		
Space group	$P\ 1\ 2_1\ 1$	$P\ 1\ 2_1\ 1$
Cell dimensions		
a, b, c (Å)	52.1 91.5 58.2	57.6 90.6 91.4
α, β, γ (°)	90 113.3 90	90 95.6. 90
Unique reflections [a]	22344 (4914)	46542 (3013)
Resolution (Å)	28.64 - 2.3 (2.5 - 2.3)	48.44 - 2.2 (2.25 - 2.2)
Copies in asymmetric unit	1	2
R_{sym} (%)	7.1 (148.8)	12.5 (172.8)
I / σI	15.26 (1.24)	9.79 (1.10)
Completeness (%)	99.6 (99.7)	97.2 (97.6)
$CC_{1/2}$ (%)	100 (78.3)	99.7 (39.8)
Redundancy	6.91 (7.11)	7.18 (7.23)
Refinement		
Resolution, Å	2.3	2.2
R_{work}[b]/R_{free}[c]	22.5/30.33	18.08/23.87
Number of atoms		
Protein	3571	7250
Ligand	5xGOL[d]	7xGOL, 1x MPO[e]
Water	30	293

[a] Values in parentheses refer to the highest-resolution bin

[b] $R_{work} = \sum hkl||F_o|–|F_c||/\sum|F_o|$, where F_o is the observed structure factor amplitude and F_c is the structure-factor amplitude calculated from the mode;

[c] R_{free} is calculated using a subset, 2%, of the data that are not included in any refinement calculations;

[d] GOL-Glycerol;

[e] MPO-3[N-Morpholino]-Propane Sulfonic Acid (MOPS);

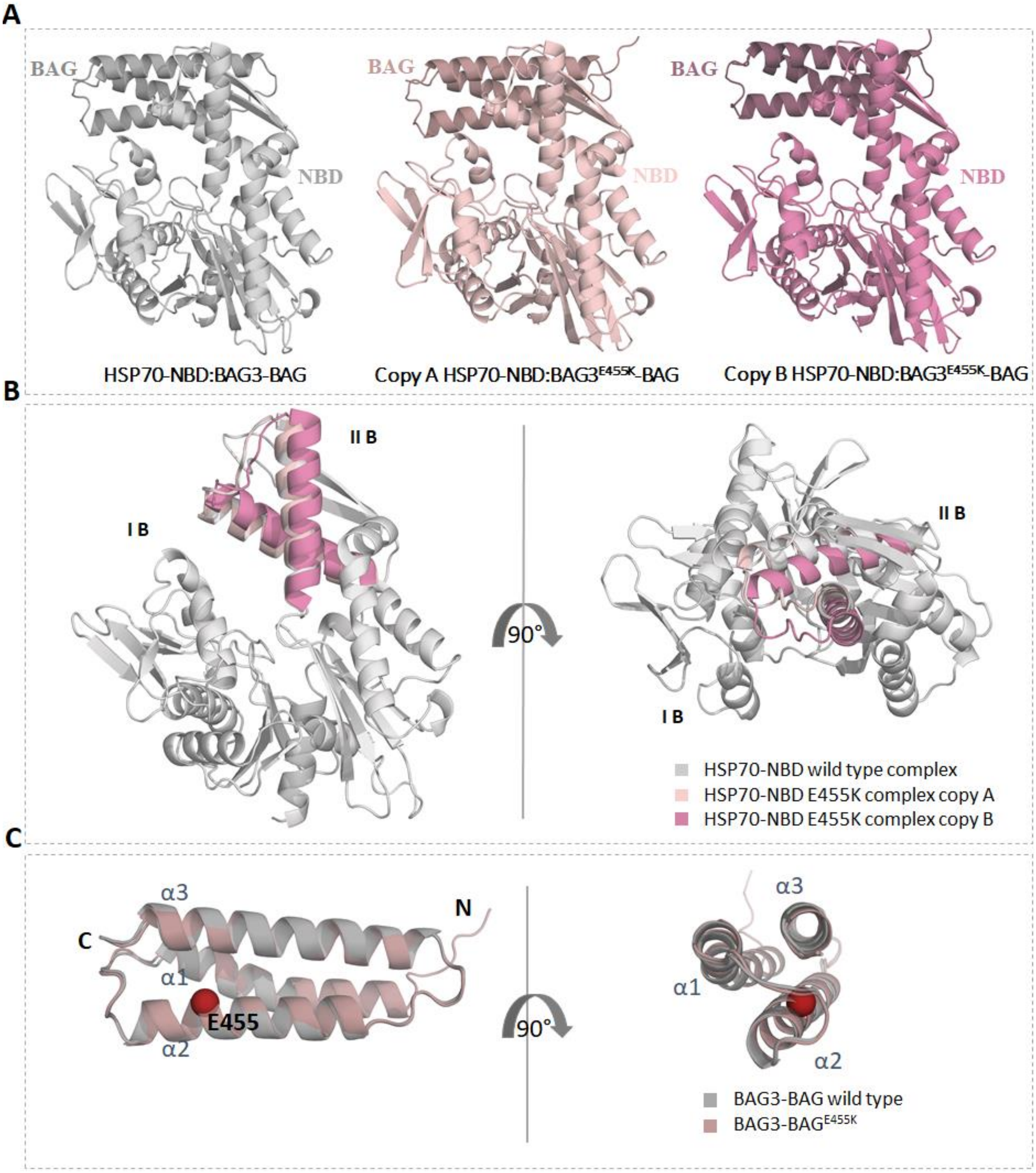


Figure 3.5 Crystal structures of the NBD-BAG complexes.
A The asymmetric unit of the wild type complex contains one copy, while the asymmetric unit of the E455K mutant contains two (Copy A and Copy B). **B** Copy A of the mutant structure has a highly similar structure to the wild type (1.4 Å, for all atoms). The main conformational change in the copy B is seen in the IIB subdomain, shown in salmon for Copy A and pink for Copy B. **C** The E455K mutation in the BAG domain of BAG3 (C alpha shown in red ball) does not change the overall structure of BAG3.

The structure shows four NBD subdomains (IA, IIA IB, IIB) arranged around the nucleotide binding site and BAG domain in proximity of IB and IIB. Relative to the IB and IIB distance and rotation of the IIB subdomain, NBD was originally reported in open (PDB ID: 1YUW; (Jiang *et al.*, 2005) and closed (PDB ID: 2KHO; (Bertelsen *et al.*, 2009)) conformation. For this reason, in the conformational analysis of the two copies in the asymmetric unit of the NBD-BAG^{E455K}, subdomains IA, IB and IIA were superimposed and the position of IIB was studied (Figure 3.5 B). The position of IIB subdomain in the wild type and Copy A is displaced compared to the Copy B of the mutant. The subdomain arrangement found in the mutant structure was compared with the existing structures of NBD. In fact, both conformational states the HSP70-NBD in the complex with BAG3^{E455K} have been reported in previous studies. Copy B of the mutant structure shows subdomain arrangement seen in the bovine NBD-SBD structure (PDB ID: 1YUW;(Jiang *et al.*, 2005)), referred to as an open state of the nucleotide-binding domain (Baumann, Liu and Thompson, 2016). The conformational state of the IIB subdomain in wild type and Copy A of the mutant structures resembles one of the NBD in complex with BAG5 (PDB ID: 3A8Y; (Arakawa *et al.*, 2010) referred to as a novel conformation, distinct from known “open” and “close”. In the meantime, many other structures of HSP70-NBD with varying conformations of a relative position of the IIB subdomain were described, demonstrating that NBD has a very dynamic nature, which is not strictly related to the presence of nucleotides or NEFs. Here elucidated structures of the NBD-BAG^{WT} and the NBD-BAG^{E455K} show that the BAG3 E455K mutation does not affect the overall structure of the BAG domain, nor cause conformational change in the NBD-BAG complex in the crystal structure.

3.3.4 Mutating E455 in the BAG domain shifts the interaction pattern with HSP70

Crystal structures of NBD-BAG complex show that the main interaction interface between the two chaperones involves helices 2 and 3 of BAG3 and IIB subdomain of HSP70-NBD, while the interaction with IB region is rather transient (Figure 3.2). As the overall fold of the complex is preserved in the mutant structure, to investigate the effect of E455K mutation, interaction patterns established between BAG3 and HSP70 were analyzed. Considering that the wild type and copy A of the mutant complex are in the same conformation, interaction interface analysis was compared for those two structures.

In both structures, the wild type and the mutant, several key interactions were identified as a core of the complex interface (Figure 3.6). In the NBD the main interacting residues are: positively charged Arg258, Arg261, Arg262; negatively charged Glu283 and Asp285; and hydrophilic Ser286. In the BAG domain of BAG3 two residues in helix2 and four residues in helix3 were found to mediate interaction with HSP70. Residues Glu456 (interacting with NBD-Arg261) and Glu466 (interacting with NBD-Arg262) form helix2 and Gln488 (interacting with NBD-Ser286), Arg477 (interacting with NBD-Glu283), Arg480 (interacting with NBD-Asp281) and Glu495 (interacting with NBD-Arg258) from the helix3 are observed in both wild type and mutant (Figure 3.6). The core interactions found are located in the central region of BAG3 (Figure 3.6 B).

The residue Glu455 is highly conserved in BAG3, BAG4 and BAG5 BD5 (Arakawa *et al.*, 2010). In the crystal structure of BAG3-BAG and HSP70-NBD, Glu455 is found to interact with Arg258 from IIB region of HSP70 (Figure 3.6 A). In the wild type structure, Arg258 establishes hydrogen bonds with two more residues form BAG3 Glu495 (helix3) and Thr459 (helix2). In the mutant structure, however, the hydrogen bond between Arg258 and Glu455 is lost, together with the hydrogen bond established with helix2, Thr459. The loss of HSP70-NBD Arg258 interaction with the helix2, mutant compensates with hydrogen bond with the Glu495 from helix3 of BAG3-BAG (Figure 3.6 C).

Another significant alteration in the interaction pattern due to the mutation is seen in the interaction between BAG3 and IB region of HSP70 (Figure 3.6 C). Moreover, in the wild type structure, weak hydrogen bond networks involve BAG3 helix2 (Lys449, Lys460 residues) and IB subdomain of NBD (Thr47, Glu48 and Asn57) (Figure 3.6 A). In the mutant structure, those interactions are no longer observed, suggesting that transient interaction with IB region does not occur anymore. In the distal region of IB, results revealed a new interaction cluster around C-terminal region of BAG3 (Arg481 of the helix3) involved in interaction with the distal loop in IIB (Tyr294 and Arp292) of HSP70, that seems to shift interaction pattern in the mutant (Figure 3.6 C).

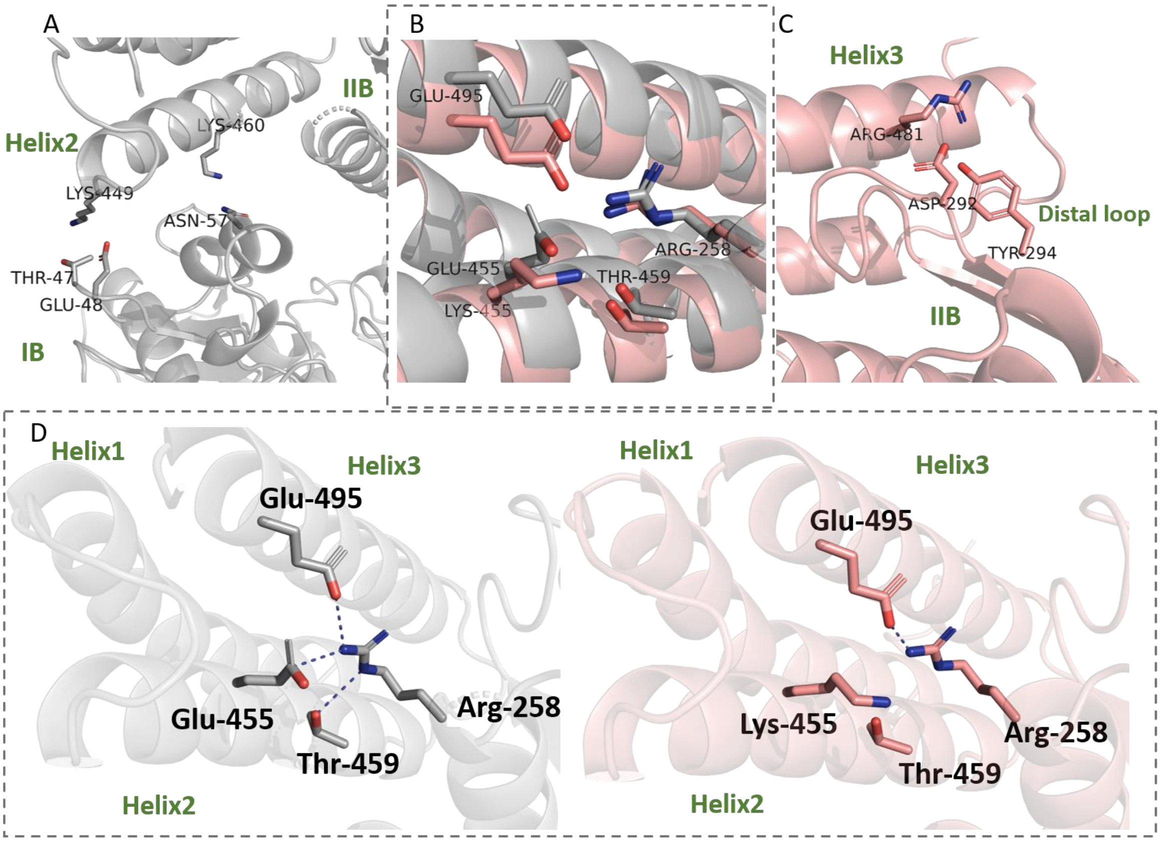


Figure 3.6 BAG3-HSP70 interaction pattern shift. Three main interacting interfaces between BAG and NBD. **A** Helix2 in BAG3 wild type maintains contact with IB domain of NBD, while this interaction is lost in the mutant; **B** Central residues cluster around NBD ARG258 and this interaction are shifted towards the helix3 in mutant structure; **D** Close up of the central residues interaction pattern; **C** in the C terminal domain of BAG3 new interactions establish with the distal loop of NBD. Main contributors for BAG/NBD interaction are shown in grey for the wild type and salmon for the mutant. Blue dashed line represents hydrogen bond;

The hydrogen bond network of the BAG3 and HSP70 involves the interaction of helixes 2 and 3 with the IIB and IB region of HSP70. It seems that Glu455 mediates helix2 stabilization required for the interaction with IB region (Figure 3.6 D). Moreover, helix2 residues Glu455 and Thr459 interaction with NBD Arg258, together with Glu456:Arg261 establishes the core interaction of the central BAG3 region. Based on the crystal structure it can be concluded that the Glu455 mutation to Lys alters the binding interface of BAG3. However, in order to better understand interactions described and their role in modulating mechanism and molecular dynamics of HSP70, molecular dynamic simulations were conducted.

3.3.5 Glu455 has a role in allosteric regulation of BAG3 interaction with IB domain in HSP70 nucleotide binding domain.

To investigate if the mutation in the BAG3-BAG domain interferes with its ability to regulate HSP70-NBD conformational dynamics, the molecular dynamics simulations (MDS) were designed and performed by Dr. Jennifer Fleming and Pier-Maria Bauer. GROMACS (GROningen MAchine for Chemical Simulations) package 2018.3 (Van Der Spoel *et al.*, 2005) was used for carrying out simulations with the AMBER99SB-ILDN force field with periodic boundary conditions. The two copies of the complex in the asymmetric system of HSP70-NBD-BAG^{E445K} were used to generate six systems for investigation, consisting of: HSP70-NBD, HSP70-NBD-BAG^{wt}, Copy A HSP70-NBD-BAG^{E445K} and Copy B HSP70-NBD-BAG^{E445K}. Due to the greater suitability of mutant complex as a model for molecular dynamics (resolution 2.2 A, R/R$_{free}$ of 0.18/0.23), the crystal structure of HSP70-NBD-BAG^{E445K} was used to create BAG^{wt} by *in silico* mutation of the position back to the original residue. The HSP70-NBD alone was generated by deleting the complexed BAG domain.

As a result, not unexpectedly the E455K mutation did not alter BAG3-BAG domain conformational dynamics. BAG3-BAG^{E445K} and BAG3-BAG3^{wt} both remained within a rmsd of 17.6 ± 0.5 Å each other over the course of all simulations.

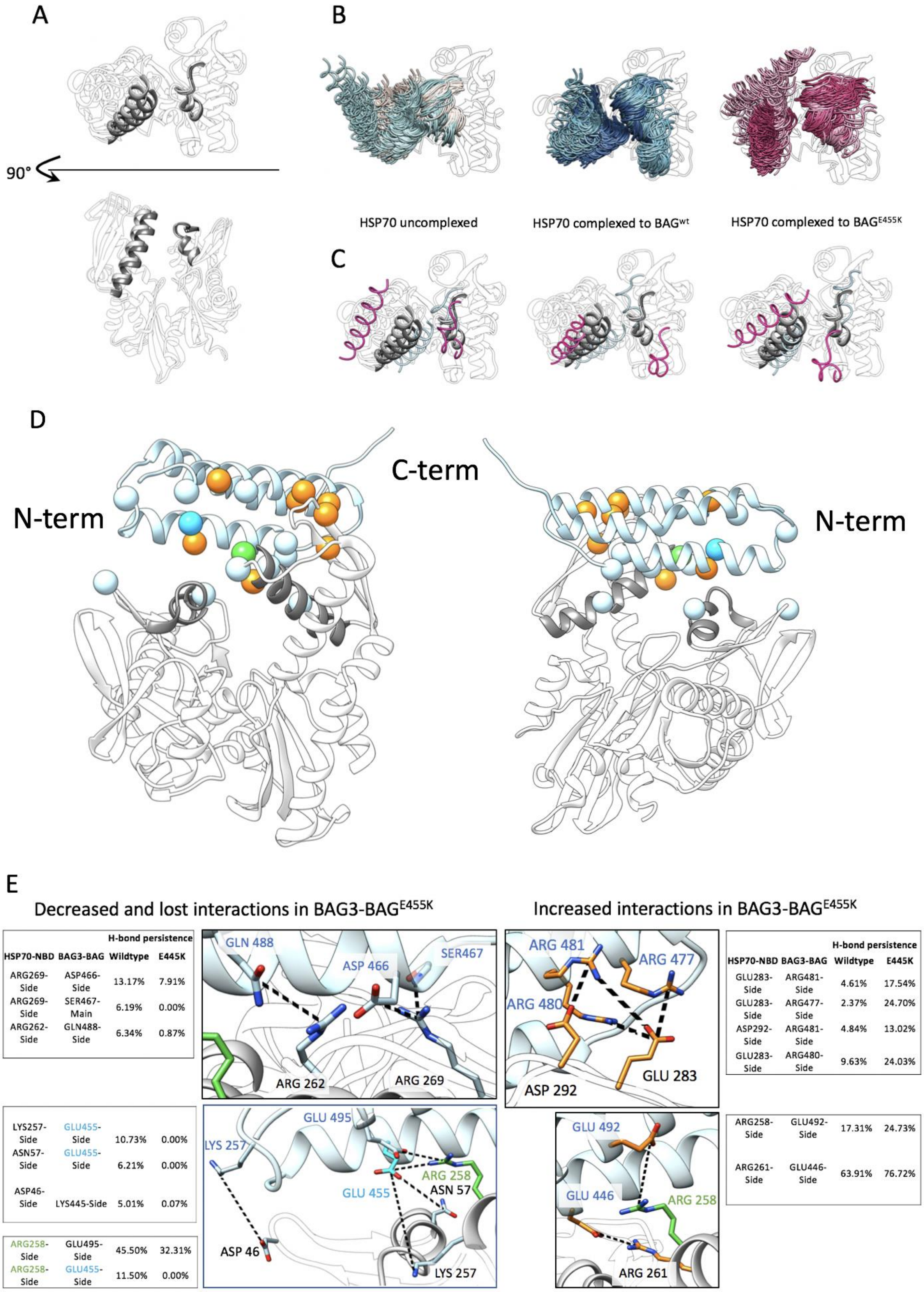
Root mean square fluctuation (RMSF) analysis of all system trajectories HSP70-NBD reveals that it is subdomain IIB (residues 245-295) experiences the greatest fluctuations during all simulations.

To identify differences in the three systems the last 100ns of the trajectories were analyzed, as this is when the simulations had reached their metastable states. This revealed that HSP70-NBD alone moves freely, sampling many conformations, in particular, IIB samples very wide conformations whilst the movement of IB is more conservative. In the BAG3-BAG bound forms the HSP70-NBD IIB domain movement is rather restricted. The BAG^{wt} bound HSP70-NBD IIB domain can still be seen to pinch closely towards the IB domain, but this movement is not seen in the BAG^{E445K} simulations at all. In the BAG^{E445K} bound HSP70-NBD a gap remains between the domains that is not seen in the free form of HSP70-NBD and is much smaller in the BAG3-BAG^{WT} bound chaperone (Figure 3.7 A B C). The results suggest that BAG-BAG3 binding alters HSP70-NBD conformational dynamics and in the complex the HSP70-NBD IIB domain movement is restrained but it also appears responsible for causing larger moments of IB (Figure 3.7 B).

Analysis of the persistent residue pair H-bonds between the BAG3-BAG domain and HSP70-NBD of the HSP70-NBD-BAG^{wt} and HSP70-NBD-BAG^{E445K} systems (again from the last

100ns of the trajectories) could explain the conformational difference seen in HSP70-NBD. Although during all of the simulations a similar number of H-bonds are observed (119 for HSP70-NBD-BAG^{wt} and 118 for HSP70-NBD-BAG^{E445K}), many occur for a very low percentage of the time. Therefore, only H-bond interactions that persisted more than 5% of the time within either of the systems were selected for analysis and the H-bonds that showed a greater than 10 percentage point change were examined (Figure 7.6 D). This showed that the loss of E445 H-bonds shift the H-bond network of HSP70-NBD-BAG so that the C-terminus of BAG3-BAG favors interactions with the distal loop in IIB (residues 285-295) and interactions with the N-terminal BAG and IB, although rare before, never occur in the mutant. The mutation of Glu445 to Lys eliminates several H-bonds in the BAG-NBD interacting interface, between this residue and K257, N57 and R258 (Figure 3.7 D, green residue). The loss of the Glu445:Arg258 interaction has several knock-on effects, allowing R258 to increase the number of interactions with E495 from helix3. Furthermore, a cluster of interactions with the C-terminal end of BAG Glu283:Arg481, become more persistent, stabilizing the complex association mediated by helix3 of BAG3 and IIB subdomain of HSP70-NBD.

Another interesting interaction which is no longer possible in the HSP70-NBD-BAG^{E445K} systems is the H-bonding of Glu455 to Lys257 (IIB). During the wild type simulations Lys257 can be seen to move in between two orientations, one facing up towards the BAG domain H-bonding to Glu455 and a second orientation facing towards the center of HSP70. When Lys257 connects with Glu455 it encourages the formation of the ‘tight’ conformation where the IB and IIB subdomains come together. In contrast, the loss of the BAG Glu455 and IIB Lys257 H-bond results in this lysine moving freely and creating a wedge where K257 turns towards the NBD center pushing the subdomains apart. This results in the HSP70-NBD bound to BAG^{E445K} adopting the ‘loose’ conformation seen in Figure 3. 7.



		H-bond persistence	
HSP70-NBD	BAG3-BAG	Wildtype	E445K
ARG269-Side	ASP466-Side	13.17%	7.91%
ARG269-Side	SER467-Main	6.19%	0.00%
ARG262-Side	GLN488-Side	6.34%	0.87%

LYS257-Side	GLU455-Side	10.73%	0.00%
ASN57-Side	GLU455-Side	6.21%	0.00%
ASP46-Side	LYS445-Side	5.01%	0.07%

ARG258-Side	GLU495-Side	45.50%	32.31%
ARG258-Side	GLU455-Side	11.50%	0.00%

		H-bond persistence	
HSP70-NBD	BAG3-BAG	Wildtype	E445K
GLU283-Side	ARG481-Side	4.61%	17.54%
GLU283-Side	ARG477-Side	2.37%	24.70%
ASP292-Side	ARG481-Side	4.84%	13.02%
GLU283-Side	ARG480-Side	9.63%	24.03%

ARG258-Side	GLU492-Side	17.31%	24.73%
ARG261-Side	GLU446-Side	63.91%	76.72%

Figure 3.7 Conformational freedom of HSP70-NBD when free or bound to wild type or E445K mutated BAG

A HSP70 with helix of IIB (residues 257-276) and helix-loop of IB (residues 53-64) shown in grey. Dark grey for chain A and light grey for chain B. **B** These selections are then shown as ribbon of the last 100 ns of each simulation with a time step of 4ns, highlighting the relative movements of the I and IIB subdomains. **C** Extreme projections of the major movement of each system. Light blue, most negative eigenvalue position; pink, most positive eigenvalue position.

Major changes in H-bond persistence between HSP70-NBD and BAG3-BAG wild type and mutant.

D Cartoon representation of Chain A with the locations of residues with altered interactions in the HSP70-NBD-BAG3-BAG^{E455K} shown as spheres: cyan, GLU455; pale blue, decreased or lost interactions; orange, increased interactions; orange; changed interaction partners; green, changed interaction partners

E The last 100ns of each trajectory was used only H-bonds with >5% persistence in wild type or mutant trajectories considered then only those which were lost or have >10% change shown. Chain A of crystal structure altered to wild type shown as cartoon: light blue, BAG3-BAG; translucent, HSP70-NBD helix of IIB (residues 257-276) and helix-loop of IB (residues 53-64) shown in grey. Residues with altered interactions shown as sticks colored by heteroatom: O, red; N, blue; C coloured as panel **D**. Interactions shown as dashed black lines- note the ground crystal state is shown so many are not H-bonded in the depicted model. (Figure provided by Dr. Jennifer Fleming).

3.4 Discussion

The E455K mutation in the BAG domain of BAG3 was found in a number of patients suffering from a familiar inheritable dilated cardiomyopathy (Villard et al. 2011). The mutation alteration of E455 residue in BAG3 was further related to a mild decrease in binding affinity with HSP70 nucleotide-binding domain in its APO form measured *in vitro* (Rauch, Zuiderweg and Gestwicki, 2016; Fang *et al.*, 2017). The similar effect has been observed in the case of R481A mutation located in helix3 of the BAG domain in BAG1 (Rauch, Zuiderweg and Gestwicki, 2016). It however remained obscure how a point mutation which showed moderate effect on the binding affinities could result in a loss of function of HSP70 machinery *in vivo*.

In order to reveal the underlying molecular mechanism by which co-chaperone BAG3 modulates nucleotide exchange in HSP70 and how E455K mutation impairs its function, the stop-flow experiments were used to measure nucleotide exchange rate of NBD of HSP70. The results showed an acute effect of E455K mutation on the nucleotide exchange activity of BAG3 by decreasing its efficiency up to 60 times. It was further investigated how the presence of nucleotides in solution affects HSP70-BAG3^{E455K} binding affinity. By using BLI method for studying protein-protein interactions BAG-NBD binding was monitored in APO form and in presence of ADP and ATP. It was revealed that the stalling effect of E455K mutation might be caused by a dramatically decreased binding affinity of BAG3 for HSP70 in presence of ATP (~10 fold). Binding affinity in the absence of nucleotides, even though decreased, remained in the range of detectable signal. This phenomena of moderate binding in the APO form allowed for complex formation and crystallization of the NBD-BAG^{WT} and NBD-BAG^{E455K} in absence of nucleotides. Thus, our structural knowledge of the mutant structure is limited to the APO form. Nevertheless, the APO structures of NBD are widely used as an important tool towards understanding the conformational dynamics of the HSP70 nucleotide binding domain (Mayer, 2013). Combining structural analysis and molecular dynamics simulations the role of E455 in the interaction pattern of NBD-BAG was investigated.

Obtained results proposed that BAG3 might act as a nucleotide exchange factor not by inducing a conformational change in HSP70 NBD, but through locking NBD dynamics in a narrow range of existing conformations that is favorable for nucleotide exchange to occur (Figure 3.8). It could be hypothesized that BAG3 locks a dynamic range of NBD via the tightly regulated hydrogen bond network established in the interacting region of BAG3 with

IIB and IB subdomains of NBD and the E455K mutation acts to shift the HSP70 conformational equilibrium away from the conformational space required to bind nucleotides efficiently, changing the BAG's efficacy as a NEF.

ATP hydrolysis in the NBD cycle is dependent on the assistance of co-chaperone protein J. In the ATP-bound form NBD exhibits conformational changes, mostly observed as a change of distance and rotation angle between subdomains IB and IIB. However, dynamics of conformational change is proposed to increase in the ADP-bound and APO states (Figure 3.7 A) (Woo *et al.*, 2009). For an effective ADP release and nucleotide exchange to occur, it is suggested that NBD requires a narrow range of conformational changes. The role of nucleotide exchange factors might be in arresting titling of NBD subdomains in a favorable range and E455K mutation perturbs this action. Structural data indicates that BAG3 interacts with HSP70 NBD through IIB subdomain mainly, and IB transiently. BAG transient interaction with NBD IB subdomain seems to be critical for the effective nucleotide exchange and it is enabled only if the BAG helix2 is in proximity to NBD IB. BAG3 residue Glu455 interacts with the Arg258 of IIB region in HSP70 and seems to mediate positioning of a BAG3 helix2 for a contact with IB subdomain during the molecular dynamic simulation. In the BAG^{E455K} the loss of key hydrogen bonds of helix2 with NBD K257, N57 and R258 results in a complete loss of interacting interface between IB region in NBD and helix2 in BAG3.

Newly established interactions, observed only in the mutant structure, are located in the distal loop of IIB subdomain clustered around Glu283:Arg481 may act to stabilize binding of C-terminal BAG^{E455K}. It might be foreseen that the E455 acts as a silent switch of the BAG3 binding with IIB and IB subdomains of NBD. The mutation E455K overbalanced the interacting interface towards the IIB, preserving the binding, but arresting NBD in the "loose" conformations, unfavorable for nucleotide exchange (Figure 3.8 B), further causing the stalling effect of the HSP70 cycle. This mechanistic model could explain detected binding in the APO form of NBD and BAG^{E455K} and lack of nucleotide exchange activity.

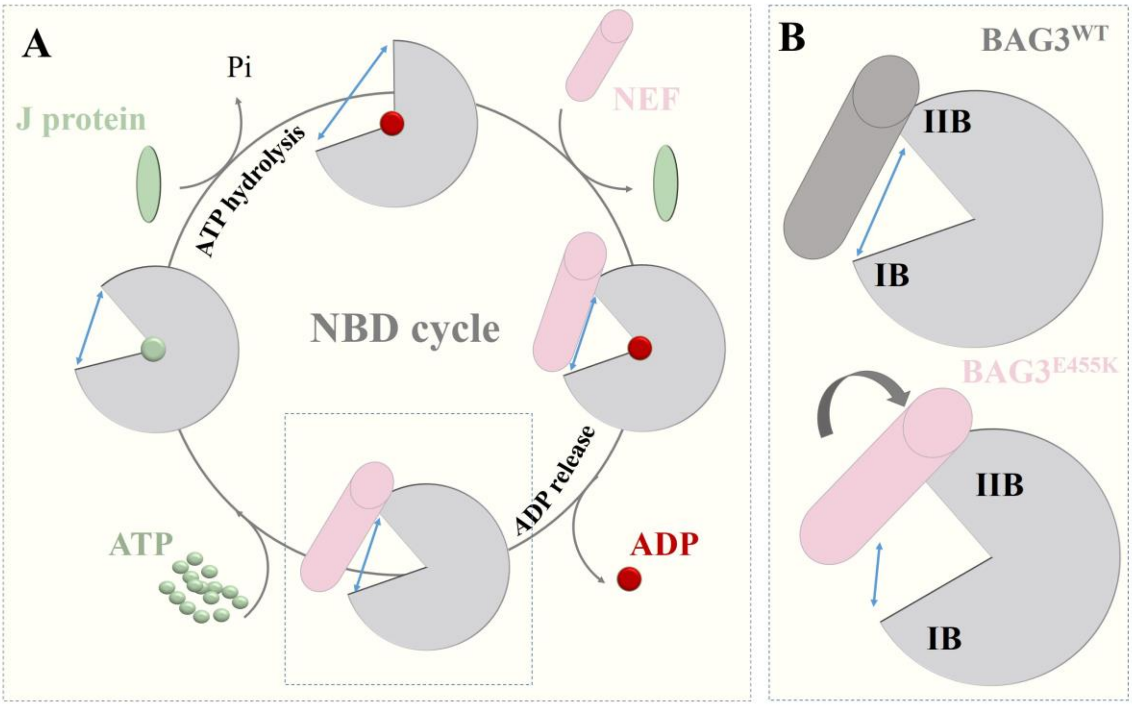


Figure 3.8 Proposed mechanism of E455K mutation in BAG3 underlying malfunction of NBD cycle. **A** Nucleotide binding domain goes through conformational change in a dynamic cycle regulated by ATP hydrolysis. ATPase activity of NBD is at a low basal level and requires protein J assistance. Conformational range seen in ADP-bound state and APO state is broader than the one seen in the ATP-bound state (Woo *et al.*, 2009). Release of the ADP is a key step for NBD to re-enter a cycle and is mediated by nucleotide exchange factors (NEF). **B** The molecular co-chaperone BAG3 binds to the IIB subdomain of NBD and by transient interactions with IB subdomain it could restrict conformational freedom of NBD. Here it is proposed that the E455 residue acts as a switch, which balances interactions with distal IIB loop and IB domain. The mutant BAG3 interaction pattern is overbalanced towards the IIB distal loop, creating a “loose” conformational space, unfavourable for nucleotide exchange. Consequently, BAG3 remains bound to HSP70 preventing it from re-entering a new cycle by keeping it arrested. Blue arrows indicate differences in conformational dynamics in NBD.

Nucleotide release as the rate-limiting step, is a critical point for fine tuning regulation of molecular chaperone cycle (Palleros *et al.*, 1993; Schmid *et al.*, 1994). Here obtained results present a step towards understanding the main triggers in nucleotide exchange regulation excreted by BAG3 co-chaperone. The BAG family of proteins is a prospective target for drug development in the field of cardiac myopathies and cancer, therefore structural information and the role of E455 residue in the interaction interface with HSP70 could guide the future work in this field.

References

Anandapadamanaban, M. *et al.* (2019) ‘E3 ubiquitin-protein ligase TRIM21-mediated lysine capture by UBE2E1 reveals substrate-targeting mode of a ubiquitin-conjugating E2.’, *The Journal of biological chemistry*, 294(30), pp. 11404–11419. doi: 10.1074/jbc.RA119.008485.

Arakawa, A. *et al.* (2010) ‘The C-Terminal BAG Domain of BAG5 Induces Conformational Changes of the Hsp70 Nucleotide- Binding Domain for ADP-ATP Exchange’, *Structure*, 18(3), pp. 309–319. doi: 10.1016/j.str.2010.01.004.

Bagchi, I. C., Kemp, B. E. and Means, A. R. (1989) ‘Myosin light chain kinase structure function analysis using bacterial expression.’, *The Journal of biological chemistry*, 264(27), pp. 15843–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2674119.

Barral, J. M. *et al.* (2002) ‘Role of the myosin assembly protein UNC-45 as a molecular chaperone for myosin.’, *Science (New York, N.Y.)*. United States, 295(5555), pp. 669–671. doi: 10.1126/science.1066648.

Baumann, C. W., Liu, H. M. and Thompson, L. D. V. (2016) ‘Denervation-induced activation of the ubiquitin-proteasome system reduces skeletal muscle quantity not quality’, *PLoS ONE*, 11(8), pp. 1–18. doi: 10.1371/journal.pone.0160839.

Behl, C. (2016) ‘Breaking BAG: The Co-Chaperone BAG3 in Health and Disease’, *Trends in Pharmacological Sciences*. Elsevier Ltd, 37(8), pp. 672–688. doi: 10.1016/j.tips.2016.04.007.

Berndsen, C. E. and Wolberger, C. (2014) ‘New insights into ubiquitin E3 ligase mechanism.’, *Nature structural & molecular biology*. United States, 21(4), pp. 301–307. doi: 10.1038/nsmb.2780.

Bertelsen, E. B. *et al.* (2009) ‘Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate.’, *Proceedings of the National Academy of Sciences of the United States of America*, 106(21), pp. 8471–8476. doi: 10.1073/pnas.0903503106.

Bodine, S. C. *et al.* (2001) ‘Identification of ubiquitin ligases required for skeletal muscle atrophy.’, *Science (New York, N.Y.)*. United States, 294(5547), pp. 1704–1708. doi: 10.1126/science.1065874.

Bodine, S. C. *et al.* (2007) ‘Skeletal Muscle Atrophy Identification of Ubiquitin Ligases Required for Skeletal Muscle Atrophy’, *Science*, 1704(2001), pp. 1704–1708. doi: 10.1126/science.1065874.

Bodine, S. C. and Baehr, L. M. (2014) ‘Skeletal muscle atrophy and the E3 ubiquitin ligases MuRF1 and MAFbx/atrogin-1’, *AJP: Endocrinology and Metabolism*, 307(6), pp. E469–E484. doi: 10.1152/ajpendo.00204.2014.

Bogomolovas, J. *et al.* (2014) ‘Titin kinase is an inactive pseudokinase scaffold that supports MuRF1 recruitment to the sarcomeric M-line’, *Open Biology*, 4(5), pp. 140041–140041. doi: 10.1098/rsob.140041.

Braten, O. *et al.* (2016) ‘Numerous proteins with unique characteristics are degraded by the 26S proteasome following monoubiquitination.’, *Proceedings of the National Academy of Sciences of the United States of America*, 113(32), pp. E4639-47. doi: 10.1073/pnas.1608644113.

Bukau, B., Weissman, J. and Horwich, A. (2006) ‘Molecular Chaperones and Protein Quality Control’, *Cell*. Elsevier Inc., 125(3), pp. 443–451. doi: 10.1016/j.cell.2006.04.014.

Campos, Y. *et al.* (2010) ‘Ozz-E3 ubiquitin ligase targets sarcomeric embryonic myosin heavy chain during muscle development’, *PLoS ONE*, 5(3). doi: 10.1371/journal.pone.0009866.

Carra, S., Seguin, S. J. and Landry, J. (2008) ‘HspB8 and Bag3: a new chaperone complex targeting misfolded proteins to macroautophagy.’, *Autophagy*. United States, 4(2), pp. 237–239. doi: 10.4161/auto.5407.

