UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Zentrum für Geburtshilfe, Kinder- und Jugendmedizin

Klinik und Poliklinik für Pädiatrische Hämatologie und Onkologie

Prof. Dr. med. Stefan Rutkowski

Structural investigation and characterisation of the sarcomeric Z-Disk Titin-Obscurin-Complex and the M-Band Myomesin My5-Dimer

Dissertation

zur Erlangung des Doktorgrades Dr. rer. biol. hum.

an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:

Philipp Hornburg

aus Magdeburg

Hamburg 2018

Angenommen von der Medizinischen Fakultät am: 23.01.2020

Veröffentlicht mit Genehmigung der Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Reinhard Schneppenheim

Prüfungsausschuss, 2. Gutachter/in: Dr. Dmitri Svergun

This thesis describes work performed under the joint supervision of Prof. Dr. Matthias Wilmanns at the European Molecular Biology Laboratory (EMBL) in Hamburg, Germany, and Prof. Dr. Reinhard Schneppenheim at the Universitätsklinikum Hamburg-Eppendorf (UKE).

“The most exciting phrase to hear in science, the one that heralds the most discoveries, is not "Eureka!" but 'That's funny...”

— Isaac Asimov

Table of Contents

List of Figures

List of Tables

1 Abstract

The fundamental functions of muscles are enabling flexibility and dynamics on the one hand while providing stability on the other. This contradictory principle of motion and rigidity needs to be finely regulated and thus requires highly ordered sub-cellular structures. An essential part in these processes play the three muscle proteins titin, obscurin and myomesin. Through targeted protein-protein interactions, titin guides numerous proteins to their specific spatial positions within the sarcomere, thus ensuring the controlled assembly of the muscle structure. Titin interacts with obscurin on both the M-band and the Z-disk, two positions more than 1 µm apart. While the interaction at the M band has already been structurally described and functionally investigated, little is known about the interaction at the Z-disk involving the titin domains Z8 and Z9 and the obscurin domains O58 and O59. In addition to titin and obscurin, myomesin is another structural protein responsible for the anchoring of the myosin filament in the M band. A crystal structure has shown that the fifth domain of myomesin may have a special type of dimerization that could further increase the structural stability of myomesin anchoring.

The work presented here, therefore, addresses i) the yet unsolved molecular structure of the titin-obscurin complex at the Z-disk and ii) the characterization of the dimerization of the My5 domain.

By solving the high-resolution X-ray crystal structure of the titin-obscurin complex and using complementary biophysical methods such as size exclusion chromatography, isothermal titration calorimetry, or small-angle X-ray scattering, it was, for the first time, possible to present a comprehensive model of the Z-disk interaction between titin and obscurin. Including extensive database analysis of single nucleotide polymorphisms within these structures and the interpretation of their impact on the molecular structure of the complex, attractive hypotheses for the involvement of four instead of two domains in complex formation have been proposed and the first model for the complete regulatory mechanism of the complex was developed.

In the second part of the thesis, the mechanism of My5 domain dimerization is structurally analysed and shown to be of physiological importance as well as central for the stability of the

sarcomere. In addition, this work presents a composite model combining all available high-resolution structures of myomesin, representing the most complete structural model of myomesin to date and further supporting the proposed myomesin M-band assembly.

In summary, this dissertation demonstrates in two different systems that the structural information plays an important role to understand the nature and behaviour of proteins and their regulatory mechanisms. However, only by considering the native environment and in combination with other methods, it will be possible to draw valid conclusions about the molecules behaviour in a broader context.

2 Zusammenfassung

Die fundamentale Funktion von Muskeln ist es Flexibilität und Dynamik zu ermöglichen, und gleichzeitig Stabilität und Halt zu gewähren. Diese gegensätzlichen Prinzipien von Beweglichkeit und Steifheit müssen genau reguliert und präzise koordiniert werden. Von zentraler Bedeutung für diese Prozesse sind die Muskelproteine Titin, Obscurin und Myomesin. Durch gezielte Protein-Protein-Interaktionen führt Titin zahlreiche Proteine auf ihre spezifischen räumlichen Positionen innerhalb des Sarkomeres und gewährleistet somit den geregelten Aufbau der Muskelstruktur. Titin interagiert mit Obscurin sowohl an der M-Bande als auch an der Z-Scheibe, zwei Positionen die mehr als 1 µm voneinander entfernt sind. Während die Interaktion an der M-Bande bereits strukturell aufgeklärt und funktional untersucht wurde, ist noch immer wenig über die Interaktion an der Z-Scheibe bekannt, an der die Titindomänen Z8 und Z9 sowie die Obscurindomänen O58 und O59 beteiligt sind. Neben Titin und Obscurin ist Myomesin als weiteres strukturgebendes Protein verantwortlich für die Verankerung des Myosinfilaments in der M-Bande. Eine Kristallstruktur hat gezeigt, dass die fünfte Domäne von Myomesin eine spezielle Art der Dimerisierung aufweist, welche die strukturelle Stabilität der Myomesin-verankerung noch weiter erhöhen könnte.

Die hier präsentierte Arbeit beschäftigt sich deshalb zum einen mit der noch ungelösten Molekularstruktur des Titin-Obscurin Komplexes an der Z-Scheibe und zum anderen mit der Charakterisierung der Dimerisierung der My5 Domäne.

Mit der Entschlüsselung der hochaufgelösten Röntgenkristallstruktur des Titin-Obscurin-Komplexes und der Anwendung ergänzender biophysikalischer Methoden wie Größenausschlusschromatographie, Isothermaler Titrationskalorimetrie oder Kleinwinkel-Röntgenstreuung war es zum ersten Mal möglich, ein umfassendes Modell der Z-Scheiben-Interaktion zwischen Titin und Obscurin zu präsentieren. Unter Einbeziehung umfassender Datenbankanalysen von Einzelnukleotid-Polymorphismen innerhalb dieser Strukturen und der Interpretation ihrer Auswirkung auf die Molekularstruktur des Komplexes, konnten spezifische Hypothesen für die Beteiligung von vier anstelle von zwei Domänen an der Komplexbildung

aufgestellt, und das bislang erste Modell für den vollständigen Regulierungsmechanismus des Komplexes entwickelt werden.

Im zweiten Teil der Dissertation wird der Mechanismus der My5-Domänen-Dimerisierung strukturell analysiert und gezeigt, dass dieser sowohl eine physiologische Bedeutung haben, wie auch eine wichtige Rolle für die Stabilität des Sarkomers spielen kann. Darüber hinaus präsentiert diese Arbeit ein integriertes Modell aller verfügbaren hochauflösenden Strukturen von Myomesin, welches das bisher vollständigste Strukturmodell von Myomesin darstellt und die vorgeschlagene Myomesin-M-Band-Anordnung weiter unterstützt.

Zusammenfassend zeigt diese Dissertation an zwei unterschiedlichen Systemen, dass die strukturellen Informationen einen wichtigen Beitrag leisten, um die Natur und das Verhalten von Proteinen weiter zu verstehen und regulatorische Mechanismen abzuleiten. Jedoch ist es nur unter Betrachtung der nativen Umgebung und in Kombination mit weiteren Methoden möglich detailliertere Aussagen über das Verhalten der Moleküle in ihrem komplexen Gesamtbild zu treffen.

3 Introduction

3.1 The muscle

Muscles are responsible for preserving and altering posture and locomotion of an organism, but also generate motion of internal organs, such as contraction of the heart and movements of the digestive system. They are enabling flexibility and dynamics on the one hand while providing stability on the other. One can differentiate between different types of muscles depending on the presence or absence of regularly repeated arrangements of contractile compartments characterising smooth or striated muscle tissue. Smooth muscles are mainly responsible for the involuntary controlled contraction of hollow organs such as the vascular system or the gastrointestinal tract by applying long lasting but weak forces. In contrast, striated muscles are able to create strong and directed forces over short periods of time and can be further subdivided into skeletal and cardiac muscles. While the skeletal muscles, which are responsible for body movement and skeletal support (so-called micromovement), are under voluntary control, stimulated by the central nervous system, the cardiac muscle is only present in the heart, self-contracting and modulated by the autonomous nervous system.

Muscles are highly ordered, yet dynamic and hierarchical organised structures that are composed of muscle cells (myocytes) which have a length of several millimetres up to 10 cm and a diameter between 10 and 100 µm. The striated cells of cardiac and skeletal muscles are referred to as muscle fibres, formed in a process known as myogenesis. While skeletal myocytes are multinucleated, the majority of cardiomyocytes contain only one nucleus (Olivetti *et al.,* 1996). Muscle fibres are composed of membranous systems like the sarcoplasmic reticulum which stores calcium, mitochondria, or one or more nuclei, and most important myofibrils, the basic rod-like unit of a muscle cell (Figure 1). Myofibrils are specialised organelles consisting of long consecutive arrays of the fundamental contractile unit of muscle cells, the sarcomere.

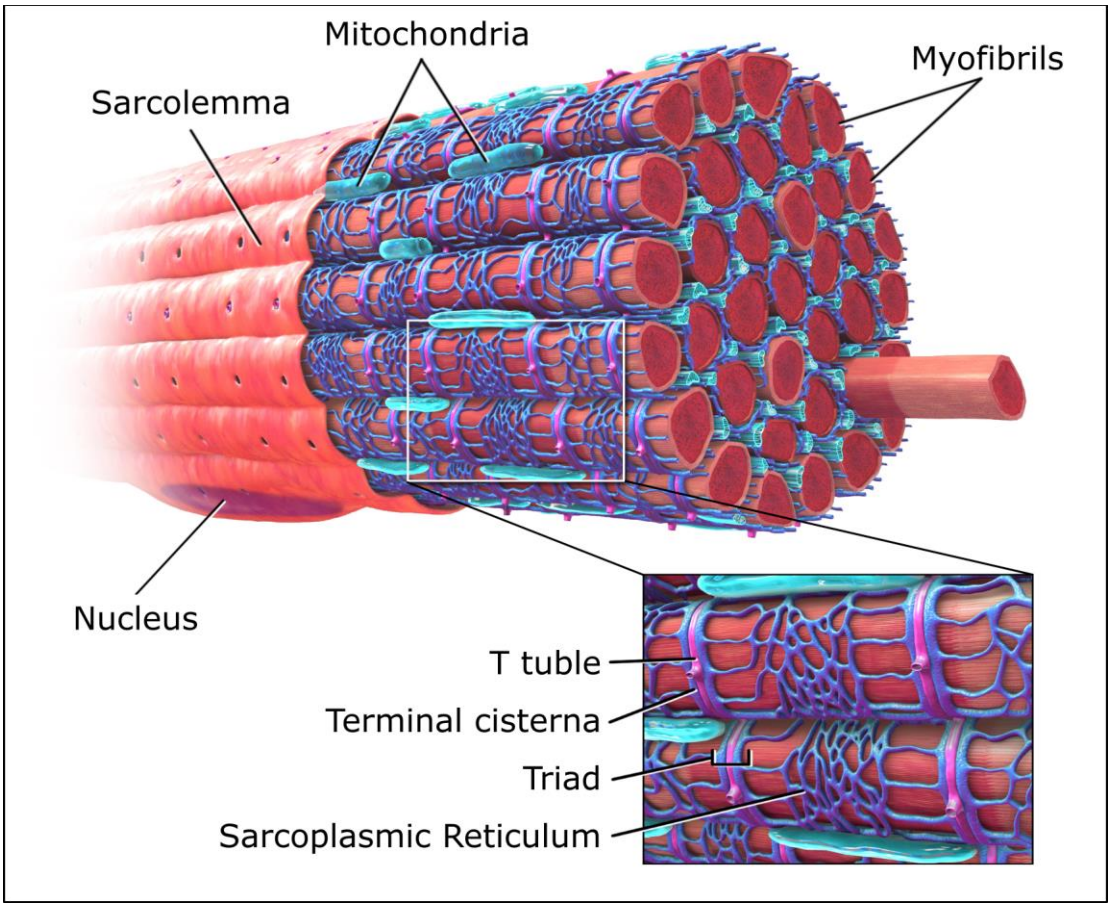

Figure 1 Illustration of a skeletal muscle fibre adapted from (Blausen.com staff, 2014). Skeletal muscle fibres are composed of bundles of myofibrils, mitochondria, one or more nuclei, the sarcoplasmic reticulum forming a tight network around the myofibrils, T tubules, facilitating fast transmission of the action potential, and the triad, a structure formed by a T tubule with a sarcoplasmic reticulum.

3.2 The sarcomere – a highly ordered molecular machinery

3.2.1 The general organisation of the sarcomere

Being one of the most complex cellular networks known the sarcomere possesses an almost crystalline order, mainly but not exclusively composed of the so-called thin and thick filaments. A huge variety of other proteins such as titin, obscurin and myomesin contribute to the structural integrity and the fine regulation of this highly dynamic structure. The thin filament is formed by polar double-helical assemblies of actin surrounded by tropomyosin and troponin-complexes which are covering the active binding sites for myosin (Ebashi, Endo and Ohtsuki, 1969). Actin filaments have a narrow length distribution that depends on the fibre type and normally exceeds

1 µm in length. The thick filament is a bipolar molecule precisely regulated to a length of 1.59 µm, containing 294 myosin molecules, or 147 per half thick filament (Tskhovrebova and Trinick, 2003). While numerous types of myosin molecules are known, each performing distinct functions such as vesicle transport (myosin type I) (Adams and Pollard, 1986) or cell division (myosin type VIII) (Reddy, A S., Day, 2001), the thick filament is assembled by myosin type II. Myosin type II which is also known as the conventional myosin is responsible for the force generation in muscle cells. Each myosin molecule is a hexamer composed of two heavy and four light chains which have regulatory and stabilising functions. Each heavy chain consists of three domains: i) an N-terminal head binding to actin and is responsible for the force generation by ATP hydrolysis, ii) a neck region mainly acting as a linker, introducing an angle between the head and the tail domain and interacting with the myosin light chains, and iii) the tail domain facilitating interaction with other myosin heavy chains.

Both, the thin and the thick filament pervade the myofibrils in highly ordered parallel arrays, partly overlapping with each other and thus creating the characteristic and repetitive compartments of the sarcomere. One sarcomere is defined as the segment between two neighbouring Z-disks (from German 'Zwischenscheibe') forming a

Z–I–A–H–M–H–A–I–Z

symmetry pattern ranging a distance of approx. 2.2 µm in a resting muscle (Figure 2). The actin filaments are anchored in the Z-disk which appears as a sequence of dark lines in micrographs of cross-striated muscle tissue. The Z-disk is surrounded by the I-band (for *isotropic* under a polarising microscope) a region of lower (optical) density where actin does not overlap with the thick filament. The actin filaments extend from the I-band into the adjacent A-band (for *anisotropic* under a polarising microscope), where it overlaps with the thick filament and contraction force is generated. The A-band spans the full length of a single thick filament and is bisected by the H-zone (from German 'hell' (bright) or Hensen after Victor Hensen) that appears brighter due to the lack of myosin head domains at the myosin stem region. In the centre of the H-zone is the M-band (from German 'Mittelscheibe'), forming the centre of the sarcomere. At the M-band myosin filaments of two neighbouring sarcomeres are anchored.

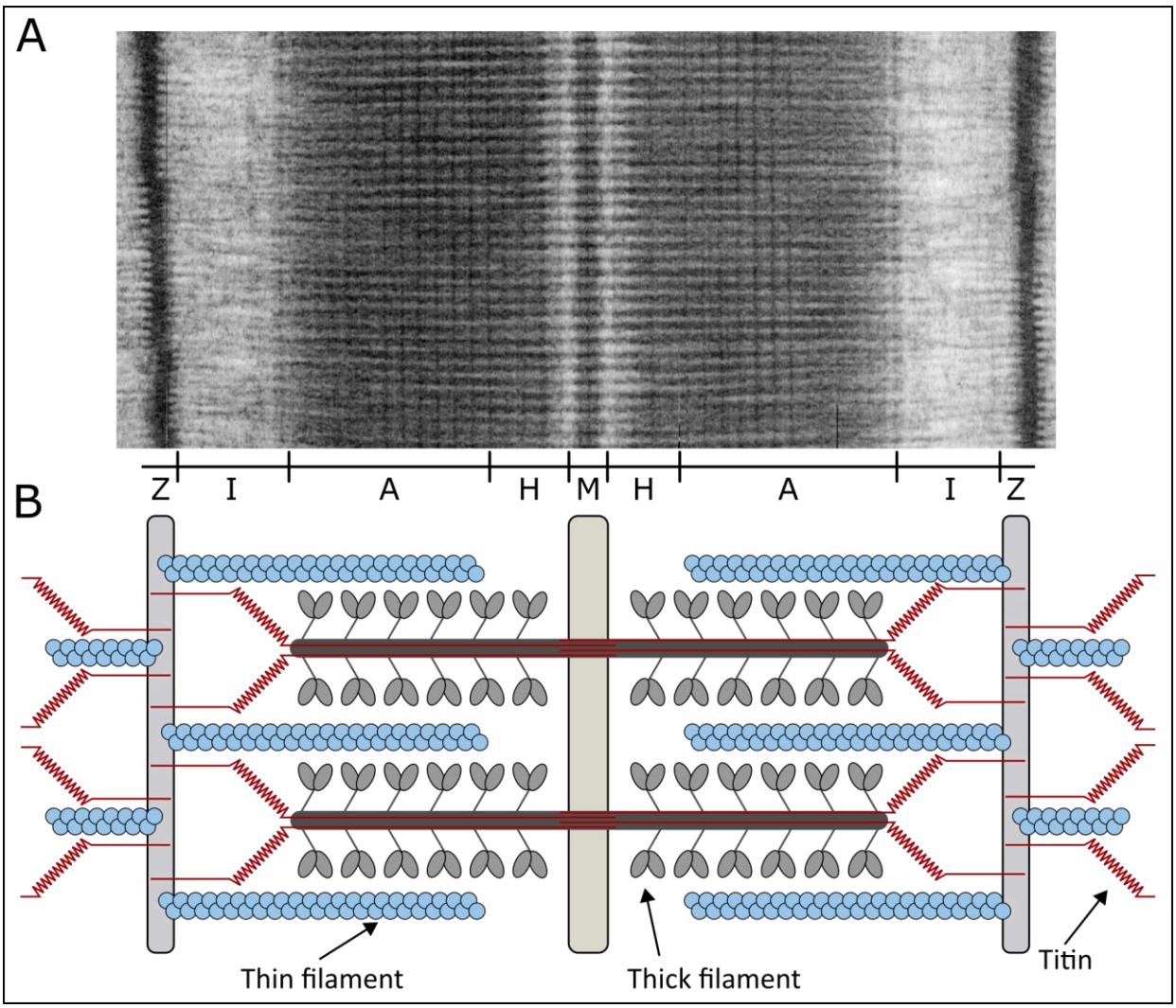

Figure 2 Sarcomeric structure. A) Representative electron micrograph of a sarcomere in relaxed conditions. Clearly visible is the cross-striated pattern of Z-disks, I-, A-and M-bands and the H-zone (adapted from (Squire *et al.*, 2005)*). B) Schematic illustration of the molecular interplay of the three main sarcomeric filaments actin (thin filament, light blue), myosin (thick filament, grey) and titin (red). Myosin head groups (light grey) are interacting with actin in a coordinated manner enabling force production (adapted from* (Sauer, 2011)*).*

3.2.2 Sarcomere contraction - the sliding filament model

Force generation in striated and cardiac muscle cells is driven by both active and passive modules. Actin and myosin as main components of the thin and thick filament allow active contraction of the sarcomeric structure. Force is generated by the simultaneous contraction of the muscle fibres, initiated by electric impulses derived from the nervous system or auto-excitation of electrically coupled cardiac muscle cells. Two different types of contraction can be differentiated: i) isotonic contraction where the length of the muscle changes, while the tension remains

constant and ii) isometric contraction where the overall muscle length remains the same, while the tension changes.

With the development of better microscopes experiments could demonstrate that during muscle contraction only the length of the I-band changes while the length of the A-band remains constant, leading to the 'sliding filament model' (Huxley and Hanson, 1954; Huxley and Niedergerke, 1954). The model describes contraction as a process in which the thin and the thick filament slide past each other, generating force by shortening of the sarcomere without changing the length of each filament. The process facilitating this movement is called the 'cross-bridging cycle': First, the head domains of myosin are elevated into a high-energy configuration. Therefore, adenosine triphosphate (ATP) is bound at the myosin heads and hydrolysed to adenosine diphosphate (ADP) and inorganic phosphate (P_i) by the ATPase activity of the myosin head domain. Energy is released and changes the angle of the myosin head into a so-called 'cocked' position, while the ADP and the inorganic phosphate remain attached. The contraction of the muscle is triggered by Ca^{2+}-release into the sarcomere. Ca^{2+} binds to the protein complex troponin leading to a conformational change of troponin which repositions another protein named tropomyosin - which in turn unprotects the active-myosin-binding site on the filamentous actin. Accordingly, the energetically excited myosin head binds to the binding site of actin, forming the actomyosin complex and establishing a cross-bridge between both filaments while releasing the bound P_i. Due to the P_i-release myosin undergoes a conformational change and pulls the actin filament approximately 10 nm towards the M-line, shortening the sarcomere and thus contracting the muscle (Reedy, Holmes and Tregear, 1965). After this so-called 'power stroke', ADP is released from the myosin head domain making it accessible for a new ATP molecule. Upon binding of a new ATP molecule, the actomyosin complex dissociates and positions myosin in its original orientation being able to start a new cycle. This finely regulated cascade of molecular processes, generating force and motion, is highly dependent on the stabilising properties of the anchor-points for the thin and the thick filaments, the Z-disk and the M-band. Additionally, the so-called third filament titin that spans half of the sarcomere connecting the Z-disk and the M-band, preserves the overall sarcomeric structure and maintains the resting tension of the muscle.

3.2.3 Anchoring regions of the sarcomere

During active contraction and passive stretching myofibrils must withstand high lateral and axial forces. At the same time, the axial distances between actin and myosin filaments, the interfilamentous distances and the length of the individual filaments need to be kept constant to enable highly coordinated force generation. This para-crystalline lattice is stabilised by multi-protein crosslinks at the Z-disk and the M-band.

3.2.3.1 The Z-disk

Discussed as the initial starting point of protein assembly, the Z-disk plays a pivotal role already in early myofibrillogenesis. In mature myofibrils, tight crosslinks *via* α-actinin form a stable meshwork anchoring the thin filaments and the N-terminal titin molecules of neighbouring sarcomeres and links sarcomeres laterally to the cell membrane *via* costameric proteins (Danowski *et al.,* 1992; Ervasti, 2003). Apart from these cytoskeletal and contracting filament complexes, numerous other molecules such as signalling molecules, enzymes or channel-proteins are known to assemble, permanently or transitionally, at the Z-disk exchanging and renewing in a remarkably dynamic manner (Wang *et al.*, 2005; Stout *et al.*, 2008; Ribeiro *et al.*, 2014). The thickness of the Z-disks is highly variable ranging between 30-100 nm (Luther, 2009) and controlled by the titin Z-repeats determining the number of α-actinin crosslinks (Gautel *et al.,* 1996; Sorimachi *et al.,* 1997). As it forms the longitudinal borders of the sarcomere the Z-disk is considered the stiffest of the anchoring planes in contracted and relaxed myofibrils, keeping the overall sarcomeric structure intact (Wakayama *et al.*, 2000; Yamada *et al.*, 2003; Akiyama *et al.*, 2006). This multifunctional role (Clark *et al.,* 2002; Frank *et al.,* 2006; Lange, Ehler and Gautel, 2006) renders the importance of the Z-disk within the sarcomeric structure.

3.2.3.2 The M-band

Located in the centre of the sarcomere, the M-band is the second anchoring point of the contracting filament complexes, crosslinking adjacent myosin filaments, maintaining the thick filaments hexagonal lattice and thus keeping the myosin motor domains aligned. With its higher elasticity compared to the Z-disk, the M-band is considered a structural protector of sarcomeric integrity upon force-generation (Agarkova *et al.*, 2003; Agarkova and Perriard, 2005). As revealed

by electron micrographs the M-band appears as an array of up to five parallel electron-dense lines, so-called M-bridge lines M6', M4', M1, M4, and M6 which are approximately 22 µm apart from one another (Sjöström and Squire, 1977; Pask *et al.*, 1994). The observed thickening of the M-bridges has been linked to another structure called M-filament which runs parallel to the thick filament (Luther and Squire, 1978).

Myosin filaments are transversely cross-linked by a ternary complex involving myomesin, and the long filamentous proteins titin and obscurin, maintaining the axial alignment and providing strong candidates to explain the M-bridge and M-filament observations. Myomesin directly crosslinks parallel myosin filaments by forming a C-terminal homodimer and interacting with myosin at the N-terminus. Additionally, it forms a complex with obscurin or its smaller analogue obscurin-like 1 (Pernigo *et al.*, 2017) which again forms a complex with titin (Pernigo *et al.*, 2010, 2015; Sauer *et al.*, 2010). These interactions could facilitate the binding of other ligands at the M-band (Pernigo *et al.*, 2015).