Castillero, E. *et al.* (2013) ‘Suppression of atrogin-1 and MuRF1 prevents dexamethasone-induced atrophy of cultured myotubes’, *Metabolism: Clinical and Experimental*. Elsevier Inc., 62(10), pp. 1495–1502. doi: 10.1016/j.metabol.2013.05.018.

Centner, T. *et al.* (2001) ‘Identification of muscle specific ring finger proteins as potential regulators of the titin kinase domain', *Journal of Molecular Biology*, 306(4), pp. 717–726. doi: 10.1006/jmbi.2001.4448.

Chen, T. *et al.* (2014) 'mUbiSiDa: A comprehensive database for protein ubiquitination sites in mammals', *PLoS ONE*, 9(1), pp. 1–6. doi: 10.1371/journal.pone.0085744.

Chen, Z. S. *et al.* (2019) 'FipoQ/FBXO33, a Cullin-1-based ubiquitin ligase complex component modulates ubiquitination and solubility of polyglutamine disease protein.', *Journal of neurochemistry*. England, 149(6), pp. 781–798. doi: 10.1111/jnc.14669.

Ciechanover, A. (1998) 'The ubiquitin-proteasome pathway: on protein death and cell life.', *The EMBO journal*, 17(24), pp. 7151–7160. doi: 10.1093/emboj/17.24.7151.

Clarke, B. A. *et al.* (2007) 'The E3 Ligase MuRF1 Degrades Myosin Heavy Chain Protein in Dexamethasone-Treated Skeletal Muscle', *Cell Metabolism*, 6(5), pp. 376–385. doi: 10.1016/j.cmet.2007.09.009.

Cohen, S. *et al.* (2009) 'During muscle atrophy, thick, but not thin, filament components are degraded by MuRF1-dependent ubiquitylation', *Journal of Cell Biology*, 185(6), pp. 1083–1095. doi: 10.1083/jcb.200901052.

Colombini, B. *et al.* (2010) 'Is the cross-bridge stiffness proportional to tension during muscle fiber activation?', *Biophysical journal*, 98(11), pp. 2582–2590. doi: 10.1016/j.bpj.2010.02.014.

Concepcion, J. *et al.* (2009) 'Label-free detection of biomolecular interactions using BioLayer interferometry for kinetic characterization.', *Combinatorial chemistry & high throughput screening*. United Arab Emirates, 12(8), pp. 791–800. doi: 10.2174/138620709789104915.

Cooke, R. and Holmes, K. C. (1986) *The mechanism of muscle contractio, Critical Reviews in Biochemistry and Molecular Biology*. doi: 10.3109/10409238609113609.

Craig, R. and Megerman, J. (1977) 'Assembly of smooth muscle myosin into side-polar filaments.', *The Journal of cell biology*, 75(3), pp. 990–996. doi: 10.1083/jcb.75.3.990.

Craig, R. and Woodhead, J. L. (2006) 'Structure and function of myosin filaments.', *Current opinion in structural biology*. England, 16(2), pp. 204–212. doi: 10.1016/j.sbi.2006.03.006.

CRICK, F. H. C. (1952) 'Is α-Keratin a Coiled Coil?', *Nature*, 170(4334), pp. 882–883. doi: 10.1038/170882b0.

Delorenzi, M. and Speed, T. (2002) 'An HMM model for coiled-coil domains and a comparison with PSSM-based predictions.', *Bioinformatics (Oxford, England)*. England, 18(4), pp. 617–625. doi: 10.1093/bioinformatics/18.4.617.

Deshaies, R. J. and Joazeiro, C. A. P. (2009) 'RING domain E3 ubiquitin ligases.', *Annual review of biochemistry*. United States, 78, pp. 399–434. doi: 10.1146/annurev.biochem.78.101807.093809.

Dikic, I. (2017) 'Proteasomal and Autophagic Degradation Systems.', *Annual review of biochemistry*. United States, 86, pp. 193–224. doi: 10.1146/annurev-biochem-061516-044908.

Ding, S. *et al.* (2018) 'An Overview of Muscle Atrophy.', *Advances in experimental medicine and biology*. United States, 1088, pp. 3–19. doi: 10.1007/978-981-13-1435-3_1.

Dobson, C. M. (2003) 'Protein folding and misfolding', *Nature*, 426(6968), pp. 884–890. doi: 10.1038/nature02261.

Domínguez, F. *et al.* (2018) 'Dilated Cardiomyopathy Due to BLC2-Associated Athanogene 3 (BAG3) Mutations.', *Journal of the American College of Cardiology*, 72(20), pp. 2471–2481. doi: 10.1016/j.jacc.2018.08.2181.

Dou, H. *et al.* (2012) 'BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer.', *Nature structural & molecular biology*, 19(9), pp. 876–883. doi: 10.1038/nsmb.2379.

Eddins, M. J. *et al.* (2011) 'Targeting the Ubiquitin E3 Ligase MuRF1 to Inhibit Muscle Atrophy', *Cell Biochemistry and Biophysics*, 60(1–2), pp. 113–118. doi: 10.1007/s12013-011-9175-7.

Emsley, P. *et al.* (2010) 'Features and development of Coot.', *Acta crystallographica. Section D,* *Biological crystallography*, 66(Pt 4), pp. 486–501. doi: 10.1107/S0907444910007493.

Fang, X. *et al.* (2017) 'Loss-of-function mutations in co-chaperone BAG3 destabilize small HSPs and cause cardiomyopathy', *Journal of Clinical Investigation*, 127(8), pp. 3189–3200. doi: 10.1172/JCI94310.

Felzen, V. *et al.* (2015) 'Estrogen receptor α regulates non-canonical autophagy that provides stress resistance to neuroblastoma and breast cancer cells and involves BAG3 function.', *Cell death & disease*, 6(7), p. e1812. doi: 10.1038/cddis.2015.181.

Fernández-Fernández, M. R. and Valpuesta, J. M. (2018) 'Hsp70 chaperone: a master player in protein homeostasis', *F1000Research*. NLM (Medline). doi: 10.12688/f1000research.15528.1.

Fielitz, J. *et al.* (2007) 'Loss of muscle-specific RING-finger 3 predisposes the heart to cardiac rupture after myocardial infarction', *Proceedings of the National Academy of Sciences*, 104(11), pp. 4377–4382. doi: 10.1073/pnas.0611726104.

Files, D. C. *et al.* (2012) 'A critical role for muscle ring finger-1 in acute lung injury-associated skeletal muscle wasting', *American Journal of Respiratory and Critical Care Medicine*, 185(8), pp. 825–834. doi: 10.1164/rccm.201106-1150OC.

Fili, N. and Toseland, C. P. (2019) 'Unconventional Myosins: How Regulation Meets Function.', *International journal of molecular sciences*, 21(1). doi: 10.3390/ijms21010067.

Franceschelli, S. *et al.* (2008) 'bag3 gene expression is regulated by heat shock factor 1', *Journal of Cellular Physiology*, 215(3), pp. 575–577. doi: 10.1002/jcp.21397.

Franke, B. *et al.* (2014) 'Molecular basis for the fold organization and sarcomeric targeting of the muscle atrogin MuRF1.', *Open biology*, 4(3), p. 130172. doi: 10.1098/rsob.130172.

Freemont, P. S. (2000) 'RING for destruction?', *Current biology : CB*. England, 10(2), pp. R84-7. doi: 10.1016/s0960-9822(00)00287-6.

Freemont, P. S., Hanson, I. M. and Trowsdale, J. (1991) 'A novel cysteine-rich sequence motif.', *Cell*. United States, pp. 483–484. doi: 10.1016/0092-8674(91)90229-r.

Frontera, W. R. and Ochala, J. (2015) 'Skeletal muscle: a brief review of structure and function.', *Calcified tissue international*. United States, 96(3), pp. 183–195. doi: 10.1007/s00223-014-9915-y.

Frydman, J. (2001) 'Folding of newly translated proteins in vivo: the role of molecular chaperones.', *Annual review of biochemistry*. United States, 70, pp. 603–647. doi: 10.1146/annurev.biochem.70.1.603.

Fuchs, M. *et al.* (2009) 'Identification of the key structural motifs involved in HspB8/HspB6-Bag3 interaction.', *The Biochemical journal*. England, 425(1), pp. 245–255. doi: 10.1042/BJ20090907.

Ganassi, M. *et al.* (2016) 'A Surveillance Function of the HSPB8-BAG3-HSP70 Chaperone Complex Ensures Stress Granule Integrity and Dynamism', *Molecular Cell*. Elsevier Inc., 63(5), pp. 796–810. doi: 10.1016/j.molcel.2016.07.021.

Gentilella, A. *et al.* (2008) 'Activation of BAG3 by Egr-1 in response to FGF-2 in neuroblastoma cells', *Oncogene*, 27(37), pp. 5011–5018. doi: 10.1038/onc.2008.142.

Glotzer, M., Murray, A. W. and Kirschner, M. W. (1991) 'Cyclin is degraded by the ubiquitin pathway.', *Nature*. England, 349(6305), pp. 132–138. doi: 10.1038/349132a0.

Goldstone, D. C. *et al.* (2014) 'Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(26), pp. 9609–9614. doi: 10.1073/pnas.1402448111.

Gordon, A., Huxley, A. and Julian, F. (1966) 'The variation in isometric tension with sarcomere', *The Journal of Physiology*, 184(1), pp. 170–192. Available at: http://e.guigon.free.fr/rsc/article/GordonAMEtAl66a.pdf.

Granzier, H. L. and Labeit, S. (2006) 'The Giant Muscle Protein Titin is an Adjustable Molecular Spring', *Exercise and Sport Sciences Reviews*, 34(2). Available at: https://journals.lww.com/acsm-essr/Fulltext/2006/04000/The_Giant_Muscle_Protein_Titin_is_an_Adjustable.2.aspx.

Green, J. R., Korenberg, M. J. and Aboul-Magd, M. O. (2009) ‘PCI-SS: MISO dynamic nonlinear protein secondary structure prediction.’, *BMC bioinformatics*, 10, p. 222. doi: 10.1186/1471-2105-10-222.

Haas, K. F., Woodruff, E. 3rd and Broadie, K. (2007) ‘Proteasome function is required to maintain muscle cellular architecture.’, *Biology of the cell*, 99(11), pp. 615–626. doi: 10.1042/BC20070019.

Hartl, F. U. and Hayer-Hartl, M. (2002) ‘Protein folding. Molecular chaperones in the cytosol: From nascent chain to folded protein’, *Science*, 295(5561), pp. 1852–1858. doi: 10.1126/science.1068408.

Hartman, M. A. and Spudich, J. A. (2012) ‘The myosin superfamily at a glance.’, *Journal of cell science*, 125(Pt 7), pp. 1627–1632. doi: 10.1242/jcs.094300.

Hershko, A. and Ciechanover, A. (1998) ‘THE UBIQUITIN SYSTEM’, *Annual Review of Biochemistry*, 67(1), pp. 425–479. doi: 10.1146/annurev.biochem.67.1.425.

Hnia, K., Clausen, T. and Moog-Lutz, C. (2019) ‘Shaping Striated Muscles with Ubiquitin Proteasome System in Health and Disease’, *Trends in Molecular Medicine*. Elsevier Ltd, 25(9), pp. 760–774. doi: 10.1016/j.molmed.2019.05.008.

Hornbeck, P. V. *et al.* (2012) ‘PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse’, *Nucleic Acids Research*, 40(D1), pp. 261–270. doi: 10.1093/nar/gkr1122.

Hu, Z. *et al.* (2016) ‘Structure of myosin filaments from relaxed Lethocerus flight muscle by cryo-EM at 6 Å resolution’, *Science Advances*, 2(9). doi: 10.1126/sciadv.1600058.

Hutagalung, A. H. *et al.* (2002) ‘The UCS family of myosin chaperones.’, *Journal of cell science*. England, 115(Pt 21), pp. 3983–3990. doi: 10.1242/jcs.00107.

Jadhav, T. S., Wooten, M. W. and Wooten, M. C. (2011) ‘Mining the TRAF6/p62 interactome for a selective ubiquitination motif’, *BMC Proceedings*, 5(2), p. S4. doi: 10.1186/1753-6561-5-S2-S4.

Jiang, J. *et al.* (2005) ‘Structural basis of interdomain communication in the Hsc70 chaperone’, *Molecular cell*, 20(4), pp. 513–524. doi: 10.1016/j.molcel.2005.09.028.

Kabsch, W. (2010) ‘XDS’, *Acta Crystallographica Section D: Biological Crystallography*, 66(2), pp. 125–132. doi: 10.1107/S0907444909047337.

Kampinga, H. H. and Bergink, S. (2016) ‘Heat shock proteins as potential targets for protective strategies in neurodegeneration’, *The Lancet Neurology*. Elsevier Ltd, 15(7), pp. 748–759. doi: 10.1016/S1474-4422(16)00099-5.

Kedar, V. *et al.* (2004) ‘Muscle-specific RING finger 1 is a bona fide ubiquitin ligase that degrades cardiac troponin I.’, *Proceedings of the National Academy of Sciences of the United States of America*, 101(52), pp. 18135–18140. doi: 10.1073/pnas.0404341102.

Kim, J., Löwe, T. and Hoppe, T. (2008) ‘Protein quality control gets muscle into shape’, *Trends in Cell Biology*. doi: 10.1016/j.tcb.2008.03.007.

Kiss, L. *et al.* (2019) ‘A tri-ionic anchor mechanism drives Ube2N-specific recruitment and K63-chain ubiquitination in TRIM ligases.’, *Nature communications*, 10(1), p. 4502. doi: 10.1038/s41467-019-12388-y.

Koliopoulos, M. G. *et al.* (2016) ‘Functional role of TRIM E3 ligase oligomerization and regulation of catalytic activity.’, *The EMBO journal*, 35(11), pp. 1204–1218. doi: 10.15252/embj.201593741.

Korkmaz, E. N. *et al.* (2016) ‘A composite approach towards a complete model of the myosin rod’, *Proteins: Structure, Function and Bioinformatics*, 84(1), pp. 172–189. doi: 10.1002/prot.24964.

Koyama, S. *et al.* (2008) ‘Muscle RING-Finger Protein-1 (MuRF1) as a Connector of Muscle Energy Metabolism and Protein Synthesis’, *Journal of Molecular Biology*, 376(5), pp. 1224–1236. doi: 10.1016/j.jmb.2007.11.049.

Krissinel, E. and Henrick, K. (2007) ‘Inference of Macromolecular Assemblies from Crystalline State’, *Journal of Molecular Biology*, 372(3), pp. 774–797. doi: https://doi.org/10.1016/j.jmb.2007.05.022.

Kroemer, G., Mariño, G. and Levine, B. (2010) ‘Autophagy and the integrated stress response.’, *Molecular cell*, 40(2), pp. 280–293. doi: 10.1016/j.molcel.2010.09.023.

Labeit, S. *et al.* (2010) ‘Modulation of muscle atrophy, fatigue and MLC phosphorylation by MuRF1 as indicated by hindlimb suspension studies on MuRF1-KO mice.’, *Journal of biomedicine & biotechnology*, 2010, p. 693741. doi: 10.1155/2010/693741.

Lang, F. *et al.* (2017) ‘Dynamic changes in the mouse skeletal muscle proteome during denervation-induced atrophy’, *Disease Models & Mechanisms*, 10(7), pp. 881–896. doi: 10.1242/dmm.028910.

Lange, S. *et al.* (2005) ‘The kinase domain of titin controls muscle gene expression and protein turnover.’, *Science (New York, N.Y.)*. United States, 308(5728), pp. 1599–1603. doi: 10.1126/science.1110463.

Lecker, S. H. *et al.* (1999) ‘Muscle Protein Breakdown and the Critical Role of the Ubiquitin-Proteasome Pathway in Normal and Disease States’, *The Journal of Nutrition*, 129(1), pp. 227S-237S. doi: 10.1093/jn/129.1.227S.

Lecker, S. H., Goldberg, A. L. and Mitch, W. E. (2006) ‘Protein degradation by the ubiquitin-proteasome pathway in normal and disease states.’, *Journal of the American Society of Nephrology : JASN*. United States, 17(7), pp. 1807–1819. doi: 10.1681/ASN.2006010083.

Lee, L. A. *et al.* (2019) ‘The ancient sarcomeric myosins found in specialized muscles.’, *Skeletal muscle*, 9(1), p. 7. doi: 10.1186/s13395-019-0192-3.

Li, Y. *et al.* (2014) ‘Structural insights into the TRIM family of ubiquitin E3 ligases.’, *Cell research*, pp. 762–765. doi: 10.1038/cr.2014.46.

Liao, Q. *et al.* (2001) ‘The anti-apoptotic protein BAG-3 is overexpressed in pancreatic cancer and induced by heat stress in pancreatic cancer cell lines.’, *FEBS letters*. England, 503(2–3), pp. 151–157. doi: 10.1016/s0014-5793(01)02728-4.

Lupas, A., Van Dyke, M. and Stock, J. (1991) ‘Predicting coiled coils from protein sequences.’, *Science (New York, N.Y.)*. United States, 252(5009), pp. 1162–1164. doi: 10.1126/science.252.5009.1162.

Lupas, B. A. N. and Gruber, M. (2005) ‘THE STRUCTURE OF a-HELICAL COILED COILS - Helical coiled coils are versatile protein domains , supporting a I . Historical Introduction The first investigations into the structure of coiled coils were made by William Astbury at the University of Leeds in t’, *Advances in Protein Chemistry*, 70(04), pp. 37–78. doi: 10.1016/S0065-3233(04)70003-0.

Marblestone, J. G. *et al.* (2012) ‘Analysis of ubiquitin E3 ligase activity using selective polyubiquitin binding proteins’, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. Elsevier B.V., 1823(11), pp. 2094–2097. doi: 10.1016/j.bbamcr.2012.06.013.

Matthews, B. W. (1968) ‘Solvent content of protein crystals’, *Journal of Molecular Biology*, 33(2), pp. 491–497. doi: https://doi.org/10.1016/0022-2836(68)90205-2.

Mayans, O. and Labeit, S. (2012) ‘MuRFs specialized members of the TRIM/RBCC family with roles in the regulation of the trophic state of muscle and its metabolism’, *Advances in Experimental Medicine and Biology*, pp. 119–129. doi: 10.1007/978-1-4614-5398-7_9.

Mayer, M. P. (2013) ‘Hsp70 chaperone dynamics and molecular mechanism’, *Trends in Biochemical Sciences*. Elsevier, 38(10), pp. 507–514. doi: 10.1016/j.tibs.2013.08.001.

Mayer, M. P. and Bukau, B. (2005) ‘Hsp70 chaperones: Cellular functions and molecular mechanism’, *Cellular and Molecular Life Sciences*, 62(6), pp. 670–684. doi: 10.1007/s00018-004-4464-6.

McCoy, A. J. *et al.* (2007) ‘Phaser crystallographic software.’, *Journal of applied crystallography*, 40(Pt 4), pp. 658–674. doi: 10.1107/S0021889807021206.

McDowell, G. S. and Philpott, A. (2013) ‘Non-canonical ubiquitylation: Mechanisms and consequences’, *The International Journal of Biochemistry & Cell Biology*, 45(8), pp. 1833–1842.

doi: https://doi.org/10.1016/j.biocel.2013.05.026.

McElhinny, A. S. *et al.* (2002) ‘Muscle-specific RING finger-1 interacts with titin to regulate sarcomeric M-line and thick filament structure and may have nuclear functions via its interaction with glucocorticoid modulatory element binding protein-1’, *Journal of Cell Biology*, 157(1), pp. 125–136. doi: 10.1083/jcb.200108089.

McElhinny, A. S. *et al.* (2004) ‘Muscle-specific RING finger-2 (MURF-2) is important for microtubule, intermediate filament and sarcomeric M-line maintenance in striated muscle development.’, *Journal of cell science*. England, 117(Pt 15), pp. 3175–3188. doi: 10.1242/jcs.01158.

McLachlan, A. D. and Karn, J. (1982) ‘Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle’, *Nature*, 299(5880), pp. 226–231. doi: 10.1038/299226a0.

Meister-Broekema, M. *et al.* (2018) ‘Myopathy associated BAG3 mutations lead to protein aggregation by stalling Hsp70 networks’, *Nature Communications*. Springer US, 9(1), pp. 1–14. doi: 10.1038/s41467-018-07718-5.

Meroni, G. (2020) ‘TRIM E3 Ubiquitin Ligases in Rare Genetic Disorders.’, *Advances in experimental medicine and biology*. United States, 1233, pp. 311–325. doi: 10.1007/978-3-030-38266-7_14.

Moraes, G. L. *et al.* (2015) ‘Structural and functional features of enzymes of Mycobacterium tuberculosis peptidoglycan biosynthesis as targets for drug development.’, *Tuberculosis (Edinburgh, Scotland)*, 95(2), pp. 95–111. doi: 10.1016/j.tube.2015.01.006.

Mrosek, M. *et al.* (2007) ‘Molecular determinants for the recruitment of the ubiquitin-ligase MuRF-1 onto M-line titin.’, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*. United States, 21(7), pp. 1383–1392. doi: 10.1096/fj.06-7644com.

Mrosek, M. *et al.* (2008) ‘Structural analysis of B-Box 2 from MuRF1: identification of a novel self-association pattern in a RING-like fold.’, *Biochemistry*. United States, 47(40), pp. 10722–10730. doi: 10.1021/bi800733z.

Nguyen, T. *et al.* (2020) ‘Expression of MuRF1 or MuRF2 is essential for the induction of skeletal muscle atrophy and dysfunction in a murine pulmonary hypertension model’, *Skeletal Muscle*, 10(1), p. 12. doi: 10.1186/s13395-020-00229-2.

Nowak, M. *et al.* (2019) ‘DCAF8, a novel MuRF1 interaction partner, promotes muscle atrophy’, *Journal of Cell Science*. The Company of Biologists, 132(17), p. jcs233395. doi: 10.1242/jcs.233395.

O’Brien, M. C. and McKay, D. B. (1993) ‘Threonine 204 of the chaperone protein Hsc70 influences the structure of the active site, but is not essential for ATP hydrolysis’, *Journal of Biological Chemistry*, 268(32), pp. 24323–24329.

O’Shea, E. K. *et al.* (1991) ‘X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil.’, *Science (New York, N.Y.)*. United States, 254(5031), pp. 539–544. doi: 10.1126/science.1948029.

Oh, E., Akopian, D. and Rape, M. (2018) ‘Principles of Ubiquitin-Dependent Signaling.’, *Annual review of cell and developmental biology*. United States, 34, pp. 137–162. doi: 10.1146/annurev-cellbio-100617-062802.

Ojima, K. *et al.* (2015) ‘Dynamics of myosin replacement in skeletal muscle cells.’, *American journal of physiology. Cell physiology*. United States, 309(10), pp. C669-79. doi: 10.1152/ajpcell.00170.2015.

Palleros, D. R. *et al.* (1993) ‘K + but not ATP hydrolysis’, 365(OCTOBER), pp. 664–666.

Passmore, L. A. and Barford, D. (2004) ‘Getting into position: the catalytic mechanisms of protein ubiquitylation.’, *The Biochemical journal*, 379(Pt 3), pp. 513–525. doi: 10.1042/BJ20040198.

Peckham, M. (2016) ‘How myosin organization of the actin cytoskeleton contributes to the cancer phenotype', *Biochemical Society Transactions*, 44(4), pp. 1026–1034. doi: 10.1042/BST20160034.

Pellecchia, M. *et al.* (2000) 'Structural insights into substrate binding by the molecular chaperone DnaK.', *Nature structural biology*. United States, 7(4), pp. 298–303. doi: 10.1038/74062.

Perera, S., Mankoo, B. and Gautel, M. (2012) 'Developmental regulation of MURF E3 ubiquitin ligases in skeletal muscle.', *Journal of muscle research and cell motility*, 33(2), pp. 107–122. doi: 10.1007/s10974-012-9288-7.

Pfleger, C. M. and Kirschner, M. W. (2000) 'The KEN box: an APC recognition signal distinct from the D box targeted by Cdh1.', *Genes & development*, 14(6), pp. 655–665.

Plant, P. J. *et al.* (2009) 'Absence of caspase-3 protects against denervation-induced skeletal muscle atrophy', *Journal of Applied Physiology*, 107(1), pp. 224–234. doi: 10.1152/japplphysiol.90932.2008.

Plechanovová, A. *et al.* (2012) 'Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis.', *Nature*, 489(7414), pp. 115–120. doi: 10.1038/nature11376.

Pollard, T. D. and Korn, E. D. (1973) 'Acanthamoeba myosin. I. Isolation from Acanthamoeba castellanii of an enzyme similar to muscle myosin.', *The Journal of biological chemistry*. United States, 248(13), pp. 4682–4690.

Rauch, J. N. *et al.* (2017) 'BAG3 Is a Modular, Scaffolding Protein that physically Links Heat Shock Protein 70 (Hsp70) to the Small Heat Shock Proteins', *Journal of molecular biology*. 2016/11/21, 429(1), pp. 128–141. doi: 10.1016/j.jmb.2016.11.013.

Rauch, Jennifer N and Gestwicki, J. E. (2014) 'Binding of human nucleotide exchange factors to heat shock protein 70 (Hsp70) generates functionally distinct complexes in vitro', *The Journal of biological chemistry*. 2013/12/05. American Society for Biochemistry and Molecular Biology, 289(3), pp. 1402–1414. doi: 10.1074/jbc.M113.521997.

Rauch, Jennifer N. and Gestwicki, J. E. (2014) 'Binding of human nucleotide exchange factors to heat shock protein 70 (Hsp70) generates functionally distinct complexes in vitro', *Journal of Biological Chemistry*, 289(3), pp. 1402–1414. doi: 10.1074/jbc.M113.521997.

Rauch, J. N., Zuiderweg, E. R. P. and Gestwicki, J. E. (2016) 'Non-canonical Interactions between Heat Shock Cognate Protein 70 (Hsc70) and Bcl2-associated Anthanogene (BAG) Co-Chaperones Are Important for Client Release', *The Journal of biological chemistry*. 2016/07/29. American Society for Biochemistry and Molecular Biology, 291(38), pp. 19848–19857. doi: 10.1074/jbc.M116.742502.

Reymond, A. *et al.* (2001) 'The tripartite motif family identifies cell compartments.', *The EMBO journal*, 20(9), pp. 2140–2151. doi: 10.1093/emboj/20.9.2140.

Rogers, S., Wells, R. and Rechsteiner, M. (1986) 'Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis', *Science*. American Association for the Advancement of Science, 234(4774), pp. 364–368. doi: 10.1126/science.2876518.

Root, D. D. *et al.* (2006) 'Coiled-coil nanomechanics and uncoiling and unfolding of the superhelix and alpha-helices of myosin.', *Biophysical journal*, 90(8), pp. 2852–2866. doi: 10.1529/biophysj.105.071597.

Rosenberg, M. *et al.* (2008) 'MHC-IIB filament assembly and cellular localization are governed by the rod net charge', *PLoS ONE*, 3(1). doi: 10.1371/journal.pone.0001496.

Roy, A., Kucukural, A. and Zhang, Y. (2010) 'I-TASSER: a unified platform for automated protein structure and function prediction.', *Nature protocols*, 5(4), pp. 725–738. doi: 10.1038/nprot.2010.5.

Ruppel, K. M. and Spudich, J. A. (1996) 'Structure-function studies of the myosin motor domain: Importance of the 50-kDa cleft', *Molecular Biology of the Cell*, 7(7), pp. 1123–1136. doi: 10.1091/mbc.7.7.1123.

Sanchez, J. G. *et al.* (2014) 'The tripartite motif coiled-coil is an elongated antiparallel hairpin dimer.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(7), pp. 2494–2499. doi: 10.1073/pnas.1318962111.

Sanger, J. W. *et al.* (2005) ‘How to build a myofibril’, *Journal of Muscle Research and Cell Motility*, 26(6–8), pp. 343–354. doi: 10.1007/s10974-005-9016-7.

Scheffner, M., Nuber, U. and Huibregtse, J. M. (1995) ‘Protein ubiquitination involving an E1–E2–E3 enzyme ubiquitin thioester cascade’, *Nature*, 373(6509), pp. 81–83. doi: 10.1038/373081a0.

Schindelin, J. *et al.* (2012) ‘Fiji: an open-source platform for biological-image analysis.’, *Nature methods*, 9(7), pp. 676–682. doi: 10.1038/nmeth.2019.

Schmid, D. *et al.* (1994) ‘Kinetics of molecular chaperone action’, *Science*, 263(5149), pp. 971–973. doi: 10.1126/science.8310296.

Schneider, T. D. and Stephens, R. M. (1990) ‘Sequence logos: a new way to display consensus sequences.’, *Nucleic acids research*, 18(20), pp. 6097–6100. doi: 10.1093/nar/18.20.6097.

Seo, J. and Cohen, C. (1993) ‘Pitch diversity in α-helical coiled coils’, *Proteins: Structure, Function, and Bioinformatics*, 15(3), pp. 223–234. doi: 10.1002/prot.340150302.

Short, K. M. and Cox, T. C. (2006) ‘Subclassification of the RBCC/TRIM superfamily reveals a novel motif necessary for microtubule binding.’, *The Journal of biological chemistry*. United States, 281(13), pp. 8970–8980. doi: 10.1074/jbc.M512755200.

Silvestre, J. G. *et al.* (2019) ‘The e3 ligase murf2 plays a key role in the functional capacity of skeletal muscle fibroblasts’, *Brazilian Journal of Medical and Biological Research*, 52(9), pp. 1–10. doi: 10.1590/1414-431x20198551.

Sohn, R. L. *et al.* (1997) ‘A 29 residue region of the sarcomeric myosin rod is necessary for filament formation11Edited by J. Karn’, *Journal of Molecular Biology*, 266(2), pp. 317–330. doi: https://doi.org/10.1006/jmbi.1996.0790.

Solomon, V. and Goldberg, A. L. (1996) ‘Importance of the ATP-ubiquitin-proteasome pathway in the degradation of soluble and myofibrillar proteins in rabbit muscle extracts.’, *The Journal of biological chemistry*. United States, 271(43), pp. 26690–26697. doi: 10.1074/jbc.271.43.26690.

Sondermann, H. *et al.* (2001) ‘Structure of a Bag/Hsc70 complex: Convergent functional evolution of Hsp70 nucleotide exchange factors’, *Science*, 291(5508), pp. 1553–1557. doi: 10.1126/science.1057268.

Spencer, J. A. *et al.* (2000) ‘Regulation of microtubule dynamics and myogenic differentiation by MURF, a striated muscle RING-finger protein’, *Journal of Cell Biology*, 150(4), pp. 771–784. doi: 10.1083/jcb.150.4.771.

Van Der Spoel, D. *et al.* (2005) ‘GROMACS: Fast, flexible, and free’, *Journal of Computational Chemistry*. John Wiley & Sons, Ltd, 26(16), pp. 1701–1718. doi: 10.1002/jcc.20291.

Spudich, J. A. (2001) ‘The myosin swinging cross-bridge model’, *Nature Reviews Molecular Cell Biology*, 2(5), pp. 387–392. doi: 10.1038/35073086.

Squire, J. M., Paul, D. M. and Morris, E. P. (2017) ‘Myosin and Actin Filaments in Muscle: Structures and Interactions.’, *Sub-cellular biochemistry*. United States, 82, pp. 319–371. doi: 10.1007/978-3-319-49674-0_11.

Srikakulam, R. and Winkelmann, D. A. (2004) ‘Chaperone-mediated folding and assembly of myosin in striated muscle’, *Journal of Cell Science*. The Company of Biologists Ltd, 117(4), pp. 641–652. doi: 10.1242/jcs.00899.

Stevens, M. *et al.* (2019) ‘Exploration of the TRIM Fold of MuRF1 Using EPR Reveals a Canonical Antiparallel Structure and Extended COS-Box’, *Journal of Molecular Biology*. The Author(s), 431(15), pp. 2900–2909. doi: 10.1016/j.jmb.2019.05.025.

Stone, S. L. *et al.* (2005) ‘Functional analysis of the RING-type ubiquitin ligase family of Arabidopsis.’, *Plant physiology*, 137(1), pp. 13–30. doi: 10.1104/pp.104.052423.

Streich, F. C. J. *et al.* (2013) ‘Tripartite motif ligases catalyze polyubiquitin chain formation through a cooperative allosteric mechanism.’, *The Journal of biological chemistry*, 288(12), pp. 8209–8221. doi: 10.1074/jbc.M113.451567.

Strelkov, S. V and Burkhard, P. (2002) ‘Analysis of α-Helical Coiled Coils with the Program

TWISTER Reveals a Structural Mechanism for Stutter Compensation', *Journal of Structural Biology*, 137(1), pp. 54–64. doi: https://doi.org/10.1006/jsbi.2002.4454.

Stürner, E. and Behl, C. (2017) 'The Role of the Multifunctional BAG3 Protein in Cellular Protein Quality Control and in Disease', *Frontiers in Molecular Neuroscience*, 10(June), pp. 1–18. doi: 10.3389/fnmol.2017.00177.

Takayama, S. and Reed, J. C. (2001) 'Molecular chaperone targeting and regulation by BAG family proteins', *Nature Cell Biology*, 3(10), pp. E237–E241. doi: 10.1038/ncb1001-e237.

Takayama, S., Xie, Z. and Reed, J. C. (1999) 'An Evolutionarily Conserved Family of Hsp70/Hsc70 Molecular Chaperone Regulators', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 274(2), pp. 781–786. doi: 10.1074/JBC.274.2.781.

Taylor, K. C. *et al.* (2015) 'Skip residues modulate the structural properties of the myosin rod and guide thick filament assembly', *Proceedings of the National Academy of Sciences*, 112(29), pp. E3806–E3815. doi: 10.1073/pnas.1505813112.

Theyssen, H. *et al.* (1996) 'The second step of ATP binding to DnaK induces peptide release.', *Journal of molecular biology*. England, 263(5), pp. 657–670. doi: 10.1006/jmbi.1996.0606.

Villard, E. *et al.* (2011) 'A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy', *European heart journal*. 2011/04/01. Oxford University Press, 32(9), pp. 1065–1076. doi: 10.1093/eurheartj/ehr105.

Walsh, P. *et al.* (2004) 'The J-protein family: Modulating protein assembly, disassembly and translocation', *EMBO Reports*, 5(6), pp. 567–571. doi: 10.1038/sj.embor.7400172.

Weinert, C. *et al.* (2015) 'Crystal structure of TRIM20 C-terminal coiled-coil/B30.2 fragment: implications for the recognition of higher order oligomers.', *Scientific reports*, 5, p. 10819. doi: 10.1038/srep10819.

Weinert, S. *et al.* (2006) 'M line-deficient titin causes cardiac lethality through impaired maturation of the sarcomere.', *The Journal of cell biology*, 173(4), pp. 559–570. doi: 10.1083/jcb.200601014.

Witt, C. C. *et al.* (2008) 'Cooperative control of striated muscle mass and metabolism by MuRF1 and MuRF2', *EMBO Journal*, 27(2), pp. 350–360. doi: 10.1038/sj.emboj.7601952.

Witt, S. H. *et al.* (2005) 'MURF-1 and MURF-2 target a specific subset of myofibrillar proteins redundantly: Towards understanding MURF-dependent muscle ubiquitination', *Journal of Molecular Biology*, 350(4), pp. 713–722. doi: 10.1016/j.jmb.2005.05.021.

Wollersheim, T. *et al.* (2014) 'Dynamics of myosin degradation in intensive care unit-acquired weakness during severe critical illness', *Intensive Care Medicine*, 40(4), pp. 528–538. doi: 10.1007/s00134-014-3224-9.

Wolny, M. *et al.* (2013) 'Cardiomyopathy mutations in the tail of β-cardiac myosin modify the coiled-coil structure and affect integration into thick filaments in muscle sarcomeres in adult cardiomyocytes', *Journal of Biological Chemistry*, 288(44), pp. 31952–31962. doi: 10.1074/jbc.M113.513291.

Woo, H.-J. *et al.* (2009) 'ATP-induced conformational changes in Hsp70: molecular dynamics and experimental validation of an in silico predicted conformation.', *Biochemistry*, 48(48), pp. 11470–11477. doi: 10.1021/bi901256y.

Yudina, Z. *et al.* (2015) 'RING Dimerization Links Higher-Order Assembly of TRIM5α to Synthesis of K63-Linked Polyubiquitin.', *Cell reports*, 12(5), pp. 788–797. doi: 10.1016/j.celrep.2015.06.072.

Zenker, M. *et al.* (2005) 'Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome).', *Nature genetics*. United States, 37(12), pp. 1345–1350. doi: 10.1038/ng1681.

Zhou, T. *et al.* (2016) 'CARP interacts with titin at a unique helical N2A sequence and at the domain Ig81 to form a structured complex.', *FEBS letters*. England, pp. 3098–3110. doi:

Appendices

Appendix A

All protein constructs used in this study were recombinantly produced in *E. coli* if not indicated differently. In this Appendix are provided all information about used constructs, including plasmid description and corresponding protein sequence.

A.1 Full-length mouse MuRF1 wild type and F148D mutant

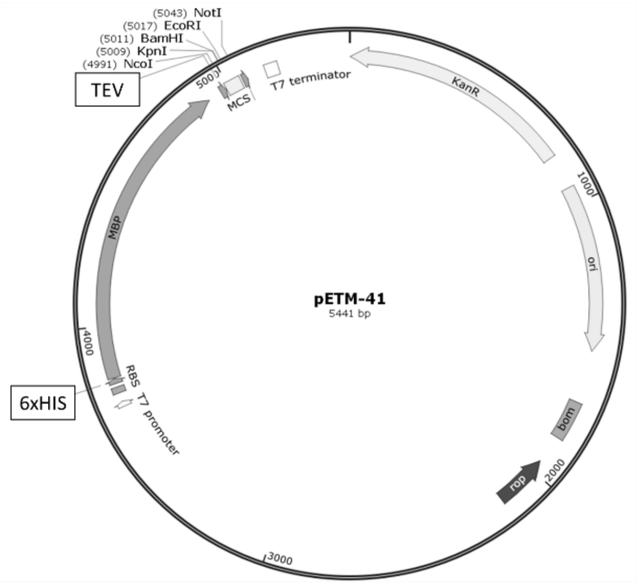


Figure A.1 Full length MuRF1 was inserted in pETM-41 expression vector via NcoI and KpnI restriction sites which includes His$_6$-MBP tag and TEV (tobacco etch virus) protease cleavage site prior to the inserted gene. Construct is C293S variant. F148D mutation was introduced via site directed mutagenesis.