3.2.4 Assembly of the sliding filament – the sarcomerogenesis

The assembly of this highly complex network of hundreds of molecules is considered no less complex and still not fully understood. Numerous filament proteins need to assemble and arrange into the sarcomeric subunits of the myofibril, and additional proteins need to associate to these filaments. Due to the vast number and uncertain function of the identified components, intermediate states of the sarcomeric assembly are difficult to study and identify, and thus only partially characterised (Sanger *et al.*, 2010). The myofibrillogenesis of skeletal muscles follows a chronological programme of sequential activation of genes and integration of these proteins into the sarcomeric structures (Lin *et al.,* 1994). In contrast, most of the sarcomeric components in cardiac muscle cells are co-expressed before the first myofibrils are detectable, probably due to the very early development of the cardiac muscle compared to skeletal muscle (Ehler *et al.,* 1999). Different models exist to explain this complex process of myofibril formation:

(1) **The pre-myofibril model:** pre-myofibrils consisting of mini-sarcomeres form at the borders of muscle cells containing so-called Z-bodies with actin filaments, α-actinin and other sarcomeric proteins as well as non-muscle myosin II filaments. Next, titin and muscle myosin II

start to penetrate the pre-myofibrils which then start to align at the Z-body boundaries, forming Z-disks. The thick filaments assemble, and the myofibril maturates (Rhee, Sanger and Sanger, 1994; Du *et al.,* 2003; Sanger *et al.,* 2005).

(2) **The stitching model:** myosin filaments and I-Z-I-bodies (assemblies of primitive Z-disks and actin filaments) assemble independently. Accordingly, titin interacts with both structures and 'stitches' the thick filaments and I-Z-I bodies together to promote their assembly into mature myofibril (Schultheiss *et al.,* 1990; Holtzer *et al.,* 1997; Ojima *et al.,* 1999).

(3) **The template model:** the main components of the Z-disk, as well as the thin and the thick filament, assemble at actin stress fibre-like structures (SFLS) which serves as a temporary template (Dlugosz *et al.,* 1984).

(4) **Direct assembly model:** myofibrils form without intermediate states (Ehler *et al.,* 1999; Ehler, Fowler and Perriard, 2004).

(5) **Combined model:** this model combines observations made for the different models and will be explained in more detail:

Initially, proteins associated with the Z-disk, mainly the muscle-specific α-actinin, FATZ (filamin-, α-actinin-, and telethonin-binding protein at the Z-disk) and myotilin are expressed and assemble as Z-bodies on actin stress-fibre like structures (SFLS), which in turn form so-called premyofibrils together with the non-muscle isoform of myosin II (Sanger, Nicklen and Coulson, 1977; Dlugosz *et al.*, 1984; Rhee, Sanger and Sanger, 1994; Sanger *et al.*, 2002; Stout *et al.*, 2008). At this stage, titin epitopes associated with the Z-disk can be detected and are presumably responsible for anchoring alpha-actinin in the Z-bodies and aligning the Z-bodies (Tokuyasu and Maher, 1987; Schultheiss *et al.*, 1990; Turnacioglu *et al.*, 1996; Sorimachi *et al.*, 1997; Young *et al.*, 1998; Stout *et al.*, 2008). These assemblies of primitive Z-disks and actin filaments (IZI-bodies) are approximately 1 µm apart from each other and lack sarcomeric myosin.

The expression of thick filament components such as the myosin heavy and light chain and myosin-binding protein C is transcriptionally regulated by the helix-loop-helix factor Mef2 (Hinits and Hughes, 2007) and at this stage of sarcomerogenesis still inactivated. In a second step, the

premyofibrils are elongated forming an initial Z-M-Z structure with clearly separated Z- and M-epitopes of titin (van der Loop *et al.,* 1996; Yang, Obinata and Shimada, 2000). The stretching of titin exposes binding sites for other myofibrillar components and is seen as a crucial step for the assembly of the sarcomere as it initiates the assembly of myomesin at the M-bands (van der Ven and Fürst, 1997; Van der Ven *et al.,* 1999), allowing myosin filaments to integrate into the sarcomeric A-band. Finally, additional proteins migrate into the sarcomeric structure, completing the maturation of the myofibril.

3.3 Beads on a string - proteins of the sarcomeric cytoskeleton

Most proteins involved in the sarcomeric cytoskeleton consist primarily of long arrays of approximately 100 amino acid long immunoglobulin (Ig) and fibronectin type-III (FN-III) domains, both mainly possessing a β-sandwich structure. Ig domains are formed by seven to nine antiparallel β-strands arranged in two β-sheets with a main architecture of βA-βB-βE-βD and βG-βF-βC (Bork, Holm and Sander, 1994). Different types of Ig domains differ by their Greek-key arrangement as well as the number and size of β-strands and hairpin regions. Muscle protein associated Ig domains usually belong to the I (intermediate)-type (Harpaz and Chothia, 1994), which have a split A-strand, resulting in very short A and A' strands that belong to the two different β-sheets in the fold (Figure 3). FN-III domains are mainly formed by seven antiparallel β-strands arranged in two β-sheets with a main architecture of βA-βB-βE and βG-βF-βC-βD (Leahy, Aukhil and Erickson, 1996).

The Ig and FN-III domains are arranged in a 'beads on a string'-like way initially identified for twitchin, titin and Unc-89 (Benian *et al.,* 1989, 1996). Even though the similarity between several identified sarcomeric proteins is high (e.g. twitchin and unc-89 are considered as analogues of titin and obscurin) substantial differences in the interactions and the activity of signalling domains exist. Notably, striated muscles of all known species contain sarcomeric cytoskeleton proteins composed of Ig and FN-III superfamily domains indicating that ancestral molecules evolved as parts of the first striated muscles even before the Cambrian explosion (Benian *et al.,* 1989, 1996; Higgins *et al.,* 1994; Fürst and Gautel, 1995; Tskhovrebova and Trinick, 2003; Burkart *et al.,* 2007).

Even though Ig and FN-III domains show low sequence identities in general (typically <30%) the secondary structure is highly conserved (Leahy, Aukhil and Erickson, 1996; Fattorusso *et al.,* 1999). Taking together, Ig and FN-III domains are able to provide stability that is required for mechanical integration while being able to form individual and highly specific protein-ligand interactions (Chatzifthimiou, Hornburg and Sauer *et al.,* submitted). Titin, obscurin, and myomesin are three prominent representatives of sarcomeric proteins and will be discussed in detail in the following chapters.

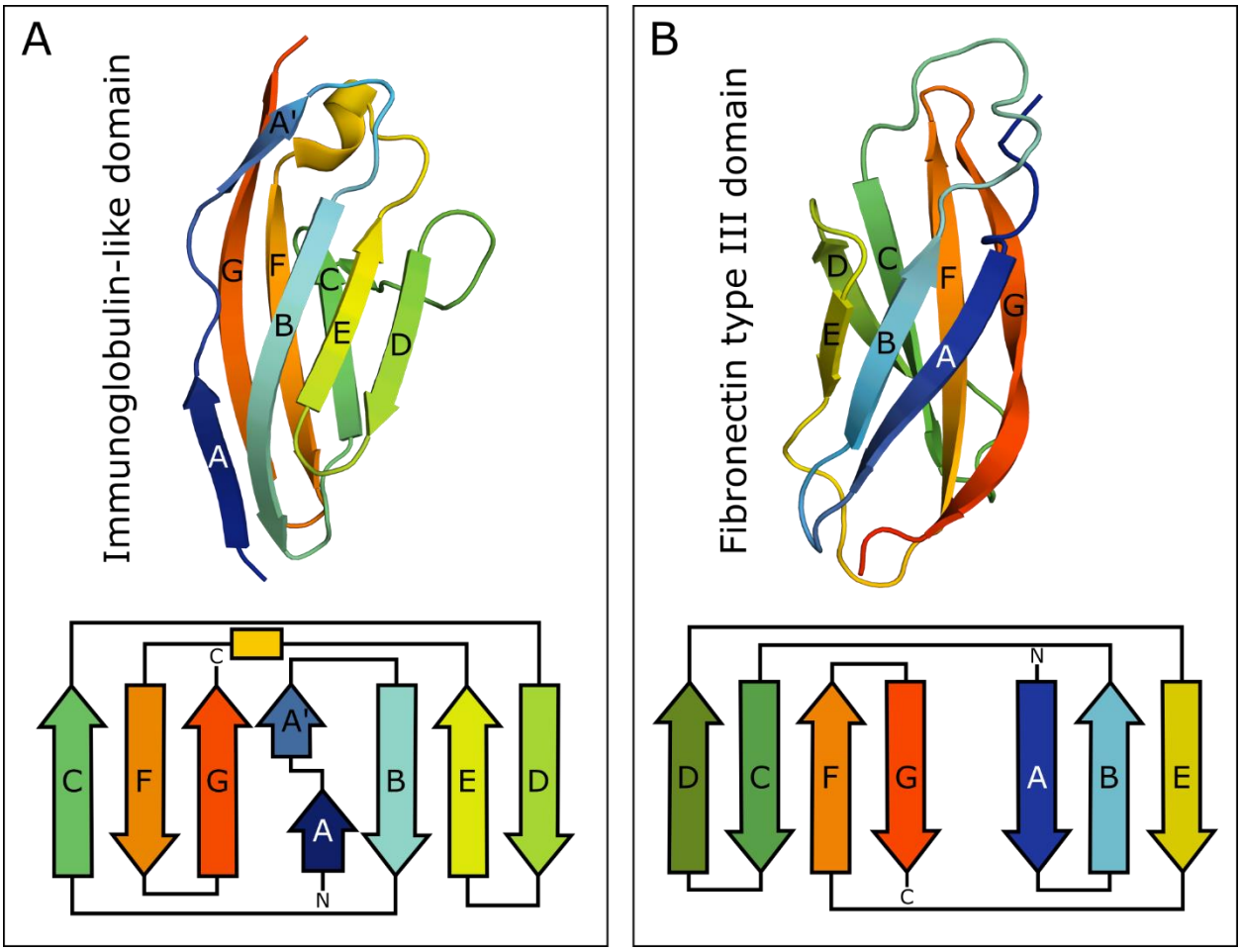

Figure 3 Architecture of Ig and FN-III domains. Cartoon (top) and schematic (bottom) representation of a typical A) I-type immunoglobulin-like domain (titin M7, PDB: 3PUC) and B) fibronectin type III domain (titin A77, PDB: 3LPW). Colour scheme from blue as the N-terminus to red as the C-terminus. Arrows indicate β-strands, box indicate α-helices. β-sheets are formed by neighbouring β-strands.

3.3.1 Titin holds the sarcomere together

Even though there is evidence that the so-called third filament, the long filamentous protein titin, also has a role in active contraction of the muscle (Horowits *et al.*, 1986; Konhilas, Irving and de Tombe, 2002) it is mainly considered a 'blueprint'-protein and molecular spring. It passively preserves the overall sarcomeric structure by maintaining the resting tension of the muscle on the one hand and functioning as a ruler that guides numerous proteins to their specific spatial location on the other. A complete deletion of the titin gene leads to a disruption of the whole myofibril formation in which sarcomeric myosin is not regularly assembled and randomly distributed (Person *et al.,* 2000; van der Ven *et al.,* 2000). Titin has a modular organisation consisting of up to 244 individually folded Immunoglobulin (Ig) and Fibronectin-3 (FN-III) domains (Figure 4) and - with a length of more than 1.0 µm - spans half of the sarcomere connecting the peripheral Z-disk and the central M-band (Labeit, Kolmerer and Linke, 1997; Tskhovrebova and Trinick, 2003; Ehler and Gautel, 2008). It is encoded by one of the largest known genes: *TTN*, which contains 364 exons, translating to numerous tissue-specific protein isoforms (Bang *et al.*, 2001).

From the Z-disk, where the N-terminus of titin localises, to the M-band, where the C-terminal part integrates, titin is responsible for a large variety of tasks (Figure 4). At the Z-disk titin functions as an anchor permitting force transmission by binding to multiple proteins, for example the Z-disk-cross-linking α-actinin (Young *et al.*, 1998) and Telethonin (Pinotsis *et al.*, 2006), it binds to regulatory molecules like the Ca^{2+}-dependent protease calpain-1 (Raynaud *et al.*, 2005) and proteins of the sarcomeric cytoskeleton, such as actin and tropomyosin (Linke *et al.*, 1997; Trombitás, Greaser and Pollack, 1997; Raynaud, Astier and Benyamin, 2004) and obscurin (Young, Ehler and Gautel, 2001), and also regulates the thickness of the Z-disk through up to seven alternatively-spliced Z-repeats (Gautel *et al.*, 1996; Sorimachi *et al.*, 1997). Notably, the function of the interaction between titin and the long myofibrillar protein obscurin at the junction between Z-disk and I-band is still unclear. The I-band region of titin contains highly repetitive PEVK domains (named due to the high content of proline (P), glutamate (E), valine (V), and lysine (K)) that provide elastic properties and passive tension and help to maintain the Z-disk and M-band anchoring during contraction and elongation of the muscle (Linke *et al.*, 2002; Huebsch *et*

al., 2005). In the A-band region titin is responsible for the correct assembly of the thick filament by interacting with myosin (Bang *et al.*, 2001) and myosin-binding protein C (MyBPC) (Maruyama *et al.*, 1987; Labeit *et al.*, 1992; Okagaki *et al.*, 1993; Freiburg and Gautel, 1996) and thus stabilizing the sarcomere. Finally, at the M-band titin molecules from two neighbouring sarcomeres overlap antiparallel orientation (Fürst *et al.*, 1988; Sauer *et al.*, 2010), forming a complex network of interactions with the main crosslinking M-band molecule myomesin, and obscurin (Fukuzawa *et al.*, 2008; Pernigo *et al.*, 2010, 2015, 2017; Sauer *et al.*, 2010). Additionally, a titin kinase domain at the periphery of the M-band is thought to modulate titin expression and turnover, and to function as a sensory and signalling mediator (Labeit and Kolmerer, 1995; Mayans *et al.*, 1998; Trinick and Tskhovrebova, 1999; Bang *et al.*, 2001; Miller *et al.*, 2003; Lange, Xiang, *et al.*, 2005).

Remarkably, while the majority of splicing events affect regions that encode for the sarcomeric I-band region of titin and thus, altering its elastic spring function (Freiburg *et al.*, 2000), constitutively expression of nearly all titin exons encoding for the N-terminal Z-disk, the A-band, and the C-terminal M-band regions can be observed. With its enormous size, so far 127 *TTN* disease-causing mutations have been identified and linked to at least ten different pathological conditions such as childhood diseases affecting both types of striated muscles, purely skeletal muscle phenotypes and isolated cardiomyopathies (Chauveau, Rowell and Ferreiro, 2014). Due to the huge amount of titin related disease phenotypes, a new category was introduced describing titin related diseases, called titinopathies (Gerull *et al.*, 2002). This indicates the importance of titin related research to understand titins role in the complex sarcomeric interaction network.

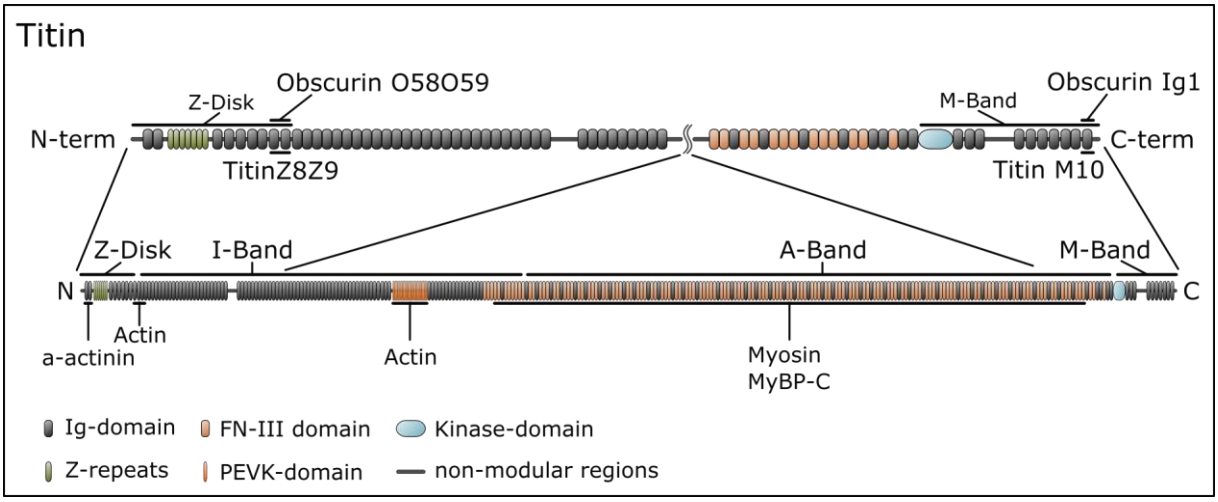

Figure 4 Schematic overview of titin. The top part shows a zoomed-in illustration of the Z-disk and M-band region of titin indicating the two interaction interfaces formed by titin and obscurin. The bottom part shows the main interactions with proteins of the sarcomeric cytoskeleton actin, myosin, α-actinin, and myosin binding protein C (MyBP-C).

3.3.2 Obscurin – a multitasking giant

As described above, one of the interaction partners of titin is obscurin. Obscurin - like titin - is a long filamentous sarcomeric protein. The human OBSCN gene spans 170 kb and contains 119 exons, encoding for several 'giant' (approx. 800 kDa) and numerous smaller (40–500 kDa) obscurin isoforms (Young, Ehler and Gautel, 2001; Fukuzawa, Idowu and Gautel, 2005; Kontrogianni-Konstantopoulos *et al.,* 2009; Perry *et al.,* 2013; Ackermann *et al.,* 2014).

Named after the adjective 'obscure' due to the initial difficulties in detection and characterisation, nowadays, obscurin is connected to a whole family of polypeptides expressed from a single gene. Two giant isoforms are known, only differing in the composition of the C-terminus that either contains a non-modular region with an ankyrin-binding domain (isoform A; ~720 kDa) or an elongated 'signalling-end' with two active Ser/Thr kinase motifs, two additional Ig domains, and another FN-III domain (isoform B~800 kDa), respectively (Figure 5). Both isoforms share a common N-terminal part consisting of 53 Ig and three FN-III domains, a calmodulin-binding IQ motif, a SRC (sarcoma Tyrosinkinase) homology-3 domain (SH3) motif, a tandem Rho-guanine nucleotide exchange factor (RhoGEF) and a pleckstrin homology (PH) domain.

Initially identified as a ligand for the titin domains Z8 and Z9 located at the sarcomeric Z-disk, obscurin has been shown to localise to many other cellular locations such as the M-band, A/I junctions, costameres, which connect the sarcomere to the sarcolemma, to the sarcoplasmic reticulum (SR), responsible for Ca^{2+} homeostasis, and, in cardiomyocytes, to intercalated discs, membrane structures enabling electrical and mechanical communication to synchronise contraction (Kontrogianni-Konstantopoulos, 2006; Bowman *et al.,* 2007; Fukuzawa *et al.,* 2008; Shriver *et al.,* 2015). Numerous ligands have been identified to interact with obscurin at these diverse cell compartments (Figure 5). Obscurin was preferentially investigated at the M-band level and has been reported to interact with C-terminal titin, myomesin, ankyrin-B, RhoA, and sMyBP-Cv1 (slow skeletal myosin binding protein C variant 1) (Perry *et al.,* 2013). At the Z-disk obscurin has been shown to interact with the N-terminus of titin and RanBP9 (Ran-binding protein 9), which is involved in the titin incorporation into the Z-disk (Bowman *et al.,* 2008). Furthermore, obscurin interacts with ankyrin-B at costameres (Randazzo *et al.,* 2013), with small ankyrin-1 at the SR (Kontrogianni-Konstantopoulos *et al.,* 2003), and with N-cadherin and Na+/K+ ATPase at the intercalated discs of cardiac muscle cells (Hu and Kontrogianni-Konstantopoulos, 2013).

With these multiple locations and various ligands, obscurin presents itself as a highly multitasking molecule fulfilling various functions that involve mediating sarcomeric assembly and stability (Borisov *et al.,* 2006; Kontrogianni-Konstantopoulos, 2006; Spooner *et al.,* 2012), preserving the integrity of the sarcolemma (Randazzo *et al.,* 2013), connecting the sarcomeric cytoskeleton and the SR, possibly contributing to Ca^{2+} homeostasis and intercellular adhesion (Hu and Kontrogianni-Konstantopoulos, 2013), and signal transduction (Young, Ehler and Gautel, 2001).

Interestingly, the localisation of obscurin within the sarcomere is controversial, as different reports show a change of obscurin localisation during myofibrillogenesis, however, with contradictory observations. While some studies show a shift from the Z-disk to the M-band during myofibrillogenesis (Young, Ehler and Gautel, 2001; Carlsson, Yu and Thornell, 2008) others report that obscurin is first observed at M-bands and later localised to Z-disks (Borisov *et al.,* 2004; Kontrogianni-Konstantopoulos and Bloch, 2005; Borisov, Martynova and Russell, 2008). This once more underlines the obscure nature of obscurin. Remarkably, during development

obscurin isoform A, lacking the kinase domains, localises preferentially at the M-bands probably taking over scaffolding functions such as connecting the sarcomere with the SR, other plasma membranes or neighbouring myofibrils (Agarkova and Perriard, 2005). The kinase-containing isoform B of obscurin localises preferentially at the Z-disk (Raeker *et al.,* 2006), which could indicate a more regulatory function at this side of the sarcomere. It is discussed that the Z-disk interaction might either link signalling pathways established by C-terminal SH3-, calcium-calmodulin- and small GTPase domains with the forming Z-disk during myofibrillogenesis or regulate the reorganisation of the sarcomere according to stress, development or diseases as reviewed by Kontrogianni-Konstantopoulos et al. (2009). Genomic analysis identified multiple variants within the OBSCN gene directly linked to the development of hypertrophic (HCM) and dilated (DCM) cardiomyopathy as well as left ventricular noncompaction (Arimura *et al.,* 2007; Marston *et al.,* 2015; Xu *et al.,* 2015; Rowland *et al.,* 2016; Hu *et al.,* 2017).

It remains elusive how the localisation of obscurin is regulated. Whether obscurin, as being a scaffolding and sarcomeric ruler protein itself, is guided by other molecules like titin has not been finally assessed. Remarkably, two known interaction interfaces between obscurin and titin located at the Z-disk and the M-band might indicate guidance of one another (Young, Ehler and Gautel, 2001; Fukuzawa *et al.*, 2008).

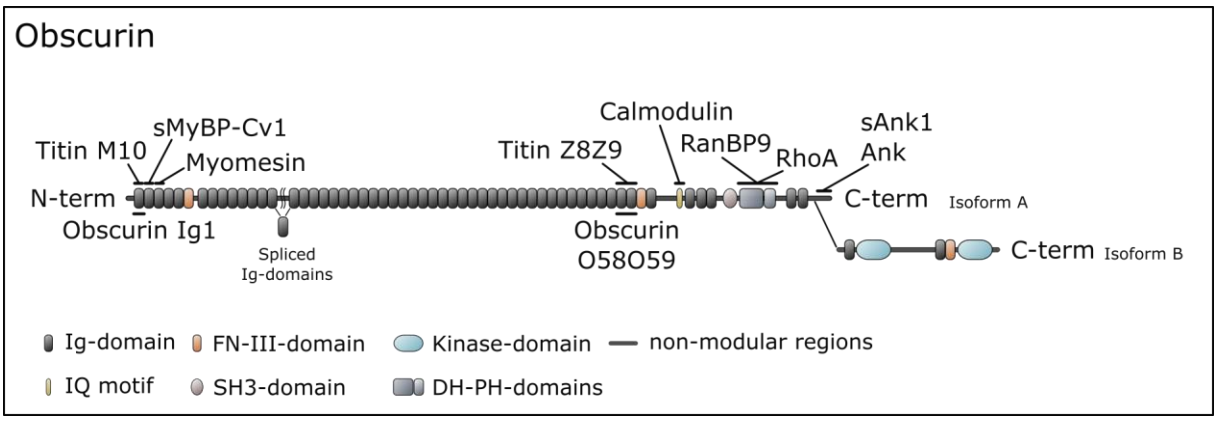

Figure 5 Schematic overview of obscurin showing the two giant isoforms A and B and the main interactions of obscurin. Isoform A contains an ankyrin binding site whereas isoform B comprises two kinase domains. The two interaction sites with titin at the Z-disk and the M-band, as well as the interaction with myomesin, are indicated.

3.3.3 Titin obscurin complex at different sites of the sarcomere

Titin and obscurin interact with each other at two distinct parts of the sarcomere. One interaction takes place at the M-band, one at the Z-disk, two locations more than 1.0 µm apart from each other (Figure 5). The interaction at the M-band between the very C-terminal titin domain M10 and the very N-terminal obscurin domain Ig1 or obscurin-like 1, respectively, has already been structurally determined and functionally described (Pernigo *et al.*, 2010, 2015; Sauer *et al.*, 2010). It has been shown that M-band titin and obscurin interact *via* β-sheet augmentation – a widespread Ig-Ig protein interaction (Remaut and Waksman, 2006) – in which the β-strand G of obscurin Ig 1 interacts with the β-strand B of titin M10 (Pernigo *et al.*, 2010, 2015; Sauer *et al.*, 2010). The interaction is part of the triangular interaction between titin, obscurin and myomesin, tightly interlaced, cross-linking neighbouring thick filaments and presumably stabilising the M-band structure. Interestingly, while obscurin, which contains multiple signalling and kinase domains or an ankyrin binding site, localises favourably at the M-band periphery (Bagnato *et al.,* 2003; Kontrogianni-Konstantopoulos *et al.,* 2003), its smaller isoform obscurin-like 1 localises at the myofibril core (Fukuzawa *et al.,* 2008) and does not contain these specialised domains, fulfilling mainly scaffolding and architectural functions. Interestingly, this is the only known case where a sarcomeric cytoskeleton domain forms an interaction with two different proteins with a different function in different localisations (Benian and Mayans, 2015).

In contrast to this, only little is known about the function and interaction between titin and obscurin at the Z-disk in which obscurin Ig domains 58 and 59 (O58 and O58) form a complex with titin Ig domains 8 and 9 (Z8 and Z9) (Young, Ehler and Gautel, 2001). Ultrastructurally, the titin Z8 and Z9 domains are located at a comb-like transition region of the peripheral Z-disk, close to the I-band (Gautel *et al.,* 1996; Yajima *et al.,* 1996). It has been shown that all four domains are necessary for the complex formation. While individual domains of titin Z8Z9 show a strongly reduced interaction with obscurin O58O58, the individual obscurin Ig domains were shown to not at all interact with titin Z8Z9 (Young, Ehler and Gautel, 2001; Arimura *et al.*, 2007; Rossi *et al.*, 2017). Interestingly, a recent publication presents NMR data suggesting that only titin Z8 is involved in the complex formation and proposing that titin Z9 is likely necessary to stabilise the Ig fold of titin Z8 (Wright and Letourneau, 2018). Even though being the very first interaction

described for obscurin and even more the origin of obscurins identification, the true nature and structure of the Z-disk titin–obscurin complex remains unknown. Numerous studies focusing on this interaction only were able to elucidate smaller parts of the whole complex. Three out of four single domains of the titin-obscurin complex have been structurally solved, namely obscurin Ig 58 (Hu *et al.*, 2017), obscurin Ig 59 (Rossi *et al.,* 2017), and titin Z9 (Wright and Letourneau, 2018), but neither a clear functional role nor the structure of the whole titin-obscurin complex has been described.

3.3.4 The titin-obscurin complex has been connected to a wide range of myopathies

The interaction between protein domains is often dependent on highly specific residue-residue interactions. Already small changes within these interfaces might lead to severe weakening effects or even a total disruption of the interaction interfaces and might, therefore, be causative for the development of various disease phenotypes. Variants located in domains involved in both titin-obscurin complexes have been connected to the development of myopathies. At the M-band several reports demonstrate a connection between variants that are directly located within the titin-obscurin binding site to tibial muscular dystrophy (TMD) and limb-girdle muscular dystrophy type 2J (LGMD2J) possibly abolishing the interaction between titin and obscurin (Udd and Griggs, 2001; Pollazzon *et al.*, 2010; Udd, 2012; Evilä *et al.*, 2016).

For the titin-obscurin complex at the level of the Z-disk, two variants R4344Q and R4444W, both located on the interacting domains of obscurin, have been reported to be involved in myopathies (Arimura *et al.*, 2007; Hu *et al.*, 2017; Rossi *et al.*, 2017). Arimura et al. (2007) and Hu et al. (2017) report the mutation R4344Q that is located in obscurin O58. The variant was in a 19-year-old male with hypertrophic cardiomyopathy (HCM) but not in 576 control-chromosomes or the unaffected father and thus was presumably inherited from the mother who also suffered from HCM (Arimura *et al.*, 2007). A modelling approach of the obscurin domains O58 and O59 indicated a conformational change on the surface of O58 due to the R4344Q variant. Furthermore, it was demonstrated that the binding of obscurin O58O59 to titin Z8Z9 was decreased in both, mammalian two-hybrid and co-immunoprecipitation assays. Therefore, the

authors concluded that the R4344Q variant is able to affect the incorporation of obscurin into the Z-disk ultimately resulting in an impaired recruitment of the 'signalling-end' of obscurin at the Z-disk – I-Band junction and the development of HCM (Arimura *et al.*, 2007).

In contrast to this conclusion, a second study demonstrates a gain-of-function mechanism by an increased binding of obscurin M-band O58 to PLN, a protein that regulates the Ca^{2+} homeostasis by binding to the Ca^{2+}-pump SERCA2 resulting in the development of dilated cardiomyopathy (DCM) (Hu *et al.*, 2017). In a constitutive knock-in mouse model expressing full-length obscurin carrying the R4344Q mutation, initially, no clear cardiomyopathic phenotype could be observed. Evidence, however, was found that the R4344Q variant leads to an up-regulation and over-activation of the sarco-/endoplasmic reticulum Ca^{2+}-ATPase (SERCA2) increasing the transfer of Ca^{2+} from the cytosol of the cell to the lumen of the SR. This results in accelerated kinetics of the contraction and relaxation cycle, increased pressure, and consequently to the development of spontaneous ventricular arrhythmia and DCM by the age of one year, when subjected to pressure overload (Hu *et al.*, 2017). The authors claim that the dysregulation of SERCA2 is the result of a disinhibition of the same, due to an increased binding of monomeric phospholamban (PLN) to the mutant obscurin. Presenting a high-resolution structure of O58, they were able to map the R4344Q variant on the domain surface and demonstrated in molecular modelling simulations improved electrostatic interactions and tighter binding between obscurin and PLN due to the variant. In pull-down assays using obscurin O58 and cell-homogenate, they demonstrated a 10-fold increased binding of PLN to the mutated obscurin. These findings indicate an additional function of the obscurin domain O58 apart from binding to titin at the Z-disk.

The second variant R4444W described by Rossi et al. has been found in patients affected by distal myopathy that also carry a frameshift mutation in the FLNC gene encoding filamin-C, an actin-binding-like protein located at the Z-disk (Razinia et al., 2012; Rossi et al., 2017). The authors show that the variant has no impact on the correct domain folding and demonstrate in an *in vitro* pull-down assay and surface plasmon resonance spectroscopy (SPR) that the mutation of obscurin O58O59 decreases the binding to titin by ~15-fold. A frameshift mutation only one nucleotide upstream of the mentioned one is known to be linked with distal myopathy (Guergueltcheva *et al.,* 2011). Since the FLNC mutation reported by Rossi et al. (2017) was also

present in one healthy individual not carrying the R4444W variant the authors conclude that the obscurin mutation might contribute to full expression of the myopathic phenotype.

3.3.5 Myomesin – the elastic crosslinker of the M-band

Members of the myomesin family are integral proteins of the sarcomeric cytoskeleton of the M-band and highly responsible for crosslinking the thick filament at the M-band (Grove *et al.*, 1984). The family consists of three isoforms: myomesin-1, M-protein (myomesin-2) and myomesin-3, all encoded by the MYOM genes (Schoenauer *et al.,* 2008). Myomesin-1 is constitutively expressed in all striated muscles and, in contrast to the other isoforms, thought to be crucial for myofibril integrity, as cellular knockdown results in a complete disruption of the sarcomeric structure (Fukuzawa *et al.,* 2008). Myomesin-2 and -3 are expressed only in the adult heart and fast fibres, or in skeletal muscle intermediate fibre types, respectively (Agarkova *et al.,* 2004; Schoenauer *et al.,* 2008, 2011). The crosslinking of myosin by myomesin-1, henceforth called myomesin, is established by two interaction interfaces: i) a homotypic Ig–Ig interaction of two C-terminal myomesin domains (My13) forming an antiparallel dimer (Lange, Himmel, *et al.*, 2005; Pinotsis *et al.*, 2008, 2012) and ii) an interaction of the very N-terminal myosin-binding domain with myosin (Obermann *et al.*, 1997). While the first interaction has been studied extensively, the N-terminal myosin binding is not yet fully understood.

The highly divergent N-termini of the different myomesin isoforms and the numerous phosphorylation sites within these domains indicate different interaction sites with myosin on the one hand and a highly dynamic regulation on the other. Remarkably, while crosslinking myosin filaments, myomesin also shows a unique architecture providing highly elastic properties at the C-terminal region, enabling the dimeric myomesin-1 molecule to extend by approx. 50 nm and possibly allowing reversible deformation of the sarcomeric M-band (Berkemeier *et al.,* 2011; Pinotsis *et al.,* 2012). To enable this, Ig domains My9 to My13 are connected by short helical linkers which can fast and reversibly un- and refold at forces between 15 to 40 pN. These forces are much lower than what is necessary for the unfolding of individual Ig domains (>80 pN) (Schoenauer *et al.,* 2005), or the dissociation of the My13 dimer (>130 pN) (Berkemeier *et al.,* 2011; Pinotsis *et al.,* 2012; Xiao and Gräter, 2014). As mentioned before, myomesin (but not

myomesin-2 and -3) is additionally part of a triangular interaction with titin and obscurin at the M-band. Myomesin FN-III domain 4 interacts with the third Ig domain of obscurin or the smaller homologue obscurin-like-1 (Fukuzawa *et al.,* 2008; Pernigo *et al.,* 2017) (Figure 6). This mechanically stable complex, which is able to resist forces of up to 135 pN, is formed by a trans-complementation mechanism in which the interdomain sequence between FN-III domains My4 and My5 of myomesin are incorporated into an incomplete immunoglobulin-like domain of obscurin (Pernigo *et al.*, 2017). Furthermore, Pernigo et al. (2017) present evidence for a tetrameric complex involving twice the My4-O3 heterodimer indicating an additional junction within the M-band meshwork (Figure 6).

Interestingly, unpublished high-resolution structural data acquired by Sauer *et al.* (EMBL-Hamburg) additionally suggest yet another junction by homo-dimerisation of the fifth domain of myomesin. Again, an unusual mechanism called three-dimensional domain swapping first described by (Bennett, Schlunegger and Eisenberg, 1995; Heringa and Taylor, 1997) was observed. Three-dimensional domain swapping is a process in which two or more identical proteins exchange either a whole folded domain or a smaller secondary structural part, trans-complementing one another's structure (Rousseau, Schymkowitz and Itzhaki, 2012). Except for the hinge region, which links the exchanging domains, all phi and psi backbone angles remain unaltered. Besides the so-call primary interface that represents the interaction that is complemented by the swapped domain, a secondary interface that is not present in the monomeric form often forms between the two interchanging domains due to their close proximity (Gronenborn, 2009). For the My5 dimer, the N-terminal β-strands of both monomers are exchanged forming an elongated β-sheet, each complementing the FN-III domain of the one another (Sauer *et al.,* unpublished).

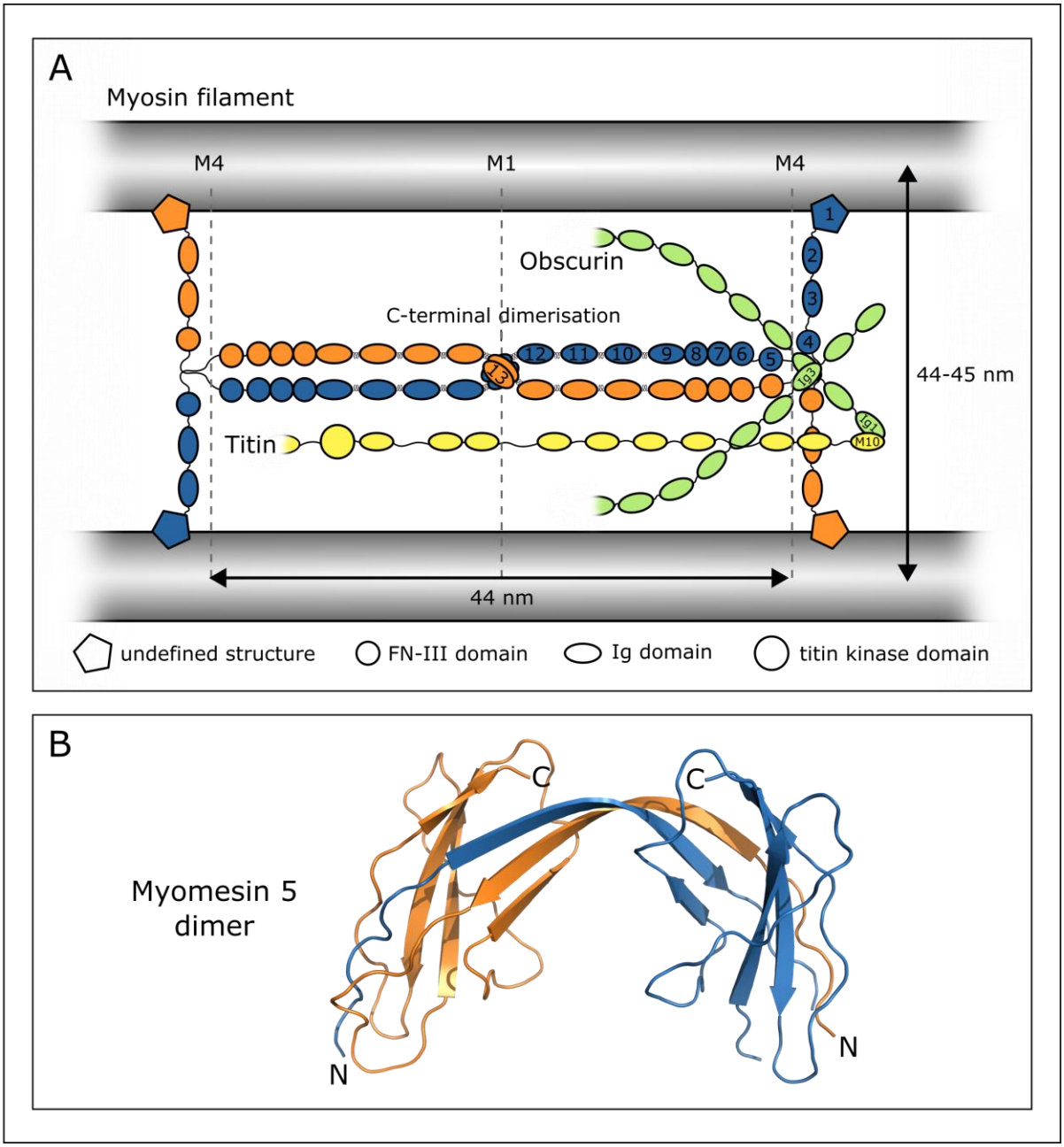

Figure 6 Myomesin interactions. A) the current model of the myomesin arrangement in the sarcomeric M-band derived from localisation studies (Obermann *et al.*, 1997) *and interaction interfaces found for M-band proteins* (Fukuzawa *et al.*, 2008; Pernigo *et al.*, 2010, 2015, 2017; Sauer *et al.*, 2010). *Myomesin domain 13 dimerises at the M1 line. Interaction between titin M10 and obscurin Ig1 and the complex forming between myomesin My4 and obscurin Ig3 (or obscurin-like 1 Ig3) are located around M-band line M4. (adapted from* (Sauer, 2011; Pernigo *et al.*, 2017)*) B) Homo-dimerisation myomesin FN-III domain 5. The N-terminal β-strands of both monomers are exchanged forming an elongated β-sheet. (Sauer et al., unpublished).*

3.4 The sarcomere as a common target for hereditary myopathies

The sarcomeric contraction machinery is a highly ordered and finely regulated system. Proteins of the sarcomeric cytoskeleton are common causes of hereditary myopathies. It is a common concept that myopathies are diseases of mendelian rather than of complex origin, and thus controlled by a single locus (Arbustini *et al.,* 2000; Maron, 2002; Guglieri *et al.,* 2008). In a Mendelian disease, a single gene can cause a pathogenic phenotype, while complex diseases are influenced by a combination of multiple genes and environmental factors. As detailed above, studies link numerous sarcomeric proteins like titin, obscurin and myomesin but also α-actinin, filamin-C, telethonin, myopalladin and myosin-binding proteins to the development of hereditary diseases (Chauveau, Bönnemann, *et al.*, 2014; Lyon *et al.*, 2015; Tardiff *et al.*, 2015). Especially for long filamentous proteins the effect of truncating mutations, which introduce an artificial stop codon in the gene sequence and thus induce a premature transcriptional termination, has often been linked to myopathies. Nevertheless, also single mutations are directly connected to myopathies. Depending on the protein involved and the nature of the inherited change, different types of muscle tissue, namely cardiac, skeletal or both, may be affected.

As being part of cardiovascular diseases that are causative for one of every three deaths worldwide (Mozaffarian *et al.,* 2016), the exploration and identification of factors causing cardiomyopathies is one of the big challenges to date. Cardiomyopathies ultimately affect the capability of the heart to contract and/or fill with blood (Coronel, de Groot and van Lieshout, 2001; Tan *et al.*, 2010; Savarese and Lund, 2017). Different types of cardiomyopathies are known and can be differentiated into the common hypertrophic and dilated cardiomyopathy, and the less common restrictive and unclassified cardiomyopathies as well as the rare arrhythmogenic right ventricular dysplasia. While hypertrophic cardiomyopathy (HCM) is mostly inherited and indicated by enlargement and thickening of the heart muscle causing blockage, narrowing, or increased stiffness of the ventricles, dilated cardiomyopathy (DCM) results in an enlargement and weakening of the ventricles and a wall thinning. DCM is inherited for about one-third of the patients but can also be induced by different other factors like a heart attack, high blood

pressure, diabetes, thyroid disease but also substances like alcohol, heavy metals, cocaine or amphetamines.

Less common is the group of restrictive cardiomyopathies, caused by a stiffening of the ventricles without thickening of the walls. Arrhythmogenic right ventricular dysplasia describes a condition where the muscle tissue in the right ventricle is replaced with fatty or fibrous tissue, disrupting electrical signal transduction, which causes arrhythmias.

Besides cardiac muscles, hereditary myopathies can also affect skeletal muscle tissue leading to reduced function and loss (atrophy) of skeletal muscle tissue in different compartments of the body. Two skeletal myopathies frequently linked to variants located in muscle proteins like titin and obscurin are tibial muscular dystrophy (TMD) and limb-girdle muscular dystrophy type 2J (LGMD2J). TMD describes a distal myopathy with late-onset, typically after the age of 35 (Udd, 2012) and is indicated by a weakening and atrophy of the lower leg muscles, especially the anterior shinbone (tibial) muscle. Interestingly, even though affecting sarcomeric scaffold proteins like titin, the sarcomeric structure remains intact (Udd *et al.,* 2005). Interestingly, TMD has never been reported to coincide with any type of cardiomyopathy. LGMD2J is developed within the first two decades of life and affects the pelvic or the shoulder girdle muscles by weakening and wasting of the muscle tissue (Hackman *et al.,* 2008). As big data is emerging, and tremendous efforts are made to collect whole genome sequences, the amount of information on gene variants and their impact on the primary protein structure is rising dramatically. Databases like the Exome Aggregation Consortium (ExAC) (Exome Aggregation Consortium *et al.,* 2015) are collecting and presenting vast amounts of data, enabling the screening of sequences of interest, identifying interesting and possibly disease-causing variants and allowing in combination with functional and/or structural studies a detailed analysis of their impact.

3.5 Work performed prior to this thesis

Initial experiments on the complex formation of titin Z8Z9 and obscurin O58O59 were performed by Chatziefthimiou and Pelissier (EMBL-Hamburg) including cloning of the first DNA constructs, expression, and protein purification. While the purification of individual titin and obscurin was

successful, it was not possible to successfully purify the complex and perform the protein-complex crystallisation.

The crystal structure of the homodimer formed by two myomesin FN-III domain My5 has been solved by Sauer et al. (Figure 6 B). To further investigate the dimerisation behaviour, different experiments were performed. First, myomesin constructs with different length with and without domain My5 were cloned, expressed and purified to map the dimerisation site. It could be demonstrated that constructs including My5 showed dimerisation while constructs without did not, indicating My5 as the causal domain for dimerisation (Sauer, 2011, dissertation). In an attempt to break the dimer a series of different mutations located within the beta strands forming the dimer, the so-called secondary interface formed by the dimer swap. However, these mutations were not able to prevent the dimer formation. In a second attempt, mutations were chosen that are located in the primary domain swap interface, the interface that is complemented by the swapped N-terminus of the other protein. In this interface Chatziefthimiou and Temmerman (unpublished work) were able to identify three mutants able to substantially decrease in signal in a fluorescence-based *in vivo* assay - Bimolecular fluorescence complementation (BiFC). This indicated either a disruption of the complex formation or a rearrangement between the two domains.

3.6 Aim of the work

This thesis will shed light on the function and regulation of the titin-obscurin complex and draw conclusions about the physiological relevance of domain swapping mechanism of myomesin domain My5. Previous studies revealed two separate interaction sites between titin and obscurin: one at the M-band and one at the Z-disk. The interaction at the M-band has already been structurally determined and functionally described (Pernigo *et al.*, 2010; Sauer *et al.*, 2010). In contrast, the complete molecular structure of the titin-obscurin interaction at the Z-disk remains elusive. The aim of the first part of this thesis was to crystallise the titin-obscurin Z-disk complex and to solve its high-resolution crystal structure. In combination with biophysical approaches like isothermal titration calorimetry (ITC) and low-resolution structural methods like SAXS, the crystal structure helped to investigate the complex domain interplay between the four titin Z8Z9 and obscurin O58O59 domains and allowed to identify the interfaces forming upon complex formation. Furthermore, it was a prerequisite to elucidate the impact of variants found within the complex domain sequences and to explain the structural impact of putative disease-related mutations like obscurin R4344Q or R4444W. Therefore, the following research questions will be addressed:

- What is the high-resolution crystal structure of the titin-obscurin Z-disk complex?
- Which interfaces form upon complex formation?
- Are there known variants within the titin Z8Z9 and obscurin O58O59 domain interface that could induce large structural effects?
- What is the structural role of variants linked to myopathies, especially the obscurin variants R4344Q and R4444W?
- Why does the titin-obscurin complex consist of four, instead of two domains like its counterpart at the sarcomeric M-band?
- How is the titin-obscurin Z-disk complex regulated?

The second part of the thesis will focus on the myomesin My5 dimerisation. Previous work has demonstrated that myomesin is a major contributor to the cross-linking of the thick filament at the sarcomeric M-band. An unpublished high-resolution structure suggests a homo-dimerisation

of the fifth domain of myomesin, with an unusual mechanism called three-dimensional domain swapping. This structure was used to further investigate the domain swapping mechanism. SAXS experiments on the My5 wild-type and mutants that were shown to disrupt the dimerisation were used to assess the in-solution behaviour of myomesin My5. Furthermore, sequence alignments and variant databases were used to assess the unique features that sets the myomesin My5 domain apart from other non-swapping FN-III domains. This will enable to draw conclusions about the physiological relevance of domain swapping. The following research questions will be addressed:

- What is the structural mechanism behind the domain swapping of the My5 dimer?
- Does the swapped dimer exist in a physiological context?
- Does the swapped dimer have a physiological function?