His$_6$-MBP-MuRF1 mouse (Mus musculus) sequence (UniProtKB - Q38HM41)

```
        10         20         30         40         50         60
MKHHHHHHPM KIEEGKLVIW INGDKGYNGL AEVGKKFEKD TGIKVTVEHP DKLEEKFPQV

        70         80         90        100        110        120
AATGDGPDII FWAHDRFGGY AQSGLLAEIT PDKAFQDKLY PFTWDAVRYN GKLIAYPIAV

       130        140        150        160        170        180
EALSLIYNKD LLPNPPKTWE EIPALDKELK AKGKSALMFN LQEPYFTWPL IAADGGYAFK

       190        200        210        220        230        240
YENGKYDIKD VGVDNAGAKA GLTFLVDLIK NKHMNADTDY SIAEAAFNKG ETAMTINGPW

       250        260        270        280        290        300
AWSNIDTSKV NYGVTVLPTF KGQPSKPFVG VLSAGINAAS PNKELAKEFL ENYLLTDEGL

       310        320        330        340        350        360
EAVNKDKPLG AVALKSYEEE LAKDPRIAAT MENAQKGEIM PNIPQMSAFW YAVRTAVINA

       370        380        390        400        410        420
ASGRQTVDEA LKDAQTNSSS NNNNNNNNNN PMSENLYFQG AMMDYKSSLI PDGNAMENLE

       430        440        450        460        470        480
KQLICPICLE MFTKPVVILP CQHNLCRKCA NDIFQAANPY WTNRGGSVSM SGGRFRCPSC

       490        500        510        520        530        540
RHEVIMDRHG VYGLQRNLLV ENIIDIYKQE CSSRPLQKGS HPMCKEHEDE KINIYCLTCE

       550        560        570        580        590        600
VPTCSLCKVF GAHQACEVAP LQSIFQGQKT ELSNCISMLV AGNDRVQTII SQLEDSCRVT

       610        620        630        640        650        660
KENSHQVKEE LSQKFDTLYA ILDEKKSELL QRITQEQEEK LGFIEALILQ YREQLEKSTK

       670        680        690        700        710        720
LVETAIQSLD EPGGATFLSS AKQLIKSIVE ASKGCQLGKT EQGFENMDYF TLDLEHIAEA

       730        740        750
LRAIDFGTDE EEEEEEFTEE EADEEEGVTT EGKEEHQ
```

*Mutated residue is labeled in red

A.2 MuRF1-BHD and B-box

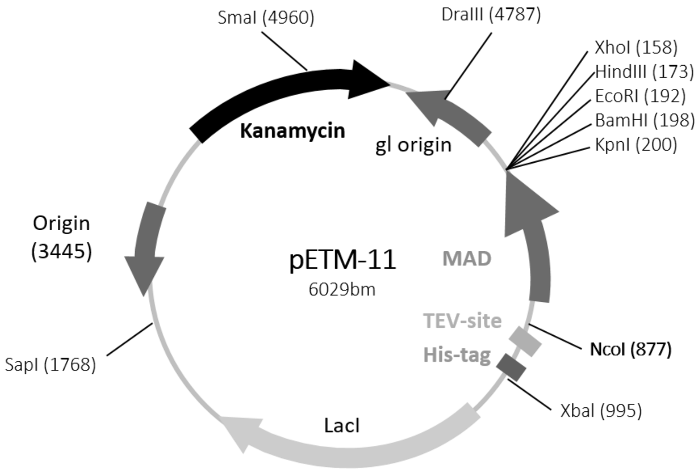


Figure A.2 BHD and B-box fragments of MuRF1 were inserted in pETM-11 via restriction sites NcoI and KpnI which includes an N-terminal His$_6$- tag and a TEV (tobacco etch virus) protease cleavage site prior to the inserted gene.

His$_6$-BHD human (Homo sapiens) sequence (UniProtKB - Q969Q1, residues 117-327)

```
        10         20         30         40         50         60
MKHHHHHHPM SDYDIPTTEN LYFQGAMGSH PMCKEHEDEK INIYCLTCEV PTCSMCKVFG

        70         80         90        100        110        120
IHKACEVAPL QSVFQGQKTE LNNCISMLVA GNDRVQTIIT QLEDSRRVTK ENSHQVKEEL

       130        140        150        160        170        180
SQKFDTLYAI LDEKKSELLQ RITQEQEEKL SFIEALIQQY QEQLDKSTKL VETAIQSLDE

       190        200        210        220        230
PGGATFLLTA KQLIKSIVEA SKGSQLGKTE QGFENMDFFT LDLEHIADAL RAIDFGTD
```

*F148 residue (labeled in red) was mutated to D using site directed mutagenesis.

His$_6$-Bbox human (Homo sapiens) sequence (UniProtKB - Q969Q1, residues 117-161)

```
        10         20         30         40         50         60
MKHHHHHHPM SDYDIPTTEN LYFQGAMGSH PMCKEHEDEK INIYCLTCEV PTCSMCKVFG

        70
IHKACEVAPL QS
```

A.3 Myosin heavy chain fragments

For this study all myosin heavy chain fragments (MYH7^{x-y}) expressed using pOPINJ expression vectors (Figure A.1.3) and in fusion cloning method.

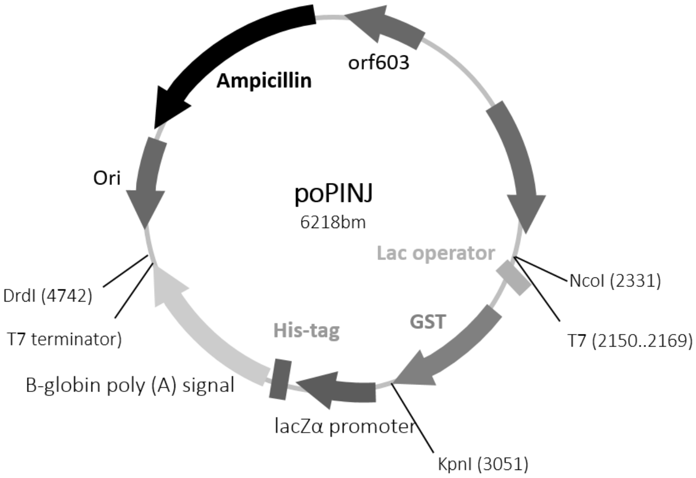


Figure A.3 pOPINJ expression vector containing a His$_6$-GST tag and a 3C protease cleavage site N-terminal to the target insert.

His$_6$-GST-MYH7$^{838-964}$ human (Homo sapiens) sequence (UniProtKB P12883, residues 838-964)

```
        10         20         30         40         50         60
MAHHHHHHMS PILGYWKIKG LVQPTRLLLE YLEEKYEEHL YERDEGDKWR NKKFELGLEF

        70         80         90        100        110        120
PNLPYYIDGD VKLTQSMAII RYIADKHNML GGCPKERAEI SMLEGAVLDI RYGVSRIAYS

       130        140        150        160        170        180
KDFETLKVDF LSKLPEMLKM FEDRLCHKTY LNGDHVTHPD FMLYDALDVV LYMDPMCLDA

       190        200        210        220        230        240
FPKLVCFKKR IEAIPQIDKY LKSSKYIAWP LQGWQATFGG GDHPPKSDLS SGLEVLFQGP

       250        260        270        280        290        300
LLKSAEREKE MASMKEEFTR LKEALEKSEA RRKELEEKMV SLLQEKNDLQ LQVQAEQDNL

       310        320        330        340        350        360
ADAEERCDQL IKNKIQLEAK VKEMNERLED EEEMNAELTA KKRKLEDECS ELKRDIDDLE
```

His$_6$-GST-MYH7$^{964-1085}$ human sequence (UniProtKB P12883, residues 964-1088)

```
        10         20         30         40         50         60
MAHHHHHHMS PILGYWKIKG LVQPTRLLLE YLEEKYEEHL YERDEGDKWR NKKFELGLEF

        70         80         90        100        110        120
PNLPYYIDGD VKLTQSMAII RYIADKHNML GGCPKERAEI SMLEGAVLDI RYGVSRIAYS

       130        140        150        160        170        180
KDFETLKVDF LSKLPEMLKM FEDRLCHKTY LNGDHVTHPD FMLYDALDVV LYMDPMCLDA

       190        200        210        220        230        240
FPKLVCFKKR IEAIPQIDKY LKSSKYIAWP LQGWQATFGG GDHPPKSDLS SGLEVLFQGP

       250        260        270        280        290        300
VEKEKHATEN KVKNLTEEME TAGLDEIIAK LTKEKKALQE AHQQALDDLQ AEEDKVNTLT

       310        320        330        340        350        360
KAKVKLEQQV DDLEGSLEQE KKVRMETDLE RAKRKLEGDL KLTQESIMET DLENDKQQLD
```

His$_6$-GST-MYH7$^{1088-1218}$ human sequence (UniProtKB P12883, residues 1088-1218)

```
        10         20         30         40         50         60
MAHHHHHHMS PILGYWKIKG LVQPTRLLLE YLEEKYEEHL YERDEGDKWR NKKFELGLEF

        70         80         90        100        110        120
PNLPYYIDGD VKLTQSMAII RYIADKHNML GGCPKERAEI SMLEGAVLDI RYGVSRIAYS

       130        140        150        160        170        180
KDFETLKVDF LSKLPEMLKM FEDRLCHKTY LNGDHVTHPD FMLYDALDVV LYMDPMCLDA

       190        200        210        220        230        240
FPKLVCFKKR IEAIPQIDKY LKSSKYIAWP LQGWQATFGG GDHPPKSDLS SGLEVLFQGP

       250        260        270        280        290        300
ALNARIEDEQ RLEEAGGATS NLQRVKQKLE ALGSQLQKKL KELQARIEEL EEELEAERTA

       310        320        330        340        350        360
RAKVEKLRSD LSRELEEISE VQIEMNKKRE AEFQKMRRDL EEATLQHEAT AAALRKKHAD
```

His$_6$-GST-MYH7$^{1218-1301}$ human sequence (UniProtKB P12883, residues 1219-1301)

```
        10         20         30         40         50         60
MAHHHHHHMS PILGYWKIKG LVQPTRLLLE YLEEKYEEHL YERDEGDKWR NKKFELGLEF

        70         80         90        100        110        120
PNLPYYIDGD VKLTQSMAII RYIADKHNML GGCPKERAEI SMLEGAVLDI RYGVSRIAYS

       130        140        150        160        170        180
KDFETLKVDF LSKLPEMLKM FEDRLCHKTY LNGDHVTHPD FMLYDALDVV LYMDPMCLDA


       190        200        210        220        230        240
FPKLVCFKKR IEAIPQIDKY LKSSKYIAWP LQGWQATFGG GDHPPKSDLS SGLEVLFQGP

       250        260        270        280        290        300
KEKSEFKLEL DDVTSNMETE QIIKAKANLE KMETCRTLED QMETNEHRSK AEETQRSVND

       310        320
LTSQRAKLQT ENGELSRQLD EKEALISQ
```

His$_6$-GST-MYH7$^{1301-1378}$ human sequence (UniProtKB P12883, residues 1301-1378)

```
        10         20         30         40         50         60
MAHHHHHHMS PILGYWKIKG LVQPTRLLLE YLEEKYEEHL YERDEGDKWR NKKFELGLEF

        70         80         90        100        110        120
PNLPYYIDGD VKLTQSMAII RYIADKHNML GGCPKERAEI SMLEGAVLDI RYGVSRIAYS

       130        140        150        160        170        180
KDFETLKVDF LSKLPEMLKM FEDRLCHKTY LNGDHVTHPD FMLYDALDVV LYMDPMCLDA

       190        200        210        220        230        240
FPKLVCFKKR IEAIPQIDKY LKSSKYIAWP LQGWQATFGG GDHPPKSDLS SGLEVLFQGP

       250        260        270        280        290        300
LTRGKLTYTQ QLEDLKRQLE EEVKAKNALA HALQSARHDC DLLREQYEEE TEAKAELQRV

       310
LSKANSEVAQ WRTKYETD
```

His$_6$-GST-MYH7$^{1357-1407}$ human sequence (UniProtKB P12883, residues 1357-1407)

```
        10         20         30         40         50         60
MAHHHHHHMS PILGYWKIKG LVQPTRLLLE YLEEKYEEHL YERDEGDKWR NKKFELGLEF

        70         80         90        100        110        120
PNLPYYIDGD VKLTQSMAII RYIADKHNML GGCPKERAEI SMLEGAVLDI RYGVSRIAYS

       130        140        150        160        170        180
KDFETLKVDF LSKLPEMLKM FEDRLCHKTY LNGDHVTHPD FMLYDALDVV LYMDPMCLDA

       190        200        210        220        230        240
FPKLVCFKKR IEAIPQIDKY LKSSKYIAWP LQGWQATFGG GDHPPKSDLS SGLEVLFQGP

       250        260        270        280        290
LQRVLSKANS EVAQWRTKYE TDAIQRTEEL EEAKKKLAQR LQEAEEAVEA V
```

His$_6$-GST-MYH7$^{1477-1793}$ human sequence (UniProtKB P12883, residues 1477-1793)

```
        10         20         30         40         50         60
MAHHHHHHMS PILGYWKIKG LVQPTRLLLE YLEEKYEEHL YERDEGDKWR NKKFELGLEF

        70         80         90        100        110        120
PNLPYYIDGD VKLTQSMAII RYIADKHNML GGCPKERAEI SMLEGAVLDI RYGVSRIAYS

       130        140        150        160        170        180
KDFETLKVDF LSKLPEMLKM FEDRLCHKTY LNGDHVTHPD FMLYDALDVV LYMDPMCLDA

       190        200        210        220        230        240
FPKLVCFKKR IEAIPQIDKY LKSSKYIAWP LQGWQATFGG GDHPPKSDLS SGLEVLFQGP
       250        260        270        280        290        300
LSTELFKLKN AYEESLEHLE TFKRENKNLQ EEISDLTEQL GSSGKTIHEL EKVRKQLEAE

       310        320        330        340        350        360
KMELQSALEE AEASLEHEEG KILRAQLEFN QIKAEIERKL AEKDEEMEQA KRNHLRVVDS

       370        380        390        400        410        420
LQTSLDAETR SRNEALRVKK KMEGDLNEME IQLSHANRMA AEAQKQVKSL QSLLKDTQIQ

       430        440        450        460        470        480
LDDAVRANDD LKENIAIVER RNNLLQAELE ELRAVVEQTE RSRKLAEQEL IETSERVQLL

       490        500        510        520        530        540
HSQNTSLINQ KKKMDADLSQ LQTEVEEAVQ ECRNAEEKAK KAITDAAMMA EELKKEQDTS
```

A.4 Titin A168-170 and A168-M1 constructs

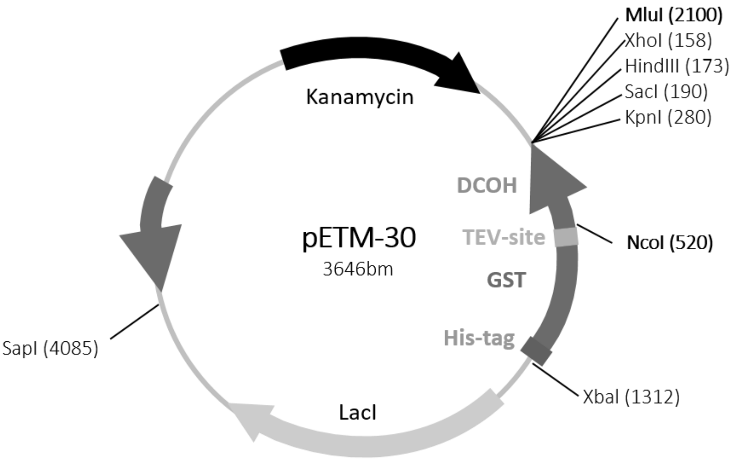


Figure A.4 Titin A168-170 and titin A170-M1 were inserted at restriction sites NcoI and MIuI in pETM-30 which includes an N-terminal His$_6$-GST tag and a TEV (tobacco etch virus) protease cleavage site prior to the inserted gene.

His$_6$-GST-Titin A168-170 human (Homo sapiens) sequence (UniProtKB Q8WZ42, residues 24431–24731)

```
        10         20         30         40         50         60
MKHHHHHHPM SDKIIHLTDD SFDTDVLKAD GAILVDFWAE WCGPCKMIAP ILDEIADEYQ

        70         80         90        100        110        120
GKLTVAKLNI DQNPGTAPKY GIRGIPTLLL FKNGEVAATK VGALSKGQLK EFLDANLAGS

       130        140        150        160        170        180
GSGSENLYFQ GVGKASKNSE CYVARDPCDP PGTPEPIMVK RNEITLQWTK PVYDGGSMIT

       190        200        210        220        230        240
GYIVEKRDLP DGRWMKASFT NVIETQFTVS GLTEDQRYEF RVIAKNAAGA ISKPSDSTGP

       250        260        270        280        290        300
ITAKDEVELP RISMDPKFRD TIVVNAGETF RLEADVHGKP LPTIEWLRGD KEIEESARCE

       310        320        330        340        350        360
IKNTDFKALL IVKDAIRIDG GQYILRASNV AGSKSFPVNV KVLDRPGPPE GPVQVTGVTS

       370        380        390        400        410        420
EKCSLTWSPP LQDGGSDISH YVVEKRETSR LAWTVVASEV VTNSLKVTKL LEGNEYVFRI

       430
MAVNKYGVGE P
```

His$_6$-GST-TitinA170-M1 human (Homo sapiens) sequence (UniProtKB Q8WZ42, residues 24622-25170)

```
        10         20         30         40         50         60
MKHHHHHHPM SDKIIHLTDD SFDTDVLKAD GAILVDFWAE WCGPCKMIAP ILDEIADEYQ

        70         80         90        100        110        120
GKLTVAKLNI DQNPGTAPKY GIRGIPTLLL FKNGEVAATK VGALSKGQLK EFLDANLAGS

       130        140        150        160        170        180
GSGSENLYFQ GAMVADVPDP PRGVKVSDVS RDSVNLTWTE PASDGGSKIT NYIVEKCATT

       190        200        210        220        230        240
AERWLRVGQA RETRYTVINL FGKTSYQFRV IAENKFGLSK PSEPSEPTIT KEDKTRAMNY

       250        260        270        280        290        300
DEEVDETREV SMTKASHSST KELYEKYMIA EDLGRGEFGI VHRCVETSSK KTYMAKFVKV

       310        320        330        340        350        360
KGTDQVLVKK EISILNIARH RNILHLHESF ESMEELVMIF EFISGLDIFE RINTSAFELN

       370        380        390        400        410        420
EREIVSYVHQ VCEALQFLHS HNIGHFDIRP ENIIYQTRRS STIKIIEFGQ ARQLKPGDNF

       430        440        450        460        470        480
RLLFTAPEYY APEVHQHDVV STATDMWSLG TLVYVLLSGI NPFLAETNQQ IIENIMNAEY

       490        500        510        520        530        540
TFDEEAFKEI SIEAMDFVDR LLVKERKSRM TASEALQHPW LKQKIERVST KVIRTLKHRR

       550        560        570        580        590        600
YYHTLIKKDL NMVVSAARIS CGGAIRSQKG VSVAKVKVAS IEIGPVSGQI MHAVGEEGGH

       610        620        630        640        650        660
VKYVCKIENY DQSTQVTWYF GVRQLENSEK YEITYEDGVA ILYVKDITKL DDGTYRCKVV

       670        680
NDYGEDSSYA ELFVKGVREV YDY
```

A.5 CARP$^{106-319}$

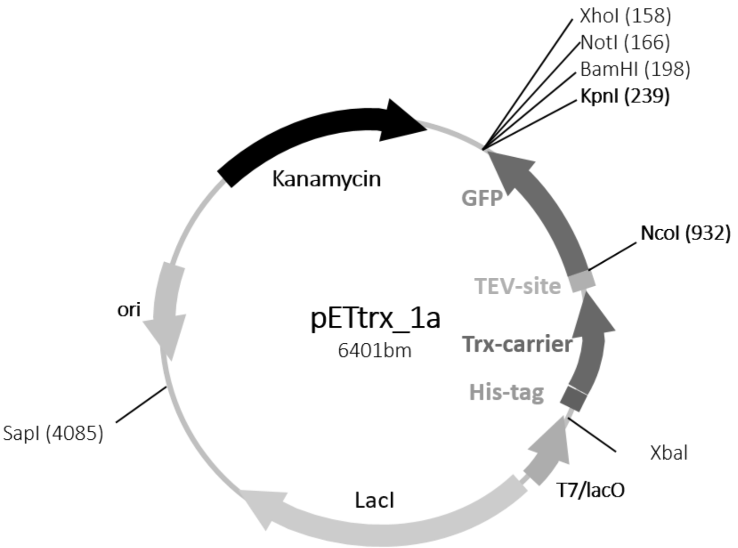


Figure A.5 CARP$^{106-319}$ coding sequence was inserted via NcoI and KpnI insertion sites into the pET-trx1a vector that adds a His$_6$- Tag, a thioredoxin domain and a tobacco etch virus (TEV) protease recognition sequence Nterminal to the target gen.

His$_6$-Trx-CARP 106-319 human (Homo sapiens) Sequence (UniProtKB Q15327, residues 106-319)

```
        10         20         30         40         50         60
HHHHHHGTSD KIIHLTDDSF DTDVLKADGA ILVDFWAEWC GPCKMIAPIL DEIADEYGIR

        70         80         90        100        110        120
GIPTLLLFKN GEVAATKVGA LSKGQLKEFL DANLAVVKEP EPEIITEPVD VPTFLKAALE

       130        140        150        160        170        180
NKLPVVEKFL SDKNNPDVCD EYKRTALHRA CLEGHLAIVE KLMEAGAQIE FRDMLESTAI

       190        200        210        220        230        240
HWASRGGNLD VLKLLLNKGA KISARDKLLS TALHVAVRTG HYECAEHLIA CEADLNAKDR

       250        260        270        280        290        300
EGDTPLHDAV RLNRYKMIRL LIMYGADLNI KNCAGKTPMD LVLHWQNGTK AIFDSLRENS


YKTSRIATF
```

A.6 Ubiquitin

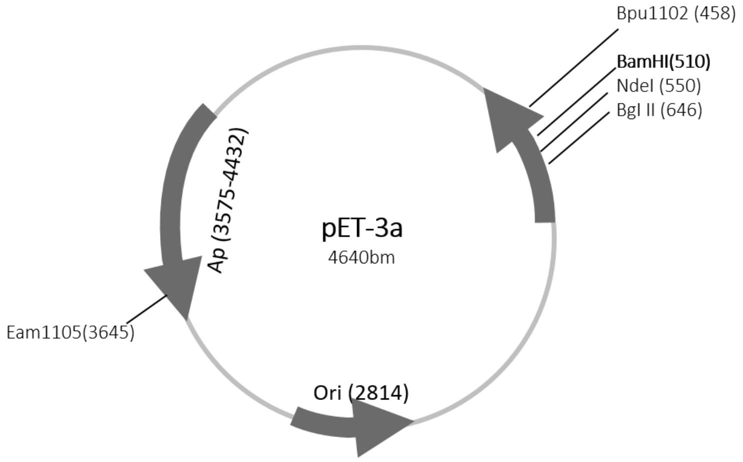


Figure A.6 Ubiquitin was inserted in pET-3a via restriction site BamHI which does not add any tag to the inserted sequence.

Ubiquitin human (Homo Sapiens) sequence (UniProtKB P0CG47)

```
        10         20         30         40         50         60
MQIFVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL EDGRTLSDYN
        70
IQKESTLHLV LRLRGG
```

A.7 UbcH5b (E2 conjugating enzyme)

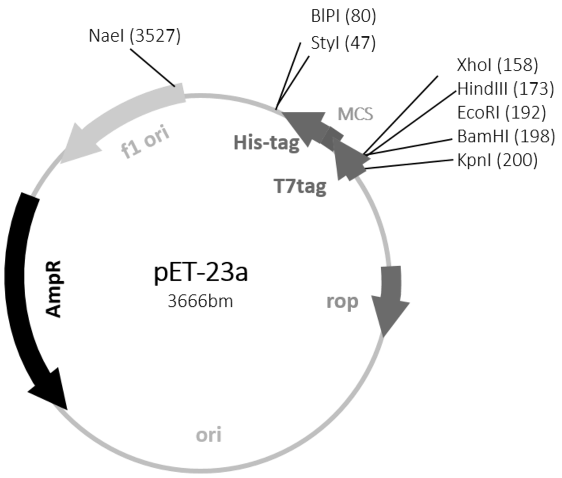


Figure A.7 UbcH5b was inserted in pET-23a via restriction site BamHI which adds His$_6$-tagto the insert.

His$_6$-Ubiquitin conjugating enzyme UbcH5b human (Homo Sapiens) sequence (UniProtKB - P62837)

```
        10         20         30         40         50         60
MKHHHHHHPM SDYDIPTTEN LYFQGAMALK RIHKELNDLA RDPPAQCSAG PVGDDMFHWQ

        70         80         90        100        110        120
ATIMGPNDSP YQGGVFFLTI HFPTDYPFKP PKVAFTTRIY HPNINSNGSI CLDILRSQWS

       130        140        150        160        170
PALTISKVLL SICSLLCDPN PDDPLVPEIA RIYKTDREKY NRIAREWTQK YAM
```

A.8 HSP70-nucleotide binding domain

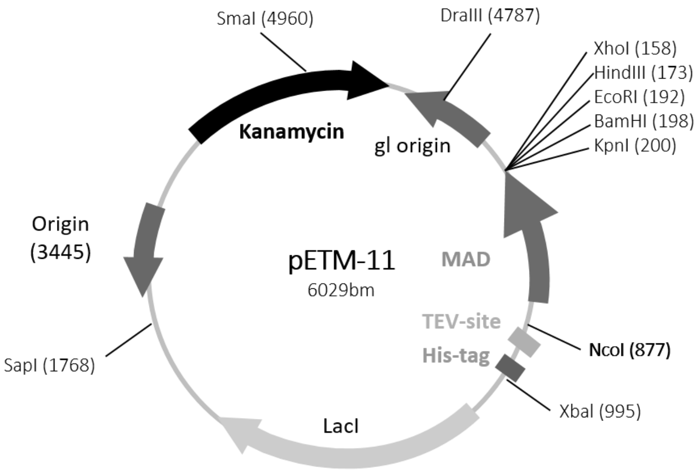


Figure A.8 HSP70-NBD was inserted in pETM-11 via restriction sites NcoI and KpnI which includes an N-terminal His$_6$- tag and a TEV (tobacco etch virus) protease cleavage site prior to the inserted gene.

His-HSP70-NBD mouse sequence (UniProtKB - Q61696, residues 1–388)

```
        10         20         30         40         50         60
MKHHHHHHPM SDYDIPTTEN LYFQGAMAKN TAIGIDLGTT YSCVGVFQHG KVEIIANDQG

        70         80         90        100        110        120
NRTTPSYVAF TDTERLIGDA AKNQVALNPQ NTVFDAKRLI GRKFGDAVVQ SDMKHWPFQV

       130        140        150        160        170        180
VNDGDKPKVQ VNYKGESRSF FPEEISSMVL TKMKEIAEAY LGHPVTNAVI TVPAYFNDSQ

       190        200        210        220        230        240
RQATKDAGVI AGLNVLRIIN EPTAAAIAYG LDRTGKGERN VLIFDLGGGT FDVSILTIDD

       250        260        270        280        290        300
GIFEVKATAG DTHLGGEDFD NRLVSHFVEE FKRKHKKDIS QNKRAVRRLR TACERAKRTL

       310        320        330        340        350        360
SSSTQASLEI DSLFEGIDFY TSITRARFEE LCSDLFRGTL EPVEKALRDA KMDKAQIHDL

       370        380        390        400        410
VLVGGSTRIP KVQKLLQDFF NGRDLNKSIN PDEAVAYGAA VQAAILMGDK SENV
```

A.9 BAG3-BAG domain wild type and E455K mutation

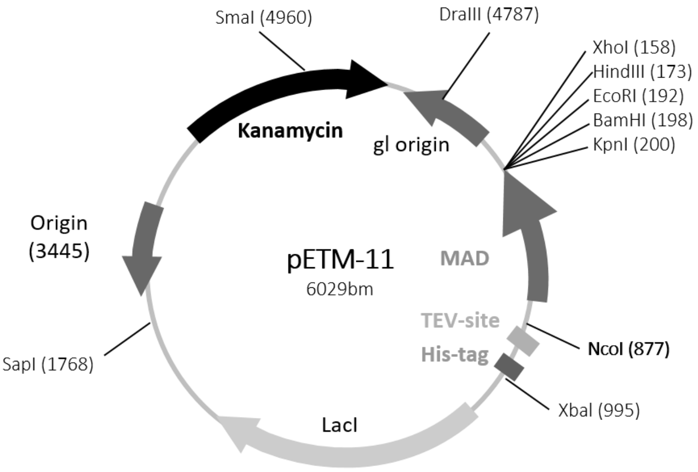


Figure A.9 BAG3-BAG were inserted in pETM-11 via restriction sites NcoI and KpnI which includes an N-terminal His$_6$- tag and a TEV (tobacco etch virus) protease cleavage site prior to the inserted gene

BAG3-BAG mouse (Mus musclus) sequence (UniProtKB - O95817, residues 419–509)

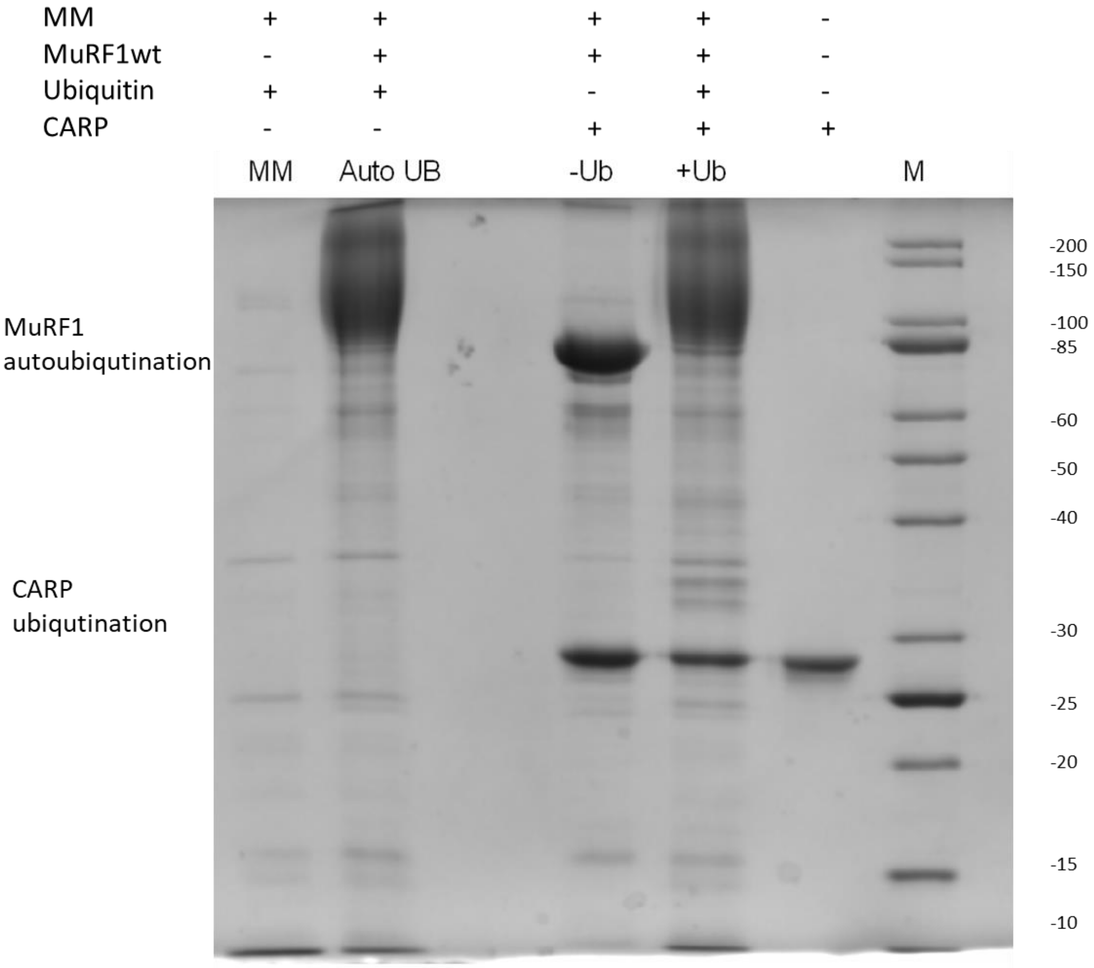
```
         10         20         30         40         50         60
MKHHHHHHPM SDYDIPTTEN LYFQGAMAET PPKHPGVLKV EAILEKVQGL EQAVDSFEGK

         70         80         90        100        110
KTDKKYLMIK EYLTKELLAL DSVDPEGRAD VRQARRDGVR KVQTILEKLE QKAIDVPGQ
```

*Mutated residue is labeled in red

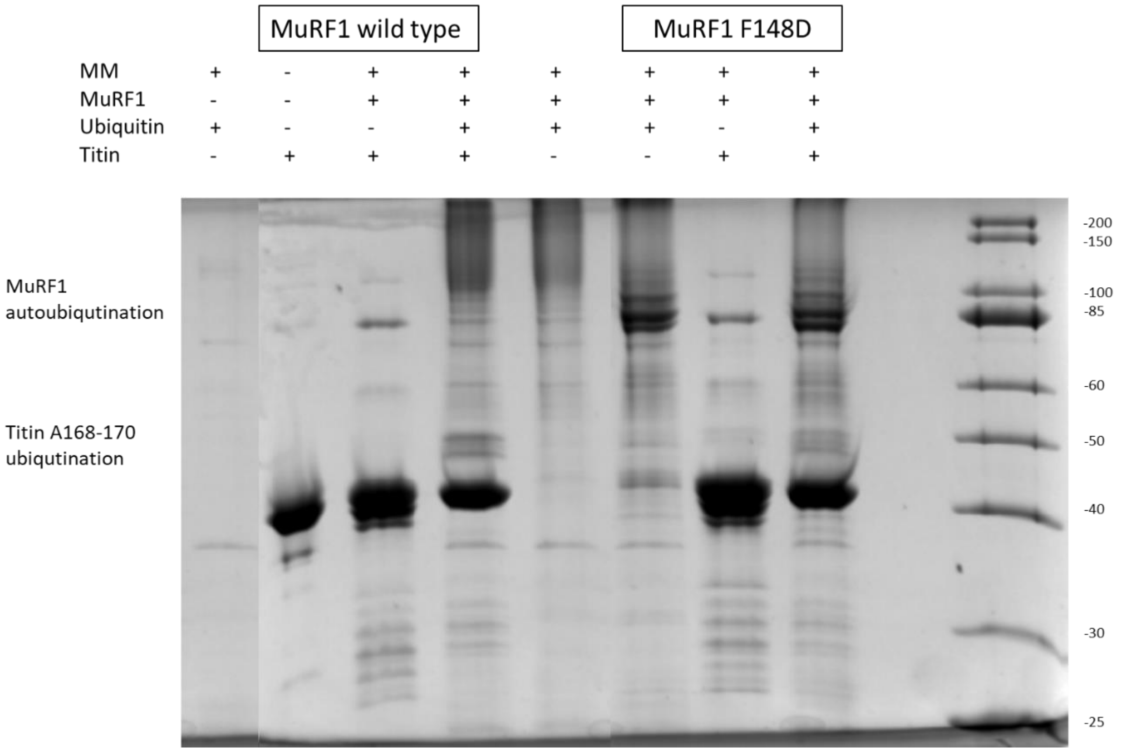
Appendix B

For the purpose of representation of important result, not the full SDS-PAGE gels of ubiquitination assays described in Chapters 1 and 2 were shown, but only selected parts. In this appendix, full SDS-PAGE gels obtained after ubiquitination gels are shown with the description shown outside of the gel area. In all the gels MM is “master mix” consisting of assay buffer, ubiquitin activating enzyme (E1), ubiquitin conjugating enzyme (E2), ATP and $MgCl_2$.

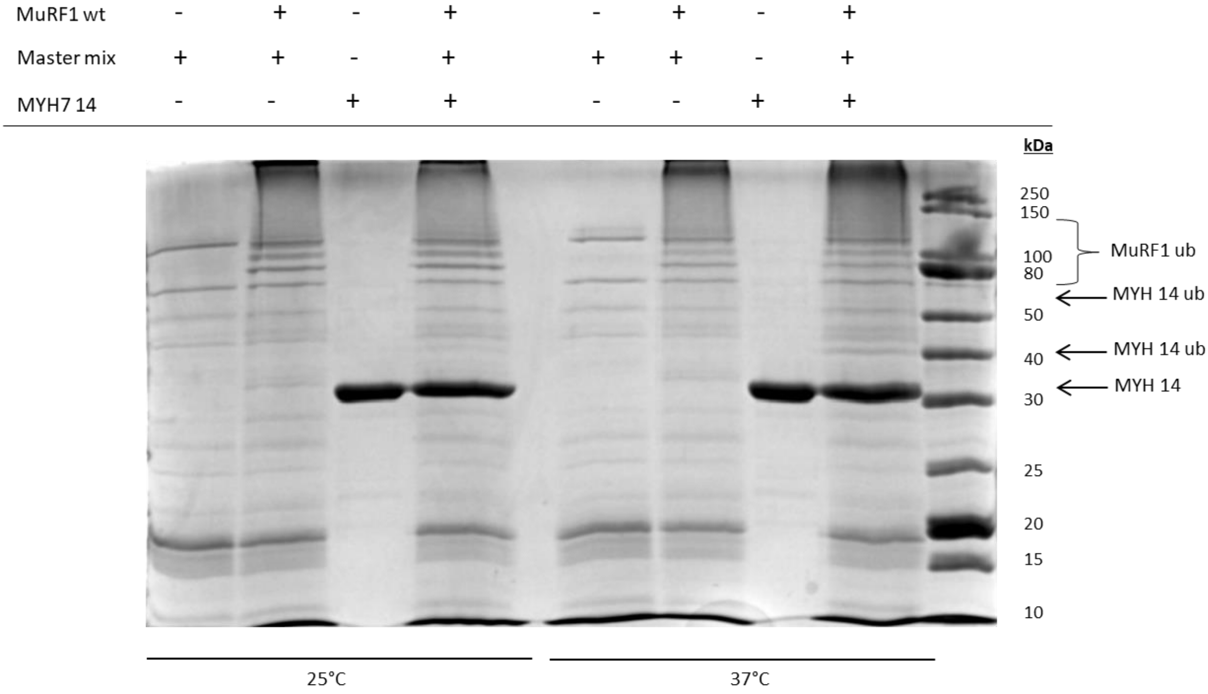
B.1. MuRF1 ubiquitination of CARP



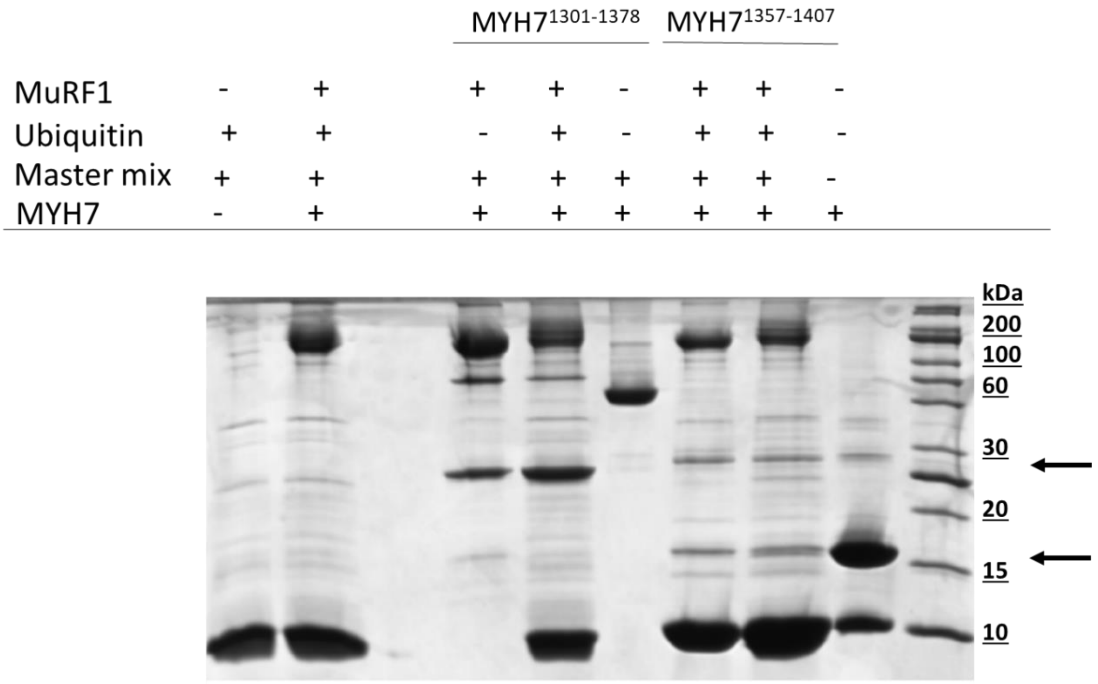
B.2 MuRF1 wt vs mutant ubiquitination of TitinA168-170



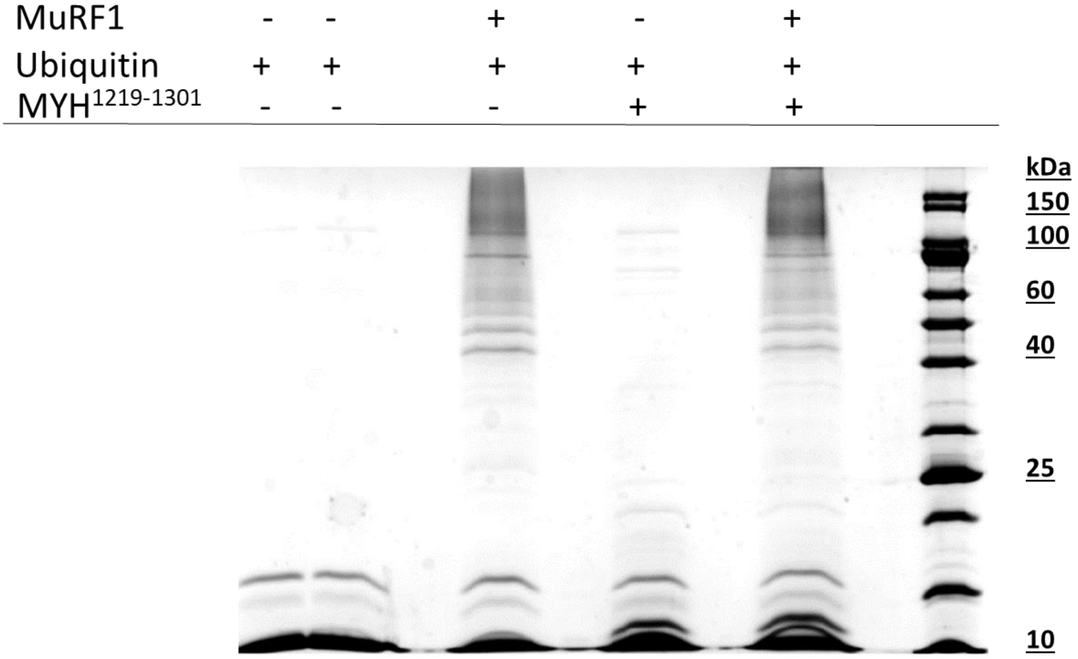
B.3 MuRF1 ubiquitination of MYH7$^{1477-1793}$



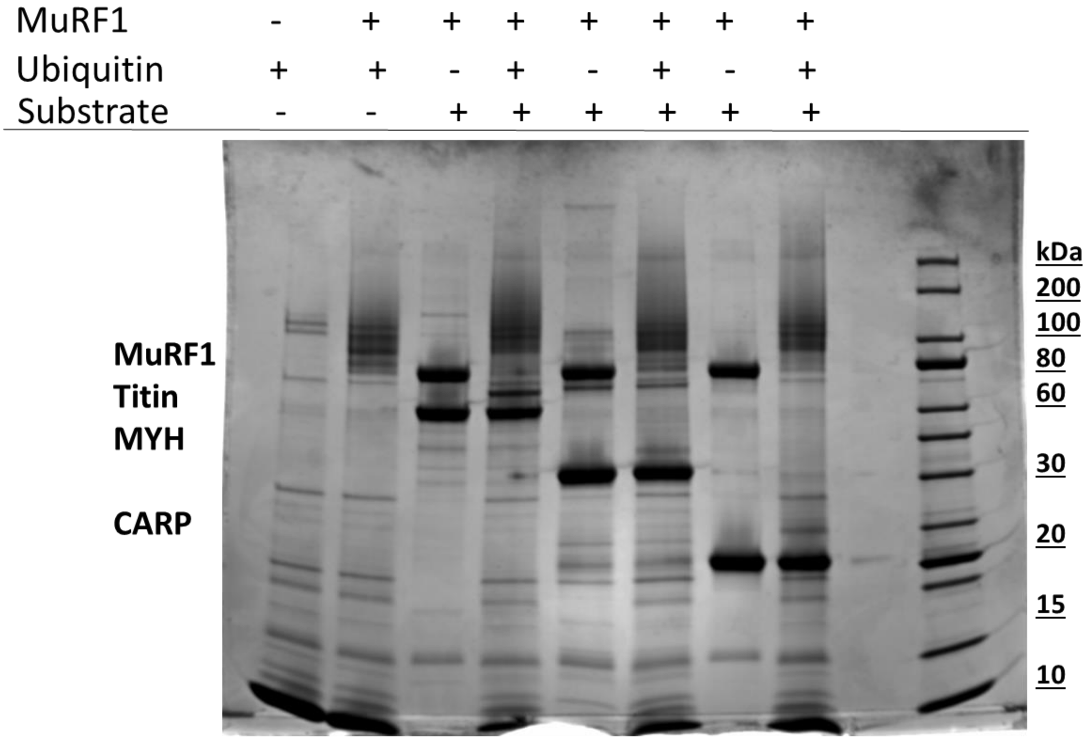
B.4 MuRF1 ubiquitination of MYH7$^{1301-1378}$ and MYH7$^{1357-1401}$



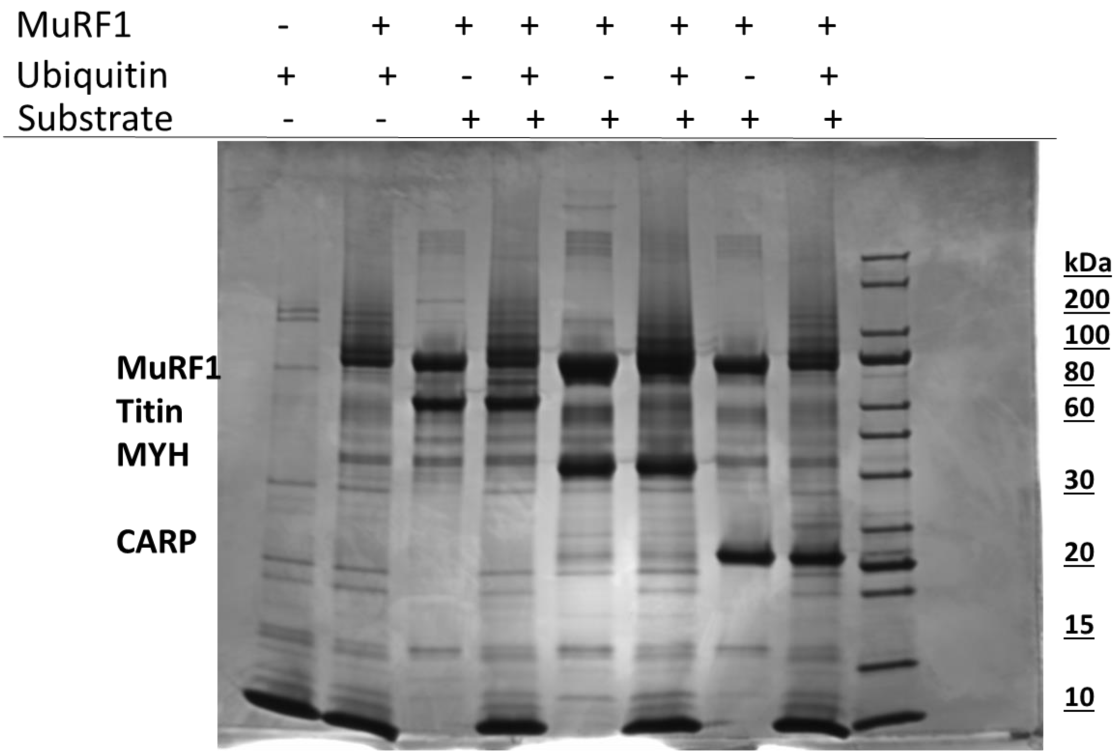
B.5 MuRF1 ubiquitination of MYH7$^{1219-1301}$



B.6 Ubiquitination activity of MuRF1^{wt} with its substrates titin A168-170, CARP and MYH$^{964-1088}$ MuRF1^{F148D}



B.7 Ubiquitination activity of MuRF1^{F148D} with its substrates titin A168-170, CARP and MYH$^{964-1088}$



Appendix C

C.1 MARCOIL analysis of myosin heavy chain sequence

MARCOIL input file UniProtKB - P12883 (MYH7_HUMAN), residues 838-1935

```
PLL KSAEREKEMA SMKEEFTRLK EALEKSEARR KELEEKMVSL LQEKNDLQLQ VQAEQDNLAD
AEERCDQLIK NKIQLEAKVK EMNERLEDEE EMNAELTAKK RKLEDECSEL KRDIDDLELT
LAKVEKEKHA TENKVKNLTE EMAGLDEIIA KLTKEKKALQ EAHQQALDDL QAEEDKVNTL
TKAKVKLEQQ VDDLEGSLEQ EKKVRMDLER AKRKLEGDLK LTQESIMDLE NDKQQLDERL
KKKDFELNAL NARIEDEQAL GSQLQKKLKE LQARIEELEE ELEAERTARA KVEKLRSDLS
RELEEISERL EEAGGATSVQ IEMNKKREAE FQKMRRDLEE ATLQHEATAA ALRKKHADSV
AELGEQIDNL QRVKQKLEKE KSEFKLELDD VTSNMEQIIK AKANLEKMCR TLEDQMNEHR
SKAEETQRSV NDLTSQRAKL QTENGELSRQ LDEKEALISQ LTRGKLTYTQ QLEDLKRQLE
EEVKAKNALA HALQSARHDC DLLREQYEEE TEAKAELQRV LSKANSEVAQ WRTKYETDAI
QRTEELEEAK KKLAQRLQEA EEAVEAVNAK CSSLEKTKHR LQNEIEDLMV DVERSNAAAA
ALDKKQRNFD KILAEWKQKY EESQSELESS QKEARSLSTE LFKLKNAYEE SLEHLETFKR
ENKNLQEEIS DLTEQLGSSG KTIHELEKVR KQLEAEKMEL QSALEEAEAS LEHEEGKILR
AQLEFNQIKA EIERKLAEKD EEMEQAKRNH LRVVDSLQTS LDAETRSRNE ALRVKKKMEG
DLNEMEIQLS HANRMAAEAQ KQVKSLQSLL KDTQIQLDDA VRANDDLKEN IAIVERRNNL
LQAELEELRA VVEQTERSRK LAEQELIETS ERVQLLHSQN TSLINQKKKM DADLSQLQTE
VEEAVQECRN AEEKAKKAIT DAAMMAEELK KEQDTSAHLE RMKKNMEQTI KDLQHRLDEA
EQIALKGGKK QLQKLEARVR ELENELEAEQ KRNAESVKGM RKSERRIKEL TYQTEEDRKN
LLRLQDLVDK LQLKVKAYKR QAEEAEEQAN TNLSKFRKVQ HELDEAEERA DIAESQVNKL
RAKSRDIGTK GLNEE
```

MARCOIL output

```
PREDICTED COILED-COIL DOMAINS: OVERVIEW
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
NUMBER PREDICTED COILED-COIL DOMAINS WITH THRESHOLD 1.0 : 1

  1. from 1 to 1111 (length = 1111) with max = 95647400.0

NUMBER PREDICTED COILED-COIL DOMAINS WITH THRESHOLD 10.0 : 1

  1. from 2 to 1111 (length = 1110) with max = 95647400.0

NUMBER PREDICTED COILED-COIL DOMAINS WITH THRESHOLD 50.0 : 2

  1. from 2 to 321 (length = 320) with max = 100.0
  2. from 331 to 1094 (length = 764) with max = 100.0

NUMBER PREDICTED COILED-COIL DOMAINS WITH THRESHOLD 90.0 : 2

  1. from 3 to 319 (length = 317) with max = 100.0

  2. from 354 to 1093 (length = 740) with max = 100.0

NUMBER PREDICTED COILED-COIL DOMAINS WITH THRESHOLD 99.0 : 6

  1. from 9 to 316 (length = 308) with max = 100.0

  2. from 356 to 597 (length = 242) with max = 100.0

  3. from 613 to 737 (length = 125) with max = 100.0

  4. from 776 to 874 (length = 99) with max = 100.0

  5. from 879 to 1013 (length = 135) with max = 100.0

  6. from 1018 to 1093 (length = 76) with max = 100.0

COILED-COIL PROBABILITY PER RESIDUE, COMPACT REPRESENTATION


Using a calibration from one of the implementations of the Lupas et al. method
Using: cc: Normalpdf(1.63, 0.22, score)   others: Normalpdf(0.77, 0.20, score)
```

```
>8651217   ## 1
PLLKSAEREK EMASMKEEFT RLKEALEKSE ARRKELEEKM VSLLQEKNDL QLQVQAEQDN
LADAEERCDQ LIKNKIQLEA KVKEMNERLE DEEEMNAELT AKKRKLEDEC SELKRDIDDL
ELTLAKVEKE KHATENKVKN LTEEMAGLDE IIAKLTKEKK ALQEAHQQAL DDLQAEEDKV
NTLTKAKVKL EQQVDDLEGS LEQEKKVRMD LERAKRKLEG DLKLTQESIM DLENDKQQLD
ERLKKKDFEL NALNARIEDE QALGSQLQKK LKELQARIEE LEEELEAERT ARAKVEKLRS
DLSRELEEIS ERLEEAGGAT SVQIEMNKKR EAEFQKMRRD LEEATLQHEA TAAALRKKHA
DSVAELGEQI DNLQRVKQKL EKEKSEFKLE LDDVTSNMEQ IIKAKANLEK MCRTLEDQMN
EHRSKAEETQ RSVNDLTSQR AKLQTENGEL SRQLDEKEAL ISQLTRGKLT YTQQLEDLKR
QLEEEVKAKN ALAHALQSAR HDCDLLREQY EEETEAKAEL QRVLSKANSE VAQWRTKYET
DAIQRTEELE EAKKKLAQRL QEAEEAVEAV NAKCSSLEKT KHRLQNEIED LMVDVERSNA
AAAALDKKQR NFDKILAEWK QKYEESQSEL ESSQKEARSL STELFKLKNA YEESLEHLET
FKRENKNLQE EISDLTEQLG SSGKTIHELE KVRKQLEAEK MELQSALEEA EASLEHEEGK
ILRAQLEFNQ IKAEIERKLA EKDEEMEQAK RNHLRVVDSL QTSLDAETRS RNEALRVKKK
MEGDLNEMEI QLSHANRMAA EAQKQVKSLQ SLLKDTQIQL DDAVRANDDL KENIAIVERR
NNLLQAELEE LRAVVEQTER SRKLAEQELI ETSERVQLLH SQNTSLINQK KKMDADLSQL
QTEVEEAVQE CRNAEEKAKK AITDAAMMAE ELKKEQDTSA HLERMKKNME QTIKDLQHRL
DEAEQIALKG GKKQLQKLEA RVRELENELE AEQKRNAESV KGMRKSERRI KELTYQTEED
RKNLLRLQDL VDKLQLKVKA YKRQAEEAEE QANTNLSKFR KVQHELDEAE ERADIAESQV
NKLRAKSRDI GTKGLNEE*
```

```
coiled-coil probability in percent and heptad position with highest probability
   1P  5.3  1d    2L 57.7  1e    3L 93.7  1f    4K 97.6  1g    5S 97.6  2a    6A 97.9  1b    7E 98.9
1c
   8R 98.9  2d    9E 100   1e   10K 100   2f   11E 100   1g   12M 100   2a   13A 100   3b   14S 100
4c
  15M 100   5d   16K 100   6e   17E 100   7f   18E 100   8g   19F 100   9a   20T 100  10b   21R 100
11c
  22L 100   1d   23K 100   1e   24E 100   1f   25A 100   2g   26L 100   3a   27E 100   4b   28K 100
5c
  29S 100   6d   30E 100   7e   31A 100   8f   32R 100   9g   33R 100  10a   34K 100  11b   35E 100
12c
  36L 100  13d   37E 100  14e   38E 100  15f   39K 100  16g   40M 100  17a   41V 100  18b   42S 100
19c
  43L 100  20d   44L 100  21e   45Q 100  22f   46E 100  23g   47K 100  24a   48N 100  25b   49D 100
26c
  50L 100  27d   51Q 100  28e   52L 100  19f   53Q 100  20g   54V 100  21a   55Q 100  22b   56A 100
23c
  57E 100  24d   58Q 100  25e   59D 100  26f   60N 100  27g   61L 100  28a   62A 100   3b   63D 100
1c
  64A 100   2d   65E 100   3e   66E 100   4f   67R 100   5g   68C 100   6a   69D 100   7b   70Q 100
8c
  71L 100   9d   72I 100  10e   73K 100  11f   74N 100  12g   75K 100  13a   76I 100  14b   77Q 100
1c
  78L 100   2d   79E 100   3e   80A 100   4f   81K 100   1g   82V 100   1a   83K 100   2b   84E 100
3c
  85M 100   4d   86N 100   5e   87E 100   6f   88R 100   7g   89L 100   8a   90E 100   9b   91D 100
10c
  92E 100  11d   93E 100  12e   94E 100  13f   95M 100  14g   96N 100  15a   97A 100  16b   98E 100
17c
  99L 100  18d  100T 100  19e  101A 100  20f  102K 100  21g  103K 100  22a  104R 100  23b  105K 100
24c
 106L 100  25d  107E 100  26e  108D 100  27f  109E 100  28g  110C 100   8a  111S 100   9b  112E 100
10c
 113L 100  11d  114K 100  12e  115R 100  13f  116D 100  14g  117I 100  15a  118D 100  16b  119D 100
17c
 120L 100  18d  121E 100  19e  122L 100  20f  123T 100  21g  124L 100  22a  125A 100  23b  126K 100
24c
 127V 100  25d  128E 100  26e  129K 100  27f  130E 100  28g  131K 100  28a  132H 100  21b  133A 100
22c
 134T 100  23d  135E 100  24e  136N 100  25f  137K 100  26g  138V 100  27a  139K 100  28b  140N 100
6c
 141L 100   7d  142T 100   8e  143E 100   9f  144E 100  10g  145M 100  11a  146A 100  12b  147G 100
13c
 148L 100  14d  149D 100  15e  150E 100  16f  151I 100  17g  152I 100  18a  153A 100  19b  154K 100
20c
 155L 100  21d  156T 100  22e  157K 100  23f  158E 100  24g  159K 100  25a  160K 100  26b  161A 100
27c
 162L 100  28d  163Q 100  27e  164E 100  28f  165A 100  28g  166H 100  15a  167Q 100  16b  168Q 100
17c
 169A 100  18d  170L 100  19e  171D 100   1f  172D 100   2g  173L 100   3a  174Q 100   4b  175A 100
5c
```

```
 176E 100   6d  177E 100   7e  178D 100   8f  179K 100   9g  180V 100  10a  181N 100  11b  182T 100
12c
 183L 100  13d  184T 100  14e  185K 100  15f  186A 100  16g  187K 100  17a  188V 100  18b  189K 100
19c
 190L 100  20d  191E 100  21e  192Q 100  22f  193Q 100  23g  194V 100  24a  195D 100  25b  196D 100
26c
 197L 100  27d  198E 100  28e  199G 100  24f  200S 100  25g  201L 100  26a  202E 100  27b  203Q 100
28c
 204E 100  28d  205K 100  27e  206K 100  28f  207V 100  17g  208R 100  18a  209M 100  19b  210D 100
1c
 211L 100   2d  212E 100   3e  213R 100   4f  214A 100   5g  215K 100   6a  216R 100   7b  217K 100
1c
 218L 100   2d  219E 100   3e  220G 100   4f  221D 100   5g  222L 100   6a  223K 100   7b  224L 100
8c
 225T 100   9d  226Q 100  10e  227E 100  11f  228S 100  12g  229I 100  13a  230M 100  14b  231D 100
15c
 232L 100  16d  233E 100  17e  234N 100  18f  235D 100  19g  236K 100  20a  237Q 100  21b  238Q 100
22c
 239L 100  23d  240D 100  24e  241E 100  25f  242R 100  26g  243L 100  27a  244K 100  28b  245K 100
28c
 246K 100  16d  247D 100  17e  248F 100  18f  249E 100   1g  250L 100   2a  251N 100   3b  252A 100
1c
 253L 100   1d  254N 100   2e  255A 100   1f  256R 100   1g  257I 100   1a  258E 100   1b  259D 100
2c
 260E 100   3d  261Q 100   4e  262A 100   5f  263L 100   6g  264G 100   7a  265S 100   1b  266Q 100
1c
 267L 100   2d  268Q 100   3e  269K 100   4f  270K 100   5g  271L 100   6a  272K 100   7b  273E 100
8c
 274L 100   9d  275Q 100  10e  276A 100  11f  277R 100  12g  278I 100  13a  279E 100  14b  280E 100
15c
 281L 100  16d  282E 100  17e  283E 100  18f  284E 100  19g  285L 100  20a  286E 100  21b  287A 100
22c
 288E 100  23d  289R 100  24e  290T 100  25f  291A 100  26g  292R 100  27a  293A 100  28b  294K 100
28c
 295V 100  27d  296E 100  28e  297K 100  28f  298L 100  28g  299R 100  28a  300S 100  28b  301D 100
28c
 302L 100  28d  303S 100  25e  304R 100  26f  305E 100  27g  306L 100  28a  307E 100  27b  308E 100
28c
 309I 99.9 28d  310S 99.5 22e  311E 99.5 23f  312R 99.5 24g  313L 99.5 25a  314E 99.5 26b  315E 99.5
27c
 316A 99.5 28d  317G 96.1 28e  318G 92.0 27f  319A 92.0 28g  320T 73.1 28a  321S 50.5 28b  322V 19.4
28c
 323Q 28.0  1d  324I 28.6  1e  325E 28.6  2f  326M 28.6  3g  327N 28.6  4a  328K 41.8  1b  329K 41.8
2c
 330R 41.8  3d  331E 65.8  1e  332A 65.8  2f  333E 65.8  3g  334F 65.8  4a  335Q 65.8  5b  336K 65.8
6c
 337M 65.8  7d  338R 65.8  8e  339R 65.8  9f  340D 65.8 10g  341L 65.8 11a  342E 65.8 12b  343E 65.8
13c
 344A 65.8 14d  345T 65.8 15e  346L 65.8 16f  347Q 65.8 17g  348H 65.8 18a  349E 65.8 19b  350A 65.8
20c
 351T 65.8 21d  352A 65.8 22e  353A 73.0  1b  354A 95.5  1f  355L 97.2  1g  356R 99.7  1a  357K 99.9
1b
 358K 99.9  2c  359H 99.9  3d  360A 99.9  4e  361D 100   1f  362S 100   2g  363V 100   1a  364A 100
1b
 365E 100   1c  366L 100   2d  367G 100   3e  368E 100   4f  369Q 100   5g  370I 100   6a  371D 100
7b
 372N 100   8c  373L 100   9d  374Q 100  10e  375R 100  11f  376V 100  12g  377K 100  13a  378Q 100
14b
 379K 100  15c  380L 100  16d  381E 100  17e  382K 100  18f  383E 100  19g  384K 100  20a  385S 100
21b
 386E 100  22c  387F 100  23d  388K 100  24e  389L 100  25f  390E 100  26g  391L 100  27a  392D 100
28b
 393D 100  28c  394V 100  27d  395T 100  28e  396S 100  24f  397N 100  25g  398M 100  26a  399E 100
27b
 400Q 100  28c  401I 99.9 12d  402I 99.9 13e  403K 100   1f  404A 100   2g  405K 100   3a  406A 100
4b
 407N 100   5c  408L 100   1d  409E 100   2e  410K 100   3f  411M 100   4g  412C 100   5a  413R 100
6b
 414T 100   7c  415L 100   8d  416E 100   9e  417D 100  10f  418Q 100  11g  419M 100  12a  420N 100
13b
 421E 100  14c  422H 100  15d  423R 100  16e  424S 100  17f  425K 100  18g  426A 100  19a  427E 100
20b
 428E 100  21c  429T 100  22d  430Q 100  23e  431R 100  24f  432S 100  25g  433V 100  26a  434N 100
27b
 435D 100  28c  436L 100  28d  437T 99.9 22e  438S 99.9 23f  439Q 99.9 24g  440R 99.9 25a  441A 99.9
26b
```

```
 442K 99.9 27c  443L 99.9 28d  444Q 99.8 28e  445T 99.7 19f  446E 99.7 20g  447N 99.7 21a  448G 99.7
22b
 449E 99.7 23c  450L 99.7 24d  451S 99.7 25e  452R 99.7 26f  453Q 99.7 27g  454L 99.7 28a  455D 99.5
27b
 456E 99.5 28c  457K 99.2 27d  458E 99.4  1e  459A 99.4  2f  460L 99.4  3g  461I 99.4  4a  462S 99.4
5b
 463Q 99.4  6c  464L 99.4  7d  465T 99.4  8e  466R 99.4  9f  467G 99.4 10g  468K 99.4 11a  469L 99.4
12b
 470T 99.5  1c  471Y 99.6  1d  472T 99.8  1e  473Q 99.9  1f  474Q 99.9  2g  475L 99.9  3a  476E 99.9
4b
 477D 99.9  5c  478L 99.9  6d  479K 99.9  7e  480R 99.9  8f  481Q 99.9  9g  482L 99.9 10a  483E 99.9
11b
 484E 99.9 12c  485E 99.9 13d  486V 99.9 14e  487K 99.9 15f  488A 99.9 16g  489K 99.9 17a  490N 99.9
18b
 491A 99.9 19c  492L 99.9 20d  493A 99.9 21e  494H 99.9 22f  495A 99.9 23g  496L 99.9 24a  497Q 99.9
25b
 498S 99.9 26c  499A 99.9 27d  500R 99.9 28e  501H 99.9  7f  502D 99.9  8g  503C 99.9  9a  504D 99.9
1b
 505L 99.9  2c  506L 100   1d  507R 100   2e  508E 100   3f  509Q 100   4g  510Y 100   5a  511E 100
6b
 512E 100   7c  513E 100   8d  514T 100   9e  515E 100  10f  516A 100  11g  517K 100  12a  518A 100
13b
 519E 100  14c  520L 100  15d  521Q 100  16e  522R 100  17f  523V 100  18g  524L 100  19a  525S 100
20b
 526K 100  21c  527A 100  22d  528N 100  23e  529S 100  24f  530E 100  25g  531V 100  26a  532A 100
27b
 533Q 100  28c  534W 99.7 28d  535R 99.7 28e  536T 99.0 27f  537K 99.1  1f  538Y 99.3  1g  539E 99.9
1a
 540T 100   1b  541D 100   1c  542A 100   2d  543I 100   3e  544Q 100   4f  545R 100   5g  546T 100
6a
 547E 100   1b  548E 100   2c  549L 100   3d  550E 100   4e  551E 100   5f  552A 100   6g  553K 100
7a
 554K 100   8b  555K 100   9c  556L 100  10d  557A 100  11e  558Q 100  12f  559R 100  13g  560L 100
14a
 561Q 100  15b  562E 100  16c  563A 100  17d  564E 100  18e  565E 100  19f  566A 100  20g  567V 100
21a
 568E 100  22b  569A 100  23c  570V 100  24d  571N 100  25e  572A 100  26f  573K 100  27g  574C 100
28a
 575S 100  28b  576S 100  13c  577L 100  14d  578E 100  15e  579K 100  16f  580T 100  17g  581K 100
18a
 582H 100  19b  583R 100  20c  584L 100  21d  585Q 100  22e  586N 100  23f  587E 100  24g  588I 100
25a
 589E 100  26b  590D 100  27c  591L 100  28d  592M 99.9 28e  593V 99.3 24f  594D 99.3 25g  595V 99.3
26a
 596E 99.3 27b  597R 99.3 28c  598S 98.6 22d  599N 98.6 23e  600A 98.6 24f  601A 98.6 25g  602A 98.6
26a
 603A 98.6 27b  604A 98.6 28c  605L 98.6 28d  606D 97.0 26e  607K 97.0 27f  608K 97.0 28g  609Q 95.6
26a
 610R 97.8  1b  611N 97.8  2c  612F 97.8  3d  613D 99.6  1e  614K 99.6  2f  615I 99.6  3g  616L 99.9
1a
 617A 99.9  2b  618E 99.9  3c  619W 99.9  4d  620K 99.9  5e  621Q 99.9  6f  622K 99.9  7g  623Y 99.9
8a
 624E 99.9  9b  625E 99.9 10c  626S 99.9 11d  627Q 99.9 12e  628S 99.9 13f  629E 99.9 14g  630L 99.9
15a
 631E 99.9 16b  632S 99.9 17c  633S 99.9 18d  634Q 99.9 19e  635K 99.9 20f  636E 99.9 21g  637A 99.9
22a
 638R 99.9 23b  639S 99.9 24c  640L 99.9 25d  641S 99.9 26e  642T 99.9 27f  643E 99.9 28g  644L 99.9
28a
 645F 99.5 24b  646K 99.9  1c  647L 99.9  1d  648K 99.9  2e  649N 99.9  3f  650A 99.9  4g  651Y 99.9
1a
 652E 100   1b  653E 100   2c  654S 100   3d  655L 100   4e  656E 100   5f  657H 100   6g  658L 100
7a
 659E 100   8b  660T 100   9c  661F 100  10d  662K 100  11e  663R 100  12f  664E 100  13g  665N 100
14a
 666K 100  15b  667N 100  16c  668L 100  17d  669Q 100  18e  670E 100  19f  671E 100  20g  672I 100
21a
 673S 100  22b  674D 100  23c  675L 100  24d  676T 100  25e  677E 100  26f  678Q 100  27g  679L 100
28a
 680G 99.4 17b  681S 99.4 18c  682S 99.4 19d  683G 99.9  1e  684K 100   1f  685T 100   2g  686I 100
3a
 687H 100   4b  688E 100   1c  689L 100   2d  690E 100   3e  691K 100   4f  692V 100   5g  693R 100
6a
 694K 100   7b  695Q 100   8c  696L 100   9d  697E 100  10e  698A 100  11f  699E 100  12g  700K 100
13a
 701M 100  14b  702E 100  15c  703L 100  16d  704Q 100  17e  705S 100  18f  706A 100  19g  707L 100
20a
```

```
 708E 100  21b  709E 100  22c  710A 100  23d  711E 100  24e  712A 100  25f  713S 100  26g  714L 100
27a
 715E 100  28b  716H 100  23c  717E 100  24d  718E 100  25e  719G 100  26f  720K 100  27g  721I 100
28a
 722L 100  28b  723R 100  28c  724A 100  28d  725Q 100  28e  726L 100  25f  727E 100  26g  728F 100
27a
 729N 100  28b  730Q 99.9 28c  731I 99.9 25d  732K 99.9 26e  733A 99.9 27f  734E 99.9 28g  735I 99.8
27a
 736E 99.8 28b  737R 99.4 28c  738K 98.2 13g  739L 98.2 14a  740A 98.2 15b  741E 98.2 16c  742K 98.2
17d
 743D 98.2 18e  744E 98.2 19f  745E 98.2 20g  746M 98.2 21a  747E 98.2 22b  748Q 98.2 23c  749A 98.2
24d
 750K 98.2 25e  751R 98.2 26f  752N 98.2 27g  753H 98.2 28a  754L 98.2 28b  755R 96.5 27c  756V 96.5
28d
 757V 93.6 24e  758D 93.6 25f  759S 93.6 26g  760L 93.6 27a  761Q 93.6 28b  762T 93.4  3f  763S 93.4
4g
 764L 96.5  1a  765D 96.5  2b  766A 96.5  3c  767E 96.5  4d  768T 96.5  5e  769R 96.5  6f  770S 96.5
7g
 771R 96.5  8a  772N 98.2  1b  773E 98.3  1c  774A 98.3  1d  775L 98.3  2e  776R 99.7  1f  777V 99.7
1g
 778K 100   1a  779K 100   1b  780K 100   1c  781M 100   2d  782E 100   1e  783G 100   2f  784D 100
3g
 785L 100   4a  786N 100   5b  787E 100   6c  788M 100   7d  789E 100   8e  790I 100   9f  791Q 100
10g
 792L 100  11a  793S 100  12b  794H 100  13c  795A 100   1d  796N 100   2e  797R 100   3f  798M 100
4g
 799A 100   5a  800A 100   6b  801E 100   7c  802A 100   8d  803Q 100   9e  804K 100  10f  805Q 100
11g
 806V 100  12a  807K 100  13b  808S 100  14c  809L 100  15d  810Q 100  16e  811S 100  17f  812L 100
18g
 813L 100  19a  814K 100  20b  815D 100  21c  816T 100  22d  817Q 100  23e  818I 100  24f  819Q 100
25g
 820L 100  26a  821D 100  27b  822D 100  28c  823A 100  28d  824V 100   1e  825R 100   1f  826A 100
2g
 827N 100   3a  828D 100   4b  829D 100   5c  830L 100   6d  831K 100   7e  832E 100   8f  833N 100
9g
 834I 100  10a  835A 100  11b  836I 100  12c  837V 100  13d  838E 100  14e  839R 100  15f  840R 100
16g
 841N 100  17a  842N 100  18b  843L 100  19c  844L 100  20d  845Q 100  21e  846A 100  22f  847E 100
23g
 848L 100  24a  849E 100  25b  850E 100  26c  851L 100  27d  852R 100  28e  853A 100  25f  854V 100
26g
 855V 100  27a  856E 100  28b  857Q 100  28c  858T 100  17d  859E 100  18e  860R 100  19f  861S 100
20g
 862R 100  21a  863K 100  22b  864L 100  23c  865A 100  24d  866E 100  25e  867Q 100  26f  868E 100
27g
 869L 100  28a  870I 100  27b  871E 100  28c  872T 99.8 28d  873S 99.4 27e  874E 99.4 28f  875R 98.8
28g
 876V 97.5 28a  877Q 95.7 28b  878L 91.2  5c  879L 99.0  1d  880H 99.0  2e  881S 99.0  3f  882Q 99.0
4g
 883N 99.0  5a  884T 99.0  6b  885S 99.5  1c  886L 99.7  1d  887I 99.7  2e  888N 100   1f  889Q 100
1g
 890K 100   1a  891K 100   2b  892K 100   3c  893M 100   4d  894D 100   1e  895A 100   2f  896D 100
3g
 897L 100   4a  898S 100   5b  899Q 100   6c  900L 100   7d  901Q 100   8e  902T 100   9f  903E 100
10g
 904V 100  11a  905E 100  12b  906E 100  13c  907A 100  14d  908V 100  15e  909Q 100  16f  910E 100
17g
 911C 100  18a  912R 100  19b  913N 100  20c  914A 100  21d  915E 100  22e  916E 100  23f  917K 100
24g
 918A 100  25a  919K 100  26b  920K 100  27c  921A 100  28d  922I 100  28e  923T 100  25f  924D 100
26g
 925A 100  27a  926A 100  28b  927M 99.9 28c  928M 99.8  1d  929A 100   1e  930E 100   1f  931E 100
2g
 932L 100   3a  933K 100   4b  934K 100   5c  935E 100   6d  936Q 100   7e  937D 100   1f  938T 100
2g
 939S 100   3a  940A 100   1b  941H 100   2c  942L 100   3d  943E 100   4e  944R 100   5f  945M 100
6g
 946K 100   7a  947K 100   8b  948N 100   9c  949M 100  10d  950E 100  11e  951Q 100  12f  952T 100
13g
 953I 100  14a  954K 100  15b  955D 100  16c  956L 100  17d  957Q 100  18e  958H 100  19f  959R 100
20g
 960L 100  21a  961D 100  22b  962E 100   1b  963A 100   2c  964E 100   3d  965Q 100   4e  966I 100
5f
 967A 100   1g  968L 100   1a  969K 100   2b  970G 100   3c  971G 100   4d  972K 100   1e  973K 100
2f
```

```
 974Q 100   3g  975L 100   4a  976Q 100   5b  977K 100   6c  978L 100   7d  979E 100   8e  980A 100
9f
 981R 100  10g  982V 100  11a  983R 100  12b  984E 100  13c  985L 100  14d  986E 100  15e  987N 100
16f
 988E 100  17g  989L 100  18a  990E 100  19b  991A 100  20c  992E 100  21d  993Q 100  22e  994K 100
23f
 995R 100  24g  996N 100  25a  997A 100  26b  998E 100  27c  999S 100  28d 1000V 100  27e 1001K 100
28f
1002G 100  28g 1003M 100  28a 1004R 99.9 28b 1005K 99.9 28c 1006S 99.9 22d 1007E 99.9 23e 1008R 99.9
24f
1009R 99.9 25g 1010I 99.9 26a 1011K 99.9 27b 1012E 99.9 28c 1013L 99.9 28d 1014T 98.7  8e 1015Y 98.7
9f
1016Q 98.7 10g 1017T 98.7 11a 1018E 99.8  1b 1019E 99.8  2c 1020D 99.8  3d 1021R 100   1e 1022K 100
1f
1023N 100   1g 1024L 100   1a 1025L 100   2b 1026R 100   1c 1027L 100   2d 1028Q 100   3e 1029D 100
4f
1030L 100   5g 1031V 100   6a 1032D 100   7b 1033K 100   8c 1034L 100   9d 1035Q 100  10e 1036L 100
11f
1037K 100  12g 1038V 100  13a 1039K 100  14b 1040A 100  15c 1041Y 100  16d 1042K 100  17e 1043R 100
18f
1044Q 100  19g 1045A 100  20a 1046E 100  21b 1047E 100  22c 1048A 100  23d 1049E 100  24e 1050E 100
25f
1051Q 100  26g 1052A 100  27a 1053N 100  28b 1054T 100  28c 1055N 99.8 12d 1056L 100   1e 1057S 100
1f
1058K 100   1g 1059F 100   2a 1060R 100   3b 1061K 100   4c 1062V 100   5d 1063Q 100   6e 1064H 100
7f
1065E 100   8g 1066L 100   9a 1067D 100  10b 1068E 100  11c 1069A 100  12d 1070E 100  13e 1071E 100
14f
1072R 100  15g 1073A 100  16a 1074D 100  17b 1075I 100  18c 1076A 100  19d 1077E 100  20e 1078S 100
21f
1079Q 100  22g 1080V 100  23a 1081N 100  24b 1082K 100  25c 1083L 100  26d 1084R 100  27e 1085A 100
28f
1086K 100  28g 1087S 100  26a 1088R 100  27b 1089D 100  28c 1090I 100  28d 1091G 99.9 27e 1092T 99.9
28f
1093K 99.9 28g 1094G 86.2 28a 1095L 43.4 28b 1096N 28.1 28c 1097E 13.9 28d 1098E 13.9 28e  *
```