4 Materials and Methods

4.1 Materials

4.1.1 Chemicals and consumables

Table 1 List of chemicals and consumables

Chemicals	**Supplier**	**Catalogue Number**
Acetic acid 100%, Rotipuran® 100%, p.a.	Roth	Cat# 3738
Acrylamide/Bis stock sol. 29:1 (40% w/v)	Roth	Cat# A515
Agarose for DNA electrophoresis	SERVA	Cat# 11404
Ampicillin disodium salt	Roth	Cat# K029
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	Cat# D2650
Desthiobiotin	IBA	Cat# 2-1000-001
EDTA Tetrasodiumsalt	Roth	Cat# 3619
Ethanol ≥ 99,8%	Roth	Cat# 9065
Ethidium bromide Solution 0,025%	Roth	Cat# HP47
Glucose	Roth	Cat# X997
Glycerol, Rotipuran® ≥ 99,5%, p.a.	Roth	Cat# 3783
Guanidine hydrochloride ≥ 99,5%	Roth	Cat# 0037
HEPES, Pufferan® ≥ 99,5%, p.a.	Roth	Cat# 9105
Imidazole, Pufferan® ≥ 99%, p.a.	Roth	Cat# X998
Isopropyl β-D-1-thiogalactopyranoside	Roth	Cat# 2316
Kanamycin sulphate	Roth	Cat# T832
LB Agar (Lennox)	Roth	Cat# X965
LB Broth Low Salt Granulated	Melford	Cat# GL1703
MES, Pufferan® ≥ 99%	Roth	Cat# 4256
Methanol Rotipuran® ≥ 99,9%, p.a., ACS, ISO	Roth	Cat# 4627
Methoxypolyethylene glycol maleimide 5000	Sigma-Aldrich	Cat# 63187
Nickel(II) chloride hexahydrate ≥ 98%, p.a.	Roth	Cat# 4489
Sekusept Plus	Ecolab	Cat# 104372E
Sodium chloride, >99,5%, p.a., ACS, ISO	Roth	Cat# 3957
Sodium hydroxide ≥ 99	Roth	Cat# 9356
TEMED ≥ 99%, p.a.	Roth	Cat# 2367
Tris Acetate-EDTA (TAE) buffer 10X	Sigma-Aldrich	Cat# T8280
TRIS hydrochloride, Pufferan® ≥ 99%, p.a.	Roth	Cat# 9090
Urea ≥ 99,5%, p.a.	Roth	Cat# 3941
Molecular cloning		
ATP	Sigma-Aldrich	Cat# A9187
Customised lyophilised primers	Eurofins Genomics	
DNase, RNase-free PCR tubes	Sarstedt	
dNTPs	New England BioLabs® Inc.	Cat# N0447S
Gene Ruler 1 kb DNA Ladder	Thermo Fisher Scientific	Cat# SM0311
Nuclease-free water	Qiagen	Cat# 129115
6x DNA Loading Dye	Thermo Fisher Scientific	Cat# R0611

Enzymes		
DpnI restriction enzyme	New England BioLabs® Inc.	Cat# R0176S
Taq DNA Polymerase	New England BioLabs® Inc.	Cat# M0267S
Phusion ® High-Fidelity DNA Polymerase	New England BioLabs® Inc.	Cat# M0530S
PCR Using Q5® High-Fidelity DNA Polymerase	New England BioLabs® Inc.	Cat# M0491
T4 DNA Polymerase	New England BioLabs® Inc.	Cat# M0203S
T4 DNA Ligase	New England BioLabs® Inc.	Cat# M0202S
Taq 2X Master Mix	New England BioLabs® Inc.	Cat# M0270
T4 Polynucleotide Kinase	New England BioLabs® Inc.	Cat# M0201S
Lysozyme	Roth	Cat# 8259
Commercial kits		
Monarch® PCR & DNA Cleanup Kit	New England BioLabs® Inc.	Cat# T1030S
QIAquick Miniprep Kit	Qiagen	Cat# 27104
Monarch® DNA Gel Extraction Kit	New England BioLabs® Inc	Cat# T1020L
CloneJET PCR Cloning Kit	Thermo Fisher Scientific	Cat# K1231
Protein purification and characterization consumables		
Ni-NTA Agarose	Invitrogen	Cat# R901-15
Roti®-Mark 10-150 Protein-Marker	Roth	Cat# T850
NuPAGE® LDS Sample Buffer (4X)	Thermo Fisher Scientific	Cat# NP008
NuPAGE® Novex® 4-12% gels	Thermo Fisher Scientific	Cat# 83.1826
InstantBlue™	Expedeon	Cat# ISB1L
Monolith NT.115 MST Standard Treated Capillaries	Nanotemper	Cat# MO-K022
Alexafluor 532 C5-maleimide	Thermo Fisher Scientific	Cat# A10255
Filtropur S 0.45	Sarstedt	Cat# 83.1826

4.1.2 Equipment

Table 2 List of Equipment

Equipment	**SOURCE**
Molecular biology	
PCR cycler	Eppendorf
Electrophoresis chamber for agarose gels	NeoLab
Power supply	Consort
Gel imaging and documentation system	Bio-Rad
Protein production	
Water bath at 42˚C	VWR
Centrifuge Avanti JXN-26	Beckman-Coulter
JLA 8.1000 rotor	Beckman-Coulter
New Brunswick™ Innova® 42 small shaking incubator	Eppendorf
New Brunswick™ Innova® 44 shaking incubator for large-scale cultures	Eppendorf
Protein purification and characterisation	
ÄKTA Pure 25 L1 with fraction collector F9C	GE Healthcare
Analytical gel filtration system	Agilent technologies
HiLoad 16/600 Superdex200 (gel filtration column)	GE Healthcare
HiLoad 16/600 Superdex75 (gel filtration column)	GE Healthcare
Electrophoresis chamber for SDS-PAGE gels	Invitrogen

Monolith NT.115 microscale thermophoresis (MST)	Nanotemper
Crystallisation	
Rock Imager	Formulatrix
Scorpion dispenser	ARI-Art Robbins Instruments
Microscope	Nikon
General laboratory equipment	
pH meter	Mettler Toledo
Balances	Sartorius
Microwave	
MilliQ water machine	Millipore
Magnetic stirrer	Roth
Thermoblock	Eppendorf
Refrigerated bench-top microcentrifuge (5424R)	Eppendorf
Refrigerated bench-top centrifuge (5810R)	Eppendorf
Nanodrop UV-Vis spectrophotometer	Thermo Fisher Scientific
VP-ITC calorimeter	Microcal / GE Healthcare
Tecan infinite M1000	Tecan
Chirascan V100 Circular Dichroism Spectrometer	Applied Photophysics

4.1.3 Plasmids and bacterial strains

Table 3 List of Plasmids and bacterial strains

Resource	Source	Antibiotic resistance	Comments
Plasmids			
pJET1.2	Thermo Scientific (Fermentas)	Ampicillin	Cloning vector
pCDF-13	A. Geerlof	Streptomycin	Expression vector
pETM-11	G. Stier	Kanamycin	Expression vector
pETM-14	A. Geerlof	Kanamycin	Expression vector
pnEK	C. Romier	Kanamycin	Expression vector
pET-151D	Invitrogen	Ampicillin	Expression vector
Bacterial cell strains			
E.coli DH5α	EMBL Hamburg	-	Cloning strain
E.coli BL(21) Codon Plus RIL	Stratagene	Cam	Expression strain

4.1.4 Computational resources

Table 4 List of computational resources

Resource	Source
Unicorn 7	GE Healthcare
XDS / XSCALE / XDSCONV (version January 26, 2018)	(Kabsch, 2010)
PHENIX (version 1.11.1-2575-000)	(Adams *et al.,* 2010)
PHASER (version 2.1)	(McCoy *et al.,* 2007)
COOT (version 0.8.2)	(Emsley *et al.,* 2010)
MolProbity (in PHENIX package version 1.11.1-2575-000)	(Chen *et al.,* 2010)

PyMol™ (version 1.7.x)	Schrödinger LLC; www.pymol.org
Protein Data Bank (PDB)	www.pdb.org
GraphPad Prism (version 5.0f)	www.graphpad.com
SnapGene® Viewer (version 4.1.4)	www.snapgene.com
Origin	OriginLab

4.2 Buffers and media

4.2.1 Bacterial growth culture media and antibiotic stocks

4.2.1.1 Luria-Bertani (LB) broth media

Table 5 LB-media preparation

LB broth low salt granulated	200 g
MilliQ water	Up to 1 L
Each liter of LB agar contains:	
10 g Tryptone	
5 Sodium chloride	
5 g Yeast extract	

Bacterial growth medium was autoclaved prior to usage. To grow bacterial cultures, LB medium was supplemented with the required antibiotics prior to inoculation of the cells. Antibiotics used in this work were kanamycin (30 µg/ml), streptomycin (50 µg/ml), ampicillin (100 µg/ml), and chloramphenicol (34 µg/ml). Antibiotic stocks were prepared in 1000x concentration. Lyophilised antibiotic powder was weighted using an analytical balance and dissolved in the required volume of ultrapure water or molecular grade ethanol (for chloramphenicol), respectively. The solution was sterile filtered using a 0.22 µm pore size filter, aliquoted and stored at -20˚C until further use.

4.2.1.2 LB-agar (Lennox)

Table 6 LB-agar preparation

LB-agar	35 g
MilliQ water	Up to 1 L
Each liter of LB agar contains:	
10 g Tryptone	
5 g Yeast extract	
5 g Sodium chloride	
15 g Agar-agar	

To prepare agarose plates autoclaved and solidified LB-agar was heated in the microwave until fully melted and allowed to cool down to approx. 50˚C prior to antibiotic supplementation. Antibiotics were added under a sterile atmosphere to a final concentration equal to the bacterial growth medium. Subsequently, approx. 20 ml LB-agar were poured into Petri-dishes and allowed to solidify. Plates were stored at +4˚C until further use. Note: LB-agar supplemented with ampicillin should not be stored longer than a week.

4.2.2 Buffers for DNA electrophoresis

Table 7 TAE preparation

50x Tris-Acetic-EDTA (TAE) buffer	
Tris base (MW: 121.1 g/ mol)	242 g
EDTA	18.6 g
Acetic acid (MW: 60.05 g/ mol)	57.1 ml
MilliQ water	Up to 1 L
Filter using an 0.45 µm pore size filter	

Table 8 TBE preparation

5x Tris-Borate-EDTA (TBE) buffer	
Tris base (MW: 121.1 g/ mol)	54 g
0.5 M EDTA pH 8.0	20 ml
Boric acid (MW: 61.83 g/ mol)	27.5 ml
MilliQ water	Up to 1 L
Filter using an 0.45 µm pore size filter	

For preparation of 1% (w/v) agarose gels for DNA electrophoresis 1 g of agarose was dissolved in 100 ml 1x TAE or TBE buffer and boiled in the microwave. Dissolved agarose was under continuous stirring allowed to cool down to approx. 60˚. Subsequently, agarose was poured into a gel casting chamber adding one drop of ethidium bromide or gel green and allowed to solidify for 30 minutes. 3 µl samples were mixed with 0.6 µl of 6x DNA Loading Dye and together with the Gene Ruler 1 kb DNA Ladder as a reference pipetted into the gel pockets. Gels were run at 100 V for 40-50 minutes at room temperature. Visualisation of the gels was done using ChemiDoc imager.

Table 9 1% agarose gel preparation

1% agarose gels (w/ v)	
Agarose	1 g
1x TAE or TBE buffer	100 ml

4.2.3 Buffers for protein acrylamide gel electrophoresis (PAGE)

Table 10 20x MES buffer preparation

20x MES running buffer	
MES	195.2 g
Tris base	121.2 g
SDS	20 g
EDTA-free acid	6 g
MilliQ water	Up to 1 L
pH was adjusted to 7.3	

Table 11 SDS-PAGE loading dye preparation

For 500 µl of loading dye (8% 2-Mercaptoethanol)	
NuPAGE® LDS Sample Buffer (4X)	460 µl
2-Mercaptoethanol	40 µl

Table 12 SDS-PAGE gel preparation

Required chemicals	Resolving gel			Stacking gel
	10%	12%	15%	
5x Bis-Tris buffer (pH 6.5- 6.8), ml	4	4	4	1.6
MilliQ water, ml	10.9	9.9	8.4	5.4
40% acrylamide, ml	5	6	7.5	1
20% APS, µl	100	100	100	40
TEMED, µl	20	20	20	20

18 µl protein sample was mixed with 6 µl 4x loading dye and boiled at 75˚C for ten minutes. High concentrated protein samples (e.g. IMAC load and flow through sample) were diluted 1:10 with ultra-pure H_2O prior to mixing with the loading dye. Gels were runs at 180 V for 40-45 minutes at room temperature. For staining the gel was covered in InstantBlue™ and incubated for 30 minutes on a shaking platform. Subsequently, staining solution was removed, and gel was incubated for another 30 minutes in water.

4.3 Methods

4.3.1 Molecular cloning

DNA sequence for all cloned constructs were already available in the lab and have previously been amplified *via* PCR from skeletal muscle cDNA cloned into expression vectors. For the sake of construct optimisation or adaptation to a different type of experiment, the constructs for titinZ8Z9, My5 wt, My5 P638R, My5 P641R, My5 D643P, and the My5 construct for AFM measurements were re-cloned into new expression vectors using SLiCE cloning (Zhang, Werling and Edelmann, 2012). On all vectors, expression is controlled by the T7 promotor and the *lac* operon and can be induced by supplementation of isopropyl β-D-1-thiogalactopyranoside (IPTG) to the culture media. All used constructs contained an N-terminal hexa or deca Histidine- or a Streptavidin derived Strep-purification tag II that can be cleaved off at a Tobacco Etch virus (TEV) or a Human Rhinovirus (HRV) 3C Protease cleavage site. As a selection marker, all vectors carried either a kanamycin, streptomycin, or ampicillin coding cassette. For generation of mutants, site-

directed mutagenesis was performed using either the NEBaseChanger (New England BioLabs Inc.) protocol or a customised single primer mutagenesis protocol was used.

Table 13 List of constructs created in this work

Protein (Uniprot sequence ID)	Construct	Residue range	N-terminal tag	Vector backbone	Resistance
Titin (Q8WZ42)	Z8Z9-wt	1703-1930	10xHis-3C	pCoofy-18	Kanamycin
	Z8Z9-K1782A		10xHis-3C	pCoofy-18	Kanamycin
	Z8Z9-S1711D		10xHis-3C	pCoofy-18	Kanamycin
	Z8Z9-T1727D		10xHis-3C	pCoofy-18	Kanamycin
	Z8Z9-T1791V		10xHis-3C	pCoofy-18	Kanamycin
	Z8Z9-D1906N		10xHis-3C	pCoofy-18	Kanamycin
	Z8Z9-D1906R		10xHis-3C	pCoofy-18	Kanamycin
	Z8	1703-1793	6xHis-TEV	pETM-11	Kanamycin
	Z9	1841-1928	6xHis-TEV	pETM-11	Kanamycin
Obscurin (Q5VST9)	O58O59-wt	4334-4522	6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-E4387R		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-E4356R/E4387R		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-D4377R		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-R4411E		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-D4485N		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-D4485R		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-R4344Q		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-D4377R/D4485R		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
	O58O59-E4356R/D4377R/E4387R		6xHis-SUMO-3C	pCDF-SUMO-3C	Streptomycin
Myomesin (P52179)	My5-wt	635-734	6xHis-3C	pETM-14	Kanamycin
	My5-P638R		6xHis-3C	pETM-14	Kanamycin
	My5-P641R		6xHis-3C	pETM-14	Kanamycin
	My5-D643P		6xHis-3C	pETM-14	Kanamycin
	My5-wt AFM A		6xHis	pETM-14	Kanamycin
	My5-wt AFM B		Strep	pCDF-13	Streptomycin
	My5-wt AFM A/B		6xHis-Strep	pnEK	Kanamycin
	My7-My13 FRET	933-1666	6xHis-V5-TEV	pET-151D	Ampicillin
	My7-My13 FRET negative control		6xHis-V5-TEV	pET-151D	Ampicillin

4.3.1.1 Ligation-independent cloning using Seamless Ligation Cloning Extract (SLiCE)

Traditional molecular cloning methods require the use of specific restriction enzymes for the introduction of a gene of interest into the target vector and are thus dependent on sequence

limitations. Ligation-independent cloning (LIC) technics use the principle of homologue recombination of complementary DNA sequence regions to overcome these limitations and enable the introduction of a gene of interest at virtually any position in a plasmid. The target protein is amplified using primers containing 5'-overhangs complementary to the insertion region within the target vector. The target vector is linearised by enzymatic digestion or PCR amplification. PCR is performed using Q5 polymerase or Phusion polymerase using standard manufacturer PCR protocols with conditions chosen according to primer properties. Methylated original template vectors of the PCR products of the target gene and target vector are subsequently removed by DpnI digestion (20U DpnI) at 37°C for 3 h. PCR products are purified using Monarch® DNA Gel Extraction Kit (New England BioLabs). For the enzymatic digestion, 2 µg DNA are digested at 37°C for >4 h with one or ideally two restriction enzymes in optimal buffer conditions at a total volume of 30 µl. 3 µl of Antarctic Phosphatase buffer and 1 µl Antarctic Phosphatase are added and incubated for 1 h at 37°C for dephosphorylation of the vector ends. Linearised vector is then gel purified using NEB Monarch kit. For the SLiCE reaction 1 µl 10xT4 DNA ligase buffer, 1 µl of the SLiCE extract (Zhang, Werling and Edelmann, 2014), 50 ng of the linearised and purified target vector and the purified insert DNA in a vector:target ratio between 1:1 and 1:10 are mixed and diluted with ultra-pure H_2O to a final volume of 10 µl. The mixture is incubated for 1 h at 37˚C. 1 µl of the reaction mix is used for transformation of chemical competent DH5α cells.

Table 14 50 µl PCR protocol using Q5 Polymerase

Component	Volume / final concentration
5X Q5 Reaction Buffer	10 µl / 1x
10 mM dNTPs	1 µl / 200 µM
10 µM Forward Primer	2.5 µl / 0.5 µM
10 µM Reverse Primer	2.5 µl / 0.5 µM
Template DNA	<1,000 ng / <250 ng
Q5 DNA Polymerase	0.5 µl / 0.02 U/µl
5X Q5 High GC Enhancer (optional)	10 µl / 1x
Nuclease-free water	to 50 µl

Table 15 Thermocycling Conditions for a Routine Q5 PCR

STEP	TEMP	TIME
Initial Denaturation	98°C	30 seconds
	98°C	5–10 seconds
25–35 Cycles	50–72°C	10–30 seconds
	72°C	20–30 seconds/kb
Final Extension	72°C	2 minutes
Hold	4–10°C	

Table 16 50 µl PCR protocol using Phusion Polymerase

Component	Volume / final concentration
5X Phusion HF or GC Buffer	10 µl / 1x
10 mM dNTPs	1 µl / 200 µM
10 µM Forward Primer	2.5 µl / 0.5 µM
10 µM Reverse Primer	2.5 µl / 0.5 µM
Template DNA	<250 ng
Phusion DNA Polymerase	0.5 µl / 0.02 U/µl
DMSO (optional)	1.5 µl / 3%
Nuclease-free water	to 50 µl

Table 17 Thermocycling Conditions for a Routine Phusion PCR

STEP	TEMP	TIME
Initial Denaturation	98°C	30 seconds
	98°C	5–10 seconds
25–35 Cycles	42–72°C	10–30 seconds
	72°C	15–30 seconds/kb
Final Extension	72°C	5-10 minutes
Hold	4–10°C	

4.3.1.2 Heat-shock transformation of E. coli

Homemade chemically-competent *E. coli* cells (prepared with CaCl2) (Swords, 2003) were slowly thawed on ice, carefully mixed with 1 µl of approx. 5 ng/µl of DNA, left on ice for another 15 min and subsequently placed for 30 sec in a water bath at 42°C. Cell were cooled on ice for additional 2 min, mixed with 1 ml of SOC medium, incubated at 37°C for 1 h and plated out on LB-Agar with appropriate antibiotics. Plates were incubated at 37°C and colonies could be picked after 12-16 h. Only the cells successfully transformed with the vector are expected to grow overnight.

4.3.1.3 PCR for colony screening

Positive gene insertion is tested by colony PCR. Half of a colony grown on agar plate after transformation is resuspended in 20 µl ultra-pure H_2O, lysed by heating for 10 min at 75°C, centrifuged for 10 minutes at maximum speed and subsequently used as template for the colony PCR. PCR is performed with Taq 2x Master Mix (NEB) following the supplier's standard procedure using T7 promotor (5'-TAATACGACTCACTATAGGG-3') and T7 terminator primers (5'-CTAGTTATTGCTCAGCGGT-3'). The size of the amplified DNA fragment is estimated by agarose gel electrophoresis. Colonies with positive insertions are further prepared for DNA sequencing.

Table 18 25 µl PCR with Taq 2x Master Mix

Component	Volume / final concentration
10 µM Forward Primer (T7 promotor)	0.5 µl / 0.2 µM
10 µM Reverse Primer (T7 terminator)	0.5 µl / 0.2 µM
Template DNA	5 µl
Taq 2X Master Mix	12.5 µl
Nuclease-free water	to 11.5 µl

Table 19 Thermocycling Conditions colony PCR with Taq 2x Master Mix

STEP	TEMP	TIME
Initial Denaturation	95°C	30 seconds
	95°C	30 seconds
25–35 Cycles	55°C (Default 55°C)	30 seconds
	68°C	60 seconds
Final Extension	68°C	10 minutes
Hold	4–10°C	

4.3.1.4 Sequencing of the clones and storage

Colonies positively tested in colony PCR are inoculated in in 5 ml LB-medium supplemented with the appropriate antibiotic and incubated overnight at 37 °C. Plasmid is purified with the QIAquick Miniprep Kit (Quiagen) and send for sequencing using the Eurofins Genomics sequencing service. Clones carrying the correct expression cassette are stored in 25% glycerol at -80 °C.

4.3.1.5 NEBaseChanger PCR protocol for site-directed mutagenesis

For the site-directed mutagenesis introducing point mutations within the protein sequence of target proteins, two different approaches were applied. The NEBasechanger protocol is based on non-overlapping Primers designed with 5´-ends annealing back-to-back enabling exponential amplification. The nucleotide mismatch is introduced in the centre of one mutagenic primer while the other primer can be 100% complementary. Primer were designed using the NEBasechanger primer design web tool. PCR was performed according to the standard Q5 protocol and the suggested annealing temperatures for the used primer. In contrast to the manufacturers protocol, the PCR product was not treated with the commercial KDL (Kinase, Ligase and DpnI) mix but with the home-made SLiCE extract according to the SLiCE protocol (Zhang, Werling and Edelmann, 2014) (internal test experiments showed no difference between the usage of the commercial KDL mix and the SLiCE extract in terms of positive colonies). Transformation and sample preparation for sequencing was performed as described in chapter 4.3.1.2 and 4.3.1.4.

4.3.1.6 Single primer PCR protocol for site-directed mutagenesis

As in some cases, site-directed mutagenesis using NEBasechanger protocol was not successful and resulted in no positive colonies. In these cases, an alternative method was used only using the forward primer with a length of about 40 nucleotides with the mutation in the middle. To find the optimal primer length, the following formula was applied: $T_m = 81.5 + 0.41 * (GC(\%)) - 675/N - missmatch(\%)$, with N – primer length in bases, GC (%) and mismatch (%) given in whole numbers. The calculated Tm was aimed to be around 78°C. PCR was performed using Phusion High-Fidelity DNA Polymerase. As only one primer is used the PCR does not result in an exponential amplification of the target sequence resulting in a low amplicon concentration. Intensive DpnI digest of the methylated template DNA is hence highly important. 2.5 µl DpnI and 7.5 µl 10x Cut Smart buffer (NEB) are added to a 50 µl PCR and incubated at 37°C for 3 h. 2 µl of the reaction mix is directly transformed in chemically competent cell following the protocol described in chapter 4.3.1.2. Correct sequence was assessed by sequencing following the procedure in chapter 4.3.1.4.

Table 20 50 µl single primer PCR using Phusion Polymerase

Component	Volume / final concentration
5X Phusion HF or GC Buffer	10 µl / 1x
10 mM dNTPs	1 µl / 200 µM
10 µM Primer	0.1 µl / 20 nM
Template DNA	0.2 ul (0.05 ug to 0.2 ug)
Phusion DNA Polymerase	0.4 µl
DMSO (optional)	1.5 µl / 3%
Nuclease-free water	to 50 µl
Component	Volume / final concentration

Table 21 Thermocycling Conditions for single primer mutagenesis Phusion PCR using a single primer

STEP	TEMP	TIME
Initial Denaturation	98°C	30 seconds
	98°C	5–10 seconds
25–35 Cycles	55–65°C (Default 55°C)	10–30 seconds
	65°C	10 minutes
Final Extension	65°C	10 minutes
Hold	4–10°C	

4.3.1.7 Titin construct cloning

As the initial titin Z8Z9 construct present at EMBL-Hamburg could not be purified in sufficient amounts necessary for crystallisation the construct was re-cloned into pCoofy-18 carrying a deca histidine tag. At the same time, a short two amino acid long linker sequence between the protein sequence and the purification tag was introduced to increase accessibility of the purification tag. The titin Z8Z9 sequence (UniProt ID: Q8WZ42; residues 1703-1930) was amplified from cDNA using primers containing overhangs to the multiple cloning site of the target vector pCoof-18 (forward primer: 5'-AAGTTCTGTTCCAGGGGCCCAGCGCGCCATTTTTCAAGAAAAAACTCACTTCC-3'; reverse primer: 5'-CCCCAGAACATCAGGTTAATGGCGTTATTGAATCTCAAGCTTCACTTTATGC-3'). For linearisation of the target vector, it was amplified using primers complementary to the overhang regions used for insert primers (forward primer: 5'-CGCCATTAACCTGATGTTCTGGGG-3'; reverse primer: 5'-GGGCCCCTGGAACAGAACTTCCAG-3'). PCR was performed according to the standard protocol required for Q5 polymerase. DpnI digest, purification, SLiCE reaction and Sequencing was executed as described in chapter 4.3.1.1 - 4.3.1.4.

4.3.1.8 Obscurin construct cloning

Obscurin O58O59 (UniProt ID: Q5VST9; residues 4334-4522) was amplified *via* PCR from skeletal muscle cDNA. The sequence differs from the deposited UniProt sequence in the highly frequent natural variants H4381R (MAF: 0.7014), C4450R (MAF: 0.6949) and R4516W (MAF: 0.423). The custom oligonucleotide primers (forward primer: 5'-AGTTCTGTTCCAGGGGCCCCTGAAAGTC ACAGCCAAGAAC-3'; reverse primer: 5'-CGGAGCTCGAATTCGGATCCTTAGAGGCCCAGCACGGTGA-

3') contain appropriate 5'-overhangs to the pCDF recipient vector that was modified with a cleavable (3C cleavage site) SUMO (Small Ubiquitin-like Modifier) solubility tag C-terminal of the Hexa-Histidine purification tag and carrying a streptomycin selection cassette. The vector was linearised with primers complementary the overhangs used for the insert primers (forward primer: 5'-GAAGTTCTGTTCCAGGGGCCC-3'; reverse primer: 5'- ACGGAGCTCGAATTCGGATCC-3'). PCR was performed according to the standard protocol required for Q5 polymerase. DpnI digest, purification, SLiCE reaction and Sequencing was executed as described in chapter 4.3.1.1- 4.3.1.4.

4.3.1.9 Site-directed mutagenesis for the titin-obscurin interfaces

Site-directed mutagenesis was performed using the NEBaseChanger PCR protocol (chapter 4.3.1.5) as a standard or the single primer PCR protocol (4.3.1.6) in cases in which the NEBaseChanger PCR protocol did repetitively not result in positive colonies carrying the desired mutation. For the single primer PCR protocol, primers designed for NEBaseChanger PCR protocol were re-used. DpnI digest, purification, SLiCE reaction and Sequencing was executed as described in chapter 4.3.1.1- 4.3.1.4.