Myosin heavy chain coiled-coil assessment by MARCOIL

```
   840                   850                   860                   870                   880                   890
P L  L K S A E R E K E M  A S M K E E F T R L  K E A L E K S E A R  R K E L E E K M V S  L L Q E K N D L Q L  Q V Q A E Q D
d e  f g a b c d e f g a  b c d e f g a b c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f
   900                   910                   920                   930                   940                   950
L A  D A E E R C D Q L I  K N K I Q L E A K V  K E M N E R L E D E  E E M N A E L T A K  K R K L E D E C S E  L K R D I D D
a b  c d e f g a b c d e  f g a b c d e f g a  b c d e f g a b c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c
   960                   970                   980                   990                  1000                  1010
E L  T L A K V E K E K H  A T E N K V K N L T  E E M A G L D E I I  A K L T K E K K A L  Q E A H Q Q A L D D  L Q A E E D K
e f  g a b c d e f g a b  c d e f g a b c d e  f g a b c d e f g a  b c d e f g a b c d  e f g a b c d e f g  a b c d e f g
  1020                  1030                  1040                  1050                  1060                  1070
N T  L T K A K V K L E Q  Q V D D L E G S L E  Q E K K V R M D L E  R A K R K L E G D L  K L T Q E S I M D L  E N D K Q Q L
b c  d e f g a b c d e f  g a b c d e f g a b  c d e f g a b c d e  f g a b c d e f g a  b c d e f g a b c d  e f g a b c d
  1080                  1090                  1100                  1110                  1120                  1130
E R  L K K K D F E L N A  L N A R I E D E Q A  L G S Q L Q K K L K  E L Q A R I E E L E  E E L E A E R T A R  A K V E K L R
f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f g a b  c d e f g a b c d e  f g a b c d e f g a  b c d e f g a
  1140                  1150                  1160                  1170                  1180                  1190
D L  S R E L E E I S E R  L E E A G G A T S V  Q I E M N K K R E A  E F Q K M R R D L E  E A T L Q H E A T A  A A L R K K H
c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f g a b  c d e f g a b c d e  b f g a b c d
  1200                  1210                  1220                  1230                  1240                  1250
D S  V A E L G E Q I D N  L Q R V K Q K L E K  E K S E F K L E L D  D V T S N M E Q I I  K A K A N L E K M C  R T L E D Q M
f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f g a b  c d e f g a b c d e  f g a b c d e f g a  b c d e f g a
  1260                  1270                  1280                  1290                  1300                  1310
E H  R S K A E E T Q R S  V N D L T S Q R A K  L Q T E N G E L S R  Q L D E K E A L I S  Q L T R G K L T Y T  Q Q L E D L K
c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f g a b  c d e f g a b c d e  f g a b c d e
  1320                  1330                  1340                  1350                  1360                  1370
Q L  E E E V K A K N A L  A H A L Q S A R H D  C D L L R E Q Y E E  E T E A K A E L Q R  V L S K A N S E V A  Q W R T K Y E
g a  b c d e f g a b c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f g a b  c d e f f g a
  1380                  1390                  1400                  1410                  1420                  1430
D A  I Q R T E E L E E A  K K K L A Q R L Q E  A E E A V E A V N A  K C S S L E K T K H  R L Q N E I E D L M  V D V E R S N
c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f g a b  c d e f g a b c d e  f g a b c d e
  1440                  1450                  1460                  1470                  1480                  1490
A A  A A L D K K Q R N F  D K I L A E W K Q K  Y E E S Q S E L E S  S Q K E A R S L S T  E L F K L K N A Y E  E S L E H L E
g a  b c d e f g a b c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f g a b  c d e f g a b
  1500                  1510                  1520                  1530                  1540                  1550
F K  R E N K N L Q E E I  S D L T E Q L G S S  G K T I H E L E K V  R K Q L E A E K M E  L Q S A L E E A E A  S L E H E E G
d e  f g a b c d e f g a  b c d e f g a b c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f
  1560                  1570                  1580                  1590                  1600                  1610
I L  R A Q L E F N Q I K  A E I E R K L A E K  D E E M E Q A K R N  H L R V V D S L Q T  S L D A E T R S R N  E A L R V K K
a b  c d e f g a b c d e  f g a b c g a b c d  e f g a b c d e f g  a b c d e f g a b f  g a b c d e f g a b  c d e f g a b
  1620                  1630                  1640                  1650                  1660                  1670
M E  G D L N E M E I Q L  S H A N R M A A E A  Q K Q V K S L Q S L  L K D T Q I Q L D D  A V R A N D D L K E  N I A I V E R
d e  f g a b c d e f g a  b c d e f g a b c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c d e f  g a b c d e f
  1680                  1690                  1700                  1710                  1720                  1730
N N  L L Q A E L E E L R  A V V E Q T E R S R  K L A E Q E L I E T  S E R V Q L L H S Q  N T S L I N Q K K K  M D A D L S Q
a b  c d e f g a b c d e  f g a b c d e f g a  b c d e f g a b c d  e f g a b c d e f g  a b c d e f g a b c  d e f g a b c
  1740                  1750                  1760                  1770                  1780                  1790
Q T  E V E E A V Q E C R  N A E E K A K K A I  T D A A M M A E E L  K K E Q D T S A H L  E R M K K N M E Q T  I K D L
e f  g a b c d e f g a b  c d e f g a b c d e  f g a b c d e f g a  b c d e f g a b c d  e f g a b c d e f g  a b c d
```

C.2 Twister analysis of myosin heavy chain available crystal structures

MYH7$^{850-961}$

```
*********************************************************************
* TWISTER: Analysis of coiled-coil geometry   (c) S.Strelkov 2000-8 *
* E-mail: sergei.strelkov@pharm.kuleuven.be        Version 130508   *
* Reference: Strelkov,S.V., and Burkhard,P. (2002) J. Struct. Biol. *
*********************************************************************

Title 2fxm

Input pdb_file 2fxm.pdb
2-stranded coiled coil, chains A B
First_res 850,  last_res 961
Output pdb_file 2fxm_axis.pdb

********* Sequence and heptad assignment *********

850        860        870        880        890        900
ASMKEEFTRL KEALEKSEAR RKELEEKMVS LLQEKNDLQL QVQAEQDNLA DAEERCDQLI
bcdefgabcd efgabcdefg abcdefgabc defgabcdef gabcdefgab cdefgabcde

 910        920        930        940        950        960
KNKIQLEAKV KEMNERLEDE EEMNAELTAK KRKLEDECSE LKRDIDDLEL TL
fgabcdefga bcdefgabcd efgabcdefg abcdefgabc defgabcdef gz
```

```
********* Coiled-coil parameters per residue *********

Res cc_phase cc_rad cc_rise cc_pit  cc_dang  pos  Cr_ang  a_rad a_ris res/tur a_dang
axis_cur

851   -9.13   7.38   1.46    57.6    -9.13 S c  162.7   1.63   2.05   3.61   99.80
0.0000
852  -15.07   6.48   1.54    93.2    -5.94 M d  -24.3   2.31   1.85   3.60  100.04
0.0000
853  -16.54   6.25   1.50   368.7    -1.47 K e   79.2   2.28   1.53   3.65   98.62
0.0000
854  -18.12   6.04   1.48   336.1    -1.58 E f  167.1   2.30   1.50   3.64   99.02
0.0000
855  -19.82   5.80   1.45   306.1    -1.70 E g  -81.1   2.32   1.48   3.62   99.52
0.0000
856  -21.83   5.59   1.43   256.3    -2.01 F a   19.6   2.26   1.46   3.63   99.23
0.0205
857  -24.37   5.39   1.50   212.8    -2.54 T b  121.7   2.31   1.53   3.56  101.02
0.0051
858  -27.20   5.20   1.47   187.4    -2.83 R c -134.0   2.28   1.51   3.63   99.26
0.0058
859  -30.12   5.02   1.50   185.2    -2.92 L d  -34.2   2.29   1.53   3.66   98.40
0.0070
860  -33.00   4.87   1.48   185.1    -2.88 K e   68.6   2.33   1.51   3.65   98.71
0.0068
```

```
861  -36.06   4.78   1.49   175.6    -3.06 E f  166.8   2.25   1.52   3.60  100.05
0.0056
862  -39.19   4.71   1.47   169.6    -3.13 A g  -85.9   2.36   1.50   3.60   99.91
0.0041
863  -42.43   4.68   1.45   161.5    -3.24 L a   15.2   2.23   1.48   3.67   98.14
0.0028
864  -45.95   4.67   1.47   149.8    -3.52 E b  116.8   2.35   1.50   3.66   98.41
0.0038
865  -49.42   4.62   1.44   149.0    -3.47 K c -141.2   2.29   1.46   3.67   97.97
0.0044
866  -52.83   4.56   1.45   153.2    -3.42 S d  -40.5   2.30   1.48   3.65   98.58
0.0074
867  -56.65   4.52   1.49   141.2    -3.81 E e   62.5   2.30   1.53   3.60   99.95
0.0095
868  -60.85   4.57   1.48   126.7    -4.20 A f  166.3   2.23   1.52   3.60   99.87
0.0105
869  -65.24   4.63   1.47   120.6    -4.39 R g  -89.7   2.34   1.51   3.59  100.31
0.0112
870  -69.91   4.72   1.50   115.3    -4.67 R a   16.5   2.23   1.55   3.59  100.34
0.0116
871  -75.04   4.83   1.49   104.8    -5.14 K b  120.0   2.27   1.56   3.57  100.79
0.0101
872  -80.15   4.92   1.43   100.6    -5.11 E c -132.0   2.34   1.50   3.63   99.22
0.0105
873  -85.11   5.03   1.46   106.3    -4.96 L d  -33.5   2.24   1.53   3.64   98.89
0.0136
874  -89.50   5.21   1.45   119.1    -4.39 E e   73.8   2.38   1.51   3.65   98.59
0.0194
875  -92.93   5.29   1.42   148.7    -3.43 E f  172.5   2.30   1.45   3.72   96.82
0.0239
876  -96.36   5.39   1.56   163.7    -3.43 K g  -85.7   2.31   1.59   3.64   98.98
0.0270
877 -100.12   5.41   1.62   154.6    -3.77 M a   17.8   2.20   1.65   3.58  100.71
0.0283
878 -103.87   5.42   1.49   143.5    -3.74 V b  122.4   2.33   1.53   3.63   99.12
0.0293
879 -107.50   5.34   1.48   146.1    -3.64 S c -137.2   2.31   1.52   3.68   97.82
0.0260
880 -111.37   5.18   1.56   145.3    -3.86 L d  -36.4   2.24   1.60   3.64   99.03
0.0215
881 -115.44   5.04   1.49   131.9    -4.07 L e   67.3   2.28   1.54   3.59  100.15
0.0158
882 -119.85   4.85   1.44   117.3    -4.42 Q f  171.0   2.32   1.49   3.60  100.00
0.0103
883 -124.81   4.70   1.43   104.1    -4.95 E g  -84.3   2.32   1.49   3.69   97.51
0.0047
884 -129.95   4.59   1.44   100.9    -5.14 K a   17.1   2.31   1.50   3.72   96.68
0.0032
885 -135.08   4.52   1.47   103.4    -5.13 N b  119.0   2.34   1.53   3.65   98.72
0.0077
886 -139.73   4.54   1.48   114.6    -4.65 D c -135.8   2.27   1.52   3.62   99.35
0.0111
887 -144.22   4.52   1.52   121.8    -4.49 L d  -33.6   2.24   1.56   3.55  101.29
0.0136
888 -148.62   4.51   1.49   121.7    -4.40 Q e   75.1   2.30   1.53   3.55  101.41
0.0151
889 -152.43   4.48   1.48   139.8    -3.81 L f  176.8   2.25   1.51   3.60   99.95
0.0156
890 -155.80   4.46   1.55   165.2    -3.38 Q g  -78.0   2.24   1.57   3.56  101.10
0.0141
891 -158.34   4.51   1.47   208.4    -2.53 V a   26.2   2.31   1.48   3.59  100.31
0.0126
892 -160.37   4.53   1.47   259.4    -2.03 Q b  127.6   2.32   1.47   3.65   98.61
0.0097
893 -162.31   4.55   1.47   273.1    -1.94 A c -132.6   2.31   1.48   3.71   97.17
0.0063
```

```
894 -164.43    4.54    1.49    252.0     -2.12 E d  -34.5    2.28    1.50    3.61   99.64
0.0036
895 -167.00    4.50    1.50    210.5     -2.57 Q e   70.6    2.30    1.52    3.52  102.23
0.0043
896 -169.58    4.50    1.48    206.4     -2.58 D f  174.9    2.24    1.49    3.58  100.59
0.0048
897 -172.09    4.51    1.49    214.8     -2.50 N g  -83.1    2.26    1.51    3.60   99.94
0.0069
898 -174.47    4.56    1.51    227.7     -2.38 L a   20.0    2.28    1.52    3.57  100.93
0.0092
899 -176.76    4.60    1.44    226.2     -2.29 A b  123.2    2.30    1.45    3.62   99.40
0.0102
900 -179.19    4.59    1.43    211.6     -2.43 D c -137.0    2.33    1.44    3.65   98.76
0.0094
901 -181.97    4.56    1.48    191.1     -2.78 A d  -34.6    2.30    1.49    3.61   99.65
0.0072
902 -185.33    4.51    1.51    161.8     -3.35 E e   67.8    2.27    1.53    3.58  100.47
0.0064
903 -188.84    4.57    1.54    157.5     -3.52 E f  172.7    2.23    1.57    3.55  101.53
0.0055
904 -192.15    4.64    1.53    166.4     -3.30 R g  -82.3    2.26    1.55    3.51  102.57
0.0051
905 -195.36    4.75    1.56    175.1     -3.21 C a   24.7    2.27    1.59    3.53  102.13
0.0069
906 -198.62    4.87    1.52    167.9     -3.26 D b  127.4    2.24    1.54    3.64   99.02
0.0101
907 -202.09    4.91    1.46    151.1     -3.48 Q c -131.8    2.35    1.49    3.65   98.59
0.0124
908 -206.00    5.00    1.48    136.8     -3.91 L d  -28.8    2.28    1.52    3.62   99.39
0.0138
909 -209.89    5.07    1.44    133.1     -3.88 I e   74.4    2.28    1.48    3.62   99.34
0.0155
910 -213.26    5.10    1.40    149.0     -3.38 K f  177.0    2.36    1.43    3.63   99.24
0.0157
911 -216.52    5.13    1.47    162.0     -3.26 N g  -80.6    2.23    1.49    3.59  100.31
0.0167
912 -219.83    5.12    1.52    164.7     -3.31 K a   24.5    2.33    1.54    3.56  101.02
0.0173
913 -223.34    5.11    1.49    152.5     -3.51 I b  127.5    2.25    1.52    3.64   99.01
0.0165
914 -227.11    5.06    1.47    140.3     -3.77 Q c -131.2    2.40    1.51    3.71   97.11
0.0127
915 -231.29    4.99    1.49    128.1     -4.18 L d  -31.1    2.21    1.53    3.68   97.71
0.0091
916 -235.85    4.92    1.47    116.4     -4.56 E e   72.2    2.35    1.53    3.59  100.21
0.0046
917 -240.09    4.87    1.46    124.0     -4.24 A f  178.4    2.25    1.50    3.58  100.62
0.0042
918 -244.00    4.85    1.48    136.5     -3.91 K g  -78.3    2.28    1.52    3.60   99.86
0.0064
919 -247.61    4.92    1.48    147.6     -3.60 V a   25.6    2.31    1.51    3.63   99.24
0.0082
920 -251.00    4.97    1.46    155.0     -3.39 K b  127.3    2.30    1.49    3.66   98.46
0.0103
921 -254.22    5.03    1.47    164.3     -3.23 E c -130.8    2.33    1.50    3.67   97.99
0.0127
922 -257.47    5.04    1.53    169.8     -3.25 M d  -30.6    2.22    1.56    3.64   98.89
0.0135
923 -260.77    5.05    1.51    165.0     -3.30 N e   73.2    2.32    1.54    3.62   99.36
0.0155
924 -263.66    5.03    1.45    180.9     -2.89 E f  175.0    2.27    1.47    3.66   98.40
0.0143
925 -266.49    5.01    1.52    193.2     -2.83 R g  -84.8    2.28    1.54    3.60  100.05
0.0129
926 -269.42    4.95    1.53    188.6     -2.93 L a   20.0    2.24    1.56    3.57  100.94
0.0100
```

```
927 -272.53   4.89   1.49   172.5    -3.11 E b  122.7   2.28   1.51   3.60  100.08
0.0077
928 -276.05   4.79   1.43   146.6    -3.52 D c -133.9   2.31   1.47   3.65   98.76
0.0040
929 -279.89   4.74   1.43   134.2    -3.84 E d  -32.8   2.32   1.47   3.66   98.39
0.0010
930 -283.79   4.72   1.43   131.7    -3.90 E e   70.3   2.33   1.46   3.67   98.14
0.0019
931 -287.49   4.74   1.45   140.9    -3.70 E f  169.2   2.29   1.47   3.66   98.32
0.0049
932 -290.85   4.77   1.50   160.2    -3.36 M g  -87.0   2.28   1.52   3.60   99.89
0.0061
933 -294.02   4.80   1.49   169.0    -3.17 N a   16.2   2.28   1.51   3.59  100.16
0.0089
934 -297.21   4.83   1.49   167.8    -3.19 A b  118.9   2.28   1.51   3.62   99.33
0.0110
935 -300.33   4.83   1.46   168.9    -3.12 E c -139.5   2.31   1.49   3.65   98.55
0.0112
936 -303.61   4.82   1.51   166.0    -3.27 L d  -38.0   2.26   1.54   3.59  100.36
0.0104
937 -306.99   4.81   1.52   162.2    -3.38 T e   67.5   2.31   1.55   3.57  100.96
0.0091
938 -310.14   4.75   1.45   165.6    -3.14 A f  170.5   2.30   1.47   3.67   98.16
0.0066
939 -313.38   4.69   1.48   164.5    -3.25 K g  -90.1   2.29   1.51   3.65   98.70
0.0062
940 -317.05   4.68   1.52   148.7    -3.67 K a   14.4   2.35   1.54   3.62   99.39
0.0072
941 -320.92   4.69   1.51   140.3    -3.87 R b  115.7   2.21   1.54   3.66   98.45
0.0069
942 -324.71   4.73   1.49   141.7    -3.79 K c -141.5   2.34   1.52   3.59  100.32
0.0076
943 -329.01   4.80   1.49   124.7    -4.30 L d  -36.5   2.25   1.53   3.53  102.11
0.0084
944 -333.71   4.83   1.50   115.1    -4.71 E e   70.9   2.24   1.56   3.52  102.15
0.0083
945 -338.08   4.83   1.49   123.0    -4.37 D f  176.5   2.30   1.54   3.56  101.04
0.0100
946 -342.11   4.85   1.50   133.8    -4.03 E g  -78.6   2.20   1.54   3.58  100.51
0.0120
947 -345.43   4.94   1.49   161.5    -3.32 C a   26.2   2.36   1.52   3.54  101.60
0.0150
948 -348.17   4.99   1.42   186.5    -2.74 S b  130.9   2.32   1.44   3.68   97.92
0.0149
949 -350.95   5.00   1.45   188.2    -2.78 E c -133.1   2.35   1.47   3.76   95.68
0.0129
950 -353.75   4.96   1.53   196.2    -2.80 L d  -32.3   2.26   1.55   3.63   99.29
0.0094
951 -356.61   4.92   1.50   189.0    -2.86 K e   70.8   2.36   1.52   3.58  100.65
0.0060
952 -359.21   4.92   1.49   206.4    -2.60 R f  173.9   2.22   1.51   3.63   99.07
0.0045
953 -361.44   4.93   1.50   241.7    -2.23 D g  -86.0   2.28   1.51   3.60   99.90
0.0048
954 -363.76   5.01   1.50   233.2    -2.32 I a   18.7   2.31   1.51   3.54  101.64
0.0061
955 -366.14   5.08   1.51   228.2    -2.38 D b  121.7   2.24   1.53   3.61   99.63
0.0060
956 -368.54   5.17   1.52   228.1    -2.40 D c -137.3   2.28   1.54   3.65   98.70
0.0000
957 -371.23   5.26   1.51   201.1    -2.70 L d  -36.7   2.27   1.53   3.58  100.64
0.0000
958 -374.27   5.33   1.51   179.5    -3.03 E e   69.0   2.25   1.54   3.52  102.33
0.0000
959 -377.26   5.46   1.52   182.3    -2.99 L f  172.3   2.22   1.54   3.50  103.02
0.0000
```

```
960 -380.10   5.61   1.54   195.0    -2.84 T g  -79.6   2.22   1.56   3.48  103.54
0.0000

Ave           4.94   1.48   154.7    -3.46               2.28   1.52   3.61   99.63
0.0103
Std           0.43   0.04    44.8     1.00               0.08   0.07   0.05    1.38
0.0057

Overall ccpitch is estimated as 2Pi*<cc_rise>/<cc_angle_per_res>

Overall roc of the a-helix axis (1/<curv>)   97.46 A
with curvature smoothing base 5


********* Crick angles for heptad position a *********

Res    Crick_angle

856       19.57
863       15.22
870       16.54
877       17.79
884       17.13
891       26.16
898       20.05
905       24.75
912       24.50
919       25.63
926       20.04
933       16.22
940       14.39
947       26.24
954       18.74

Ave       20.20
Std        4.05





********* Crick angles for heptad position d *********

Res    Crick_angle

852      -24.29
859      -34.18
866      -40.54
873      -33.55
880      -36.42
887      -33.59
894      -34.51
901      -34.57
908      -28.84
915      -31.08
922      -30.59
929      -32.79
936      -37.98
943      -36.48
950      -32.30
957      -36.66

Ave      -33.65
Std        3.73


********* Table for XLOGGRAPH *********
```

```
$TABLE :2fxm :
$GRAPHS :2fxm cc_rad, cc_left_pit, cc_rig_pit :A:1,2,3,4: :2fxm cc_dang, a_dang :A:1,5,6:
$$
rnum  cc_rad cc_l_pit/100 cc_r_pit/100 cc_dang, a_dang-100  $$
$$
851     7.378     0.576     0.000     -9.13     -0.20
852     6.484     0.932     0.000     -5.94      0.04
853     6.248     3.687     0.000     -1.47     -1.38
854     6.037     3.361     0.000     -1.58     -0.98
855     5.800     3.061     0.000     -1.70     -0.48
856     5.593     2.563     0.000     -2.01     -0.77
857     5.390     2.128     0.000     -2.54      1.02
858     5.196     1.874     0.000     -2.83     -0.74
859     5.021     1.852     0.000     -2.92     -1.60
860     4.873     1.851     0.000     -2.88     -1.29
861     4.784     1.756     0.000     -3.06      0.05
862     4.711     1.696     0.000     -3.13     -0.09
863     4.679     1.615     0.000     -3.24     -1.86
864     4.666     1.498     0.000     -3.52     -1.59
865     4.619     1.490     0.000     -3.47     -2.03
866     4.561     1.532     0.000     -3.42     -1.42
867     4.521     1.412     0.000     -3.81     -0.05
868     4.566     1.267     0.000     -4.20     -0.13
869     4.633     1.206     0.000     -4.39      0.31
870     4.715     1.153     0.000     -4.67      0.34
871     4.831     1.048     0.000     -5.14      0.79
872     4.924     1.006     0.000     -5.11     -0.78
873     5.030     1.063     0.000     -4.96     -1.11
874     5.205     1.191     0.000     -4.39     -1.41
875     5.286     1.487     0.000     -3.43     -3.18
876     5.393     1.637     0.000     -3.43     -1.02
877     5.409     1.546     0.000     -3.77      0.71
878     5.421     1.435     0.000     -3.74     -0.88
879     5.341     1.461     0.000     -3.64     -2.18
880     5.180     1.453     0.000     -3.86     -0.97
881     5.036     1.319     0.000     -4.07      0.15
882     4.849     1.173     0.000     -4.42     -0.00
883     4.700     1.041     0.000     -4.95     -2.49
884     4.589     1.009     0.000     -5.14     -3.32
885     4.521     1.034     0.000     -5.13     -1.28
886     4.538     1.146     0.000     -4.65     -0.65
887     4.521     1.218     0.000     -4.49      1.29
888     4.511     1.217     0.000     -4.40      1.41
889     4.481     1.398     0.000     -3.81     -0.05
890     4.458     1.652     0.000     -3.38      1.10
891     4.511     2.084     0.000     -2.53      0.31
892     4.532     2.594     0.000     -2.03     -1.39
893     4.552     2.731     0.000     -1.94     -2.83
894     4.536     2.520     0.000     -2.12     -0.36
895     4.502     2.105     0.000     -2.57      2.23
896     4.498     2.064     0.000     -2.58      0.59
897     4.507     2.148     0.000     -2.50     -0.06
898     4.559     2.277     0.000     -2.38      0.93
899     4.601     2.262     0.000     -2.29     -0.60
900     4.585     2.116     0.000     -2.43     -1.24
901     4.560     1.911     0.000     -2.78     -0.35
902     4.514     1.618     0.000     -3.35      0.47
903     4.566     1.575     0.000     -3.52      1.53
904     4.639     1.664     0.000     -3.30      2.57
905     4.749     1.751     0.000     -3.21      2.13
906     4.869     1.679     0.000     -3.26     -0.98
907     4.914     1.511     0.000     -3.48     -1.41
908     5.001     1.368     0.000     -3.91     -0.61
909     5.074     1.331     0.000     -3.88     -0.66
910     5.096     1.490     0.000     -3.38     -0.76
911     5.126     1.620     0.000     -3.26      0.31
```

912	5.117	1.647	0.000	-3.31	1.02
913	5.114	1.525	0.000	-3.51	-0.99
914	5.059	1.403	0.000	-3.77	-2.89
915	4.994	1.281	0.000	-4.18	-2.29
916	4.924	1.164	0.000	-4.56	0.21
917	4.866	1.240	0.000	-4.24	0.62
918	4.853	1.365	0.000	-3.91	-0.14
919	4.916	1.476	0.000	-3.60	-0.76
920	4.974	1.550	0.000	-3.39	-1.54
921	5.033	1.643	0.000	-3.23	-2.01
922	5.044	1.698	0.000	-3.25	-1.11
923	5.053	1.650	0.000	-3.30	-0.64
924	5.026	1.809	0.000	-2.89	-1.60
925	5.006	1.932	0.000	-2.83	0.05
926	4.950	1.886	0.000	-2.93	0.94
927	4.887	1.725	0.000	-3.11	0.08
928	4.788	1.466	0.000	-3.52	-1.24
929	4.745	1.342	0.000	-3.84	-1.61
930	4.723	1.317	0.000	-3.90	-1.86
931	4.745	1.409	0.000	-3.70	-1.68
932	4.772	1.602	0.000	-3.36	-0.11
933	4.795	1.690	0.000	-3.17	0.16
934	4.830	1.678	0.000	-3.19	-0.67
935	4.833	1.689	0.000	-3.12	-1.45
936	4.823	1.660	0.000	-3.27	0.36
937	4.806	1.622	0.000	-3.38	0.96
938	4.748	1.656	0.000	-3.14	-1.84
939	4.694	1.645	0.000	-3.25	-1.30
940	4.675	1.487	0.000	-3.67	-0.61
941	4.688	1.403	0.000	-3.87	-1.55
942	4.733	1.417	0.000	-3.79	0.32
943	4.802	1.247	0.000	-4.30	2.11
944	4.830	1.151	0.000	-4.71	2.15
945	4.828	1.230	0.000	-4.37	1.04
946	4.848	1.338	0.000	-4.03	0.51
947	4.940	1.615	0.000	-3.32	1.60
948	4.985	1.865	0.000	-2.74	-2.08
949	5.003	1.882	0.000	-2.78	-4.32
950	4.964	1.962	0.000	-2.80	-0.71
951	4.919	1.890	0.000	-2.86	0.65
952	4.920	2.064	0.000	-2.60	-0.93
953	4.933	2.417	0.000	-2.23	-0.10
954	5.008	2.332	0.000	-2.32	1.64
955	5.076	2.282	0.000	-2.38	-0.37
956	5.166	2.281	0.000	-2.40	-1.30
957	5.262	2.011	0.000	-2.70	0.64
958	5.334	1.795	0.000	-3.03	2.33
959	5.458	1.823	0.000	-2.99	3.02
960	5.605	1.950	0.000	-2.84	3.54

MYH7$^{1551-1602}$

```
*********************************************************************
* TWISTER: Analysis of coiled-coil geometry   (c) S.Strelkov 2000-8 *
* E-mail: sergei.strelkov@pharm.kuleuven.be        Version 130508   *
* Reference: Strelkov,S.V., and Burkhard,P. (2002) J. Struct. Biol. *
*********************************************************************

Title 4xa4_ABedited

Input pdb_file 4xa4_ABedited.pdb
2-stranded coiled coil, chains A B
First_res 1551,  last_res 1602
Output pdb_file 4xa4_ABedited_axis.pdb

********* Sequence and heptad assignment *********

        1560      1570       1580       1590       1600
LEHEEGKIL RAQLEFNQIK AEIERKLAEK DEEMEQAKRN HLRVVDSLQT SLD
abcdefgab cdefgdefga bcdefgabcd efgabcdefg defgabcdef gab


********* Coiled-coil parameters per residue *********

Res cc_phase cc_rad cc_rise cc_pit  cc_dang  pos  Cr_ang  a_rad a_ris res/tur a_dang
axis_cur

552   -3.04   4.91   1.48   175.9    -3.04 E b  122.8   2.32   1.51   3.63   99.08
0.0000
553   -5.87   4.93   1.48   188.4    -2.83 H c -135.7   2.29   1.50   3.65   98.63
0.0000
554   -8.51   4.91   1.51   205.5    -2.65 E d  -34.5   2.27   1.53   3.57  100.82
0.0000
555  -11.17   4.89   1.49   201.7    -2.66 E e   70.9   2.32   1.50   3.55  101.53
0.0000
556  -13.81   4.89   1.49   203.5    -2.64 G f  173.2   2.24   1.51   3.57  100.75
0.0000
557  -16.31   4.91   1.53   220.7    -2.50 K g  -82.7   2.32   1.55   3.60   99.87
0.0156
558  -18.26   5.01   1.49   275.6    -1.95 I a   18.3   2.29   1.50   3.70   97.41
0.0172
559  -19.77   5.09   1.48   354.9    -1.51 L b  115.9   2.38   1.49   3.73   96.46
0.0168
560  -21.00   5.20   1.51   441.2    -1.24 R c -145.9   2.28   1.52   3.70   97.37
0.0154
561  -21.83   5.29   1.53   663.4    -0.83 A d  -47.5   2.35   1.53   3.66   98.47
0.0136
562  -22.59   5.39   1.49   700.1    -0.77 Q e   52.2   2.30   1.49   3.68   97.81
0.0119
563  -23.30   5.48   1.47   756.1    -0.70 L f  149.5   2.34   1.47   3.71   97.09
0.0115
564  -23.80   5.57   1.51  1083.0    -0.50 E g -112.5   2.30   1.51   3.64   98.96
0.0126
565  -24.09   5.59   1.54  1913.5    -0.29 F d  -12.5   2.29   1.54   3.58  100.50
0.0126
566  -23.73   5.60   1.51  1528.8 R   0.36 N e   88.5   2.31   1.51   3.64   98.98
0.0125
567  -24.42   5.58   1.49   772.4    -0.69 Q f -173.7   2.29   1.49   3.67   97.98
0.0119
568  -25.10   5.56   1.48   788.5    -0.68 I g  -75.1   2.35   1.49   3.65   98.52
0.0116
569  -25.68   5.51   1.49   923.0    -0.58 K a   24.2   2.26   1.49   3.61   99.66
0.0105
570  -26.32   5.44   1.52   855.4    -0.64 A b  125.3   2.37   1.52   3.60   99.99
0.0114
```

```
571  -26.80    5.39    1.51  1149.4     -0.47 E c -134.6    2.24    1.51    3.69   97.51
0.0113
572  -27.66    5.29    1.55   646.0     -0.87 I d  -39.1    2.31    1.56    3.63   99.21
0.0116
573  -29.82    5.14    1.65   275.3     -2.15 E e   65.3    2.19    1.66    3.48  103.47
0.0129
574  -32.66    5.12    1.62   204.7     -2.84 R f  171.2    2.16    1.63    3.45  104.55
0.0116
575  -35.24    5.13    1.61   224.1     -2.58 K g  -80.9    2.22    1.62    3.43  104.95
0.0131
576  -36.76    5.22    1.49   352.2     -1.52 L a   26.5    2.24    1.49    3.55  101.49
0.0146
577  -37.63    5.27    1.48   614.5     -0.87 A b  124.6    2.29    1.48    3.58  100.69
0.0142
578  -38.52    5.37    1.55   621.1     -0.90 E c -131.5    2.22    1.55    3.52  102.42
0.0110
579  -39.33    5.48    1.52   676.6     -0.81 K d  -31.0    2.29    1.52    3.56  101.23
0.0066
580  -39.97    5.61    1.51   848.0     -0.64 D e   71.0    2.21    1.52    3.59  100.33
0.0068
581  -40.49    5.66    1.48  1035.6     -0.52 E f  170.5    2.36    1.48    3.58  100.48
0.0049
582  -40.96    5.70    1.49  1133.8     -0.47 E g  -87.3    2.26    1.49    3.60  100.14
0.0048
583  -41.43    5.74    1.53  1178.6     -0.47 M a   12.1    2.34    1.53    3.64   98.89
0.0053
584  -41.81    5.79    1.50  1432.8     -0.38 E b  111.2    2.26    1.50    3.69   97.46
0.0063
585  -42.04    5.85    1.49  2301.1     -0.23 Q c -152.6    2.31    1.49    3.63   99.16
0.0080
586  -42.25    5.89    1.53  2620.2     -0.21 A d  -50.3    2.28    1.53    3.54  101.79
0.0107
587  -42.07    5.89    1.50  3032.1 R   0.18 K e   51.1    2.26    1.50    3.59  100.34
0.0126
588  -41.90    5.85    1.49  3117.7 R   0.17 R f  150.2    2.31    1.49    3.62   99.36
0.0124
589  -41.63    5.79    1.52  2020.5 R   0.27 N g -110.6    2.27    1.53    3.60  100.12
0.0115
590  -41.11    5.69    1.54  1073.6 R   0.52 H d  -10.7    2.28    1.55    3.57  100.80
0.0085
591  -40.48    5.59    1.52   863.7 R   0.63 L e   90.0    2.29    1.53    3.60  100.01
0.0084
592  -40.86    5.44    1.50  1419.2     -0.38 R f -170.6    2.31    1.51    3.68   97.93
0.0119
593  -41.13    5.30    1.46  1944.9     -0.27 V g  -74.6    2.35    1.47    3.76   95.84
0.0159
594  -42.14    5.17    1.50   532.7     -1.01 V a   21.0    2.31    1.51    3.68   97.88
0.0199
595  -43.81    5.07    1.55   334.9     -1.67 D b  122.6    2.30    1.56    3.59  100.17
0.0214
596  -45.86    4.98    1.45   255.7     -2.04 S c -136.0    2.32    1.46    3.68   97.89
0.0210
597  -48.85    4.93    1.46   176.1     -2.99 L d  -38.0    2.35    1.49    3.68   97.78
0.0000
598  -52.51    4.86    1.49   146.2     -3.66 Q e   64.8    2.25    1.51    3.63   99.27
0.0000
599  -56.52    4.86    1.45   130.3     -4.02 T f  165.3    2.31    1.49    3.54  101.79
0.0000
600  -60.55    4.87    1.49   133.6     -4.03 S g  -85.3    2.33    1.53    3.55  101.61
0.0000
601  -64.28    4.85    1.49   143.4     -3.73 L a   17.2    2.19    1.52    3.66   98.49
0.0000

Ave            5.33    1.51   422.5     -1.29                 2.29    1.52    3.61   99.66
0.0121
Std            0.33    0.04   408.6      1.24                 0.05    0.04    0.07    1.91
0.0039
```

```
Overall ccpitch is estimated as 2Pi*<cc_rise>/<cc_angle_per_res>
Overall roc of the a-helix axis (1/<curv>)   82.89 A
with curvature smoothing base 5

********* Crick angles for heptad position a *********

Res    Crick_angle

558        18.29
569        24.24
576        26.47
583        12.10
594        20.96
601        17.23

Ave        19.88
Std         4.72


********* Crick angles for heptad position d *********

Res    Crick_angle

554       -34.54
561       -47.53
565       -12.51
572       -39.14
579       -31.02
586       -50.35
590       -10.67
597       -38.04
Ave       -32.97
Std        13.69


********* Table for XLOGGRAPH *********

$TABLE :4xa4_ABedited :
$GRAPHS :4xa4_ABedited cc_rad, cc_left_pit, cc_rig_pit :A:1,2,3,4: :4xa4_ABedited
cc_dang, a_dang :A:1,5,6: $$
rnum  cc_rad cc_l_pit/100 cc_r_pit/100 cc_dang, a_dang-100  $$
$$
552     4.913     1.759     0.000     -3.04     -0.92
553     4.929     1.884     0.000     -2.83     -1.37
554     4.906     2.055     0.000     -2.65      0.82
555     4.887     2.017     0.000     -2.66      1.53
556     4.889     2.035     0.000     -2.64      0.75
557     4.910     2.207     0.000     -2.50     -0.13
558     5.010     2.756     0.000     -1.95     -2.59
559     5.095     3.549     0.000     -1.51     -3.54
560     5.198     4.412     0.000     -1.24     -2.63
561     5.285     6.634     0.000     -0.83     -1.53
562     5.389     7.001     0.000     -0.77     -2.19
563     5.481     7.561     0.000     -0.70     -2.91
564     5.573    10.830     0.000     -0.50     -1.04
565     5.591    19.135     0.000     -0.29      0.50
566     5.598     0.000    15.288      0.36     -1.02
567     5.576     7.724     0.000     -0.69     -2.02
568     5.558     7.885     0.000     -0.68     -1.48
569     5.507     9.230     0.000     -0.58     -0.34
570     5.441     8.554     0.000     -0.64     -0.01
571     5.385    11.494     0.000     -0.47     -2.49
572     5.292     6.460     0.000     -0.87     -0.79
573     5.136     2.753     0.000     -2.15      3.47
574     5.117     2.047     0.000     -2.84      4.55
575     5.126     2.241     0.000     -2.58      4.95
```

```
576     5.217     3.522     0.000     -1.52      1.49
577     5.272     6.145     0.000     -0.87      0.69
578     5.367     6.211     0.000     -0.90      2.42
579     5.476     6.766     0.000     -0.81      1.23
580     5.607     8.480     0.000     -0.64      0.33
581     5.656    10.356     0.000     -0.52      0.48
582     5.696    11.338     0.000     -0.47      0.14
583     5.738    11.786     0.000     -0.47     -1.11
584     5.788    14.328     0.000     -0.38     -2.54
585     5.846    23.011     0.000     -0.23     -0.84
586     5.893    26.202     0.000     -0.21      1.79
587     5.891     0.000    30.321      0.18      0.34
588     5.849     0.000    31.177      0.17     -0.64
589     5.795     0.000    20.205      0.27      0.12
590     5.693     0.000    10.736      0.52      0.80
591     5.590     0.000     8.637      0.63      0.01
592     5.441    14.192     0.000     -0.38     -2.07
593     5.304    19.449     0.000     -0.27     -4.16
594     5.174     5.327     0.000     -1.01     -2.12
595     5.072     3.349     0.000     -1.67      0.17
596     4.976     2.557     0.000     -2.04     -2.11
597     4.931     1.761     0.000     -2.99     -2.22
598     4.860     1.462     0.000     -3.66     -0.73
599     4.859     1.303     0.000     -4.02      1.79
600     4.874     1.336     0.000     -4.03      1.61
601     4.846     1.434     0.000     -3.73     -1.51
$$
```

MYH7$^{1526-1571}$

```
*********************************************************************
* TWISTER: Analysis of coiled-coil geometry   (c) S.Strelkov 2000-8 *
* E-mail: sergei.strelkov@pharm.kuleuven.be        Version 130508   *
* Reference: Strelkov,S.V., and Burkhard,P. (2002) J. Struct. Biol. *
*********************************************************************

Title 5cj1_AB_edited

Input pdb_file 5cj1_AB_edited.pdb
2-stranded coiled coil, chains A B
First_res 1526,  last_res 1571
Output pdb_file 5cj1_AB_edited_axis.pdb

********* Sequence and heptad assignment *********

     1530       1540       1550       1560       1570
LEKV RKQLEAEKME LQSALEEAEA SLEHEEGKIL RAQLEFNQIK AE
defg abcdefgabc defgabcdef gabcdefgab cdefgdefga bc


********* Coiled-coil parameters per residue *********

Res cc_phase cc_rad cc_rise cc_pit  cc_dang  pos  Cr_ang  a_rad a_ris res/tur a_dang
axis_cur

527   -4.41   4.99   1.36   111.2    -4.41 E e   70.3   2.36   1.42   3.63   99.10
0.0000
528   -8.83   4.94   1.43   116.4    -4.42 K f  173.0   2.26   1.48   3.60  100.12
0.0000
529  -13.22   4.93   1.49   122.7    -4.38 V g  -81.1   2.33   1.54   3.60  100.08
0.0000
530  -17.28   4.93   1.48   131.5    -4.06 R a   22.3   2.23   1.52   3.64   98.79
0.0000
531  -21.01   4.93   1.47   141.6    -3.73 K b  124.4   2.39   1.50   3.67   97.99
0.0000
532  -24.52   4.92   1.46   149.6    -3.51 Q c -134.5   2.25   1.49   3.68   97.93
0.0127
533  -27.82   4.89   1.51   164.7    -3.30 L d  -32.9   2.33   1.54   3.60  100.08
0.0110
534  -30.94   4.87   1.48   170.4    -3.12 E e   72.1   2.30   1.50   3.62   99.34
0.0097
535  -33.95   4.86   1.44   172.8    -3.01 A f  171.1   2.30   1.47   3.66   98.26
0.0089
536  -37.04   4.86   1.48   172.6    -3.09 E g  -85.7   2.30   1.50   3.59  100.29
0.0084
537  -40.14   4.82   1.47   171.0    -3.10 K a   18.1   2.27   1.49   3.56  101.21
0.0073
538  -43.36   4.79   1.48   165.0    -3.22 M b  122.6   2.31   1.50   3.59  100.20
0.0070
539  -46.72   4.75   1.47   157.7    -3.36 E c -135.3   2.27   1.50   3.67   98.23
0.0070
540  -50.13   4.72   1.51   159.2    -3.41 L d  -34.5   2.29   1.53   3.62   99.39
0.0067
541  -53.44   4.70   1.48   160.3    -3.31 Q e   70.0   2.31   1.50   3.61   99.80
0.0076
542  -56.56   4.69   1.42   163.6    -3.12 S f  169.6   2.29   1.44   3.65   98.56
0.0082
543  -59.77   4.67   1.49   166.7    -3.21 A g  -87.5   2.31   1.51   3.61   99.74
0.0083
544  -63.11   4.61   1.51   162.5    -3.34 L a   17.1   2.26   1.53   3.61   99.82
0.0065
545  -66.55   4.56   1.51   157.7    -3.44 E b  118.4   2.30   1.53   3.63   99.13
0.0057
546  -70.06   4.51   1.49   152.7    -3.51 E c -138.5   2.29   1.51   3.65   98.64
0.0056
```

```
547  -73.50    4.46   1.46   152.5     -3.44 A d  -37.6    2.30    1.48    3.63   99.21
0.0064
548  -76.93    4.45   1.47   153.5     -3.44 E e   66.5    2.34    1.49    3.60   99.92
0.0073
549  -80.31    4.49   1.45   154.4     -3.38 A f  167.1    2.25    1.48    3.63   99.24
0.0081
550  -83.77    4.55   1.51   157.1     -3.46 S g  -88.6    2.33    1.53    3.60   99.96
0.0087
551  -87.07    4.59   1.50   163.9     -3.30 L a   16.8    2.23    1.53    3.63   99.33
0.0088
552  -90.20    4.63   1.50   172.2     -3.13 E b  116.5    2.34    1.52    3.61   99.62
0.0049
553  -93.12    4.65   1.51   186.3     -2.91 H c -138.3    2.28    1.53    3.62   99.43
0.0078
554  -96.20    4.66   1.55   180.6     -3.08 E d  -39.4    2.26    1.57    3.59  100.41
0.0130
555  -99.83    4.64   1.58   157.0     -3.63 E e   68.1    2.28    1.61    3.56  101.26
0.0217
556 -104.24    4.73   1.74   141.9     -4.41 G f  166.8    2.12    1.76    3.47  103.92
0.0288
557 -107.50    4.94   1.71   188.9     -3.26 K g  -79.7    2.14    1.74    3.47  103.80
0.0334
558 -106.30    5.32   1.46   438.0 R    1.20 I a   27.6    2.32    1.51    3.62   99.52
0.0310
559 -105.39    5.67   1.44   570.5 R    0.91 L b  118.8    2.26    1.49    3.58  100.47
0.0258
560 -104.23    6.07   1.46   451.3 R    1.16 R c -135.9    2.29    1.51    3.56  101.09
0.0178
561 -103.06    6.45   1.46   448.8 R    1.17 A d  -42.0    2.24    1.51    3.57  100.76
0.0097
562 -102.22    6.86   1.47   629.6 R    0.84 Q e   61.9    2.31    1.53    3.58  100.47
0.0033
563 -101.46    7.27   1.45   686.3 R    0.76 L f  153.5    2.26    1.50    3.63   99.04
0.0061
564 -100.90    7.62   1.45   938.9 R    0.56 E g -103.6    2.32    1.50    3.63   99.24
0.0109
565 -100.27    8.03   1.48   839.5 R    0.64 F d  -12.5    2.29    1.53    3.61   99.85
0.0152
566  -99.77    8.39   1.48  1060.2 R    0.50 N e   95.9    2.27    1.52    3.61   99.69
0.0000
567 -100.11    8.71   1.43  1492.1     -0.35 Q f -156.7    2.37    1.47    3.65   98.56
0.0000
568 -100.71    9.02   1.44   864.4     -0.60 I g  -68.4    2.32    1.48    3.72   96.91
0.0000
569 -101.81    9.31   1.60   529.3     -1.09 K a   32.8    2.30    1.65    3.67   98.17
0.0000
570 -103.20    9.67   1.67   430.7     -1.39 A b  128.6    2.23    1.72    3.65   98.75
0.0000

Ave            5.57   1.49   229.1     -2.35                2.29    1.53    3.61   99.67
0.0115
Std            1.50   0.07   179.0      1.83                0.05    0.07    0.05    1.29
0.0076

Overall ccpitch is estimated as 2Pi*<cc_rise>/<cc_angle_per_res>

Overall roc of the a-helix axis (1/<curv>)   87.28 A
with curvature smoothing base 5


********* Crick angles for heptad position a *********

Res    Crick_angle

530        22.33
537        18.13
544        17.06
```

```
551        16.81
558        27.61
569        32.76

Ave        22.45
Std         5.95


********* Crick angles for heptad position d *********

Res    Crick_angle

533       -32.94
540       -34.52
547       -37.59
554       -39.39
561       -41.99
565       -12.50

Ave       -33.15
Std         9.70


********* Table for XLOGGRAPH *********

$TABLE :5cj1_AB_edited :
$GRAPHS :5cj1_AB_edited cc_rad, cc_left_pit, cc_rig_pit :A:1,2,3,4: :5cj1_AB_edited
cc_dang, a_dang :A:1,5,6: $$
rnum  cc_rad cc_l_pit/100 cc_r_pit/100 cc_dang, a_dang-100  $$
$$
527     4.990     1.112     0.000     -4.41     -0.90
528     4.938     1.164     0.000     -4.42      0.12
529     4.928     1.227     0.000     -4.38      0.08
530     4.930     1.315     0.000     -4.06     -1.21
531     4.929     1.416     0.000     -3.73     -2.01
532     4.921     1.496     0.000     -3.51     -2.07
533     4.885     1.647     0.000     -3.30      0.08
534     4.867     1.704     0.000     -3.12     -0.66
535     4.864     1.728     0.000     -3.01     -1.74
536     4.858     1.726     0.000     -3.09      0.29
537     4.824     1.710     0.000     -3.10      1.21
538     4.790     1.650     0.000     -3.22      0.20
539     4.755     1.577     0.000     -3.36     -1.77
540     4.718     1.592     0.000     -3.41     -0.61
541     4.704     1.603     0.000     -3.31     -0.20
542     4.689     1.636     0.000     -3.12     -1.44
543     4.667     1.667     0.000     -3.21     -0.26
544     4.614     1.625     0.000     -3.34     -0.18
545     4.564     1.577     0.000     -3.44     -0.87
546     4.506     1.527     0.000     -3.51     -1.36
547     4.464     1.525     0.000     -3.44     -0.79
548     4.448     1.535     0.000     -3.44     -0.08
549     4.492     1.544     0.000     -3.38     -0.76
550     4.549     1.571     0.000     -3.46     -0.04
551     4.591     1.639     0.000     -3.30     -0.67
552     4.628     1.722     0.000     -3.13     -0.38
553     4.655     1.863     0.000     -2.91     -0.57
554     4.660     1.806     0.000     -3.08      0.41
555     4.635     1.570     0.000     -3.63      1.26
556     4.731     1.419     0.000     -4.41      3.92
557     4.941     1.889     0.000     -3.26      3.80
558     5.317     0.000     4.380      1.20     -0.48
559     5.672     0.000     5.705      0.91      0.47
560     6.071     0.000     4.513      1.16      1.09
561     6.452     0.000     4.488      1.17      0.76
562     6.856     0.000     6.296      0.84      0.47
563     7.268     0.000     6.863      0.76     -0.96
```

```
564      7.623     0.000     9.389      0.56     -0.76
565      8.025     0.000     8.395      0.64     -0.15
566      8.387     0.000    10.602      0.50     -0.31
567      8.712    14.921     0.000     -0.35     -1.44
568      9.021     8.644     0.000     -0.60     -3.09
569      9.307     5.293     0.000     -1.09     -1.83
570      9.667     4.307     0.000     -1.39     -1.25
$$
```

MYH7$^{1678-1739}$

```
*********************************************************************
* TWISTER: Analysis of coiled-coil geometry   (c) S.Strelkov 2000-8 *
* E-mail: sergei.strelkov@pharm.kuleuven.be        Version 130508   *
* Reference: Strelkov,S.V., and Burkhard,P. (2002) J. Struct. Biol. *
*********************************************************************

Title 5wlz_AB_edited

Input pdb_file 5wlz_AB_edited.pdb
2-stranded coiled coil, chains A B
First_res 1678,  last_res 1739
Output pdb_file 5wlz_AB_edited_axis.pdb

********* Sequence and heptad assignment *********

   1680       1690       1700       1710       1720       1730
NN LLQAELEELR AVVEQTERSR KLAEQELIET SERVQLLHSQ NTSLINQKKK MDADLSQL QT
ab cdefgabcde fgabcdefga bcdefgabcd efgabcdefg abcdefgabc defgabcd ef




********* Coiled-coil parameters per residue *********

Res cc_phase cc_rad cc_rise cc_pit  cc_dang  pos  Cr_ang  a_rad a_ris res/tur a_dang
axis_cur

679   -3.21   4.83   1.56   175.3    -3.21 N b  122.4   2.27   1.58   3.58  100.59
0.0000
680   -6.51   4.79   1.47   160.5    -3.31 L c -133.7   2.27   1.50   3.63   99.30
0.0000
681  -10.06   4.80   1.45   147.6    -3.55 L d  -32.9   2.32   1.48   3.62   99.32
0.0000
682  -13.60   4.84   1.46   148.1    -3.54 Q e   71.6   2.29   1.49   3.62   99.46
0.0000
683  -16.96   4.92   1.45   155.0    -3.36 A f  172.8   2.27   1.48   3.59  100.16
0.0000
684  -20.21   5.02   1.50   166.5    -3.25 E g  -81.5   2.31   1.53   3.55  101.41
0.0105
685  -23.00   5.11   1.47   189.9    -2.79 L a   22.3   2.23   1.50   3.62   99.52
0.0164
```

```
686  -25.75   5.19   1.44   188.4    -2.75 E b  123.1   2.34   1.46   3.63   99.08
0.0200
687  -28.42   5.20   1.44   193.6    -2.67 E c -134.1   2.35   1.46   3.67   97.97
0.0197
688  -31.29   5.13   1.52   191.4    -2.86 L d  -35.8   2.21   1.54   3.64   98.98
0.0165
689  -34.54   5.03   1.54   170.5    -3.26 R e   69.3   2.32   1.57   3.54  101.60
0.0132
690  -37.63   4.91   1.50   174.7    -3.08 A f  172.0   2.25   1.52   3.59  100.32
0.0107
691  -40.82   4.82   1.51   170.3    -3.20 V g  -84.1   2.29   1.54   3.64   99.02
0.0101
692  -43.87   4.79   1.51   178.3    -3.04 V a   18.4   2.29   1.53   3.63   99.18
0.0113
693  -46.70   4.73   1.49   189.6    -2.83 E b  120.2   2.36   1.51   3.67   98.14
0.0113
694  -49.60   4.67   1.48   184.0    -2.89 Q c -139.9   2.30   1.50   3.74   96.25
0.0113
695  -52.76   4.60   1.50   170.3    -3.16 T d  -41.7   2.34   1.52   3.70   97.33
0.0112
696  -56.03   4.53   1.45   160.0    -3.27 E e   60.8   2.32   1.47   3.65   98.74
0.0097
697  -59.24   4.52   1.42   159.4    -3.21 R f  160.3   2.30   1.44   3.62   99.56
0.0094
698  -62.56   4.52   1.46   158.8    -3.32 S g  -94.0   2.33   1.49   3.60   99.94
0.0110
699  -66.14   4.47   1.48   148.4    -3.58 R a   11.8   2.22   1.50   3.62   99.34
0.0122
700  -70.26   4.46   1.51   131.8    -4.13 K b  111.2   2.30   1.54   3.55  101.41
0.0139
701  -74.77   4.43   1.43   114.0    -4.50 L c -141.0   2.28   1.46   3.57  100.77
0.0170
702  -79.75   4.46   1.47   106.1    -4.98 A d  -40.0   2.31   1.51   3.61   99.87
0.0172
703  -84.38   4.51   1.50   116.9    -4.63 E e   67.0   2.27   1.54   3.66   98.27
0.0158
704  -88.50   4.54   1.47   128.2    -4.12 Q f  163.8   2.33   1.50   3.69   97.53
0.0161
705  -92.80   4.59   1.55   129.8    -4.30 E g  -90.4   2.30   1.59   3.62   99.52
0.0165
706  -96.50   4.61   1.47   143.3    -3.70 L a   13.6   2.26   1.50   3.66   98.47
0.0165
```

```
707 -100.01   4.64   1.46   149.8    -3.51 I b  113.7   2.34   1.49   3.61   99.83
0.0167
708 -103.48   4.64   1.40   145.1    -3.48 E c -140.8   2.33   1.43   3.65   98.72
0.0166
709 -107.46   4.64   1.42   128.8    -3.98 T d  -42.7   2.29   1.46   3.65   98.72
0.0143
710 -112.26   4.62   1.49   111.9    -4.80 S e   63.6   2.26   1.54   3.53  101.96
0.0121
711 -116.83   4.58   1.45   114.2    -4.57 E f  170.3   2.32   1.50   3.57  100.81
0.0109
712 -121.32   4.58   1.54   123.8    -4.49 R g  -86.0   2.19   1.58   3.60   99.87
0.0088
713 -124.91   4.63   1.47   147.6    -3.60 V a   19.2   2.36   1.50   3.65   98.53
0.0077
714 -127.70   4.68   1.49   193.2    -2.78 Q b  117.4   2.28   1.51   3.64   98.95
0.0077
715 -130.51   4.78   1.43   183.3    -2.81 L c -137.5   2.36   1.45   3.69   97.62
0.0117
716 -134.69   4.85   1.44   123.8    -4.18 L d  -45.2   2.31   1.48   3.67   98.03
0.0144
717 -139.68   4.92   1.48   106.9    -4.98 H e   62.9   2.22   1.54   3.57  100.89
0.0143
718 -144.03   4.93   1.37   113.4    -4.35 S f  166.2   2.36   1.41   3.56  101.23
0.0128
719 -148.22   4.96   1.45   124.7    -4.20 Q g  -86.8   2.23   1.50   3.56  101.10
0.0130
720 -151.82   4.95   1.44   144.5    -3.60 N a   17.2   2.28   1.48   3.61   99.66
0.0127
721 -155.27   4.96   1.46   152.5    -3.45 T b  119.1   2.30   1.49   3.56  101.00
0.0132
722 -158.66   4.98   1.46   155.3    -3.38 S c -135.2   2.31   1.49   3.61   99.72
0.0124
723 -162.06   4.99   1.49   158.0    -3.40 L d  -35.3   2.24   1.52   3.64   98.89
0.0121
724 -165.99   5.02   1.58   144.2    -3.93 I e   68.9   2.30   1.61   3.56  101.06
0.0099
725 -170.06   5.03   1.47   130.1    -4.07 N f  172.4   2.28   1.51   3.65   98.75
0.0104
726 -174.37   5.04   1.43   119.7    -4.31 Q g  -87.1   2.31   1.48   3.66   98.30
0.0109
727 -178.94   5.06   1.47   115.9    -4.56 K a   18.6   2.30   1.51   3.57  100.76
0.0113
```

```
728 -183.16   5.12   1.43   121.6    -4.23 K b  120.7   2.32   1.47   3.62   99.57
0.0108
729 -186.99   5.17   1.42   134.0    -3.83 K c -136.6   2.30   1.46   3.67   98.12
0.0119
730 -191.08   5.22   1.51   133.4    -4.09 M d  -36.1   2.29   1.56   3.60  100.13
0.0161
731 -194.98   5.26   1.47   135.1    -3.91 D e   70.9   2.29   1.51   3.61   99.71
0.0235
732 -197.89   5.29   1.44   178.3    -2.91 A f  171.9   2.29   1.47   3.62   99.54
0.0292
733 -200.22   5.33   1.66   257.3    -2.33 D g  -85.5   2.18   1.68   3.44  104.80
0.0320
734 -201.81   5.26   1.60   362.9    -1.59 L a   25.2   2.24   1.61   3.50  102.95
0.0000
735 -202.69   5.16   1.50   618.6    -0.87 S b  121.0   2.27   1.51   3.65   98.61
0.0000
736 -202.19   4.98   1.47  1062.1 R   0.50 Q c -139.2   2.29   1.49   3.64   98.97
0.0000
737 -204.73   4.76   1.57   222.7    -2.54 L d  -40.6   2.20   1.64   3.42  105.17
0.0000
738 -210.18   4.25   1.76   116.2    -5.45 Q e   77.0   2.02   1.87   3.28  109.74
0.0000

Ave           4.84   1.48   152.6    -3.50               2.28   1.52   3.61   99.88
0.0139
Std           0.26   0.06    42.3     0.96               0.05   0.07   0.07    2.03
0.0047

Overall ccpitch is estimated as 2Pi*<cc_rise>/<cc_angle_per_res>

Overall roc of the a-helix axis (1/<curv>)   71.95 A
with curvature smoothing base 5


********* Crick angles for heptad position a *********

Res    Crick_angle
685        22.35
692        18.35
699        11.82
706        13.61
713        19.16
720        17.23
```

```
727        18.57
734        25.24
Ave        18.29
Std         4.04


********* Crick angles for heptad position d *********

Res    Crick_angle

681       -32.88
688       -35.81
695       -41.74
702       -39.95
709       -42.75
716       -45.20
723       -35.27
730       -36.09
737       -40.61

Ave       -38.92
Std         3.85



********* Table for XLOGGRAPH *********

$TABLE :5wlz_AB_edited :
$GRAPHS :5wlz_AB_edited cc_rad, cc_left_pit, cc_rig_pit :A:1,2,3,4: :5wlz_AB_edited
cc_dang, a_dang :A:1,5,6: $$
rnum  cc_rad cc_l_pit/100 cc_r_pit/100 cc_dang, a_dang-100  $$
$$

679     4.829     1.753     0.000     -3.21      0.59
680     4.786     1.605     0.000     -3.31     -0.70
681     4.799     1.476     0.000     -3.55     -0.68
682     4.837     1.481     0.000     -3.54     -0.54
683     4.916     1.550     0.000     -3.36      0.16
684     5.018     1.665     0.000     -3.25      1.41
685     5.109     1.899     0.000     -2.79     -0.48
686     5.191     1.884     0.000     -2.75     -0.92
687     5.197     1.936     0.000     -2.67     -2.03
688     5.129     1.914     0.000     -2.86     -1.02
689     5.030     1.705     0.000     -3.26      1.60
```

690	4.914	1.747	0.000	-3.08	0.32
691	4.821	1.703	0.000	-3.20	-0.98
692	4.790	1.783	0.000	-3.04	-0.82
693	4.734	1.896	0.000	-2.83	-1.86
694	4.668	1.840	0.000	-2.89	-3.75
695	4.599	1.703	0.000	-3.16	-2.67
696	4.529	1.600	0.000	-3.27	-1.26
697	4.515	1.594	0.000	-3.21	-0.44
698	4.520	1.588	0.000	-3.32	-0.06
699	4.472	1.484	0.000	-3.58	-0.66
700	4.457	1.318	0.000	-4.13	1.41
701	4.431	1.140	0.000	-4.50	0.77
702	4.460	1.061	0.000	-4.98	-0.13
703	4.507	1.169	0.000	-4.63	-1.73
704	4.545	1.282	0.000	-4.12	-2.47
705	4.592	1.298	0.000	-4.30	-0.48
706	4.613	1.433	0.000	-3.70	-1.53
707	4.639	1.498	0.000	-3.51	-0.17
708	4.643	1.451	0.000	-3.48	-1.28
709	4.642	1.288	0.000	-3.98	-1.28
710	4.622	1.119	0.000	-4.80	1.96
711	4.585	1.142	0.000	-4.57	0.81
712	4.583	1.238	0.000	-4.49	-0.13
713	4.628	1.476	0.000	-3.60	-1.47
714	4.678	1.932	0.000	-2.78	-1.05
715	4.776	1.833	0.000	-2.81	-2.38
716	4.850	1.238	0.000	-4.18	-1.97
717	4.921	1.069	0.000	-4.98	0.89
718	4.925	1.134	0.000	-4.35	1.23
719	4.955	1.247	0.000	-4.20	1.10
720	4.947	1.445	0.000	-3.60	-0.34
721	4.962	1.525	0.000	-3.45	1.00
722	4.977	1.553	0.000	-3.38	-0.28
723	4.992	1.580	0.000	-3.40	-1.11
724	5.018	1.442	0.000	-3.93	1.06
725	5.034	1.301	0.000	-4.07	-1.25
726	5.042	1.197	0.000	-4.31	-1.70
727	5.060	1.159	0.000	-4.56	0.76
728	5.122	1.216	0.000	-4.23	-0.43
729	5.166	1.340	0.000	-3.83	-1.88
730	5.220	1.334	0.000	-4.09	0.13
731	5.256	1.351	0.000	-3.91	-0.29
732	5.294	1.783	0.000	-2.91	-0.46

```
733     5.331     2.573     0.000     -2.33      4.80
734     5.262     3.629     0.000     -1.59      2.95
735     5.157     6.186     0.000     -0.87     -1.39
736     4.984     0.000    10.621      0.50     -1.03
737     4.761     2.227     0.000     -2.54      5.17
738     4.250     1.162     0.000     -5.45      9.74
$$
```

MYH7$^{1779-1855}$

```
*********************************************************************
* TWISTER: Analysis of coiled-coil geometry   (c) S.Strelkov 2000-8 *
* E-mail: sergei.strelkov@pharm.kuleuven.be        Version 130508   *
* Reference: Strelkov,S.V., and Burkhard,P. (2002) J. Struct. Biol. *
*********************************************************************


Title 4xa6_AB


Input pdb_file 4xa6_AB.pdb
2-stranded coiled coil, chains A B
First_res 1779,  last_res 1855
Output pdb_file 4xa6_AB_axis.pdb


********* Sequence and heptad assignment *********

  1780       1790       1800       1810       1820       1830         840
850

L ERXXXNXEQT IXDLQHRLDE AEQIALXGGX XQLQXLEARV RELENELEAE QXRNAESVX G
XRXSERRIXELTYQTE

e fgabczabcz zzzdefgabc defgdedefd efgzababca bcdefgabca
bczabcabczabczabczdededefgab


********* Coiled-coil parameters per residue *********


Res cc_phase cc_rad cc_rise cc_pit  cc_dang  pos  Cr_ang  a_rad a_ris res/tur a_dang
axis_cur


780  -19.13   5.06  12.09   227.5   -19.13 E f  -15.8   0.91  12.17   4.14   88.45
0.0000
781  -32.67   4.65  17.36   461.7   -13.54 R g -145.0  12.59  17.46   2.40  150.33
0.0000
782     nan   1.79   -nan    -nan R    nan X a   82.5  11.56   -nan   -nan    -nan
0.0000
783     nan   -nan   -nan    -nan R    nan X b   -nan   -nan   -nan   -nan    -nan
0.0000
784     nan   1.16   -nan    -nan R    nan X c   83.1   8.37   -nan   -nan    -nan
0.0000
785     nan   2.93  10.70   381.6   -10.09 N z  -96.9  29.23  10.76   2.02  178.47    -
nan
786     nan   3.34  11.29   193.2   -21.03 X a   85.3  29.79  11.35   2.05  175.42    -
nan
```

```
787     nan   5.02  11.51   212.1 R  19.54 E b   91.5    7.89  11.67    2.85  130.74    -
nan
788     nan   6.24   5.54   143.6 R  13.89 Q c  175.8    1.14   5.98    4.09   93.91    -
nan
789     nan   4.96   5.30   118.9 R  16.04 T z   93.5    1.04   5.79    3.89   94.06    -
nan
790     nan   2.85   8.42   196.3   -15.44 I z   49.6    9.57   8.59    2.77  131.18
0.0082
791     nan   3.02   9.28   277.3 R  12.05 X z  -86.4  37.09   9.35    2.06  174.92
0.0055
792     nan   4.42  10.35   234.3   -15.90 D z  -98.2    9.10  10.45    2.80  131.45
0.0238
793     nan   3.92   5.79   137.9   -15.12 L d  -43.3    1.06   6.09    3.99   94.06
0.0153
794     nan   4.84   1.42    65.4    -7.83 Q e   70.0    2.32   1.69    3.63   99.04
0.0091
795     nan   4.80   1.41   175.7    -2.89 H f  169.3    2.28   1.43    3.65   98.70
0.0224
796     nan   4.77   1.46   167.0    -3.16 R g  -87.0    2.35   1.49    3.60   99.90
0.0329
797     nan   4.69   1.53   145.8    -3.77 L a   16.8    2.22   1.56    3.63   99.05
0.0301
798     nan   4.63   1.50   125.9    -4.28 D b  118.2    2.30   1.53    3.67   98.01
0.0238
799     nan   4.54   1.45   123.0    -4.23 E c -139.1    2.30   1.48    3.65   98.63
0.0211
800     nan   4.42   1.49   136.2    -3.93 A d  -38.0    2.27   1.52    3.57  100.73
0.1371
801     nan   4.32   1.45   146.1    -3.58 E e   68.6    2.35   1.48    3.64   98.95
0.1429
802     nan   4.27   1.49   155.2    -3.45 Q f  164.1    2.21   1.51    3.67   98.20
0.1357
803     nan   4.22   1.51    88.2    -6.15 I g  -89.7    2.32   1.76    3.58  100.69
0.1036
804     nan   3.06   6.53   159.9 R  14.70 A d  -21.4    1.02   6.82    3.72   98.38
0.0884
805     nan   4.00  10.79   278.5   -13.95 L e  165.6  12.42  10.86    2.67  136.11
0.0150
806     nan   3.02   5.80   205.8   -10.15 X d  -12.0  49.02   5.96    2.04  176.44
0.0480
807     nan   2.22   3.85   117.5   -11.78 G e   92.5  25.17   3.99    3.91   92.08
0.0478
```

```
808     nan   1.97  15.74   266.0 R  21.30 G f  -92.0  32.48  15.76   3.89   92.57
0.0437
809     nan   0.91  13.42   194.3   -24.85 X d   -2.1  19.08  13.51   3.96   90.96
0.0431
810     nan   1.48  19.88   582.1   -12.30 X e   85.7  19.35  19.99   3.96   90.87
0.0411
811     nan   3.62  28.99   688.7   -15.16 Q f  -96.9  20.02  29.05   2.44  147.99
0.0033
812     nan   2.49  16.36   472.4 R  12.47 L g  -93.8   0.59  16.39   3.09  117.97
0.0207
813     nan   3.05  12.94   986.6    -4.72 Q z -140.2  12.95  12.95   2.43  148.18
0.0051
814     nan   2.86  12.99   549.2 R   8.51 X a   47.5  51.64  12.99   2.01  178.74
0.0036
815     nan   3.89  13.26   464.9   -10.27 L b -123.3  13.10  13.27   2.61  138.93
0.0080
816     nan   4.58   7.53   212.8 R  12.74 E a   89.0   0.73   7.70   3.50  103.83
0.0269
817     nan   4.81   1.50    62.4    -8.66 A b -176.4   2.39   1.68   3.56  101.11
0.0108
818     nan   4.73   1.52   130.1    -4.22 R c  -81.1   2.14   1.56   3.70   97.39
0.0377
819     nan   4.67   1.52   133.9    -4.09 V a   25.0   2.40   1.56   3.72   96.80
0.0579
820     nan   4.67   1.46   169.9    -3.09 R b  117.4   2.22   1.48   3.65   98.65
0.0646
821     nan   4.71   1.47   175.8    -3.00 E c -135.0   2.41   1.48   3.57  100.83
0.0171
822     nan   4.74   1.42   154.9    -3.30 L d  -36.8   2.14   1.44   3.61   99.75
0.0093
823     nan   4.77   1.58   157.3    -3.62 E e   69.3   2.38   1.61   3.56  101.19
0.0117
824     nan   4.72   1.41   171.4    -2.96 N f  171.2   2.22   1.43   3.75   96.14
0.0192
825     nan   4.67   1.48   153.4    -3.47 E g  -93.4   2.37   1.50   3.64   99.08
0.0874
826     nan   4.59   1.46   132.5    -3.96 L a   18.0   2.30   1.49   3.57  100.95
0.0772
827     nan   4.58   1.43   145.5    -3.54 E b  111.6   2.31   1.45   3.64   98.99
0.0486
828     nan   4.54   1.63   100.0 R   5.86 A c -142.7   2.29   1.77   3.63   99.13
0.0050
```

```
829     nan   3.80   7.67   127.1 R  21.72 E a   55.0   1.18   7.86   4.17   91.76
0.0111
830     nan   2.91  13.43   185.2   -26.10 Q b  122.4  13.88  13.51   2.81  131.72
0.0714
831     nan   2.66  13.99   266.6 R  18.89 X c  -88.8  55.69  14.02   2.04  176.70
0.0784
832     nan   3.00  14.59   203.2   -25.86 R z  -93.0  14.00  14.67   2.90  131.16
0.0334
833     nan   3.71   7.86   164.4   -17.21 N a    0.6   1.06   8.14   4.27   93.49
0.0406
834     nan   4.89   1.42   127.9    -3.99 A b  107.7   2.40   1.67   3.59  100.34
0.0399
835     nan   4.96   1.66    96.1 R   6.20 E c -148.8   2.20   1.78   3.67   98.35
0.0607
836     nan   4.34   8.15   173.2 R  16.94 S a   59.7   1.11   8.30   3.88   94.15
0.0710
837     nan   3.68  17.85   360.8   -17.81 V b  134.1  15.07  17.91   2.70  133.62
0.0425
838     nan   2.88  14.08   406.6   -12.46 X c  -87.8  51.28  14.10   2.03  177.75
0.0616
839     nan   2.85   3.80    73.4   -18.62 G z  -92.4  44.28   4.11   2.01  178.68
0.0635
840     nan   2.85   0.71    11.9   -21.45 X a   87.2  45.08   1.31   2.01  178.70
0.0634
841     nan   3.18   4.49   101.1 R  16.00 R b  154.7  45.96   4.82   2.02  178.62
0.0595
842     nan   2.82  16.17   497.3 R  11.70 X c  -30.5  53.89  16.18   2.02  177.87
0.0523
843     nan   2.88  20.11   461.8   -15.67 S z  -91.3  14.61  20.18   2.73  135.77
0.0536
844     nan   4.66   9.08   270.5 R  12.08 E a   60.3   0.87   9.20   3.61  104.24
0.0505
845     nan   5.14   1.84   131.0 R   5.05 R b  174.5   2.33   1.92   3.38  107.08
0.0140
846     nan   5.02   8.66   291.7 R  10.69 R c   97.6   1.10   8.76   4.67   86.29
0.0138
847     nan   3.24  15.52   307.6   -18.17 I z  141.4  16.25  15.60   2.97  125.14
0.0030
848     nan   3.19  16.46   454.2 R  13.05 X d  -18.6  65.58  16.51   2.03  177.36
0.0566
849     nan   4.84  17.40   966.0    -6.48 E e  165.4  16.33  17.46   2.68  134.53
0.0625
```

```
850     nan   3.84   9.36   335.2   -10.05 L d  -43.3   0.99   9.64   3.72   97.35
0.0000
851     nan   4.80   1.45    63.4    -8.22 T e   72.5   2.24   1.73   3.50  102.85
0.0000
852     nan   4.69   1.42   148.4    -3.45 Y f  178.8   2.33   1.45   3.58  100.55
0.0000
853     nan   4.58   1.53   127.9    -4.30 Q g  -80.1   2.22   1.57   3.64   98.87
0.0000
854     nan   4.39   1.66   116.0    -5.16 T a   25.0   2.23   1.72   3.62   99.34
0.0000

Ave           -nan   -nan    -nan      nan              -nan   -nan   -nan    -nan    -
nan
Std           -nan   -nan     nan      nan              -nan   -nan   -nan    -nan    -
nan

Overall ccpitch is estimated as 2Pi*<cc_rise>/<cc_angle_per_res>

Overall roc of the a-helix axis (1/<curv>)    -nan A
with curvature smoothing base 5




********* Crick angles for heptad position a *********

Res    Crick_angle
782        82.48
786        85.34
797        16.77
814        47.52
816        89.02
819        25.00
826        17.98
829        54.96
833         0.64
836        59.69
840        87.20
844        60.26
```

```
854        24.98

Ave        50.14
Std        29.43


********* Crick angles for heptad position d *********
Res    Crick_angle

793       -43.25
800       -38.03
804       -21.36
806       -12.02
809        -2.12
822       -36.80
848       -18.57
850       -43.33
Ave       -26.94
Std        14.55




********* Table for XLOGGRAPH *********

$TABLE :4xa6_AB :
$GRAPHS :4xa6_AB cc_rad, cc_left_pit, cc_rig_pit :A:1,2,3,4: :4xa6_AB cc_dang, a_dang
:A:1,5,6: $$
rnum  cc_rad cc_l_pit/100 cc_r_pit/100 cc_dang, a_dang-100  $$
$$
780     5.059     2.275     0.000    -19.13    -11.55
781     4.655     4.617     0.000    -13.54     50.33
782     1.790     0.000      -nan       nan      -nan
783      -nan     0.000      -nan       nan      -nan
784     1.165     0.000      -nan       nan      -nan
785     2.927     3.816     0.000    -10.09     78.47
786     3.342     1.932     0.000    -21.03     75.42
787     5.019     0.000     2.121     19.54     30.74
788     6.244     0.000     1.436     13.89     -6.09
789     4.961     0.000     1.189     16.04     -5.94
790     2.852     1.963     0.000    -15.44     31.18
791     3.018     0.000     2.773     12.05     74.92
```

792	4.424	2.343	0.000	-15.90	31.45
793	3.925	1.379	0.000	-15.12	-5.94
794	4.843	0.654	0.000	-7.83	-0.96
795	4.801	1.757	0.000	-2.89	-1.30
796	4.769	1.670	0.000	-3.16	-0.10
797	4.691	1.458	0.000	-3.77	-0.95
798	4.628	1.259	0.000	-4.28	-1.99
799	4.538	1.230	0.000	-4.23	-1.37
800	4.416	1.362	0.000	-3.93	0.73
801	4.318	1.461	0.000	-3.58	-1.05
802	4.274	1.552	0.000	-3.45	-1.80
803	4.222	0.882	0.000	-6.15	0.69
804	3.063	0.000	1.599	14.70	-1.62
805	3.999	2.785	0.000	-13.95	36.11
806	3.020	2.058	0.000	-10.15	76.44
807	2.216	1.175	0.000	-11.78	-7.92
808	1.969	0.000	2.660	21.30	-7.43
809	0.907	1.943	0.000	-24.85	-9.04
810	1.482	5.821	0.000	-12.30	-9.13
811	3.619	6.887	0.000	-15.16	47.99
812	2.488	0.000	4.724	12.47	17.97
813	3.055	9.866	0.000	-4.72	48.18
814	2.856	0.000	5.492	8.51	78.74
815	3.886	4.649	0.000	-10.27	38.93
816	4.583	0.000	2.128	12.74	3.83
817	4.813	0.624	0.000	-8.66	1.11
818	4.731	1.301	0.000	-4.22	-2.61
819	4.669	1.339	0.000	-4.09	-3.20
820	4.667	1.699	0.000	-3.09	-1.35
821	4.713	1.758	0.000	-3.00	0.83
822	4.742	1.549	0.000	-3.30	-0.25
823	4.768	1.573	0.000	-3.62	1.19
824	4.724	1.714	0.000	-2.96	-3.86
825	4.668	1.534	0.000	-3.47	-0.92
826	4.590	1.325	0.000	-3.96	0.95
827	4.579	1.455	0.000	-3.54	-1.01
828	4.540	0.000	1.000	5.86	-0.87
829	3.804	0.000	1.271	21.72	-8.24
830	2.909	1.852	0.000	-26.10	31.72
831	2.661	0.000	2.666	18.89	76.70
832	3.002	2.032	0.000	-25.86	31.16
833	3.707	1.644	0.000	-17.21	-6.51
834	4.893	1.279	0.000	-3.99	0.34

```
835      4.964      0.000      0.961       6.20      -1.65
836      4.336      0.000      1.732      16.94      -5.85
837      3.682      3.608      0.000     -17.81      33.62
838      2.880      4.066      0.000     -12.46      77.75
839      2.851      0.734      0.000     -18.62      78.68
840      2.845      0.119      0.000     -21.45      78.70
841      3.179      0.000      1.011      16.00      78.62
842      2.820      0.000      4.973      11.70      77.87
843      2.883      4.618      0.000     -15.67      35.77
844      4.661      0.000      2.705      12.08       4.24
845      5.144      0.000      1.310       5.05       7.08
846      5.017      0.000      2.917      10.69     -13.71
847      3.240      3.076      0.000     -18.17      25.14
848      3.192      0.000      4.542      13.05      77.36
849      4.844      9.660      0.000      -6.48      34.53
850      3.843      3.352      0.000     -10.05      -2.65
851      4.797      0.634      0.000      -8.22       2.85
852      4.686      1.484      0.000      -3.45       0.55
853      4.582      1.279      0.000      -4.30      -1.13
854      4.386      1.160      0.000      -5.16      -0.66
$$
```

Appendix D

Ubiquitination assays of myosin heavy chain fragments, titin fragments and CARP were subjected for mass spectrometry analysis for the detection of ubiquitination sites. Raw data obtained from the mass spectrometry facility at the University of Konstanz is shown in this appendix, together with sequence coverage and ubiquitination sites labeled in the sequence.