Table 22 Site-directed mutagenesis primer for titn/obscurin mutations

	residue	**5'->3' primer sequence**
Titin Z8Z9	K1782A	forward: AGCCACTAACgcgTATGGAACAGATCAC
		reverse: CTGCAAGTAATGATACCAC
	S1711D	forward: AAAAAACTCACTGATTTAAGACTTAAG
		reverse: TTTTTTGAGTGACTAAATTCTGAATTC
	T1727D	forward: ATGCAGGCTAgatCCCATTGGTGACCCAAC
		reverse: TCAAAGTGGGCAGGCCCA
	T1791V	forward: CACATCTGCTgtgCTTATTGTTAAAGATG
		reverse: TGATCTGTTCCATATTTGTTAG
	D1906N	forward: CAAATCATATaacACAGGTGAAGTGAAG
		reverse: CAGTCCACGATGTCCAGG
	D1906R	forward: CAAATCATATCGTACAGGTGAAGTGAAG
		reverse: CTTCACCTGTACGATATGATTTGCAGTC
Obscurin O58O59	E4387R	forward: GGAGAACGCCcgtGTGGTCTTCTTCGAGAACGG
		reverse: GAGGTGCGCACGGGTTCG
	E4356R/E4387R	forward: GGAGGGCGGCcgtGCGCTGTTCG
		reverse: AGCGCCTCCACATTCTCC
	D4377R	forward: GCTGCTGGACcgtGAACCCGTGCGCACC
		reverse: CAGGTGTGGGCGGCCACC
	R4411E	forward: AGACAGCTGCgaaGTGACCTTCCTG
		reverse: TGTGGCCGCAAGTTTTTG
	D4485N	forward: TGTCACCGCCAATGGCAGTCACCACGCC

		reverse: GGTGACTGCCATTGGCGGTGACAGTCCA
	D4485R	forward: TGTCACCGCCcgtGGCAGTCACC
		reverse: GTCCAGTCGCTGTCATCC
	R4344Q	forward: CACGGTGGTGCAGGGGCTGGAGAATGTG
		reverse: TCTCCAGCCCCTGCACCACCGTGTTCTT

4.3.1.10My5 construct cloning

All My5 constructs (wild-type, P638R, P641R, D643P) (UniProt ID: P52179; residues 635-734) used for BiFC experiments performed by Chatziefthimiou and Temmerman needed to be re-cloned for bacterial expression and purification without the BiFC fluorophore. Oligonucleotide primers for the insert contained 5'-overhangs to the multiple cloning site of the target vector pETM-14 carrying a 3C-protease cleavable N-terminal Hexa-His-tag and kanamycin selection cassette. Primers used for amplification of wild-type, P641R, D643P were the same (forward primer: 5'-AGTTC TGTTCCAGGGGCCCATGGACCCTTCAGAGGGTATTGTGCC-3'; reverse primer: 5'-TCGGATCCGGTAC CTCATTATTTGTCCCCTACCACAGTCA-3'). As the sequence of residue P638R was part of the primer sequence, an additional forward primer needed to be designed (5'-AGTTCTGTTCCAGGGG CCCATGGACCCTTCAGAGGGTATTGTG-3'). Target vector was amplified using primers complementary to the used overhangs (forward primer: 5'-AGTTCTGTTCCAGGGGCCCATGGAC-3'; reverse primer: 5'-TCGGATCCGGTACCTCATTA-3'). For amplification Q5 polymerase was used according to the standard protocol. DpnI digest, purification, SLiCE reaction and Sequencing was executed as described in chapter 4.3.1.1- 4.3.1.4.

4.3.1.11 My5 AFM construct cloning

For the AFM experiments, it was necessary to form and purify a dimer carrying different tags on each monomeric component. Therefore, two different monomer constructs were designed, one containing a ybbR-AFM tag and a His-tag and one containing a sortase-tag and a Strep-tag (see chapter 5.2.5, Figure 38). Different approaches for expression and purification were tested including separate expression and subsequent combination of the two constructs during purification, co-expression of two separate vectors each carrying one construct, or co-expression on one polycistronic vector. For the first two approaches, the cloning of the My5 AFM constructs was performed in several consecutive PCRs with 5'-overhang primers to introduce the additional

N- and C-terminal AFM- and purification-tags, linker regions and homologue target vector sequence regions for both constructs. Used primers are described in Table 23. For each step, PCR products amplified using Q5 polymerase, were purified with NucleoSpin plasmid mini Kit and used as template for the next PCR. First My5 (UniProt ID: P52179; residues 635-734) was amplified introducing an N-terminal His-tag (construct A) or a Strep-tag (construct B), respectively, as well as the C-terminal linker containing a very C-terminal cysteine. The second PCR introduced the N-terminal AFM affinity-tags and the C-terminal overhang to the target vector for construct A (ybbR-tag; pETM-14) and construct (sortase-tag, pCDF-13). In the third step, the N-terminal overhang to the corresponding target vector was introduced. Target vectors were linearised using primers complementary to the overhang regions introduced in the insert (pETM-14: forward primer: 5'-GAAGGAGATATACCATGAAA-3'; reverse primer: 5'-TCGGATCCGGTACCTCATTA-3'; pCDF-13: forward primer: 5'-ACTTTAATAAGGAGATATAC-3'; reverse primer: 5'- TTCATCCGAACCGCCGCCGC-3'). DpnI digest, purification, SLiCE reaction and Sequencing was executed as described in chapter 4.3.1.1- 4.3.1.4.

Table 23 Cloning strategy for My5 AFM constructs

step	construct	Primer sequence	Sequence addition
PCR 1	A	forward: 5'-CTGCTCATCACCATCACCATCACTCTGCTCCTTCAGAGGGTATTGTGCC-3'	SA-linker-6xHis-SA-linker
		5'TTAACAACCGGACCCGCCGCTGCCGCTACCCCCTGATTTGTCCCCTACCACAGTCACCT-3'	10 AA linker to Cys
	B	forward: 5'-CACCCGCAGTTCGAAAAATCTGCTCCTTCAGAGGGTATTGTGCC-3'	SA-linker-Strep-tag-SA-linker
		5'TTAACAACCGGACCCGCCGCTGCCGCTACCCCCTGATTTGTCCCCTACCACAGTCACCT-3'	10 AA linker to Cys
PRC2	A	forward: 5'ATGGACTCTCTGGAATTCATCGCTTCTAAACTGGCTTCTGCTCATCACCATCACC-3'	ybbR-tag
		5'-TCGGATCCGGTACCTCATTATTAACAACCGGACCCGCC-3'	pETM-14 overhang
	B	forward: 5'-ATGGGTGGTGGTTCTGCTTGGTCTCACCCGCAGTTCGAAAAATC-3'	Sortase-tag
		5'-TTCATCCGAACCGCCGCCGCTTAACAACCGGACCCGCC-3'	pCDF-13 overhang
PCR 3	A	forward: 5'-GAAGGAGATATACCATGAAAATGGACTCTCTGGAATTCAT-3'	pETM-14 overhang
		5'-TCGGATCCGGTACCTCATTATTAACAACCGGACCCGCC-3'	pETM-14 overhang
	B	forward: 5'-ACTTTAATAAGGAGATATACATGGGTGGTGGTTCTGCTTG-3'	pCDF-13 overhang
		5'-TTCATCCGAACCGCCGCCGCTTAACAACCGGACCCGCC-3'	pCDF-13 overhang

For the generation of the polycistronic construct enabling the sequential expression of My5 AFM constructs A and B from the same vector, a construct containing both constructs separated by an additional ribosome binding site was ordered using a commercial gene synthesis service (Intergrated DNA Technologies (IDT) - gBlocks® Gene Fragments). For integration of the ordered construct into the target expression vector pnEk, an N-terminal NdeI and a C-terminal NheI restriction site was added. The blunt end gBlocks® gene fragment was first introduced into the cloning vector pJET using the CloneJET PCR Cloning Kit following the standard protocol and transformed into chemically competent DH5α cells, a single colony was picked, and cultured overnight and the plasmid was purified according to the protocol described in chapter 4.3.1.2 and 4.3.1.4. Purified empty pnEK target vector and the purified pJET My5 polycistronic AFM vector were digested separately for one hour at 37°C using NdeI and NheI restriction enzymes.

Table 24 Enzymatic restriction protocol for polycistronic AFM construct

Restriction reaction	
Vector / Insert	Each 2 µg
cut smart buffer	5 µl
NdeI	3 µl
NheI	3 µl
Nuclease-free water	to 50 µl

The correctly cleaved fragments were run on a DNA agarose gel as described in chapter 4.2.2 and purified *via* gel extraction using NEB Monarch kit. Ligation was performed in a vector:insert ratio of 1:3 and incubated at room temperature for 30 minutes. Ligated construct was directly transformed into chemically competent DH5α cells and further treated for validation of correct insertion by sequencing as described in chapter 4.3.1.2 and 4.3.1.4.

Table 25 Ligation protocol for polycistronic AFM construct

Ligation reaction	
Vector	0.02 pmol / ~66 ng
Insert	0.06 pmol / ~38 ng
T4 ligase buffer	2 µl
T4 ligase	1 µl
Nuclease-free water	to 20 µl

4.3.1.12 My7-My13 FRET negative control construct cloning

The My7-My13 FRET construct containing the myomesin 1 sequence (UniProt ID: P52179, residue 933-1666), an N-terminal mTFP molecule, Cys to Ser mutations within My7-M10, and residue K1324 mutated to a Cys originated from S. Chatziefthimiou. For the FRET negative control, the cysteine at position 1324 needed to be cloned back to a lysine. Single primer mutagenesis using a primer containing a triplet nucleotide exchange (5'-TTCCAGCTTCAAGATGGAAAAGCAA CTAACCATTCTACT-3') was performed according to the protocol described in chapter 4.3.1.6.

4.3.2 Protein production and purification

Protein expression requires the generation of an *E. coli* expression strain carrying the plasmid containing the gene of interest, bacterial culture growth generating sufficient amount of protein and a two-step purification *via* Immobilised metal-affinity chromatography (IMAC) and Size exclusion chromatography (SEC). In the first purification step, the tendency of deprotonated histidine side-chains to complex divalent metal cations is used to separate proteins recombinantly tagged with a polyhistidine sequence from non-tagged proteins. The polyhistidine-tagged protein interacts with a solid phase exposing divalent metal cations (e.g. Ni^{2+} or Co^{2+} immobilised with NTA on agarose beads) and can be eluted using different strategies: i) protonation of the His-tag by pH-change, ii) detachment of the bound cations from the solid phase, iii) addition of a binding competitor like imidazole in high concentrations, or iv) enzymatic separation of the protein from the polyhistidine tag. In the second purification step (SEC), the molecules are separated according to their hydrodynamic volume. Smaller molecules will enter

the fine, porous column material more easily and therefore need to pass a larger volume compared to larger molecules. This results in an increased retention time for smaller molecules. Purity of the protein after the two purification steps can be assessed by SDS-PAGE.

4.3.2.1 Transformation of the bacterial expression strain

After successfully cloning and sequence verification of the protein construct, the vector was transferred to an *E. coli* expression strain. Chemically competent *E. coli* expression strain BL(21) Codon Plus RIL, a strain carrying additional genes encoding for rare codons of arginine, isoleucine, and leucine, was transformed following the heat shock protocol described in chapter 4.3.1.2. For long time storage, a single colony of the successfully transformed cells was inoculated in 5 ml LB-medium supplemented with chloramphenicol (34 µg/ml) and the construct-specific selection antibiotic and grown overnight at 37°C. 1 ml cell culture was mixed with 1 ml 50% glycerol sored at -80°C.

4.3.2.2 Protein expression

For large scale production of the target proteins, bacterial cultures were inoculated to an initial optical density (OD600) of 0.05 from a freshly grown overnight starter culture of the corresponding *E. coli* expression clone. Expression was carried out in 2 l baffled Erlenmeyer-flasks with 500 ml LB-medium supplemented with chloramphenicol (34 µg/ml) and the construct-specific selection antibiotic and incubated under continuous shaking (250 rpm) at 37°C until an OD600 of 0.7-0.9 was reached. The protein expression was induced by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) to the culture. Subsequently, the culture was allowed to grow at 20°C and shaking at 250 rpm for 16-20 h. Cells were harvested by centrifugation at 12 000 g, 15 minutes, 20˚C (JLA 8.1 rotor of Avanti JXN-26 centrifuge, Beckman Coulter). The pellet was stored at -20˚C until purification.

4.3.2.3 Protein purification

For protein purification, the frozen cell pellet was resuspended in 5 ml pre-cooled lysis/wash buffer (Table 26 & Table 27) per gram of protein pellet (wet weight) and lysed on ice by sonication. Insoluble fractions of the cell lysate were removed by centrifugation at 43 500 g, 4°C

for 1 h (Centrifuge Avanti JXN-26, Beckman-Coulter). The supernatant was filtered with a 0.45µm sterile filter and applied at room temperature to a pre-equilibrated gravity flow Ni-NTA IMAC column (3 ml bed volume per 50 ml of lysate). Washing steps were performed by applying ten column volumes (CV) of lysis/wash buffer to remove non-binding proteins, followed by 5 CV of high salt buffer (Table 26 & Table 27), followed by another 2 CV of lysis/wash buffer to remove the high salt content. The protein of interest was cleaved on column from the polyhistidine tag by adding 2 mg of 3C-protease to the gravity flow IMAC column. Cleavage was allowed to take place by incubating for 1.5 h at 4°C under mild shaking. Subsequently, the flow through containing the cleaved protein of interest was collected and concentrated (Amicon Ultra, PLGC Ultracel-PL Membran) to volumes and concentrations suitable for size exclusion chromatography (≤2 ml and ≤50 mg per SEC run). The protein was injected on a pre-equilibrated Superdex 75 16/60 pg or Superdex 200 16/60 pg column (depending on the expected molecular weight of the protein) and eluted from the column at a constant flow rate of 1.0 ml/min. Fractions containing the protein were detected by online absorption measurements at 280 nm, collected and checked for purity by SDS-PAGE. Fractions with the highest purity were pooled, where applicable concentrated to approx. 500 µM, aliquoted to 300 µl, flash frozen in liquid nitrogen and stored at -80°C for further use.

Table 26 HEPES-based protein purification buffer preparation

	Lysis/wash	High salt wash	Elution
HEPES	100 mM [47.66 g]	100 mM [4.77 g]	100 mM [4.77 g]
NaCl	125 mM [14.61 g]	1000 mM [11.69 g]	125 mM [1.46 g]
Imidazol	20 mM [2.72 g]	20 mM [0.27 g]	300 mM [4.08 g]
MilliQ water	Up to 2 L	Up to 200 mL	Up to 200 mL

pH was adjusted to 7.5

Table 27 Tris-based protein purification buffer preparation

	Lysis/wash	High salt wash	Elution
Tris	50 mM [g]	50 mM [g]	50 mM [g]
NaCl	300 mM [14.61 g]	1000 mM [11.69 g]	300 mM [1.46 g]
Imidazol	20 mM [2.72 g]	20 mM [0.27 g]	300 mM [4.08 g]
MilliQ water	Up to 2 L	Up to 200 mL	Up to 200 mL

pH was adjusted to 8.0

4.3.2.4 Titin expression and purification

Titin expression was performed according to the described method using chloramphenicol (34 µg/ml) and kanamycin (30 µg/ml). For purification HEPES-based buffers were used. Gel filtration was performed using a Superdex 75 16/60 pg column, and a concentrator with a molecular weight cut-off of 10 kDa was applied for concentrating the protein.

4.3.2.5 Obscurin expression and purification

Obscurin expression was performed according to the described method using chloramphenicol (34 µg/ml) and streptomycin (50 µg/ml). For purification HEPES-based buffers were used. Gel filtration was performed using a Superdex 75 16/60 pg column, and a concentrator with a molecular weight cut-off of 10 kDa was applied for concentrating the protein.

4.3.2.6 Myomesin My5 wild-type and mutant expression and purification

Expression and purification were performed according to described methods. For the expression of the My5 wild-type and P638R, P641R, and D643P mutants culture LB-medium was supplemented with chloramphenicol (34 µg/ml) and kanamycin (30 µg/ml). The purification was performed using Tris-based buffers. Gel filtration was performed using a Superdex 75 16/60 pg column, and a concentrator with a molecular weight cut-off of 3 kDa was applied for concentrating the protein.

4.3.2.7 Myomesin My5 AFM expression and purification

Expression of the My5 AFM constructs was performed according to described method supplementing the LB-medium with chloramphenicol (34 µg/ml) and kanamycin (30 µg/ml) for construct A and the polycistronic construct and using chloramphenicol (34 µg/ml) and streptomycin (50 µg/ml) for construct B. Changes of the purification protocol had to be made as a two-step affinity-based purification was performed to purify only dimers containing both My5 constructs and the purification tag cleavage was not performed. The frozen cell pellet was resuspended in 5 ml pre-cooled lysis/wash buffer (100 mM Tris, pH 8.0, 150 mM NaCl, 20 mM Imidazole) per gram of protein pellet (wet weight) and lysed on ice by sonication. Insoluble fractions of the cell lysate were removed by centrifugation at 43 500 g, 4°C for 1 h (Centrifuge Avanti JXN-26, Beckman-Coulter). The supernatant was filtered with a 0.45 µm sterile filter (Filtropur S 0.45, Sarstedt) and applied at room temperature to a pre-equilibrated gravity flow Ni-NTA IMAC column (3 ml bed volume per 50 ml of lysate). Washing was performed by applying 10 CV of lysis/wash buffer to remove non-binding proteins. Bound protein was eluted by adding 1.5 CV of elution buffer (100 mM Tris, pH 8.0, 150 mM NaCl, 300 mM Imidazole). Flow through was collected and applied to a pre-equilibrated Strep-Tactin® column (IBA Lifesciences, Germany). Washing was performed by applying five column volumes (CV) of a second wash buffer (100 mM Tris, pH 8.0, 150 mM NaCl). Protein was eluted by applying 2 CV of Strep-elution buffer (100 mM Tris, pH 8.0, 150 mM NaCl, 5 mM Desthiobiotin) and concentrated (Amicon Ultra, PLGC Ultracel-PL Membran) to volumes and concentrations suitable for size exclusion chromatography (≤2 ml and ≤50 mg per SEC run). The protein was injected on a pre-equilibrated Superdex 75 16/60 pg column and eluted from the column at a constant flow rate of 1.0 ml/min. Fractions containing the protein were detected by online absorption measurements at 280 nm, collected and checked for purity by SDS-PAGE. Fractions with the highest purity were pooled, where applicable concentrated to approx. 500 µM, aliquoted to 300 µl, flash frozen in liquid nitrogen and stored at -80°C for further use.

4.3.3 Titin-Obscurin complex crystallisation

For crystallisation of the complex, titin Z8Z9 and obscurin O58O59 were mixed and purified *via* gel filtration. The protein complex was collected and concentrated to 10 mg ml^{-1}. Five different commercial screens were tested for crystallisation (Qiagen classic-I, Quiagen classic-II, Quiagen Pact, Quiagen ProComplex, and Molecular Dimensions Morpheus HT-96), each screening for 96 different conditions. Crystals were grown at 19°C by vapour diffusion by mixing equal volumes of protein solution and precipitant solution.

4.3.4 Data collection, structure solution and refinement

Data were collected to a maximum resolution of 2.1 Å at the European Synchrotron Radiation Facility (ESRF, Grenoble) at the MASSIF-1 (Massively Automated Sample Selection Integrated Facility) beamline. Indexing, reduction and scaling of the diffraction data was done applying XDS program suit (Kabsch, 2010). XDS generates a list of unique reflections, identifies equivalent intensities due to the crystal symmetry, it averages them and calculates an error (R factor) for each set of intensity values. The maximum resolution of the data was determined based on completeness and the CC1/2 correlation coefficient (Karplus and Diederichs, 2012, 2015). The structure was solved using molecular replacement with Phaser from the PHENIX suite (McCoy *et al.,* 2007; Adams *et al.,* 2010) for phase estimation using the structures of obscurin O58O59 and titin Z8 as search models (unpublished in-house data). The model was completed through several cycles of manual rebuilding in Coot (Emsley *et al.,* 2010) and refined with phenix.refine (Adams *et al.,* 2010). Model validation was performed using MolProbity (Chen *et al.,* 2010).

4.3.5 Small angle X-ray scattering - SAXS

Synchrotron SAXS data were collected at EMBL P12 beamline (DESY, Hamburg) (Blanchet *et al.*, 2015) at a wavelength of 1.24 Å and recorded at 20°C using a PILATUS 2M pixel detector (DECTRIS) at a sample–detector distance of 3.1 m.

The incident beam had a size of 120 × 200µm full width at half-maximum in a 1.7 mm quartz capillary held under vacuum. The SAXS intensities were reduced to I(s) vs s using the integrated

analysis pipeline SASFLOW (Franke *et al.*, 2017). s is calculated by $4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and $\lambda = 0.124$ nm (10 keV) the wavelength.

Two different methods were employed to collect SAXS data. In a standard 'batch' mode, the samples were loaded by an automated sample changer. Twenty successive exposures of 50 ms each were collected and processed by the integrated data analysis pipeline. Data collected on BSA were subsequently used to calibrate the forward scattering intensities at zero angle I(0).

For SEC-SAXS experiment the samples was applied to an online SEC column prior to SAXS data collection. The full eluent low-flow stream coming from the SEC column was diverted through the quartz capillary and 3000–4000 individual frames were collected with 1 s exposure (Graewert *et al.*, 2015). For the background subtraction, suitable buffer frames were identified. Frames corresponding to the component of the SEC elution peak and with consistent R_g values were averaged to produce the final reduced SAXS profile which was processed in the same manner as conventional SAXS profiles obtained with batch measurements.

For calculation of the scattering intensity $I(0)$ and the radius of gyration (R_g), the Guinier approximation implemented in PRIMUS was used (Konarev *et al.*, 2003). The pair-distance distribution function $P(r)$ was evaluated with GNOM (Svergun, Semenyuk and Feigin, 1988; Svergun, 1992). Besides the automated processing pipeline SASFLOW, additional programs of the ATSAS package (Franke *et al.*, 2017) were used. To calculate model scattering profiles from high-resolution atomic structures and evaluate the fit of the model profile to experimental data CRYSOL (Svergun, Barberato and Koch, 1995) was used. The discrepancy of data-model fits is reported using the reduced χ^2 test. To fit averaged theoretical scattering intensities derived from an ensemble of conformations to experimental SAXS data EOM (Tria *et al.*, 2015) was applied. To spatially align the available crystal structures with the generated *ab initio* models SUPCOMB (Kozin and Svergun, 2001) was applied. Shape determination is performed through ten runs of DAMMIF (Franke and Svergun, 2009) from which a starting model is derived through DAMAVER (Volkov and Svergun, 2003). A final run of DAMMIN (Svergun, 1999) using this starting model leads to a well-refined model that can be studied to gain information on the shape of the molecule.

4.3.6 Protein and protein-complex characterisation

4.3.6.1 Estimation of secondary structure content by circular dichroism

To estimate the molecule secondary structure composition and therefore validate the correct folding state of generated protein constructs circular dichroism spectroscopy can be applied. Circular dichroism can be defined as the unequal absorption of right- and left-handed circular polarised light. When asymmetric molecules interact with light, they can absorb right- and left-handed circularly polarised light to different extents. In proteins, different structural elements like α-helices, β-sheets and disordered regions have characteristic CD spectra due to a change of the optical properties of the polypeptide backbone amides chromophores (Greenfield, 2007).

All titin Z8Z9 and obscurin O58O59 mutants were tested for correct folding. Proteins were dialysed overnight at 4°C against sodium phosphate buffer (125 mM Na-F, 5 mM Na-P, pH 7.4) and diluted in the same buffer to final concentrations of ~5 µM. CD spectra were obtained in triplicates using Chirascan V100 Circular Dichroism Spectrometer (Applied Photophysics), measuring in 1nm intervals between 190 nm and 260 nm using a quartz cuvette with a 1 mm path length. Sample temperature was kept at constant 20 °C using a temperature-controlled water bath. Each sample was measured in triplicates and background-subtracted. The secondary structure content was analysed by using BeStSel (Beta Structure Selection) (Micsonai *et al.,* 2018).

4.3.6.2 Complex characterisation by ITC

Isothermal titration calorimetry (ITC) is a method enabling the determination of thermodynamic reaction parameters of two reaction partners in solution by measuring the heat change upon complex formation. Therefore, one reaction partner is titrated into the other one, while the shift in temperature is measured in comparison to a reference cell with identical volume and solvent. With known concentrations of both reaction partners it is possible to directly measure the stoichiometry (n), the change in enthalpy (ΔH) and the association/dissociation constant (K_a/K_d). Using the equation for the Gibb's free energy ($\Delta G = -RT \ln K_a = \Delta H - T\Delta S$, where R is the ideal gas constant (8.314472 J mol^{-1} K^{-1}), T the absolute temperature (K), and $K_a = [AB]/[A][B] = 1/K_d$ describing the relation between the bound and unbound reactants, in a reaction with A and B.

ITC experiments were performed using a MicroCal VP-ITC calorimeter (MicroCal Inc./ GE Healthcare) with an injection syringe volume of 273 µl and a cell volume of 1 420 µl. Both proteins were dialysed overnight at 4°C against freshly degassed buffer (100 mM HEPES, pH7.5, 125 mM NaCl) to avoid buffer mismatches and measurement errors from dilution effects. Protein concentrations were set to 30 µM for obscurin as analyte in the cell, and 300 µM for titin as titrant in the syringe. ITC experiments were performed in triplicates at a temperature of 25 °C, injecting 10 µl of titrant in 27 steps. The data were corrected by the heat of injection calculated from the basal heat remaining after saturation. A one-site binding mode was applied to fit the data using a nonlinear least squares algorithm. Data processing was done using Origin software package (Origin Lab).