D.1 Mass spectrometry results for the MuRF1 depended CARP ubiquitination

ID	Score	Cov.	Pep.	PSMs	AAs	MW [kDa]	calc. pI	Master Prot.	Accession
CARP	17394.82	90.78	24	667	217	24.2	6.93		Gi0

A2	Sequence	PSMs	Prot.	Prot. Groups	Modifications	Ion Score	Exp Value	Charge	MH+ [Da]	ΔM [ppm]	RT [min]	Missed Cleavages
High	DKLLSTALHVAVR	30	1	1		101	1.75E-09	2	1.42284E+03	-0.24	21.14	1
High	LLSTALHVAVR	19	1	1		99	2.65E-09	2	1.17972E+03	-0.62	16.09	0
High	LMEAGAQIEFR	143	1	1		87	3.80E-08	2	1.26464E+03	0.80	22.53	0
High	DmLESTAIHWASR	11	1	1	M2(Oxidation)	87	4.07E-08	2	1.53272E+03	0.77	23.28	0
High	LLImYGADLNIK	63	1	1	M4(Oxidation)	77	3.61E-07	2	1.37976E+03	0.57	30.15	0
High	LmEAGAQIEFR	50	1	1	M2(Oxidation)	74	7.34E-07	2	1.28063E+03	-0.48	18.86	0
High	EGDTPLHDAVR	24	1	1		74	8.05E-07	2	1.20959E+03	-0.74	4.49	0
High	LLIMYGADLNIK	23	1	1		74	8.07E-07	2	1.36377E+03	0.79	33.97	0
High	AcLEGHLAIVEK	21	1	1	C2(Carbamidomethyl)	72	1.19E-06	2	1.33970E+03	-0.74	14.46	0
High	GAmVVKEPEPEIITEPVDVPTFLK	21	1	1	M3(Oxidation)	69	2.63E-06	3	2.65441E+03	2.57	34.39	1
High	AALENKLPVVEK	18	1	1		66	4.84E-06	2	1.31077E+03	1.56	13.38	1
High	GGNLDVLK	30	1	1		64	8.02E-06	2	8.15463E+02	1.14	10.78	0
High	DREGDTPLHDAVR	21	1	1		63	1.09E-05	2	1.48071E+03	0.37	4.57	1

High	EPEPEIITEPVDVPTFLK	17	1	1		63	1.10E-05	2	2.05308E+03	1.17	38.78	0
High	TGHYEcAEHLIAcEADLNAK	20	1	1	C6(Carbamidomethyl); C13(Carbamidomethyl)	61	1.51E-05	2	2.30202E+03	0.52	21.87	0
High	DMLESTAIHWASR	13	1	1		59	2.54E-05	2	1.51672E+03	0.24	27.46	0
High	TPmDLVLHWQNGTK	15	1	1	M3(Oxidation)	57	3.91E-05	2	1.65582E+03	-0.06	23.07	0
High	AIFDSLR	16	1	1		56	4.77E-05	2	8.21453E+02	1.07	21.11	0
High	AALENkLPVVEK	6	1	1	K6(GG)	55	6.62E-05	2	1.42481E+03	-1.88	14.71	1
High	LPVVEkFLSDK	6	1	1	K6(GG)	54	8.59E-05	2	1.38878E+03	0.20	26.56	1
High	AIFDSLRENSYK	6	1	1		52	1.15E-04	2	1.44273E+03	0.25	20.07	1
High	GAmVVkEPEPEIITEPVDVPTFLK	5	1	1	M3(Oxidation); K6(GG)	52	1.30E-04	3	2.76844E+03	-0.47	34.17	1
High	NNPDVcDEYKR	11	1	1	C6(Carbamidomethyl)	50	1.91E-04	2	1.40961E+03	0.38	3.80	1
High	DkLLSTALHVAVR	6	1	1	K2(GG)	49	2.30E-04	2	1.53689E+03	0.28	22.52	1
High	TPMDLVLHWQNGTK	10	1	1		47	3.78E-04	2	1.63983E+03	0.93	30.79	0
High	FLSDKNNPDVcDEYK	7	1	1	C11(Carbamidomethyl)	42	1.33E-03	2	1.84382E+03	0.49	13.39	1
High	GGNLDVLKLLLNK	6	1	1		41	1.71E-03	2	1.39685E+03	-0.23	36.66	1

High	NNPDVcDEYK	2	1	1	C6(Carbamidomethyl)	39	2.45E-03	2	1.25351E+03	-0.12	5.77	0
High	LLLNkGAK	8	1	1	K5(GG)	34	7.41E-03	2	9.70604E+02	-0.78	4.06	1
High	GGNLDVLkLLLNK	15	1	1	K8(GG)	33	8.99E-03	3	1.51090E+03	-0.01	36.39	1
High	GAkISAR	1	1	1	K3(GG)	32	1.24E-02	2	8.16469E+02	-0.07	3.37	1
High	LLImYGADLNIKNcAGK	5	1	1	M4(Oxidation); C14(Carbamidomethyl)	30	1.81E-02	3	1.90999E+03	0.52	24.28	1
High	LLImYGADLNIk	2	1	1	M4(Oxidation); K12(GG)	28	2.94E-02	2	1.49380E+03	-0.02	28.75	0
High	LLLNKGAK	4	1	1		27	3.81E-02	2	8.56562E+02	0.78	3.64	1
High	FLSDkNNPDVcDEYK	4	1	1	K5(GG); C11(Carbamidomethyl)	27	3.94E-02	3	1.95786E+03	-0.15	14.11	1
High	LLIMYGADLNIKNcAGK	4	1	1	C14(Carbamidomethyl)	25	5.70E-02	3	1.89399E+03	-0.49	27.06	1
High	LLImYGADLNIkNcAGK	2	1	1	M4(Oxidation); K12(GG); C14(Carbamidomethyl)	18	3.52E-01	3	2.02403E+03	0.63	24.78	1
High	NNPDVcDEYkR	2	1	1	C6(Carbamidomethyl); K10(GG)	18	3.54E-01	2	1.52366E+03	0.54	3.98	1

Sequence coverage CARP

```
         10         20         30         40         50
60         70         80
GAMVVKEPEP EIITEPVDVP TFLKAALENK LPVVEKFLSD KNNPDVCDEY
KRTALHRACL EGHLAIVEKL MEAGAQIEFR

         90        100        110        120        130
140        150        160
DMLESTAIHW ASRGGNLDVL KLLLNKGAKI SARDKLLSTA LHVAVRTGHY
ECAEHLIACE ADLNAKDREG DTPLHDAVRL

        170        180        190        200        210
NRYKMIRLLI MYGADLNIKN CAGKTPMDLV LHWQNGTKAI FDSLRENSYK TSRIATF
```

Detected ubi sites

K6, K30, K36, K41, K51, K101, K106, K109, K179

D.2 Mass spectrometry results for the MuRF1 depended MYH7$^{964\text{-}1088}$ ubiquitination

ID	Score	Cov.	Pep.	PSMs	AAs	MW [kDa]	calc. pI	Master Protein	Accession
MYH	953.66	35.38	7	37	130	15.0	4.98	FALSE	gi0191

A2	Sequence	PSMs	Prot.	Prot. Groups	Modifications	Ion Score	Exp Value	Charge	MH+ [Da]	ΔM [ppm]	RT [min]	Missed Cleavages
High	ALQEAHQQALDDLQAEEDK	2	2	1		53	1.05E-04	3	2.15202E+03	0.30	11.99	0
High	ALQEAHQQALDDLQAEEDKVNTLTk	3	2	1	K25(GlyGly)	15	6.21E-01	4	2.92244E+03	-0.94	15.00	1
High	ALQEAHQQALDDLQAEEDKVNTLTK	4	2	1		73	1.06E-06	3	2.80840E+03	-1.16	14.93	1
High	ALQEAHQQALDDLQAEEDKVNTLTkAK	2	2	1	K25(GlyGly)	23	1.06E-01	4	3.12158E+03	0.60	13.68	2
High	LEQQVDDLEGSLEQEK	5	2	1		96	4.79E-09	2	1.85989E+03	0.02	14.43	0
High	LEQQVDDLEGSLEQEKk	2	2	1	K17(GlyGly)	59	2.29E-05	3	2.10202E+03	-0.41	13.00	1
High	LEQQVDDLEGSLEQEKK	5	2	1		46	4.99E-04	2	1.98798E+03	0.02	13.01	1
High	VkLEQQVDDLEGSLEQEK	1	2	1	K2(GlyGly)	64	7.55E-06	3	2.20109E+03	0.30	14.94	1
High	VKLEQQVDDLEGSLEQEK	6	2	1		81	1.72E-07	2	2.08705E+03	-0.72	14.85	1
High	VKLEQQVDDLEGSLEQEKk	4	2	1	K19(GlyGly)	24	7.29E-02	4	2.32919E+03	0.15	13.55	2
High	VKLEQQVDDLEGSLEQEKK	3	2	1		52	1.19E-04	3	2.21515E+03	0.81	13.46	2

Sequence coverage MYH7 964-1088

```
       971        981        991       1001       1011
1020       1030       1040

GPVEKEKHAT ENKVKNLTEE METAGLDEII AKLTKEKKAL QEAHQQALDD
LQAEEDKVNT LTKAKVKLEQ QVDDLEGSLE


      1060       1060       1070        180       1090

QEKKVRMETD LERAKRKLEG DLKLTQESIM ETDLENDKQQ LDERLKKKDF
```

Detected ubiquitinated lysine residues

K1022, K1026, K1043

D.3 Mass spectrometry results for the MuRF1 depended MYH7$^{1301-1378}$ ubiquitination

ID	Score	Cov.	Pep.	PSMs	AAs	MW [kDa]	calc. pI	Master Protein	Accession
MYH6	5840.76	88.75	15	227	80	9.3	5.50	FALSE	gi0023

A2	Sequence	PSMs	Prot.	Prot. Groups	Modifications	Ion Score	Exp Value	Charge	MH+ [Da]	ΔM [ppm]	RT [min]	Missed Cleavages
High	GkLTYTQQLEDLK	8	2	1	K2(GG)	95	5.77E-09	2	1.65087E+03	0.38	19.80	1
High	LTYTQQLEDLK	9	2	1		86	4.87E-08	2	1.35171E+03	-0.30	24.11	0
High	LTYTQQLEDLKR	20	2	1		83	9.18E-08	2	1.50781E+03	-0.22	19.69	1
High	EQYEEETEAKAELQR	16	2	1		83	1.04E-07	2	1.85286E+03	1.12	14.30	1
High	EQYEEETEAkAELQR	14	2	1	K10(GG)	77	4.30E-07	2	1.96690E+03	0.89	15.34	1
High	VLSkANSEVAQWR	14	2	1	K4(GG)	74	7.36E-07	3	1.60184E+03	0.28	14.07	1
High	GKLTYTQQLEDLK	9	2	1		72	1.23E-06	2	1.53683E+03	0.78	19.28	1
High	NALAHALQSAR	33	2	1		69	2.77E-06	2	1.15163E+03	0.17	11.30	0
High	ANSEVAQWR	11	2	1		63	8.93E-06	2	1.06052E+03	-0.41	10.30	0
High	QLEEEVKAK	4	2	1		60	1.91E-05	2	1.07358E+03	0.24	3.45	1

High	LTYTQQLED LkR	12	2	1	K11(GG)	60	1.91E-05	2	1.62185E+03	-0.48	21.35	1
High	QLEEEVK	10	2	1		52	1.14E-04	2	8.74452E+02	-0.04	3.39	0
High	EQYEEETEA K	6	2	1		50	1.80E-04	2	1.25553E+03	0.88	3.45	0
High	HDcDLLR	10	2	1	C3(Carbamido methyl)	40	1.93E-03	2	9.28430E+02	-0.24	4.07	0
High	AKNALAHA LQSAR	2	2	1		39	2.80E-03	3	1.35076E+03	-0.67	4.24	1
High	HDcDLLREQ YEEETEAK	17	2	1	C3(Carbamido methyl)	38	3.26E-03	3	2.16495E+03	0.02	17.29	1
High	VLSKANSEV AQWR	7	2	1		34	8.32E-03	3	1.48780E+03	-0.02	12.30	1
High	ANSEVAQW RTK	6	2	1		31	1.70E-02	3	1.28966E+03	-0.19	4.21	1
High	QLEEEVkAK	16	2	1	K7(GG)	28	3.04E-02	3	1.18763E+03	1.17	3.50	1
High	RQLEEEVK	1	2	1		24	8.37E-02	2	1.03055E+03	-0.85	3.61	1
High	AkNALAHAL QSAR	2	2	1	K2(GG)	22	1.37E-01	3	1.46480E+03	-0.23	6.88	1

Sequence coverage MYH7 1301-1379

```
      1308       1318       1328       1338       1348
GPLTRGKLTY TQQLEDLKRQ LEEEVKAKNA LAHALQSARH DCDLLREQYE
1358       1368       1378
EETEAKAELQ RVLSKANSEV AQWRTKYETD
```

Detected ubiquitinated lysine residues

K1305, K1354, K1363, K1374

D.4 Mass spectrometry results for the MuRF1 depended MYH7$^{1234\text{-}1406}$ ubiquitination

ID	Score	Cov.	Pep.	PSMs	AAs	MW [kDa]	calc. pI	Master Protein	Accession
MYH	3951.65	82.29	24	163	175	20.2	5.20	FALSE	gi0022

A2	Sequence	PSMs	Prot.	Prot. Groups	Modifications	Ion Score	Exp Value	Charge	MH+ [Da]	ΔM [ppm]	RT [min]	Missed Cleavages
High	LQEAEEAVEAV	8	1	1		88	3.24E-08	2	1.18758E+03	0.46	22.12	0
High	LTYTQQLEDLKR	8	1	1		84	7.74E-08	2	1.50781E+03	-0.54	18.30	1
High	TKYETDAIQR	8	1	1		76	5.06E-07	2	1.22462E+03	-0.05	3.30	1
High	NALAHALQSAR	14	1	1		64	8.63E-06	2	1.15163E+03	0.27	9.86	0
High	ANSEVAQWR	8	1	1		63	9.68E-06	2	1.06052E+03	-0.52	9.30	0
High	EALISQLTR	7	1	1		60	2.19E-05	2	1.03059E+03	0.22	20.35	0
High	SVNDLTSQR	4	1	1		59	2.56E-05	2	1.01951E+03	-0.19	4.48	0
High	TLEDQmNEHR	4	1	1	M6(Oxidation)	57	3.90E-05	2	1.28856E+03	-0.61	3.31	0
High	QLDEKEALISQLTR	12	1	1		56	5.18E-05	3	1.64390E+03	0.10	23.87	1
High	TLEDQMNEHR	10	1	1		53	1.12E-04	2	1.27256E+03	-0.11	4.37	0

High	QLEEEVK	4	1	1		52	1.28E-04	2	8.74452E+02	0.10	3.43	0
High	LTYTQQLEDLK	4	1	1		51	1.60E-04	2	1.35171E+03	-0.03	22.96	0
High	EQYEEETEAK	4	1	1		47	3.56E-04	2	1.25553E+03	-0.09	3.39	0
High	EQYEEETEAKAELQR	10	1	1		47	3.91E-04	3	1.85286E+03	-0.04	13.70	1
High	LQTENGELSR	4	1	1		46	5.01E-04	2	1.14657E+03	0.01	3.77	0
High	YETDAIQR	2	1	1		46	5.14E-04	2	9.95479E+02	-0.05	4.67	0
High	VLSkANSEVAQWR	10	1	1	K4(GG)	45	6.73E-04	2	1.60184E+03	-0.29	13.17	1
High	AKLQTENGELSR	5	1	1		44	8.13E-04	2	1.34571E+03	0.54	3.56	1
High	TkYETDAIQR	2	1	1	K2(GG)	41	1.68E-03	2	1.33867E+03	0.72	3.83	1
High	GKLTYTQQLEDLK	7	1	1		41	1.77E-03	3	1.53683E+03	0.14	18.04	1
High	HDcDLLREQYEEETEAK	8	1	1	C3(Carbamidomethyl)	40	2.07E-03	3	2.16495E+03	0.36	15.65	1
High	HDcDLLR	4	1	1	C3(Carbamidomethyl)	38	2.92E-03	2	9.28431E+02	-0.04	3.88	0
High	EQYEEETEAkAELQR	4	1	1	K10(GG)	36	4.64E-03	3	1.96690E+03	1.04	14.27	1

High	TEELEEAK	4	1	1		32	1.38E-02	2	9.48452E+02	-0.27	3.38	0
High	QLEEEVKAK	3	1	1		23	1.07E-01	2	1.07358E+03	-0.21	3.42	1
High	VLSKANSEV AQWR	2	1	1		20	1.84E-01	3	1.48780E+03	0.90	11.20	1
High	AkLQTENGE LSR	1	1	1	K2(GG)	18	3.34E-01	3	1.45975E+03	0.31	3.83	1
High	YETDAIQRT EELEEAK	1	1	1		17	3.58E-01	3	1.92491E+03	0.52	18.62	1
High	TEELEEAKK	1	1	1		17	4.41E-01	2	1.07655E+03	0.11	3.34	1

Sequence coverage MYH7 1234-1406

1244 1254 1264 1274 1284 1294 1304 1314
MEQIIKAKAN LEKMCRTLED QMNEHRSKAE ETQRSVNDLT SQRAKLQTEN GELSRQLDEK EALISQLTRG KLTYTQQLED

1324 1334 1344 1354 1364 1374 1384 1394
LKRQLEEEVK AKNALAHALQ SARHDCDLLR EQYEEETEAK AELQRVLSKA NSEVAQWRTK YETDAIQRTE ELEEAKKKLA

1404
QRLQEAEEAV EAV

Detected ubiquitinated lysine residues

K1279, K1353, K1362, K1373

D.5 Mass spectrometry results for the MuRF1 depended MYH7$^{1477\text{-}1793}$ ubiquitination

ID	Score	Cov.	Pep.	PSMs	AAs	MW [kDa]	calc. pI	Master Protein	Accession
MYH	2603.74	88.40	46	123	319	36.8	5.07	FALSE	gi0189

A2	Sequence	PSMs	Prot.	Prot. Groups	Modifications	Ion Score	Exp Value	Charge	MH+ [Da]	ΔM [ppm]	RT [min]	Missed Cleavages
High	LAEQELIETSER	2	1	1		92	1.20E-08	2	1.41773E+03	8.43	9.17	0
High	VQLLHSQNTSLINQK	2	1	1		90	2.11E-08	2	1.72296E+03	6.94	8.18	0
High	AITDAAMMAEELK	2	1	1		84	8.43E-08	2	1.39368E+03	7.85	12.33	0
High	ANDDLKENIAIVER	2	1	1		83	1.05E-07	2	1.59984E+03	6.81	8.98	1
High	RNNLLQAELEELR	2	1	1		77	4.33E-07	3	1.59788E+03	7.85	12.29	1
High	KLAEKDEEMEQAK	3	1	1		73	1.05E-06	2	1.54877E+03	7.33	3.00	2
High	AITDAAMmAEELKK	2	1	1	M8(Oxidation)	68	3.27E-06	2	1.53777E+03	8.04	8.45	1
High	mDADLSQLQTEVEEAVQEcR	2	1	1	M1(Oxidation)C19(Carbamidomethyl)	66	4.72E-06	2	2.36706E+03	7.84	17.16	0

High	NNLLQAELEELR	2	1	1		63	9.88E-06	3	1.44178E+03	7.53	15.19	0
High	DTQIQLDDAVR	1	1	1		60	1.87E-05	3	1.27365E+03	7.03	9.31	0
High	SLQSLLKDTQIQLDDAVR	3	1	1		60	2.04E-05	2	2.04313E+03	8.27	15.85	1
High	LAEKDEEMEQAK	2	1	1		58	2.85E-05	2	1.42067E+03	7.91	3.76	1
High	VVDSLQTSLDAETR	2	1	1		56	5.13E-05	3	1.53379E+03	7.07	9.62	0
High	LAEKDEEmEQAKR	1	1	1	M8(Oxidation)	55	6.07E-05	3	1.59277E+03	7.87	2.89	2
High	KLAEQELIETSER	1	1	1		53	1.04E-04	2	1.54583E+03	8.60	8.08	1
High	KAITDAAMMAEELKK	4	1	1		50	2.16E-04	3	1.64987E+03	7.47	8.77	2
High	mEGDLNEmEIQLSHANR	2	1	1	M1(Oxidation)M8(Oxidation)	49	2.30E-04	3	2.01890E+03	6.77	8.24	0
High	EQDTSAHLER	1	1	1		49	2.49E-04	2	1.18556E+03	6.89	2.97	0

High	SLQSLLkDTQIQLDDAVR	1	1	1	K7(GlyGly)	49	2.56E-04	3	2.15717E+03	7.28	16.02	1
High	NmEQTIKDL	2	1	1	M2(Oxidation)	49	2.57E-04	2	1.10754E+03	6.92	7.67	1
High	KKMEGDLNEMEIQLSHANR	2	1	1		49	2.58E-04	4	2.24311E+03	8.11	8.77	2
High	NLQEEISDLTEQLGSSGK	2	1	1		48	2.87E-04	3	1.94796E+03	7.34	17.06	0
High	KAITDAAmmAEELKK	3	1	1	M8(Oxidation)M9(Oxidation)	46	5.20E-04	2	1.68186E+03	8.49	7.12	2
High	MEGDLNEmEIQLSHANR	5	1	1	M8(Oxidation)	45	6.47E-04	3	2.00291E+03	7.88	8.70	0
High	MEGDLNEMEIQLSHANR	2	1	1		44	8.09E-04	3	1.98692E+03	8.09	11.41	0
High	mELQSALEEAEASLEHEEGK	2	1	1	M1(Oxidation)	42	1.22E-03	3	2.24603E+03	7.75	14.76	0
High	AQLEFNQIKAEIER	2	1	1		42	1.23E-03	2	1.68891E+03	6.77	11.43	1

High	KmDADLSQLQTEVEEAVQEcR	2	1	1	M2(Oxidation) C20(Carbamidomethyl)	41	1.42E-03	3	2.49516E+03	7.72	15.28	1
High	GPLSTELFK	1	1	1		40	1.80E-03	2	9.91553E+02	7.53	10.65	0
High	AITDAAmmAEELKK	1	1	1	M7(Oxidation) M8(Oxidation)	39	2.41E-03	2	1.55377E+03	6.91	7.95	1
High	AITDAAmMAEELK	2	1	1	M7(Oxidation)	39	2.45E-03	2	1.40968E+03	7.38	11.55	0
High	NMEQTIKDL	1	1	1		39	2.73E-03	2	1.09155E+03	6.84	10.36	1
High	KKmDADLSQLQTEVEEAVQEcR	2	1	1	M3(Oxidation)C21(Carbamidomethyl)	38	2.90E-03	3	2.62325E+03	6.48	13.78	2
High	kAITDAAMMAEELKK	2	1	1	K1(GlyGly)	38	3.52E-03	4	1.76392E+03	7.73	9.16	2
High	KmEGDLNEmEIQLSHANR	2	1	1	M2(Oxidation); M9(Oxidation)	37	3.93E-03	3	2.14700E+03	8.22	7.29	1
High	AQLEFNQIK	2	1	1		37	4.27E-03	2	1.09060E+03	6.47	8.89	0
High	VQLLHSQNTSLINQKK	2	1	1		36	4.51E-03	3	1.85106E+03	7.87	7.17	1

High	LAEKDEEMEQAKR	2	1	1		35	6.87E-03	2	1.57678E+03	8.26	3.13	2
High	KKMEGDLNEmEIQLSHANR	7	1	1	M10(Oxidation)	32	1.25E-02	4	2.25910E+03	7.28	7.07	2
High	NMEQTIkDL	1	1	1	K7(GlyGly)	31	1.71E-02	2	1.20559E+03	7.24	10.67	1
High	AITDAAMMAEELKK	2	1	1		30	1.99E-02	2	1.52178E+03	6.71	10.12	1
High	LAEKDEEmEQAK	2	1	1	M8(Oxidation)	29	2.46E-02	2	1.43667E+03	8.81	2.93	1
High	KmEGDLNEMEIQLSHANR	7	1	1	M2(Oxidation)	28	3.22E-02	4	2.13101E+03	7.72	9.12	1
High	ENKNLQEEISDLTEQLGSSGK	1	1	1		27	4.28E-02	3	2.31915E+03	8.89	15.17	1
High	TIHELEKVR	1	1	1		26	5.28E-02	2	1.12465E+03	7.04	3.86	1
High	LAEQELIETSERVQLLHSQNTSLINQK	2	1	1		25	6.92E-02	3	3.12167E+03	7.55	12.89	1

High	ENkNLQEEISDLTEQLGSSGK	1	1	1	K3(GlyGly)	17	4.36E-01	3	2.43319E+03	7.84	15.13	1
High	kLAEQELIETSER	1	1	1	K1(GlyGly)	16	5.58E-01	3	1.65987E+03	8.28	8.29	1
High	GPLSTELFKLK	1	1	1		15	6.43E-01	3	1.23273E+03	7.79	11.27	1
High	KAITDAAMmAEELKK	7	1	1	M9(Oxidation)	15	6.56E-01	2	1.66587E+03	8.01	7.55	2
High	NAYEESLEHLETFKR	1	1	1		14	8.07E-01	3	1.86592E+03	7.88	10.62	1
High	TIHELEkVRK	1	1	1	K7(GlyGly)	14	8.15E-01	3	1.36679E+03	7.73	3.10	2
High	TIHELEkVR	1	1	1	K7(GlyGly)	14	8.39E-01	3	1.23869E+03	7.25	4.20	1
High	AITDAAMmAEELKkEQDTSAHLER	1	1	1	M8(Oxidation) K14(GlyGly)	13	9.04E-01	3	2.81836E+03	7.89	8.88	2
High	ANDDLKENIAIVERR	1	1	1		13	9.66E-01	2	1.75595E+03	7.43	8.46	2
High	kNMEQTIK	1	1	1	K1(GlyGly)	13	9.98E-01	2	1.10558E+03	8.17	3.02	1

Sequence coverage MYH7 1477-1793

```
          1485       1495       1505       1515       1525       1535       1545       1555
GPLSTELFKLK NAYEESLEHL ETFKRENKNL QEEISDLTEQ LGSSGKTIHE LEKVRKQLEA EKMELQSALE EAEASLEHEE

          1565       1575       1585       1595       1605       1615       1625       1635
GKILRAQLEF NQIKAEIERK LAEKDEEMEQ AKRNHLRVVD SLQTSLDAET RSRNEALRVK KKMEGDLNEM EIQLSHANRM

          1645       1655       1665       1675       1685       1695       1705       1715
AAEAQKQVKS LQSLLKDTQI QLDDAVRAND DLKENIAIVE RRNNLLQAEL EELRAVVEQT ERSRKLAEQE LIETSERVQL

          1725       1735       1745       1755       1765       1775       1785
LHSQNTSLIN QKKKMDADLS QLQTEVEEAV QECRNAEEKA KKAITDAAMM AEELKKEQDT SAHLERMKKN MEQTIKDL
```

Detected ubiquitinated lysine residues

K1483, K1503, K1528, K1651, K1700, K1757, K1771, K1784, K1791

Appendix E

Posttranslational modifications have a major role in cell signaling. Together with ubiquitination, described in this thesis, phosphorylation directs cell events. In order to quantify and study phosphorylation od muscle proteins and to there was a need for the efficient method to separate phosphorylated and non-phosphorylated proteins from the phosphorylation assay. The usage of well-established AKTA system and Ni-NTA columns was a starting idea towards developing a novel method for the separation of phosphorylated proteins. Successfully developed and optimized method is described in an attached paper below. Author contributed Adams *et al.*, 2019 in development and optimization of washing and regeneration of Ga^{3+} IMAC column.

- Adams M, Fleming JR, Riehle E, Zhou T, Zacharchenko T, **Markovic M**, Mayans O. (2019). Scalable, Non-denaturing Purification of Phosphoproteins Using Ga3+IMAC: N2A and M1M2 Titin Components as Study case. Protein J 38(2):181-189

As a result of this thesis there are three additional publications in preparation:

1. Bogomolovas Fleming JR, Franke B, Manso B, Simon B, Gasch A, Konzer A, **Markovic M**, Brunner T, Knöll R, Chen J, Labeit S, Scheffner M, Peter M, Mayans O (2021). Titin kinase ubiquitination aligns turn-over pathways with mechanical signals in the muscle sarcomere. *In preparation*
2. **Markovic M**, Bogomolovas J, Fleming JR, Bauer PM, Yu J, Chen J, Mayans O. (2021) The mutational alteration of BAG3 causes loss-of function in HSP70 and leads to cardiomyopathy. *In preparation*
3. Franke B, **Markovic M**, Manso B, Ward F, Berger T, Brunner T, Diederichs K, Scheffner M, Mayans O. (2021) Structural characterization of MuRF1 and its substrate targeting properties. *In preparation*

The Protein Journal (2019) 38:181–189
https://doi.org/10.1007/s10930-019-09815-w

Scalable, Non-denaturing Purification of Phosphoproteins Using Ga^{3+}-IMAC: N2A and M1M2 Titin Components as Study case

Michael Adams[1] · **Jennifer R. Fleming**[1] · **Eva Riehle**[1] · **Tiankun Zhou**[1] · **Thomas Zacharchenko**[2] · **Marija Markovic**[1] · **Olga Mayans**[1]

Abstract

The purification of phosphorylated proteins in a folded state and in large enough quantity for biochemical or biophysical analysis remains a challenging task. Here, we develop a new implementation of the method of gallium immobilized metal chromatography (Ga^{3+}-IMAC) as to permit the selective enrichment of phosphoproteins in the milligram scale and under native conditions using automated FPLC instrumentation. We apply this method to the purification of the UN2A and M1M2 components of the muscle protein titin upon being monophosphorylated in vitro by cAMP-dependent protein kinase (PKA). We found that UN2A is phosphorylated by PKA at its C-terminus in residue S9578 and M1M2 is phosphorylated in its interdomain linker sequence at position T32607. We demonstrate that the Ga^{3+}-IMAC method is efficient, economical and suitable for implementation in automated purification pipelines for recombinant proteins. The procedure can be applied both to the selective enrichment and to the removal of phosphoproteins from biochemical samples.

Keywords Phosphorylation · FPLC protein purification · Titin · PKA

Abbreviations

MS	Mass spectrometry
IMAC	Immobilized metal affinity chromatography
MOAC	Metal oxide affinity chromatography
FPLC	Fast protein liquid chromatography
IDA	Iminodiacetic acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
PKAcα	cAMP-dependent protein kinase.

1 Introduction

Phosphorylation of substrate proteins by protein kinases is a fundamental mechanism of cellular regulation [1]. Most phosphorylation studies aim to identify modification sites in target proteins using high-resolution mass spectrometry (MS) and proteomics techniques (reviewed in [2]). In phosphoproteome profiling, phosphoproteins are part of complex mixtures (commonly cell crudes or enriched broths), where incomplete phosphorylation adds to chemical heterogeneity and low protein abundance to trouble site identification. Thus, phosphopeptide enrichment is key to proteomics and considerable efforts have been directed to developing fractionation methodologies for this purpose. Immobilized metal affinity chromatography (IMAC) is the most cost-effective and successfully applied procedure to this effect [3]. However, IMAC leads to a differential recovery of peptides, having an enrichment bias towards multiply phosphorylated species. To increase the sensitivity for monophosphorylated samples and, thereby, improve phosphopeptide coverage in proteomics, novel metal oxide affinity chromatography (MOAC) matrices have been developed (*e.g.* [4–6]), the tendency being towards the combined usage of IMAC and MOAC materials [5, 7]. IMAC systems have been reported that are based on Ga^{3+}, Fe^{3+}, Ti^{4+}, Zn^{2+} and Al^{3+} ions (*e.g.* [8–13]) and include commercial products such as Fe^{3+}-based PHOS-Select™ IMAC (Sigma) and, the now discontinued, Ga^{+3}-IMAC PhosphoProfile™ (Sigma). Traditionally, Fe^{3+} has been the preferred immobilized ion for phosphor-group capture, but Ga^{3+} has been shown to be superior in its binding specificity of phosphorylated species [8, 9, 13, 14]. Beyond reaching a certain threshold of purity,

---

✉ Olga Mayans
olga.mayans@uni-konstanz.de

[1] Department of Biology, University of Konstanz, 78457 Konstanz, Germany

[2] School of Biology, University of Leeds, Leeds LS2 9JT, UK

phosphoproteome profiling does not place requirements on high yield or the 3D-fold preservation of protein samples. In fact, IMAC protocols (as well as MOAC) are commonly performed in microscale column formats and on proteolized peptide mixtures using strongly denaturing chemical elutants (*e.g*. acetonitrile, dihydroxybenzoic acid, ammonium hydroxide, trifluoracetic acid, or phosphoric acid at highly acidic pH values; [8, 10–12, 15–18]) as this yields the best performance.

In contrast to proteome research, the fractionation of intact, non-denatured phosphoproteins has seen surprisingly little development. Such fractionation is key to biochemical, biophysical and structural research that aims to study the conformational and functional consequences of phosphorylation in the target proteins in analytical terms. These studies require natively folded proteins, in high state of purity and in high quantity (often mg scale). To achieve this goal, proteins are commonly produced recombinantly, isolated, and subjected to phosphorylation *in vitro* using a known kinase. Yet, phosphotransfer reactions are often incomplete and capturing the phosphorylated species in reaction mixtures remains challenging. Enrichment protocols as employed in proteomics are here unsuitable and often cost-prohibitive. Thus, the production of phosphomimics via site-directed mutagenesis or via the introduction of non-natural phosphor-amino acids is pursued instead to achieve homogeneous sample preparations [19]. These procedures, however, suffer from their own limitations, with phosphomimics often exhibiting behaviours different to those of natively phosphorylated proteins [2, 20].