4.3.6.3 Thermofluor

To compare the thermal stability of wild-type (wt) obscurin and the HCM-related obscurin mutant R4344Q a thermal stability assay (Pantoliano *et al.*, 2001) was performed. 23 µl of the protein was mixed with 2 µl of the hydrophobic dye SYPRO Orange. Increased fluorescence signal due to binding of SYPRO Orange to the hydrophobic protein core upon denaturation was detected using a MyIQ RT-PCR instrument (BioRad) by ramping the temperature from 5°C to 95°C (1°C / min). Measurements were performed in triplicates.

4.3.6.4 Fluorescence Anisotropy

To measure the binding between obscurin O58O59 wild-type / R4344Q and the soluble part of phospholamban (PLN1-19) fluorescence anisotropy measurements were performed using a fluorescein isothiocyanate (FITC) labelled PLN1-19. Fluorescence anisotropy is an effect where the emitted light of a fluorophore shows different intensities along different axes of polarisation and is therefore polarised itself. The magnitude of polarisation is described as anisotropy. Fluorophores have a transition moment for absorption and emission that is defined as a vector aligned along a specific orientation within the fluorophore structure. In a solution with a homogeneous particle distribution the fluorophores are all randomly oriented. When exposed to polarised light, those fluorophores oriented in a specific direction are preferentially excited. If the orientation of the molecule does not change between absorption and emission, the

polarisation of the emitted light stays constant. Emitted light can become depolarised when rotational diffusion changes the direction of the transition moments within the lifetime of the excited state. The rotational diffusion is dependent on the solvent's viscosity and the shape and size of the rotating molecule. Upon binding of a smaller labelled peptide to a larger unlabeled protein the size of the optically active particle increases dramatically, thus reducing the rotational diffusion rate, which is detectable as an increased anisotropy signal. Proteins were dialyzed against 100 mM HEPES (pH 7.5), 125 mM NaCl buffer solution for 2 h at 4°C prior to measurement. A 1:4 dilution series of obscurin O58O59 wild-type or R4344Q was prepared with a final starting concentration around 150 µM. FITC-PLN1-19 peptide concentration was kept constant at 6.6 nM in all samples. Fluorescence anisotropy assays were performed using a Tecan Infinite M1000 (Tecan) regulated at 20°C. Samples were measured in a black 96 well plate (Greiner). Excitation wavelength was set to 470 nm and emission was measured at 530 nm.

4.3.6.5 Binding studies by Microscale Thermophoresis (MST)

A second attempt to measure the binding between obscurin O58O59 wild-type / R4344Q and the soluble part of phospholamban (PLN1-19) microscale thermophoresis (MST) was performed using the Monolith NT.115 devise (NanoTemper technologies) (Seidel *et al.*, 2013). Thermophoresis is an effect observed in mixtures of molecules in solution where the different molecules exhibit different responses to the force of a temperature gradient dependent of the size, charge and solvation shell of the molecule. MST uses either the intrinsic fluorescence of a protein (label-free method) or the fluorescence of a labelled protein to monitor changes in the thermophoretic behaviour upon binding to another molecule. Thermal gradient is induced by an infrared laser where molecules diffuse from the highest temperature point to the lowest. To calculate the dissociation constant, a binding curve must be measured with the fluorescently labelled sample kept at constant concentrations and different concentrations of the binding partner. Obscurin O58O59 wild-type and R4344Q mutant, as well as the fluorescein isothiocyanate (FITC), labelled peptide PLN1-19 were dialysed overnight at 4°C against 100 mM HEPES buffer (pH 7.5, 125 mM NaCl). Experiment was performed according to the manufacturer's protocol. A dilution series of 16 samples of obscurin O58O59 wild-type or R4344Q mutant was prepared with a starting concentration approx. 600 µM in a 1:1 dilution. To each sample the same

volume of labelled peptide FITC-PLN1-19 with a concentration of 400 nM was added. Samples were loaded in Monolith NT.115 Standard Treated Capillaries (Nanotemper) and measured at 20°C.

4.3.7 Variant search

In the attempt to analyse the domains of titin Z8Z9, obscurin O58O59 and myomesin My5 for variants and their possible impact on the domains behaviour different databases were used. For obscurin (transcript: ENST00000284548) and myomesin (transcript: ENST00000356443) the Exome Aggregation Consortium (ExAC) (Exome Aggregation Consortium *et al.*, 2015) was used. For the titin domains the database TITINdb (Laddach *et al.*, 2017) was used. Interface variants were searched based on the residues identified in the crystal structure to be involved in the interface. Variants with assigned reference SNP ID numbers (rs numbers) were additionally picked for further structural investigation of potential involvement in cardiomyopathies.

4.3.8 Sequence and structural alignment

To evaluate the level of sequence conservation of the titin Z8Z9 (UniProt ID: Q8WZ42; residues 1703-1930) and obscurin O58O59 (UniProt ID: Q5VST9; residues 4334-4522) domains the individual sequences were compared to the homologous titin Z8, Z9, and obscurin O58, and O59 sequences of mouse (*Mus musculus*), chicken (*Gallus gallus*), horse (*Equus caballus*), atlantic salmon (*Salmo salar*) as well as to the ten M-band titin domains M1-M10. Multiple sequence alignments were generated using Clustalo (version 1.2.4) (Sievers *et al.*, 2011).

5 Results

5.1 The titin-obscurin complex

To characterise the titin-obscurin complex at the sarcomeric Z-disk constructs of titin Z8Z9 and obscurin O58O59 were designed, cloned and purified. The formed complex was crystallised, and its atomic structure solved to a resolution of 2.1 Å revealing two possible complex arrangements. In order to verify the structure and to find the correct arrangement SAXS measurements were performed and the effect of mutations within the formed interfaces was studied using biophysical approaches. Furthermore, a screen for known variants possibly affecting the interfaces and for disease-related variants was performed using databases combining large-scale sequencing projects. Then specific disease-related variants which have been described in recent literature were investigated, and the impact on the titin-obscurin complex is shown. In collaboration with the Gautel lab in London, the structural information was used to perform *in vivo* experiments to investigate the effect of mutations on the complex formation. At the same time, it was possible to locate the position of potential phosphorylation sites within titin Z8 with respect to the complex interface.

5.1.1 Titin Z8Z9 and obscurin O58O59 form a binary complex

For the complex formation between titin Z8Z9 and obscurin O58O59, both proteins were separately expressed in *E. coli* and purified *via* IMAC and SEC to high purities. Obscurin appeared to be very stable resulting in large amounts of protein (>20 mg per litre of culture). Titin, on the other hand, was only purifiable in relatively small amounts (<2 mg per litre of culture). To assess the stability of both proteins a degradation test was performed by incubating aliquots of both proteins over a period of time at room temperature (Figure 7). While the fresh samples did not show any degradation product, strong degradation can be seen for titin after seven days. Obscurin, on the other hand, remains relatively stable.

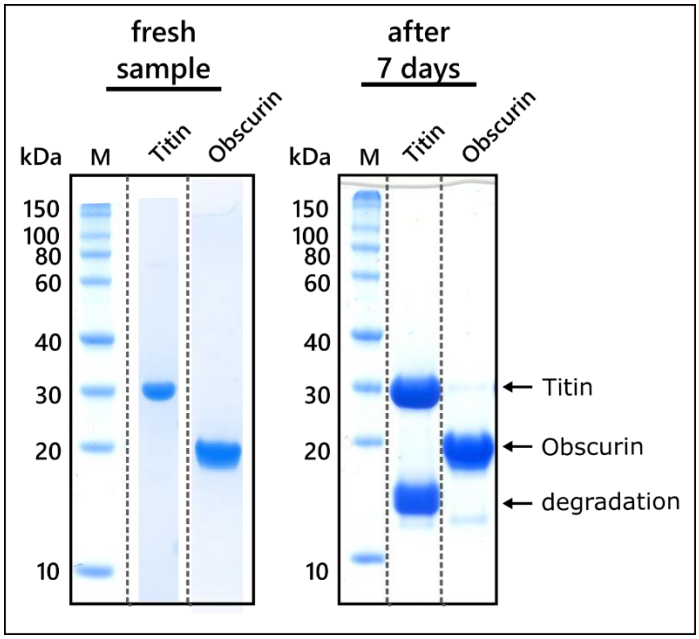

Figure 7 Degradation test of the titin and obscurin. Samples were incubated at room temperature (RT) over seven days. While obscurin appears stable, titin degrades over time into ~15 kDa large fragments, presumably single Ig domains. Theoretical molecular weight are 26.75 kDa for titin Z8Z9 and 20.69 kDa for obscurin O58O59.

For complex formation, both purified proteins titin Z8Z9 and obscurin O58O59 were combined in equal molarities at high concentrations (~10 mg ml^{-1}) which resulted in immediate precipitations. To prevent precipitation upon mixing, it was necessary to dilute both components to ~1 mg ml^{-1} prior to complex formation. While both individual proteins were stable at low temperatures, the mixture of both proteins appeared to be sensitive to low temperatures, heavily precipitating when put on ice. Concentrating the mixture at 4˚C was not possible and needed to be performed at 20˚C. The complex formation between titin Z8Z9 and obscurin O58O59 was confirmed by analytical size exclusion chromatography (SEC) in combination with multi-angle laser light scattering (MALLS) (Figure 8). The chromatogram shows the retention time of the single proteins titin Z8Z9 and obscurin O58O59 with a small shift to an earlier elution for the slightly larger titin and a clear shift for the titin-obscurin complex. This demonstrates that titin Z8Z9 and obscurin O58O58 associate spontaneously and form a stable binary complex. MALLS confirmed the theoretical molecular weight for all constructs (Table 28).

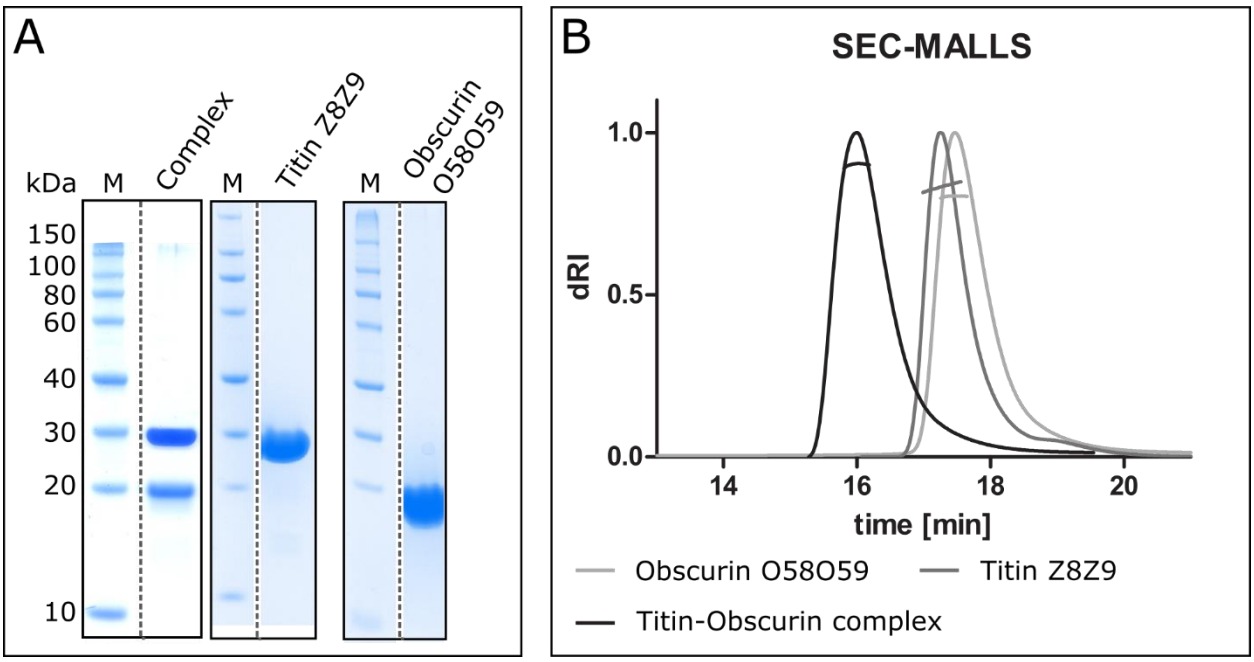

Figure 8 Formation of the titinZ8Z9-obscurinO58O59 complex. A) Comparison of SDS gels from a complex sample as well as from individual titin and obscurin protein samples. B) SEC-MALLS chromatogram of these samples shows that the complex forms. Lines in elution peaks indicate the stability of the measured molecular weight.

Table 28 SEC-MALLS derived molecular weights of the titin-obscurin complex and individual components. The theoretical and the experimental molecular weight indicate correct complex formation.

	theoretical MW (kDA)	measured Mw (kDa)
Complex	47.4	44.7 (±0.6%)
Titin	26.8	30.3 (±1.0%)
Obscurin	20.7	22.2 (±1.0%)

5.1.2 Crystallisation of the titin-obscurin complex

For crystallisation of the complex, titin Z8Z9 and obscurin O58O59 were mixed and purified *via* gel filtration. Crystals grown at 19°C by vapour diffusion were detected after 17 days. Notably, prior to the first crystal detection a temperature increase over several hours to room temperature due to technical problems in the storage room was recorded. Crystal growth was observed in a narrow condition range: 0.2 M sodium nitrate, 20% (w/v) PEG 3350, unbuffered or 0.1 M bis-tris propane pH 6.5-8.5. All crystals appeared in a very thin and plate-like shape. To further evaluate the effect of the temperature shock, the two plates were set up providing the same conditions that led to successful crystallisation. Both were stored at 19°C. After 15 days one

plate was stored at 25°C for 24 h while the other remained at 19°C. Crystals only appeared in the plate that was treated with the temperature shock.

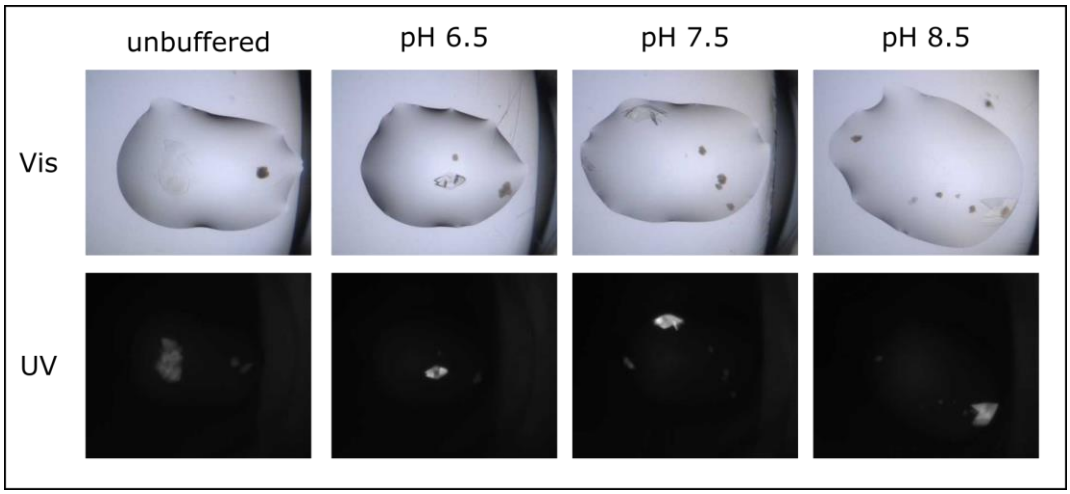

Figure 9 Crystals of the titin-obscurin complex. Crystals grown in 0.2 M sodium nitrate, 20%(w/v) PEG 3350, unbuffered and 0.1 M bis-tris propane at indicated pH values. Images were taken with visible light (Vis) and ultraviolet light (UV) to validate the protein content of the crystals. Crystals appeared in plate-like shape being highly fragile.

The crystals were fished, and the data were collected to a maximum resolution of 2.1 Å at the European Synchrotron Radiation Facility (ESRF, Grenoble) at the MASSIF-1 (Massively Automated Sample Selection Integrated Facility) beamline. The diffraction data and refinement statistics are given in Table 29.

Table 29 Data collection and refinement statistics

Data collection		Refinement	
Beamline	MASSIF-1 (ESRF)	R_{work}/R_{free} (%)	21.7 /26.98
Wavelength (Å)	0.966	Number of residues:	
Resolution range (Å)	28.91 - 2.10	Protein	740
Highest res. bin (Å)	(2.16 - 2.10)	Solvent	372
Space group	C 1 2 1	B-factors [Å^2]:	
Cell dimensions:		Protein	40.3
a, b, c (Å)	96.46 66.09 148.28	Solvent	43.1
α, β, γ (˚)	90.0 103.1 90.0	RMSD:	
Mosaicity (deg.)	0.6	Bond lengths (Å)	0.0155
Completeness (%)	98.2 (98.3)	Bond angles (°)	1.83
Multiplicity	2.9 (2.6)	Ramachandran plot (%)	
Rmerge (%)	10.2 (78.6)	favoured	98.37
I/σ(I)	6.4 (1.1)	allowed	1.49
Wilson B factor	34	outliers	0.14

5.1.3 The crystal content

The asymmetric unit contains two copies of titin Z8Z9 and obscurin O58O59, respectively, resulting in a total domain count of eight domains. They form a ring-like structure with an outer diameter of approx. 90 Å and an inner diameter of approx. 30 Å with an alternating A-B-A'-B'-A-B-A'-B' arrangement (Figure 10 A and B). While the linker between both obscurin domains O58 and O59 could be structurally determined, a 39 amino acids long region between residue Z8 E1797 and residue Z9 K1836 of the 47 amino acids long linker is not visible in the crystal structure. To investigate whether the missing structural information is due to the high flexibility of the linker rather than the result of a degradation process, the crystal composition after data collection was analyzed *via* silver-stained SDS-PAGE (Figure 11). The crystal used for data collection was washed three times by dipping the crystal loop in a small aliquot (10 µl) of buffer. The SDS-PAGE shows that the crystal is surrounded by degraded protein (Figure 11, 1st wash). However, the core of the crystal (Figure 11, 2nd wash) does only contain intact titin Z8Z9 and obscurin O58O59.

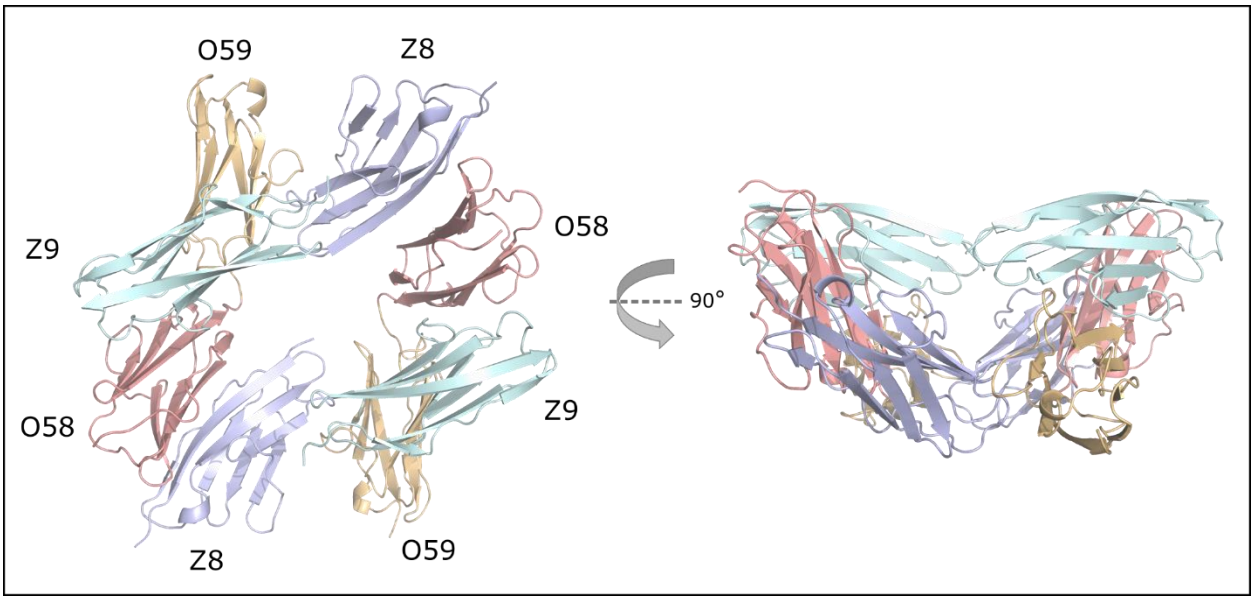

Figure 10 The structure of the titin Z8-Z9 obscurin O58-59 complex. The asymmetric unit contains a set of two heterodimers, each containing one titin Z8-Z9 and one obscurin O58-O59 molecule.

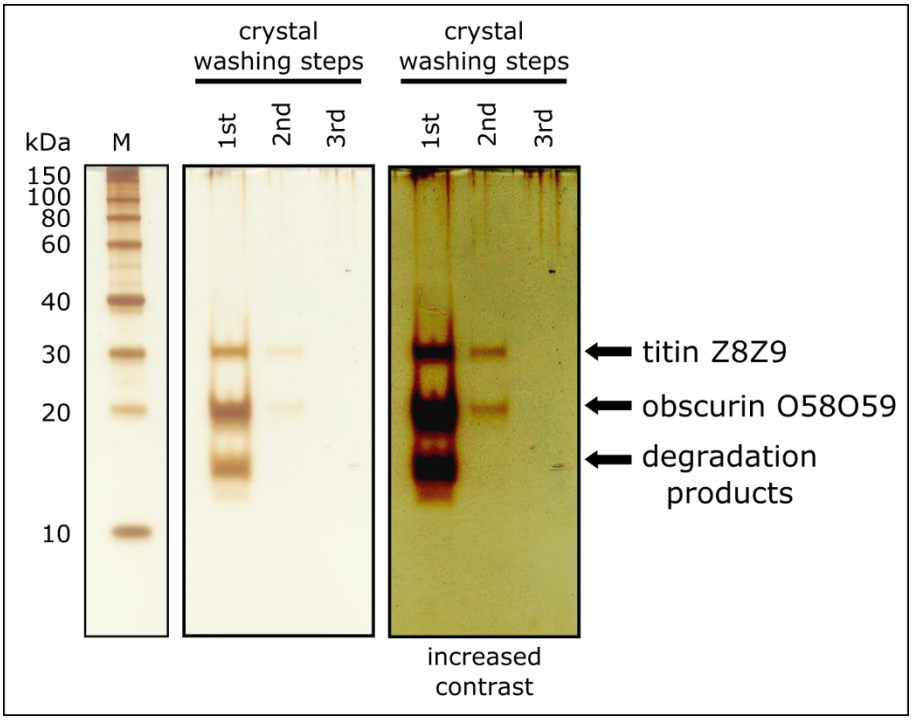

Figure 11 Silver stained SDS-PAGE of different washing steps of the protein crystal used for data collection. Lower molecular bands in the first wash step in combination with a reduced intensity for the band corresponding to titin Z8Z9 indicate degradation of titin in the solution surrounding the crystal. Protein bands in second wash step show no degradation band and intact titin Z8Z9 in the same ratio as the obscurin band. This indicates intact protein constructs in the crystal. The third wash does not show protein bands as the crystal completely dissolved during washing step 2.

5.1.4 Domain comparison of titin Z8Z9 and obscurin O58O59

All domains form a β-sandwich consisting two β-sheets. One is formed by βA-βB-βE-βD and one is formed by βA'-βG-βF-βC (Figure 12 A). The A-strand is missing in both obscurin domain structures most likely as a result of an increased flexibility of the N-terminus due to truncation of the domains from full-length obscurin. The overall architecture can be assigned to the intermediate set (I-set) of the Ig family (Figure 12 A). I-type Ig domains have a split A-strand, resulting in very short A and A' strands that belong to the two different β-sheets in the fold. Notably, the isoelectric point (pI) of the individual domains show a large distribution. While titin domains have a relatively high pI of 9.18 (Z8) and 8.71 (Z9) resulting in a positive net charge of the domains in solution at pH 7.5, the pI-values of the obscurin domains ranges between 4.47 (O58) and 5.36 (O59) indicating a negative net charge. The sequence identities range between 31% for both obscurin domains and 18% for titin Z8 and obscurin O59 (Table 30).

The structural superposition of all domains reveals a root mean square (r.m.s.) deviation between 1.1 and 1.6 Å (Table 30) and a structural diversity between 0.7 and 0.9 Å after subtraction of the estimated coordinate error and conformational diversity of surface residues. This illustrates the high level of structural conservation despite the relatively low sequence identity. The sequence alignment (Figure 12 B) shows seven highly conserved residues all located within or very close to β-strands and contributing to the hydrophobic core. A sequence comparison with ten titin Ig domains located within the sarcomeric M-band shows that especially the tryptophan in β-strand C, the leucine in β-strand E, and the leucine in β-strand G are highly conserved. As a reference, the titin-obscurin complex domains were compared with the ten titin Ig-domains from the M-band. The AxFxC motif that is highly conserved in all domains of the titin-obscurin complex does not seem to be conserved in the other Ig-domain sequences. Even though the cysteine in this motif can also be found in four of the ten M-band domains, only the titin domain M1 indicates the formation of a disulfide bridge within the hydrophobic core of the domain. Apart from this cysteine, all domains contain at least one additional cysteine. None of the cysteines is exposed or involved in an intramolecular disulfide bridge.

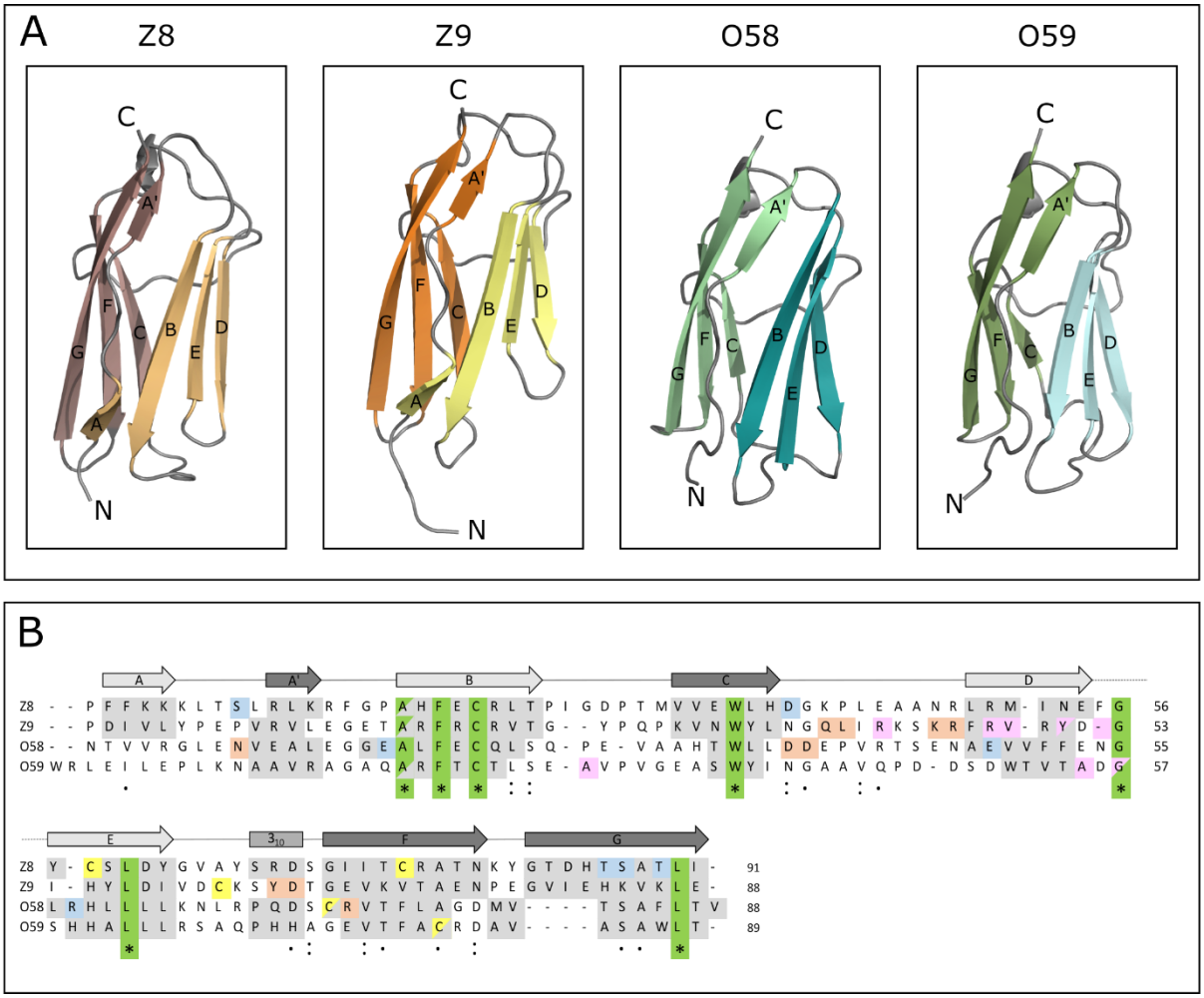

Figure 12 Individual Ig domain structures and alignment of domains involved in the titin-obscurin complex. A) Typical I-set Ig domain structures of titin domains Z8 and Z9 and obscurin domains O58 and O59. Obscurin domains are lacking the structural features of the A-strand, titin Z9 does not form the typical 3_{10}-helix at the C-terminal tip of the domain. β-strands involved in the same β-sheet are displayed in the same colour code. B) Sequence alignment of the four Ig domains. Residues marked in green are the conserved residues. Residues marked in blue are involved in the Z8O58 interaction, residues marked in orange are O58Z9 interactions, and residues marked in pink are Z9O59 interactions. Asterisks indicate positions which have a single, fully conserved residue, colons indicate conservation between groups of strongly similar properties, periods indicate conservation between groups of weakly similar properties. The positions of the secondary structural elements shown on top of the alignment correspond to titin Ig domain Z8. Light and dark grey indicate the involvement of the β-strands in the different β-sheets. Highlighted in grey is the extent of secondary structural elements for the different Ig domain sequences as extracted from the respective structures. Sequence borders have been defined according to the UniProt sequence range of the respective domains.

Table 30 Root mean square deviation and sequence identities of the titin-obscurin complex domains.

RMSD [Å]				
Domains	Z8	Z9	O58	O59
Z8	-	1.4	1.4	1.6
Z9	1.4	-	1.3	1.5
O58	1.4	1.3	-	1.1
O59	1.6	1.5	1.1	-

Sequence Identity [%]				
Domains	Z8	Z9	O58	O59
Z8	-	21.7	26.5	18.1
Z9	21.7	-	24.1	24.1
O58	26.5	24.1	-	31.3
O59	18.1	24.1	31.3	-

5.1.5 Architecture of the Z8Z9-O58O59 complex

Due to the fact that the linker between both titin domains is missing, there are two possible complex arrangements. In both cases, there is a core interaction between O58-Z9-O59. Titin Z8 can either form an interface with obscurin domain O58 resulting in a complex, referred to as complex A, or an interface with obscurin domain O59 referred to as complex B (Figure 13).

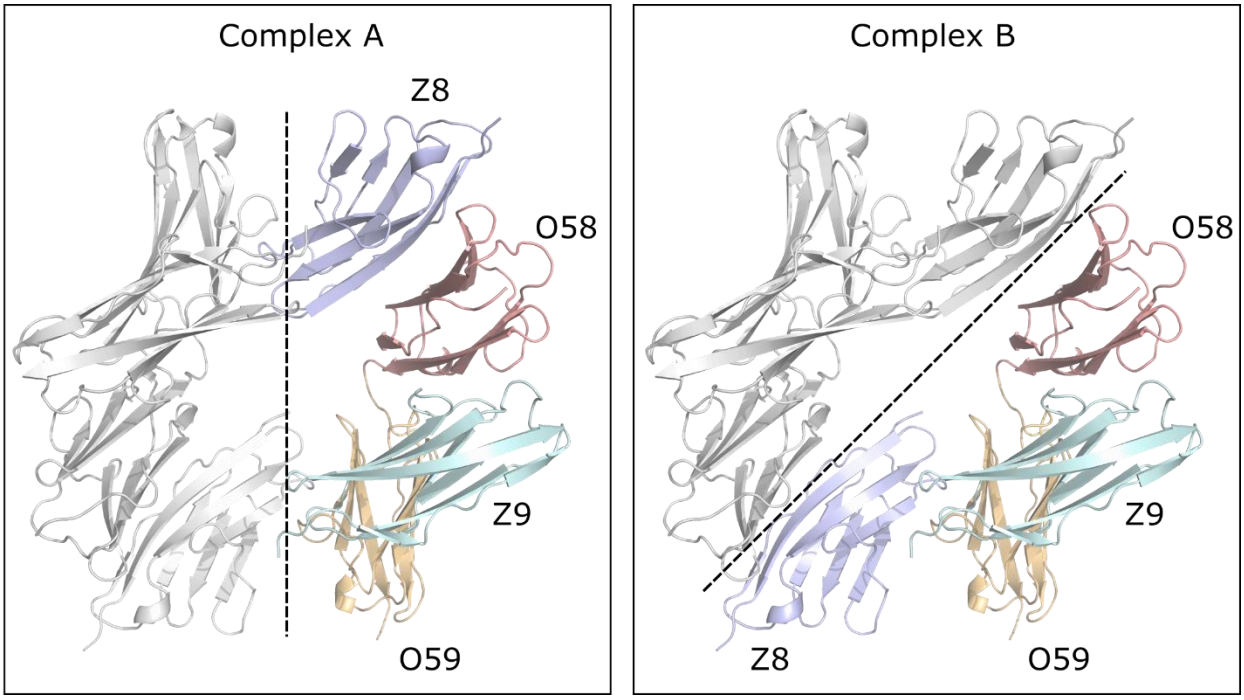

Figure 13 Different possible positions of titin Z8 in cartoon representation. Division of the asymmetric unit is indicated by the dotted line. Complex A arrangement forms an interface between Z8 and O58. Complex B forms an interface between Z8 and O59. The interface forming between O58-Z9-O59 is not affected by the two possible arrangements. Validation experiments presented in the next sections show that complex A most probably represents the arrangement in solution.

The 'Protein interfaces, surfaces and assemblies' service PISA (Krissinel and Henrick, 2007) was used to evaluate the properties of the different interfaces (Table 31). The titin Z8 obscurin O59 interface (Figure 14 A) is the smallest interface with a surface area of 323.9 $Å^2$ and only one defined interaction formed by three hydrogen bonds between Z8 Lys1782 (located in the loop region between β-strands F and G) and O59 Ser4465, Glu4463 (both located in the loop between β-strands B and C), and Ala4507 (located in β-strand F). All of these hydrogen bonds have a distance larger than 3.2 Å and, thus, must be considered as weak, electrostatic interactions (Jeffrey, 1997). The solvation free energy gain upon interface formation (ΔiG) has a value of -2.6 kcal mol^{-1} and is thus, together with the value of the Z9-O59 interaction, the lowest. Nevertheless, compared to the Z9-O59 interface the ΔiG P-value of the Z8-O59 interface is additionally close to 0.5 indicating that this interaction is likely to be an artefact of crystal packing.

In contrast to this, the alternative interface between Z8 and O58 (Figure 14 B) has an interface area of 496.8 $Å^2$, involving five hydrogen bonds and three salt bridges. All hydrogen bonds in this interface have a distance <2.95 Å and therefore can be considered moderate and mostly electrostatic (donor-acceptor distance of 2.5-3.2 Å). The ΔiG value has doubled compared to the alternative Z8 interaction, and the ΔiG P-value is notably lower indicating the higher relevance of this interaction. Specific interactions within the titin Z8 – obscurin O58 interface mainly involve residues from the obscurin O58 domain β-sheet (I) and β-strand G of β-sheet (II) of the titin Z8 domain. They are distributed over three interface clusters: i) residue E4356 from the obscurin O58 domain interact with S1789 and T1788 from the titin Z8 domain; ii) E4387 from the obscurin O58 domain interacts with T1791 and S1711 from the titin Z8 domain; iii) R4396 from the obscurin O58 domain interacts with D1741 from the titin Z8 domain. Furthermore, there are extensive hydrophobic interactions between β-strand D from the obscurin O58 domain and the C-terminal part of β-strand G from the titin Z8 domain.

The Z9-O58 interaction (Figure 14 C) has an interface area of 561.0 $Å^2$ and involves seven hydrogen bonds and six salt bridges. The ΔiG value is slightly lower than for the Z8-O58 interaction. All hydrogen bonds (with one exception for the K1885-D4376 interaction with a distance of 3.4 Å) have a donor-acceptor distance less than 3.2 Å and can, therefore, be considered moderate and mostly electrostatic. In contrast to the findings for the titin Z8 –

obscurin O58 interface, specific interactions of the second interface between the titin Z9 domain and the obscurin O58 domain involve β-strand connecting loops. A large cluster involving seven residues is formed by D4376 and D4377 from the loop connecting β-strands C and D of the obscurin O58 domain and residues K1885/R1886 from the loop connecting β-strands C and D from the titin Z9 domain, as well as residues Y1905/D1906 from the loop connecting β-strands E and F of the titin Z9 domain and R4411 located on β-strand G of the obscurin O58 domain. A second interaction cluster is formed by N4348 located N-terminally to β-strand A' of the obscurin O58 domain and Q1879/L1880 of the titin Z9 domain.

The last interface that is located between the titin Z9 domain and obscurin O59 domain has an interface area of 469.4.0 $Å^2$ and is mainly formed by an antiparallel inter-subunit β-sheet involving residues from the C-termini of β-strands D of both protein partners, Y1891 of the titin Z9 domain and A4484 of the obscurin O59 domain. In total five hydrogen bonds and two salt bridges are formed resulting in a ΔiG half as high as for the Z8-O58 interaction. Additional specific interactions are formed by V1889 from β-strand D of the titin Z9 domain and G4486 from the loop connecting β-strands D and E of the obscurin O59 domain, as well as R1882 from the loop connecting β-strands C and D of the titin Z9 domain and A4458 from the loop connecting β-strands B and C of obscurin O59. Furthermore, one additional specific interaction involving residues from β-strands D of both domains, R1888 from the titin Z9 domain and D4485 from the obscurin O59 domain can be observed.

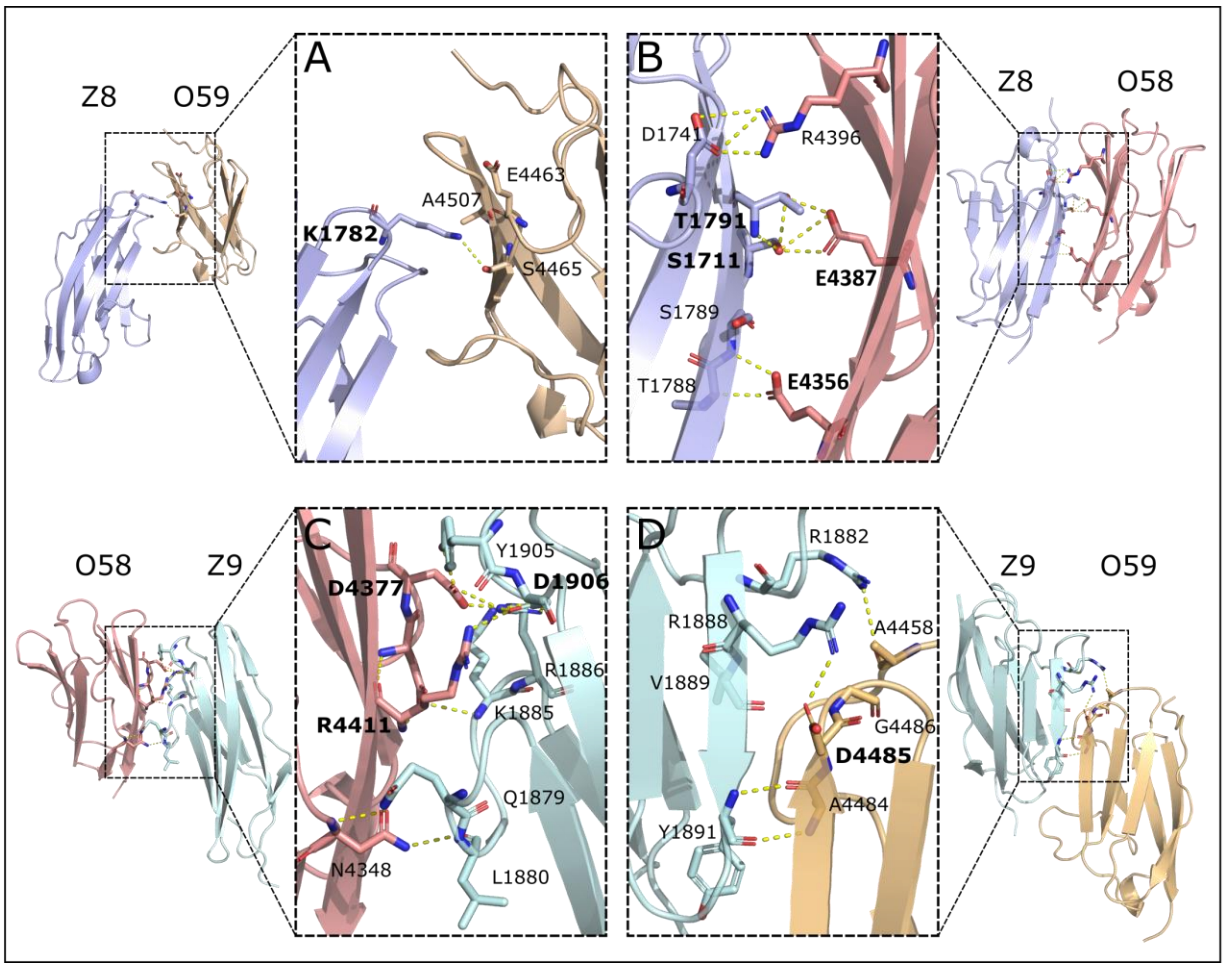

Figure 14 Interfaces of the titin obscurin complex. Panel A) and B) show the two possible interactions for titin Z8 according to the arrangement in the asymmetric unit. A) Interaction formed in complex B involving Z8 and O59. B) Interaction in complex A formed by Z8 and O58. Interactions of titin Z9 is shown in panel C) and D). Shown residues are involved in the interaction interfaces according to PISA (Protein interfaces, surfaces and assemblies) service at the European Bioinformatics Institute (Krissinel and Henrick, 2007). Residues in bold were further studied by ITC.

Table 31 List of unique interfaces formed by the four domains titin Z8, Z9, and obscurin O58, and O59. iNres indicates the number of residues involved in the corresponding interface. ΔiG indicates the solvation free energy gain upon interface formation, in kcal/mol. ΔiG P-value indicates the P-value of the observed solvation free energy gain. NHB indicates the number of potential hydrogen bonds across the interface. NSB indicates the number of potential salt bridges across the interface.

	Interaction partner 1		Interaction partner 2						
Interface	Domain	Surface [Å^2]	Domain	Surface [Å^2]	Interface area [Å^2]	ΔiG [kcal/mol]	ΔiG P-value	NHB	NSB
Z8-O58	Z8	6053	O58	5457	496.8	-5.2	0.298	8	3
Z9-O58	Z9	6007	O58	5457	561	-4.3	0.252	7	5
Z9-O59	Z9	6007	O59	5515	469.4	-2.6	0.322	4	1
Z8-O59	Z8	6044	O59	5515	323.9	-2.6	0.471	3	0

5.1.6 Complex interface validation by introduction of mutations

To validate the interfaces presented in the crystal structure, quantitative affinity measurements by isothermal titration calorimetry (ITC) with mutants of key residues contributing to the protein-protein-complex interaction interfaces were performed (Table 32). Correct folding of all purified constructs was verified by CD measurements prior to ITC experiments. The wild-type complex binds with a dissociation constant (K_d) of 4.4 µM and a stoichiometry of approximately 1:1, confirming both the *in vitro* observations with SEC-MALLS and the structural findings of a heterodimeric 1:1 titin Z8Z9 – obscurin O58O59 complex. For all tested mutants where binding was measurable, the stoichiometry remained approximately 1:1. The interaction is mainly enthalpy driven with ΔH = -7.0 kcal mol^{-1} for the wt titin Z8Z9 – obscurin O58O59 complex, while the entropic contribution (TΔS) contributes to a lesser extent. The interaction of several titin Z8Z9 or obscurin O58O59 mutants was tested and presented below (Table 32). The relationship between enthalpic and entropic factors is expressed in the free energy equation $G = H - TS$ and also was assessed for the introduced mutations. In most cases an increase of the enthalpy (H becomes less negative) can be observed indicating a weaker binding, which is coupled with a decrease of the $-TS$ term (more negative), indicating the system becomes increasingly disordered and less tightly packed.

For titin Z8 both possible interfaces were analysed. The loss-of-function mutation K1782A introduced in the smaller interface between Z8 and obscurin O59 (complex B) did not show a decrease of complex stability between both proteins once again hinting to the relevance of the alternative Z8-O58 interface.

For the Z8-O58 interaction interface found in the arrangement of complex A, five different mutants, four of which are directly involved in the interaction (Z8 S1711D, Z8 T1791V, O58 E4387R, O58 E4356R/E4387R) and a residue not involved in the interface (Z8 T1727D) were analysed. As will be further described in section 5.1.13.1 Z8 S1711D was used as a phosphomimetic mutation as it is part of the kinase recognition motive KKKXTS and might hint to a potential phosphorylation site. O58 E4387R, involved in the same interaction cluster, and the double mutation O58 E4356R/E4387R were designed as charge reversal mutations. Both mutants show a two-fold decrease in complex stability compared to the wild-type. In contrast, the Z8 mutant T1791V shows a two-fold increase in affinity by contributing to a hydrophobic residue-rich environment that can compensate for the enthalpy loss due to the loss of a hydrogen bond by hydrophobic interactions and a seven-fold entropy increase. As expected, the Z8 mutant T1727D that is located outside of the interface does not show a change in interaction between titin and obscurin.

The Z9-O58 interface was validated by the introduction of O58 D4377R, O58 R4411E, Z9 D1906N, and Z9 D1906R. The strongest influence on the titin-obscurin interaction shows the charge reversal mutant O58 D4377R with a four-fold decrease in stability. D4377 is part of and main contributor to the larger interaction cluster at the Z9-O58 interface and interacts with Y1905 and R1886 *via* hydrogen bonds. For O58 R4411E only a minor increase in K_d and thus, decrease of the stability is observable. R4411 interacts with D1906 *via* a side chain hydrogen bond pulling it toward R1886 and stabilising the former mentioned D4377 interaction cluster. Furthermore, with its backbone R4411 provides two intramolecular hydrogen bonds to D4376, one to the backbone and one to the sidechain coordinating the D4376-K1885 interaction. For Z9 D1906 first D1906R as a charge reversal mutation was aimed to be tested. Purification of this mutant was not possible due to strong aggregation, indicating the structural importance of this residue. The milder Z9 mutation D1906N shows a decrease of the complex stability, comparable to the decrease induced

by the Z8-O58 mutations. Nevertheless, all tested mutants for the interaction interface Z9-O58 show a decrease of the enthalpic factor in combination with a slight increase of the entropic factor.

The last interaction interface between Z9 and O59 shows an antiparallel β-sheet formation between both domains and only small contribution of other residues to the interaction interface. To disrupt the interaction, different mutants of O59 D4485 were tested. The mild mutation O59 D4485N only neutralises the charged group abolishing the salt bridge to D1888 which seems to contribute to the β-sheet by stabilising and coordinating. It shows a decrease comparable to mutations tested for the other interfaces. The charge reversal mutation O59 D4485R is able to decrease the protein-protein interaction by not only abolishing the salt bridge but also producing a clash of charge that might weaken the β-sheet formation.

To further investigate the structural relevance of the interfaces, cross interface mutations were introduced. The mutant D4377R/D4485R affecting the interfaces Z9-O58 and Z9-O59 shows a four-fold decrease of the stability that is comparable to the individual D4377R and higher than the D4485R mutation. The triple obscurin mutant E4356R/D4377R/E4387R that affects the Z8-O58 and the Z9-O58 interface resulted in a complete disruption of the entire complex formation.

Taking together, one can conclude that the protein complex interfaces as seen in the crystal structure represent the natural domain interplay in solution. The complex is formed by three distinct domain-domain interactions: one between titin Z8 and obscurin O58, a second between titin Z9 and obscurin O58 and a third between titin Z9 and obscurin O59. The alternative Z8-O59 interaction is most likely a crystallographic artefact. The three interaction interfaces are able to compensate for single mutations. Only by simultaneously addressing both the Z8-O58 and the Z9-O58 interface it was possible to disrupt the whole complex formation, indicating the importance of the obscurin 58 titin Z8-Z9 interaction.

Table 32 ITC measurements of titin Z8Z9 and obscurin O58O59 constructs. Introduced mutations are indicated. Stoichiometry (n) of the measured interaction shows the ratio between obscurin to titin. Enthalpic (ΔH) and entropic (-TΔS) factors indicate that the complex formation is enthalpy driven. Dissociation constant (K_d) shows the impact of the introduced mutations on the overall complex stability.