For a phosphoprotein purification technique to be applicable to downstream applications in the fields of biochemistry and biophysics, it must be highly selective, scalable up to the milligram range, non-destructive, and affordable. The most promising method to this effect is IMAC as these resins are already introduced in biochemical laboratories, are available commercially in large column formats, compatible with automated Fast Protein Liquid Chromatography (FPLC) instrumentation, rechargeable and reusable. Yet, only few studies have investigated the efficacy of IMAC in native phosphoprotein enrichment. Machida and co-workers [9] described the microscale purification of phosphoproteins from cell lysates using Ga^{3+}-IMAC on iminodiacetic acid (IDA) resin and using phosphate buffer for elution, indicating that Ga^{3+}-IMAC has promise for intact phosphoprotein purification. More recently, Ga^{3+}-IMAC on Sepharose beads has been attempted in the purification under native conditions of the protein OPN [20], but selectivity for the phosphorylated protein was not achieved in that study, likely due to the unusually acidic character of that protein.

Here, we establish a procedure for the enrichment of natively folded phosphoproteins using Ga^{3+}-IMAC resin, pre-packed columns and automated FPLC instrumentation. We also develop a protocol for the regeneration of the columns to allow reuse. We apply this method to two recombinant proteins as study cases -specifically the UN2A and M1M2 domain components of the myoprotein titin- and demonstrate that the procedure is effective on monophosphorylated species. Our results show the suitability of the method for the high-yield production of intact phosphoproteins for the study of protein structure and function.

2 Materials and Methods

2.1 Protein Production

The expression clone of UN2A from titin (residues 9472–9581; UniProtKB Q8WZ42) has been previously described [21]. The sequence coding for domains M1M2 (residues 32490–32713; UniProtKB Q8WZ42) was cloned into the pOPINB vector (Oxford Protein Production Facility, UK), which adds a His_6-tag and a protease 3C cleavage site N-terminally to the target sequence.

Recombinant overproduction of UN2A and M1M2 titin samples used *E. coli* Rosetta (DE3) cells (Merck Millipore). Cells were cultivated at 37 °C in Terrific Broth supplemented with 25 µg/ml kanamycin and 34 µg/ml chloramphenicol. Protein expression was induced at $OD_{600} = 0.9$ with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and cultures further grown overnight at 18 °C. Cells were harvested by centrifugation and lysed by sonication in ice in 25 mM HEPES pH 7.5, 300 mM NaCl, 1 mM dithiothreitol, in the presence of an EDTA-free protease inhibitor cocktail (Roche Applied Science). Lysates were clarified by centrifugation and proteins purified from the supernatant using Ni^{2+} affinity chromatography on a HisTrap™ HP column (GE Healthcare), followed by protease mediated tag removal (Tobacco Etch Virus for UN2A and 3C protease for M1M2), subtractive affinity chromatography and size-exclusion chromatography (HiLoad Superdex 75 column, GE Healthcare). Buffers used for UN2A purification were as described [21], with the final sample buffer being 25 mM HEPES pH 7.5, 100 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP). M1M2 purification used 300 mM NaCl, 50 mM Tris–HCl pH 7.5, except for the final size-exclusion chromatography that used 50 mM NaCl, 50 mM Tris–HCl pH 7.5. Purified samples were stored at 4 °C until further use.

2.2 Production of PKA Kinase

Murine PKA catalytic subunit α (PKAcα) was cloned into the pET-30 Ek/LIC vector (Novagen), which yielded an N-terminally His_6-tagged PKA sample. Expression was in *E. coli* Rosetta (DE3) cells (Merck Millipore), which were grown and harvested as described above for titin samples.

Cells were lysed in 25 mM HEPES pH 7.5, 300 mM NaCl and PKAcα purified from the supernatant by metal affinity chromatography in a HisTrap™ HP column (GE Healthcare). Binding to the HisTrap™ HP column was in lysis buffer, but elution used 25 mM HEPES pH 7.5, 300 mM NaCl, 200 mM Imidazole. Eluted fractions were dialysed against 25 mM HEPES pH 7.5, 100 mM NaCl and subjected to size exclusion chromatography (Superdex S200 16/600 column, GE Healthcare). Finally, 50% [v/v] glycerol was added to the purified sample, which was then aliquoted at a concentration of 1 mg/mL (estimated by A_{280}) and flash frozen in LN_2 in gel filtration buffer. The protein was stored at -80 °C until use.

2.3 Bulk Scale Phosphorylation Assay

Purified UN2A and M1M2 proteins were assayed for phosphorylation by PKAcα. Phosphorylation assays used ca. 30 mg of sample protein (in a 15 mL volume corresponding to a protein concentration of 2 mg/mL) in their final buffers. To this, 1.5 mL 10x magnesium buffer (100 mM $MgCl_2$, 500 mM Tris pH 7.5, 500 mM NaCl) was added. Reactions contained 6.6 µg/mL PKAcα and were initiated by addition of ATP (at a final concentration of 1.33 mM and prepared in 50 mM HEPES pH 7.5). The mixture was incubated at 37 °C and 150 rpm for 1 h, then gently mixed overnight at room temperature. Phosphorylation was confirmed via Phos-Tag™ (Wako Pure Chemical Industries; [22]) applied to 12% SDS–PAGE.

The SDS–PAGE gels were imaged using a Gel Doc™ EZ Gel imager and the supplied software Image Lab™ 5.2.1 (Bio-Rad). Gel densitometry was performed using Image Lab™ 5.2.1 by automatically detecting the gel bands at the default detection threshold. Relative abundance was then gathered from the lane profile generated by the software.

2.4 Gallium-Based Phosphoprotein Purification

Reactions mixtures from phosphorylation assays were subjected to Ga^{3+}-IMAC purification. For this, a column was prepared by stripping a commercial 5 mL HisTrap™ HP column (GE Healthcare) from Ni^{2+} ions using 20 mM sodium phosphate pH 7.4, 0.5 M NaCl, 50 mM ethylenediaminetetraacetic acid (EDTA). The column was then charged with Ga^{3+} ions by flowing 3 column volumes (CV) of 200 mM gallium chloride (Sigma-Aldrich). The column was washed in water to remove excess ions and stored in 20% ethanol until use.

For purification, protein mixtures from phosphorylation assays were loaded onto the column at a flow rate of 1 mL/min using an Äkta Start FPLC system (GE Healthcare). To minimize unspecific binding, the column containing the bound samples was washed with a high salt buffer (UN2A: 50 mM MES pH 5.5, 500 mM NaCl; M1M2: 50 mM Tris–HCl pH 6.0, 500 mM NaCl). Upon column washing (5CV), bound protein fractions were eluted using a gradient (5%-to-100%) of 500 mM sodium phosphate pH 6.0 (total 5CV) at a flow rate of 5 mL/min. Individual elution fractions were evaluated using Phos-Tag™ (Wako Pure Chemical Industries [22]) in 12% SDS–PAGE.

In all cases, upon each usage, the Ga^{3+}-IMAC column was cleaned with 50 mM citric acid pH 4.0 that eluted bound phosphate groups. The column was then used in subsequent purification rounds. After ca. 10 usage cycles, the column was stripped of metal ions using EDTA and newly charged with Ga^{3+} ions (as described above) for continued use.

2.5 Mass Spectrometry

UN2A and M1M2 samples were analyzed by reversed phase liquid chromatography (C18 resin; Acclaim PepMap100, Thermo Scientific) nanospray tandem mass spectrometry (LC-MS/MS) using an LTQ-Orbitrap mass spectrometer (Thermo Fisher) and an Eksigent nano-HPLC. After sample injection, the column was washed for 5 min with 95% phase A (0.1% formic acid) and 5% phase B (0.1% formic acid in acetonitrile), and peptides were eluted using a linear gradient of 5%-to-40% phase B, then 80% phase B at 300 nL/min. The LTQ-Orbitrap mass spectrometer was operated in a data dependent mode in which each full MS scan (30 000 resolving power) was followed by five MS/MS scans where the five most abundant molecular ions were dynamically selected and fragmented by collision-induced dissociation (CID) using a normalized collision energy of 35% in the LTQ ion trap. Dynamic exclusion was allowed. Tandem mass spectra were searched against the Swissprot human protein database using Mascot (Matrix Science) with "Trypsin/P" enzyme cleavage, static cysteine alkylation by iodoacetamide and variable methionine oxidation. Mass deconvolution was performed using mMass 5.5.0. The spectra were smoothed using two cycles of the Savitzky-Golay method and a window size of 0.3 Da/z. Peaks were picked with a signal/noise threshold of 3.0, a picking height of 75 and no deisotoping. Charges for the peaks were assigned manually. Monoisotopic deconvolution was performed with a peak grouping window of 0.01 Da/z.

2.6 Circular Dichroism

Circular dichroism (CD) spectra were recorded on a J-815 CD spectrometer (JASCO) using a quartz cuvette (Hellma) with a light path of 1 mm. Samples were studied in 50 mM NaH_2PO_4 pH 7.5 at a concentration of 0.44 mg/mL for M1M2; 1.06 mg/mL for pM1M2; 0.16 mg/mL for UN2A; and 0.26 mg/mL for pUN2A. The spectral range was $\lambda = 190$–300 nm. Spectra were acquired five times at 18 °C and averaged. The CDSSTR program [23] in the Dichroweb server [24] was used to calculate secondary structure content from spectra. The fit between experimental and reconstructed spectra was evaluated by their normalized root-mean-square deviation (NRMSD) (NRMSD > 0.05 signifies a poor fit).

3 Results

3.1 UN2A and M1M2 Components of Titin are Phosphorylation Targets of PKA

Titin is an intrasarcomeric, multi-domain protein of giant dimensions (> 3MDa) that orchestrates the response of muscle to mechanical and metabolic stress [25]. Titin is phosphorylated by a range of kinases that modulate, among other functions, its elastic properties. As study cases for the Ga^{3+}-IMAC methodology in this work, we chose to test the phosphorylation of the UN2A and M1M2 components of titin by cAMP-dependent protein kinase (PKA). We selected PKA among those kinases known to be relevant to titin physiology [25] as predictions (NetPhos 3.1 Server [26]) suggested that it might phosphorylate both samples.

UN2A is a 110-residue unique sequence (13.4 kDa), part of the essential N2A element of titin in the sarcomeric I-band [27] (Fig. 1a). UN2A is the main receptor site of the anti-apoptotic MARP/CARP proteins in titin [21, 28]. The recent biophysical characterization of a recombinant UN2A sample revealed that this domain adopts a loose, helical fold with high internal flexibility [21, 29]. Phosphorylation assays of recombinant UN2A (Fig. 1b) in our current study show that UN2A is phosphorylated by PKAcα. UN2A samples incubated with PKAcα were subjected to Phos-Tag™ SDS–PAGE [22]. The Phos-Tag™ gel contains a phosphate-binding acrylamide tag that causes phosphorylated proteins to migrate more slowly, resulting in additional SDS–PAGE bands that correspond to the level of phosphorylation of the migrating species. The Phos-Tag™ gels of PKAcα-incubated UN2A exhibited one extra band indicative of the presence of one phosphorylation site (Fig. 1d). This result was confirmed by full-mass determination mass spectrometry (MS) (Fig. 1d). Subsequently, the phosphorylation site was identified by tryptic digest MS as being located in residue S9578, which lies close to the C-terminus of UN2A in the linker sequence to the subsequent domain, Ig81. The UN2A-Ig81 domain interface has been recently shown to form the complete binding site of CARP in titin [21]. Future studies will be required to establish whether UN2A phosphorylation is physiologically relevant and whether it affects the titin/CARP interaction.

M1M2 (25,7 kDa) is a dual-Ig component of titin located immediately C-terminal to the titin kinase domain, in the sarcomeric M-line (Fig. 1a). The kinase region of titin is a multi-domain scaffold that recruits the turn-over factors MuRF1, MuRF2, Nbr1 and p62 to titin [25]. It is unknown whether the immediately adjacent M1M2 domains might participate in the scaffolding functions of this region. M1 and M2 are separated by a 32-residue long linker that is unstructured according to our sequence-based predictions: predictions of secondary structure used the Jpred4 server [30] and disorder was analyzed using the meta-server MetaDisorder [31]. Predictions were further confirmed with I-TASSER [32]. The length of the inter-domain linker in M1M2 (and in titin's M-band in general) contrasts notably with linker lengths in the rest of titin, where domains are connected by short linkers 0–4 residues in length [33]. The functionality of the extended linkers in M-line titin is unknown. In this work, we expressed recombinantly M1M2 as a soluble protein product. The protein, however, had a complex polydisperse chromatographic and electrophoretic behaviour, with SDS-resistant oligomers being observed in gels (Fig. 1c). We found that M1M2 is a target of PKAcα, with Phos-Tag™ SDS–PAGE analysis and full-length MS revealing the presence of one phosphorylation site in this segment (Fig. 1e). Site identification using tryptic digest MS revealed the targeted residue to be T32607 (Fig. 1e), which is located in the linker sequence joining M1 and M2. Remarkably, in twitchin from *Mytilus galloprovincialis* (the bivalve homologue of titin), the equivalent M1M2 linker sequence binds to myosin, with the binding being released upon phosphorylation of the linker by PKA [34, 35]. This PKA-dependent myosin/twitchin interaction is responsible for tension maintenance in the "catch" state of muscle in that animal group—a state where muscular tension is maintained for long periods of time with little energy consumption thanks to physical linkages between actin, myosin and twitchin. The M1M2 linker of vertebrate titin and bivalve twitchin do not share sequence conservation (Fig. 2) and vertebrate striated muscle does not adopt a "catch" state. However, since titin is an important scaffold in the sarcomere, it is tantalizing to envision that PKA phosphorylation at this locus might affect titin interactions with M-line associated proteins.

As negative controls in this study, the multi-domain titin kinase (A170-kinase-M1; Fig. 1(a) and the ankyrin-repeat domain of the protein CARP (binding partner of UN2A-Ig81) were also assayed for PKAcα phosphorylation. Neither protein was found to be a substrate of PKAcα (despite weak positive predictions; NetPhos 3.1 Server [26]), pointing to a possible specificity of the sites identified in UN2A and M1M2.

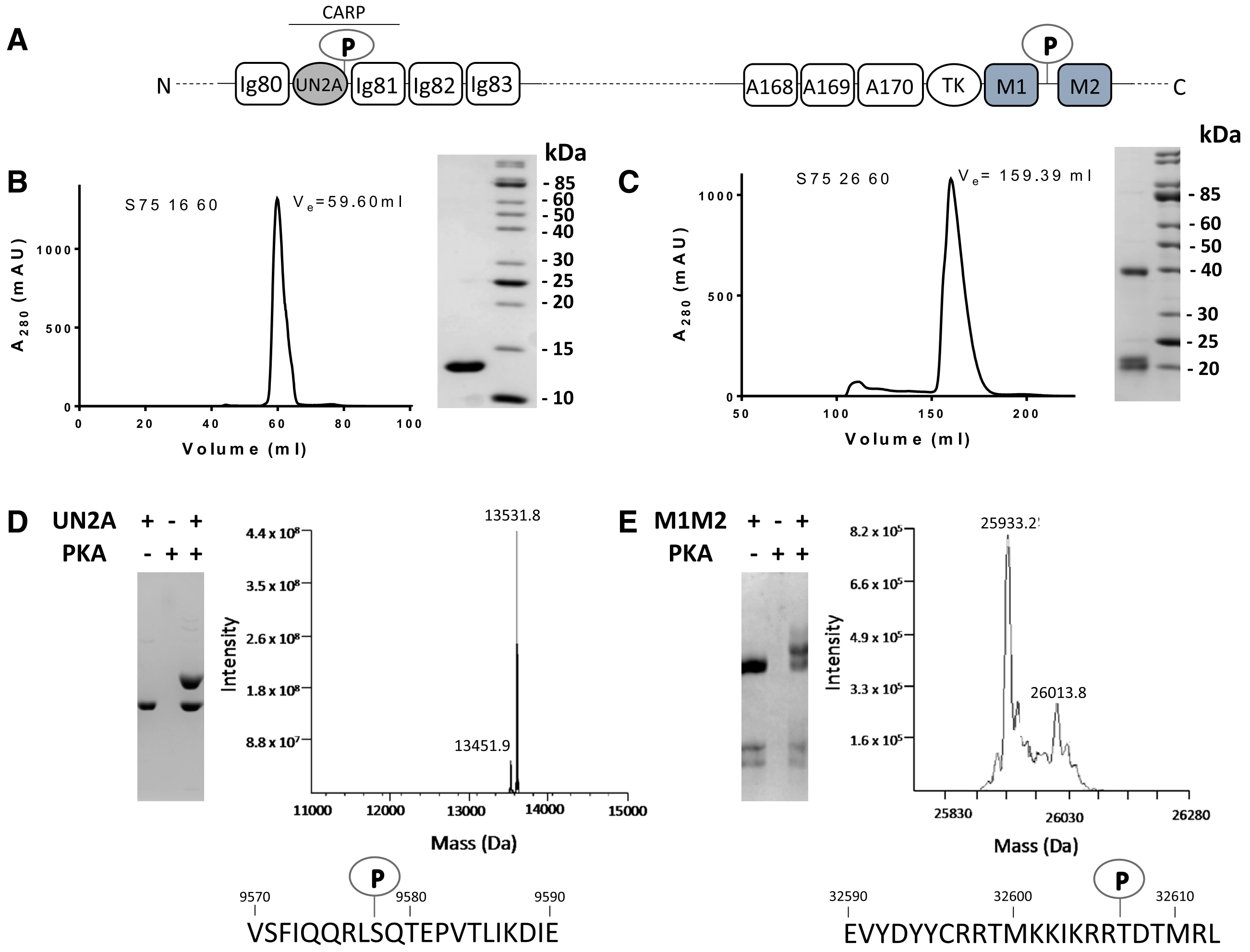


Fig. 1 Phosphorylation of UN2A and M1M2 by PKA (**a**) Schematic composition of the N2A element and the M-line multi-domain kinase region of titin. Shadowed are domains under study in this work; (**b**) and (**c**) gel filtration chromatogram and associated Coomassie-stained SDS–PAGE of purified samples. The molecular mass of the N2A construct used in this study is 13.4 kDa and that of M1M2 is 25.6 kDa. M1M2 migrates in SDS–PAGE with a lower apparent molecular mass. This is consistent with the observation of the dimeric form of the sample in the gels and with our conclusion that the sample is partly resistant to denaturation by SDS. All bands in the gels of UN2A and M1M2 were confirmed to correspond to the proteins under study using MS protein identification analysis; (**d**) and (**e**) Phos-Tag SDS–PAGE, full-mass MS analysis and sequence segments showing the phosphorylated residue for each sample. Throughout, UN2A results are shown on the left and M1M2 data on the right

3.2 Purification of Monophosphorylated UN2A and M1M2 Samples

Reaction mixtures from phosphorylation assays contained both phosphorylated and unphosphorylated species due to incomplete phosphotransfer. Based on Phos-Tag™ SDS–PAGE densitometry we estimated that UN2A and M1M2 reactions mixtures contained approx. 56% pUN2A and 65% pM1M2 phosphorylated species, respectively (Fig. 3d, f). To enrich the phosphorylated fractions, we subjected the reaction mixtures to Ga^{3+}-IMAC. For this, a Ga^{3+}-IMAC column that was compatible with automated FPLC equipment was produced by repurposing a pre-packed, commercial, metal affinity 5 mL HisTrap™ HP column (GE Healthcare) (as described in Methods). Ga^{3+} affinity chromatography was then performed on UN2A and M1M2 reaction mixtures using an Äkta Start FPLC system (GE Healthcare), with proteins being bound during in-column flow with no previous incubation or prolonged exposure to the resin. The large column format and its compatibility with FPLC instrumentation allowed here upscaling the purification procedure, so that approx. 30 mg total protein (~ 15 mL total volume) was applied to the IMAC column for each sample. In both cases, bound fractions were eluted with sodium phosphate buffer at near physiological pH (pH 6.0). Upon usage, the column was cleaned from bound phosphate groups by washing with citrate buffer pH 4.0 that protonated and eluted bound

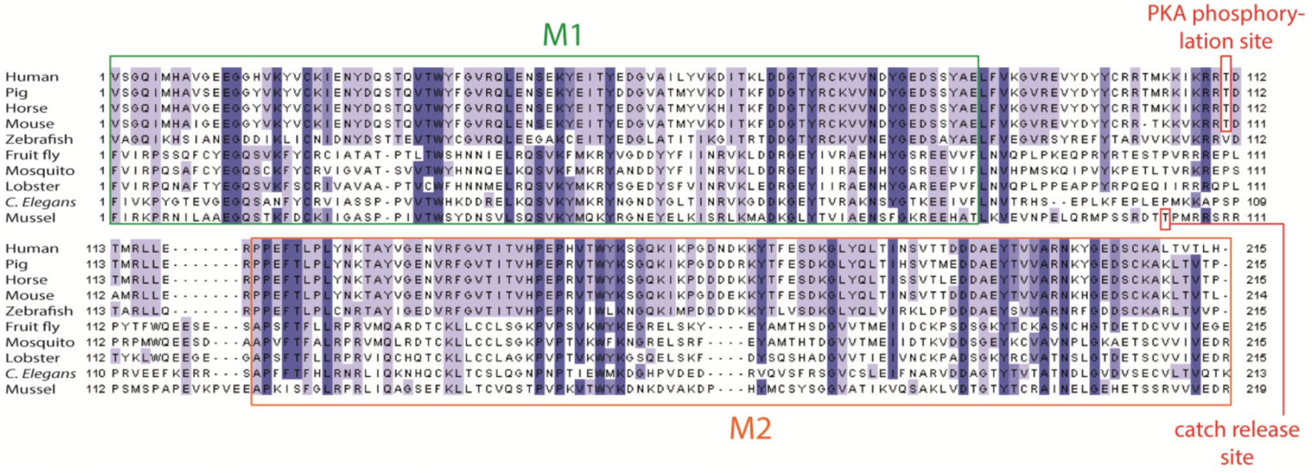


Fig. 2 Sequence alignment of M1M2 from vertebrate titin and invertebrate homologs Sequence alignment of the M1M2 linker sequence from vertebrate titin with its corresponding sequences in invertebrate homologs. The phosphorylation site in M1M2 found in this study as well as the residue found previously to be phosphorylated in Mytillus twitchin [34, 35] are indicated

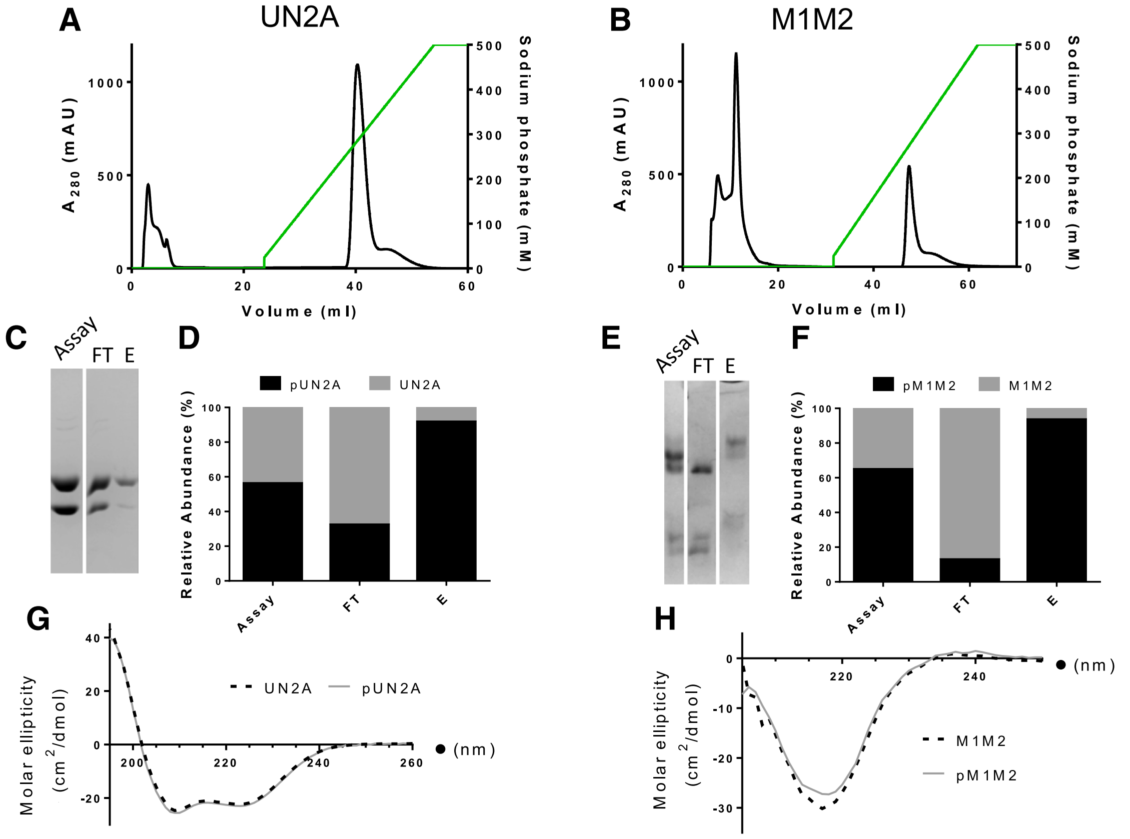


Fig. 3 Purification and characterization of phosphorylated forms of UN2A and M1M2 (**a**) and (**b**) Ga^{3+}-IMAC chromatograms showing flow-through, wash and elution fractions. Elution used a sodium phosphate gradient; (**c**) and (**e**) Phos-Tag™ SDS–PAGE corresponding to chromatograms in (**a**) and (**b**) (FT: flow-through; E: Elution); (**d**) and (**f**) Histograms of densitometry measurements for gels shown in (**c**) and (**e**).; (**g**) and (**h**) CD spectra of non-phosphorylated samples (solid line) and phosphorylated species purified by Ga^{3+}-IMAC (dashed lines). Secondary structure and disorder content was estimated from spectra as follows: UN2A (helix:70% / disorder: 21%; NRMSD: 0.003); pUN2A (helix: 66% / disorder: 22%; NRMSD: 0.002); M1M2 (strand:42% / disorder: 56%, NRMSD: 0.153); pM1M2 (strand 42% / disorder: 55%; NRMSD: 0.164)

phosphate groups; the column was then ready for further reuse.

The purification of pUN2A and pM1M2 species using this approach yielded a distinct elution peak in Ga^{3+}-IMAC chromatograms (Fig. 3a, b). Analysis of elution peaks using Phos-Tag™ SDS–PAGE densitometry confirmed enrichment of phosphorylated species, with pUN2A amounting now to 92% of the purified sample and pM1M2 being 93.5% (Fig. 3c–f). Eluted fractions contained approx. 4–5 mg of total phosphoprotein in each case, suggesting that under the conditions tested the phophoprotein binding efficiency was ca. 0.8 mg protein/mL Ga^{3+}-IMAC resin.

During this work and under the conditions described, we experienced that the Ga^{3+}-IMAC column maintained its maximal binding efficiency for approximately 5 cycles of reuse (binding/elution/cleaning). The column remained usable for up to 10–12 cycles, at which point it had to be regenerated by stripping the gallium ions with EDTA and newly reloading it with Ga^{3+} ions (as described in Methods). Although the number of reusing cycles was consistent for UN2A and M1M2 samples in this study, it should be considered that column reusability will vary depending on buffer conditions and the specific protein assayed.

3.3 Structural Integrity of Ga^{3+}-IMAC Purified pUN2A and pM1M2 Samples

To confirm the non-destructive nature of Ga^{3+}-IMAC, pre- and post- purification samples were analyzed using circular dichroism (CD) (Fig. 3g, h). CD spectra from unmodified UN2A confirmed that the secondary structure of this domain is primarily α-helical (estimated as ~70%), closely reproducing the spectral properties previously reported [21]. The CD spectrum of the phosphorylated sample, pUN2A, was in excellent agreement with that of the untreated sample, indicating that neither the introduced phosphor-group nor the purification procedure caused significant alterations to the sample. The CD spectrum of M1M2 was noisier and of lower quality, but it confirmed the expectation that this sample has an all-β secondary structure composition, with 42% β -strand content. This content is in good agreement with secondary structure estimates based on the crystal structure of M1 (PDB 2BK8) and other known Ig domains from titin [33], whose β-strand content commonly amounts to 45–50%. The β -strand content estimated from CD data supports the view that the linker sequence is unstructured. The spectrum of M1M2 was approximately reproduced by the pM1M2 sample (estimated 42% β-sheet content), suggesting that also in this case there was no significant structural difference between the treated and untreated samples.

In brief, CD spectra confirmed the gentle nature of Ga^{3+}-IMAC purification and its suitability for preparing natively folded phosphoproteins for downstream analysis. As UN2A is largely α-helical and M1M2 is β-sheet rich, we concluded that protein fold had no influence in the efficacy of the method.

4 Discussion and Conclusion

We demonstrate that the utilization of a stripped and Ga^{3+}-charged commercially pre-packed, large IMAC column allows the larger-scale purification of phosphoproteins in native conditions. In contrast to other available methods for phosphoprotein enrichment, the elution conditions are gentle and do not perturb the structural fold of proteins. We also show that the method affords scalability and automation. By developing a column regeneration procedure, we also improve affordability. Contrary to conventional studies, where proteins are incubated with the Ga^{3+}-IMAC resin in batch for a prolonged time and then run in microscale columns in gravity flow, we use in-column flow binding and automated FPLC equipment and show that the method is efficient under these conditions. Furthermore, we demonstrate that the method is sufficiently sensitive as to capture and resolve monophosphorylated species. We tested the method in two samples of contrasting characteristics. UN2A is an α-helical, monomeric and monodisperse protein. M1M2 is a dual-domain, β-rich protein that contains a long unstructured linker and that forms oligomers in solution of difficult chromatographic resolution and resistant to SDS, which leads to a complex electrophoretic migration pattern. The method successfully segregates phosphorylated and non-phosphorylated formations in both samples. The procedure is a valuable tool for research into the structure and function of phosphoproteins. In combination with ablation of undesirable phosphorylation sites using mutagenesis, it can provide singly phosphorylated species for functional and structural characterization in good yield and purity. Conversely, it can be applied to the clearing of protein samples from recombinant expression in eukaryotic systems, which often are affected by unwanted phosphorylation by kinases present in the host cell system. This could serve as an alternative methodology in cases where the addition of phosphatases is not desirable. Due to its versatility, the method shows promise for a wide variety of applications, as it offers the specificity of other protocols commonly used for phosphoprotein purification, whilst avoiding the drawbacks of denaturing conditions, low yield, or high cost.

Acknowledgements We thank the Proteomics Unit of the University of Konstanz for the contribution of mass spectrometry to this work.

Author Contributions MA, JRF and OM conceived the study; MA and TZa performed experiments and analysed data for M1M2; MA, ER and TZh performed experiments and analysed data for UN2A; MM designed protocols for Ga^{3+}-IMAC column reuse and regeneration;

OM, MA and JRF wrote the manuscript; all authors made manuscript revisions.

Funding We acknowledge the financial support of DFG SFB969 and the Leducq Foundation (TNE- 13CVD04). JRF is supported by an EU Marie Sklodowska-Curie Individual Fellowship (TTNPred, 753054).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Human and Animal Participants This article does not contain any work with human participants or animals performed by any of the authors.

References

1. Humphrey SJ, James DE, Mann M (2015) Protein phosphorylation: a major switch mechanism for metabolic regulation. Trends Endocrinol Metab 26:676–687
2. Dephoure N, Gould KL, Gygi SP, Kellogg DR (2013) Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. Mol Biol Cell 24:535–542
3. Fíla J, Honys D (2012) Enrichment techniques employed in phosphoproteomics. Amino Acids 43:1025–1047
4. Li Y, Xu X, Qi D, Deng C, Yang P, Zhang X (2008) Novel Fe3O4@TiO2 core-shell microspheres for selective enrichment of phosphopeptides in phosphoproteome analysis. J Proteome Res 7:2526–2538
5. Lin H, Deng C (2016) Development of immobilized Sn(4+) affinity chromatography material for highly selective enrichment of phosphopeptides. Proteomics 16:2733–2741
6. Zou X, Jie J, Yang B (2017) Single-step enrichment of N-glycopeptides and phosphopeptides with novel multifunctional Ti(4+)-immobilized dendritic polyglycerol coated chitosan nanomaterials. Anal Chem 89:7520–7526
7. Yang DS, Ding XY, Min HP, Li B, Su MX, Niu MM, Di B, Yan F (2017) Design and synthesis of an immobilized metal affinity chromatography and metal oxide affinity chromatography hybrid material for improved phosphopeptide enrichment. J Chromatogr A 1505:56–62
8. Posewitz MC, Tempst P (1999) Immobilized gallium(III) affinity chromatography of phosphopeptides. Anal Chem 71:2883–2892
9. Machida M, Kosako H, Shirakabe K, Kobayashi M, Ushiyama M, Inagawa J, Hirano J, Nakano T, Bando Y, Nishida E, Hattori S (2007) Purification of phosphoproteins by immobilized metal affinity chromatography and its application to phosphoproteome analysis. FEBS J 274:1576–1587
10. Lai AC, Tsai CF, Hsu CC, Sun YN, Chen YJ (2012) Complementary Fe(3+)- and Ti(4+)-immobilized metal ion affinity chromatography for purification of acidic and basic phosphopeptides. Rapid Commun Mass Spectrom 26:2186–2194
11. Zhu L, Zhang J, Guo Y (2014) Enhanced detection and desalting free protocol for phosphopeptides eluted from immobilized Fe (III) affinity chromatography in direct MALDI TOF analysis. J Proteomics 96:360–365
12. Yao Y, Dong J, Dong M, Liu F, Wang Y, Mao J, Ye M, Zou H (2017) An immobilized titanium (IV) ion affinity chromatography adsorbent for solid phase extraction of phosphopeptides for phosphoproteome analysis. J Chromatogr A 1498:22–28
13. Aryal UK, Ross AR (2010) Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry. Rapid Commun Mass Spectrom 24:219–231
14. Sykora C, Hoffmann R, Hoffmann P (2007) Enrichment of multiphosphorylated peptides by immobilized metal affinity chromatography using Ga(III)- and Fe(III)-complexes. Protein Pept Lett 14:489–496
15. Aryal UK, Olson DJ, Ross AR (2008) Optimization of immobilized gallium (III) ion affinity chromatography for selective binding and recovery of phosphopeptides from protein digests. J Biomol Tech 19:296–310
16. Albuquerque CP, Smolka MB, Payne SH, Bafna V, Eng J, Zhou H (2008) A multidimensional chromatography technology for in-depth phosphoproteome analysis. Mol Cell Proteomics 7:1389–1396
17. Yue XS, Hummon AB (2013) Combination of multistep IMAC enrichment with high-pH reverse phase separation for in-depth phosphoproteomic profiling. J Proteome Res 12:4176–4186
18. Steen H, Stensballe A, Jensen ON (2007) Phosphopeptide Purification by IMAC with Fe(III) and Ga(III). CSH Protoc https://doi.org/10.1101/pdb.prot4607
19. Liu WR, Wang YS, Wan W (2011) Synthesis of proteins with defined posttranslational modifications using the genetic noncanonical amino acid incorporation approach. Mol Biosyst 7:38–47
20. Ravi A, Guo S, Rasala B, Tran M, Mayfield S, Nikolov ZL (2018) Separation options for phosphorylated osteopontin from transgenic microalgae *Chlamydomonas reinhardtii*. Int J Mol Sci https://doi.org/10.3390/ijms19020585
21. Zhou T, Fleming JR, Franke B, Bogomolovas J, Barsukov I, Rigden DJ, Labeit S, Mayans O (2016) CARP interacts with titin at a unique helical N2A sequence and at the domain Ig81 to form a structured complex. FEBS Lett 590:3098–3110
22. Kinoshita E, Yamada A, Takeda H, Kinoshita-Kikuta E, Koike T (2005) Novel immobilized zinc(II) affinity chromatography for phosphopeptides and phosphorylated proteins. J Sep Sci 28:155–162
23. Johnson WC (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins 35:307–312
24. Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers 89:392–400
25. Krüger M, Kötter S (2016) Titin, a central mediator for hypertrophic signaling, exercise-induced mechanosignaling and skeletal muscle remodeling. Front Physiol 7:1–8. https://doi.org/10.3389/fphys.2016.00076
26. Blom N, Gammeltoft S, Brunak S (1999) Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol 294:1351–1362
27. Bang ML1, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, Witt CC, Labeit D, Gregorio CC, Granzier H, Labeit S (2001) The complete gene sequence of titin, expression of an unusual ≈ 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. Circ Res 89:1065–1072
28. Miller MK, Bang ML, Witt CC, Labeit D, Trombitas C, Watanabe K, Granzier H, McElhinny AS, Gregorio CC, Labeit S (2003) The muscle ankyrin repeat proteins: CARP, ankrd2/Arpp and DARP as a family of titin filament-based stress response molecules. J Mol Biol 333:951–964
29. Tiffany H, Sonkar K, Gage MJ (2017) The insertion sequence of the N2A region of titin exists in an extended structure with helical characteristics. Biochim Biophys Acta Proteins Proteom 1865:1–10
30. Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: A protein secondary structure prediction server. Nucleic Acids Res 43:W389–W394

31. Kozlowski LP, Bujnicki JM (2012) MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. BMC Bioinformatics 13:111
32. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinform 9:1–8
33. Zacharchenko T, von Castelmur E, Rigden DJ, Mayans O (2015) Structural advances on titin: towards an atomic understanding of multi-domain functions in myofilament mechanics and scaffolding. Biochem Soc Trans 43:850–855
34. Funabara D, Kinoshita S, Watabe S, Siegman MJ, Butler TM, Hartshorne DJ (2001) Phosphorylation of molluscan twitchin by the cAMP-dependent protein kinase. Biochemistry 40:2087–2095
35. Funabara D, Hamamoto C, Yamamoto K, Inoue A, Ueda M, Osawa R, Kanoh S, Hartshorne DJ, Suzuki S, Watabe S (2007) Unphosphorylated twitchin forms a complex with actin and myosin that may contribute to tension maintenance in catch. J Exp Biol 210:4399–4410