Titin Z8Z9	**Obscurin O58 O59**	**n**	**ΔH (kcal/mol)**	**-TΔS (kcal/mol)**	**K_d (μM)**
wt	wt	0.93±0.09	-7.0±0.7	-0.4±0.7	4.4 ± 0.1
TitinZ8 Obscurin O58 (alternative interface)					
K1782A	wt	0.94±0.12	-6.2±0.3	-1.3±0.4	3.2±0.3
Titin Z8 Obscurin O58 interface mutants					
wt	E4387R	0.90±0.12	-6.4±0.8	-0.5±0.9	8.5±1.6
wt	E4356R/E4387R	0.75±0.03	-6.4±0.6	-0.5±0.6	9.4±0.7
S1711D	wt	0.96±0.08	-5.9±0.7	-1.0±0.8	8.6±1.7
T1791V	wt	0.90±0.02	-5.5±0.4	-2.5±0.6	1.7±0.7
Titin Z9 Obscurin O58 interface mutants					
wt	D4377R	0.98±0.02	-3.5±0.2	-3.0±0.1	16.7±0.5
wt	R4411E	0.90±0.05	-5.2±0.1	-1.9±0.2	6.6±0.8
D1906N	wt	1.11±0.22	-2.7±0.4	-4.2±0.4	8.4±0.2
D1906R	wt	unstable during purification			
Titin Z9 Obscurin O59 interface mutants					
wt	D4485N	1.10±0.05	-4.9±0.3	-2.1±0.4	7.0±1.0
wt	D4485R	0.98±0.03	-2.5±0.3	-4.3±0.6	10.8±5.4
Multiple interfaces					
wt (Z9O58 +Z9O59)	D4377R/D4485R	1.01±0.04	-3.0±0.1	-3.6±0.1	16.1±1.2
wt (Z8O58 +Z9O58)	E4356R/D4377R/E4387R	ND	ND	ND	ND
Negative control (outside of the interfaces)					
T1727D	wt	1.05±0.07	-7.1±0.1	-0.1±0.1	4.5±0.3

5.1.7 Obscurin interaction with individual titin domains

For further interface validation with ITC the individual titin domains Z8 and Z9 were purified. Protein yield of Z8 was observed to be lower than for Z9, despite same purification condition. For Z8 five measurements were performed from which two were not interpretable due to high signal to noise ratio which might be attributed to a lower stability of Z8 at room temperature. The remaining three titration curves also were of low quality, as indicated by the low stoichiometry

value and high standard deviation for the K_d (1.6 µM±1.3 µM), giving only an indication that Z8 alone does interact with obscurin O58O59. For titin Z9 it was possible to measure a K_d of 15.9 µM with a standard deviation (SD) of 1.9 µM. This suggest that the interface between obscurin and titin Z8 is stronger than the interface between obscurin and titin Z9.

Table 33 ITC measurements of titin individual Z8 and Z9 domains and obscurin O58O59 constructs. Stoichiometry (n) of the measured interaction shows the ratio between obscurin to titin. Enthalpic (ΔH) and entropic (-TΔS) factors indicate that the complex formation is enthalpy driven. Dissociation constant (K_d) shows the impact of the introduced mutations on the overall complex stability.

Titin	Obscurin O58 O59	n	ΔH (kcal/mol)	-TΔS (kcal/mol)	K_d (µM)
single Z8	wt	0.74±0.11	-2.6±0.7	-5.5±1.0	1.6±1.3
single Z9	wt	0.94±0.06	-6.5±0.4	-0.1±0.4	15.9±1.9

5.1.8 Cross-species sequence comparison

As a next step, the individual domain sequences were compared to the homologous titin Z8, Z9, and obscurin O58, and O59 sequences of mouse (*Mus musculus*), chicken (*Gallus gallus*), horse (*Equus caballus*), and Atlantic salmon (*Salmo salar*). In all alignments, Atlantic salmon shows the lowest sequence identity. The sequence analysis between titin Z8 and titin Z9 of different species shows a very high overall conservation (Table 34). Sequence identities for obscurin O58 are more diverse ranging between 51% for Atlantic salmon and 94% for horse. The linker between the two titin domains is highly conserved across all species and even identical between human and horse. Obscurin O59 has the lowest sequence identities in all analysed species and is entirely missing in horse. All identified conserved residues within titin Z8 and Z9, and obscurin O58 and O59 are also highly conserved in all of the studied species with the exception of the alanine in β-strand B. Here, an amino acid exchange from alanine to valine can be observed for salmon Z8 and O59 and mouse O59.

Next, the conservation of the residues involved in the complex interfaces was analysed. First, the two possible interfaces formed by titin Z8 were evaluated. The interface between Z8 and O59 shows that three out of four residues are conserved (obscurin residues E4463 and A4507, titin residue K1782) across species even though not conserved between different Ig domains. Only

obscurin residue S4465 shows an amino acid change from serine to threonine for mouse and chicken and serine to alanine for Atlantic salmon. The main interaction in this interface is facilitated by the side chain interaction between K1782 and S4465. While a serine to threonine mutation could still be able to interact, as the hydroxyl group is preserved, a serine to alanine mutation would abolish the interaction. For the alternative Z8-O58 interaction interface arrangement of complex A, all ten involved amino acids are conserved even though not conserved across different domains. The interface between Z9 and O58 also shows highly conserved properties as ten out of ten involved residues are preserved across the analysed species. Only in salmon two amino acid substitutions, an R4411K and an D4377G can be observed. While the impact of the arginine to lysine change in residue 4411 should be small as the biochemical property is preserved, the loss of function substitution for residue 4377 might have an impact on the interaction. As reported in chapter 5.1.6, the mutation of R4411E did not have a major impact while D4377R considerably altered the complex formation. The last interface between Z9 and O59 revealed a larger variation than the previous interfaces. Here, only four out of seven residues are conserved. While the impact of the V1889L substitution found for Atlantic salmon and chicken in Z9 can be considered as minor, the A4458P exchange in O59 of salmon and A4484T in O59 of mouse might have a larger impact on the interface formation.

In summary, most of the interfaces are conserved. The majority of the variants can be assumed to have minor impacts on the interface. Only the D4377G variation seen in obscurin O58 of salmon might have a more drastic impact on the interface. As was demonstrated in performed ITC experiments the complex is able to compensate for single and partly also for double amino acid exchanges. The comparison between the two possible Z8 arrangements shows that both show a conserved pattern, nevertheless the Z8-O58 interaction interface involves more conserved residues. Considering the generally high sequence identity of the individual domains across species the conservation of the interaction interfaces is not surprising and needs to be interpreted carefully to avoid false conclusions. However, the lower sequence identity for the obscurin domain O59 and the absence of this domain in horse species does indicate that this domain might not be play a major role for the physiological function of obscurin.

Table 34 Cross-species sequence identity of the individual titin-obscurin complex domains. Titin Z8 and Z9 as well as the obscurin O58 and O59 domains were compared to the sequences of mouse (Mus musculus), chicken (Gallus gallus), horse (Equus caballus), and Atlantic salmon (Salmo salar). As reference, the human sequences were used. No equivalent for O59 was found in the horse obscurin sequence.

	Mouse	Chicken	Horse	Salmon
Z8	97.8	97.8	97.8	86.8
Z9	97.7	89.8	98.9	78.4
O58	90.9	78.4	94.3	51.1
O59	77.5	56.8	-	39.8
Z8-Z9-linker	95.7	87.2	100.0	63.8

5.1.9 Determination of the solution structure of the titin-obscurin complex and its components using SAXS

For further validation of the obtained high-resolution X-ray crystal structure and investigation of the correct complex arrangement and positioning of titin domain Z8, small angle X-ray scattering (SAXS) experiments were performed with the formed complex and the individual complex components in SAXS batch mode and with size exclusion chromatography (SEC)-SAXS mode. These measurements were performed in collaboration with Melissa Greawert (Svergun group, EMBL-Hamburg).

5.1.9.1 In solution measurements of the individual titin and obscurin complex components

First, the scattering profile of obscurin O58O59 individually in the batch and SEC-SAXS were measured. While the scattering curves measured in batch showed the presence of slightly larger particles, it was possible to improve the curves by SEC-SAXS (Figure 15 A-B). The measurements revealed that the angle of the linker between the two subunits as seen in the crystal structure remains rigid in solution (Figure 15 C).

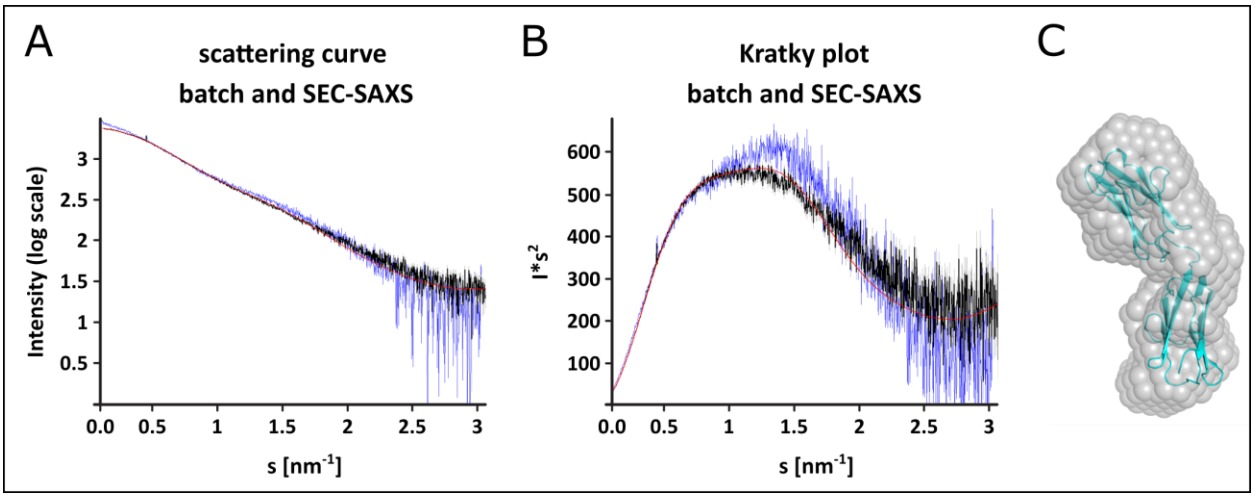

Figure 15 SAXS analysis of obscurin O58O59. Comparison between batch and SEC-SAXS approach. Fit with theoretical scattering curve shows the improvement due to the additional on-line purification. A) scattering curve with batch measurement in blue, SEC-SAXS measurement in black, theoretical scattering curve in red. The reduced and averaged frames that correspond to Obscurin58-59 fraction from the SEC run are fitted with the theoretical scattering curve of the crystal structure (X^2 =4.8). B) With the Kratky plot presentation differences between the curves in the mid s-range are emphasised. C) ab initio 3D reconstruction overlaid with the crystal structure of the double domain construct in cartoon presentation. Data analysis performed in collaboration with Melissa Graewert.

The rigidity of the linker can be attributed to stabilisation at three different locations: i) the sidechain and the backbone of residue E4353 (O58) forms a polar interaction with the backbone of W4430 and R4428 (O58O59-linker), ii) the side chain of residue R4431 (O58O59-linker) interacts with the side chain of D4510 (O59), and iii) the side chain of S4456 (O59) interacts with the backbone of E4433 (O59) linker (Figure 16).

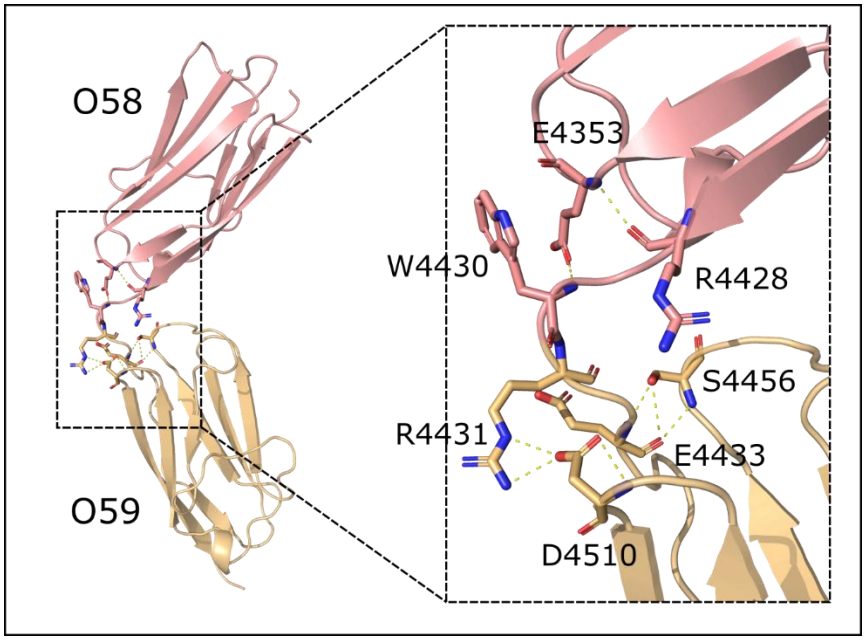

Figure 16 Structural representation of the linker region between obscurin O58 and O59. Linker is stabilised by the interaction of seven amino acids preserving the angle between both domains in solution.

Next, SAXS data of the titin construct Z8Z9 at concentrations between 1.0 and 8.0 mg ml^{-1} were collected (Figure 17). The presence of higher oligomeric species is indicated by the increase of the signal at lower angles for higher concentrations suggesting a strong concentration dependency of the sample. Thus, the scattering profiles of the lowest two concentrations were combined for further analysis (Figure 17 A). The comparison of the theoretical scattering curve from the crystal structure of titin Z8Z9 revealed a mismatch with a chi^2 of 30. This strongly indicates that even the lowest concentrations of the measured sample contained larger species. However, the close approximation to linearity point in the Guinier plot (Figure 17 B) analysis for the extrapolated scattering profile shows the absence of interference effects or heavy aggregation. While the sample does seem folded as the bell-shape of the Kratky analysis (Figure 17 C) suggests, titin seems to be prone to form higher oligomers comprising long distances. The data suggest an elongated shape.

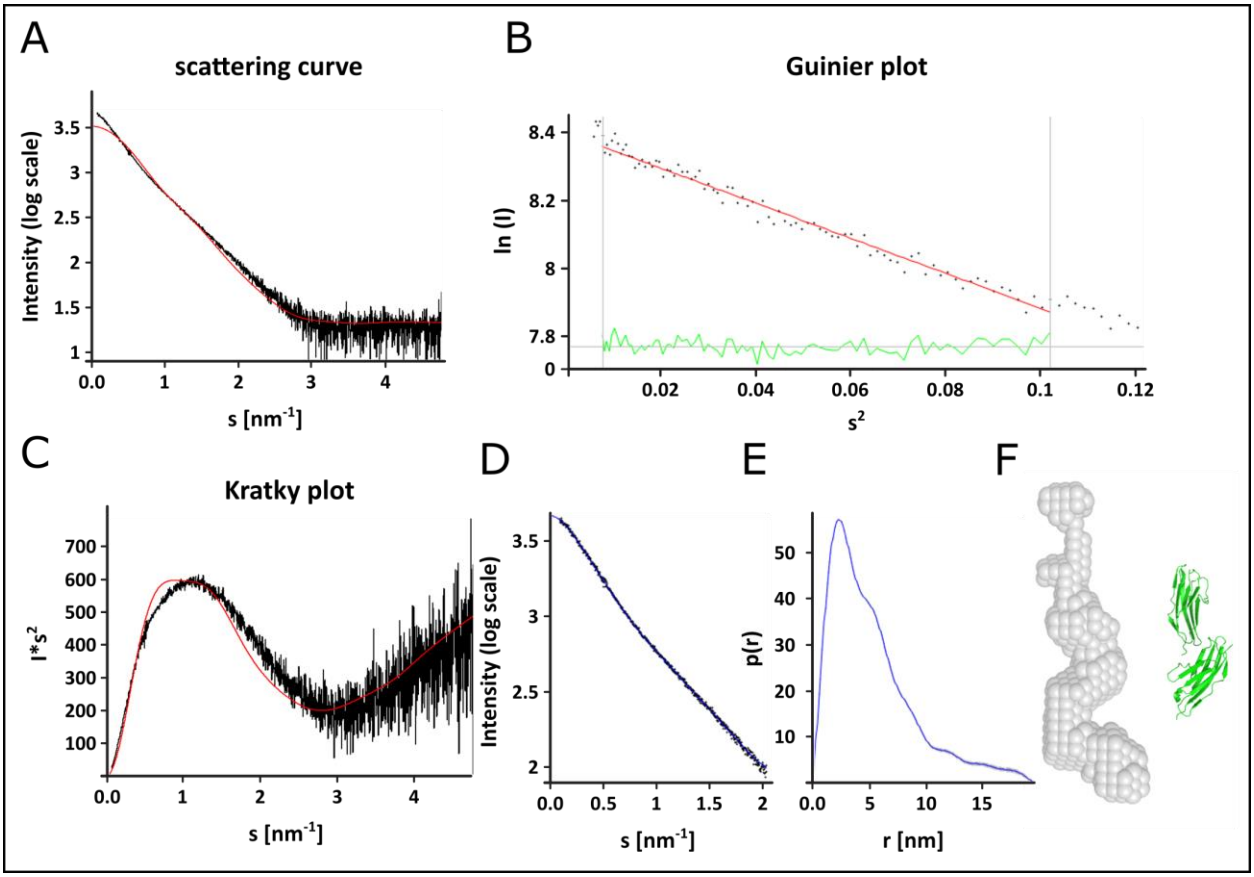

Figure 17 SAXS experiment with the individual titin Z8Z9. A) Scattering curve of titin indicates larger particles in the sample. B) Guinier Plot Analysis (ln I(s) vs s^2) for titin Z8-Z9 of the extrapolated scattering profile. The close approximation to linearity point to the absence of interference effects or heavy aggregation. However, the derived parameters for R_g (3.9 nm) and I(0) (4438) are not in accordance to the expected values of a monomeric sample C) Kratky analysis reveals a pronounced peak indicating that the sample is still folded. D) Fit of the proposed p(r) function to the scattering data. E) p(r) function indicating a maximum distance of > 15 nm. F) 3D reconstruction of titin Z8Z9 (bead model) and the crystal structure of the double domain construct in cartoon presentation. Data analysis performed in collaboration with Melissa Graewert.

5.1.9.2 SAXS measurements on the titin-obscurin complex in batch mode

To determine the correct position of titin domain Z8, first, the theoretical scattering curves of the two possible complex arrangements that result from the asymmetric unit were determined: the complex with Z8 interacting with O58 (complex A) and the complex with Z8 interacting with O59 (complex B) (Figure 18 A-B). The calculated R_g values for both complex arrangements are with 25.8 Å for complex A and 26.9 Å for complex B very similar. This is in good agreement with the well overlapping curves at low s-range (<1.5 nm^{-1}) illustrating the high similarity between both arrangements. The main difference between both complexes is pronounced in the s-range between 1.5 and 2 nm^{-1} were complex B shows a more pronounced shoulder. To resolve such a

minor difference, data has to be collected at higher resolutions. Accordingly, batch SAXS experiments with complex concentrations between 3.4 and 0.43 mg ml^{-1} were performed. The scattering curves show strong concentration dependencies (Figure 18 C-D). The scattering curves corresponding to the different concentrations were fitted with the theoretical curves calculated for both complexes. The fits for both theoretical complexes were comparable and resulted in chi^2 of up to 21.5 for high protein concentrations. The steeper descend at low angles of the high concentrated experimental scattering curves suggests that the sample comprises larger particles. At a lower concentration (0.25-0.1 mg ml^{-1}) the tendency for higher oligomeric structures could be reduced resulting in a better fit with a chi^2 of 1.2. However, the lower concentration results in a reduced resolution of the scattering curves below distances of 1.5 nm^{-1} and thus, below the s-range in which differences between the two possible structures are pronounced. Thus, a differentiation between the two possible complex arrangements was not possible using SAXS in batch mode of the titin Z8Z9 – obscurin O58O59 complex.

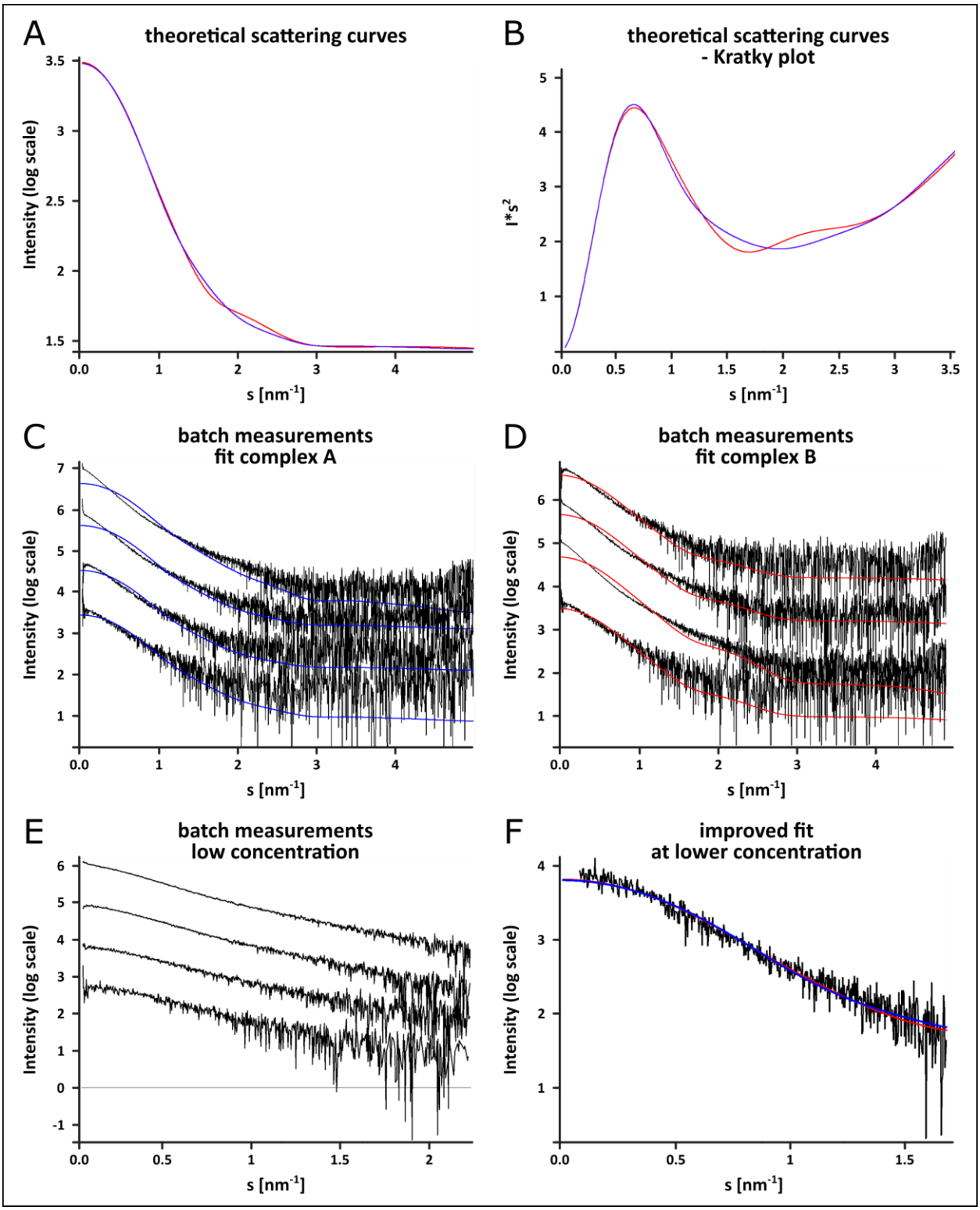

Figure 18 Theoretical and experimental scattering curves of the titin Z8Z9 - obscurin O58O59 complex. A) Theoretical scattering curves calculated from the crystal structure for the two possible complex arrangements. Scattering curve of the complex containing the Z8O58 interaction is shown in blue (complex 1) and the curve showing the weaker Z8O59 interaction is shown in red (complex 2). Curve shown on log scale. B) Kratky analysis of both theoretical scattering curves for better illustration of

differences between both curves shown in double log scale. C+D) Experimental scattering curves fitted with the theoretical scattering curves of the crystal structure (colouring see A and B). Curves represent concentrations between 3.4 and 0.43 mg/ml and are shifted along the y-axis for illustrational purpose. E) At a lower concentration range of 0.25-0.1 mg/ml the tendency for higher oligomeric structures is reduced. F) Theoretical structures of both complex arrangements can be fitted even in the low s range. The resolution of the scattering curve is only reliable until s values of 1.5 nm^{-1}. Data analysis performed in collaboration with Melissa Graewert.

5.1.9.3 SAXS measurements on the titin-obscurin complex in SEC SAXS mode

To remove the apparent larger particles at higher concentrations an additional purification step prior to the measurements using the SEC-SAXS approach was performed. The Chromixs profile (Figure 19 A) for which the intensities of measured SAXS frames between 0.8 and 1.2 nm^{-1} are integrated and plotted against the frame number, shows a slight increase in scattering intensity beginning at frame 700 before the main elution peak. This indicates that components of higher molecular weight elute prior to the complex in the void volume of the column. Furthermore, capillary fouling, a process in which molecules attach to the surface of the capillary and thus impairing the scattering signal, occurred during measurement as indicated by the shift in scattering intensity of the buffer background before and after the main elution peak. The long tailing of the peak could indicate the presence of free subunits which elute in close proximity to the complex. Due to the capillary fouling, a number of different frames were tested for background subtraction before suitable buffer frames were identified. 'AutoRG' was employed to determine the I(0) and R_g of each subtracted frame (Figure 19 B). The resulting R_g and I(0) versus frame plots suggest a mixture of molecules eluting from the column as i) the I(0) value shows a strong tailing, and ii) the R_g values are not stable throughout the whole peak. Nevertheless, frames from the beginning of the elution peak with constant R_g values were averaged. The generated scattering profile cannot be fitted with neither complex A (chi^2 =52) nor complex B (chi^2=58) (Figure 20). The scattering curve still suggests larger species which could be an artefact of the capillary fouling.

Taking these data together, it is not possible to draw conclusions about the arrangements of the titin obscurin complex from these experiments. First, the resulting difference in the theoretical scattering curves of both possible complex states are not pronounced enough to be confidently fitted with the experimental data. Second, the protein complex appeared to be difficult to handle resulting in poly-disperse samples with either larger particles or smaller fragments even after size

exclusion chromatography. Additional capillary fouling has been observed reducing the quality of the measurements.

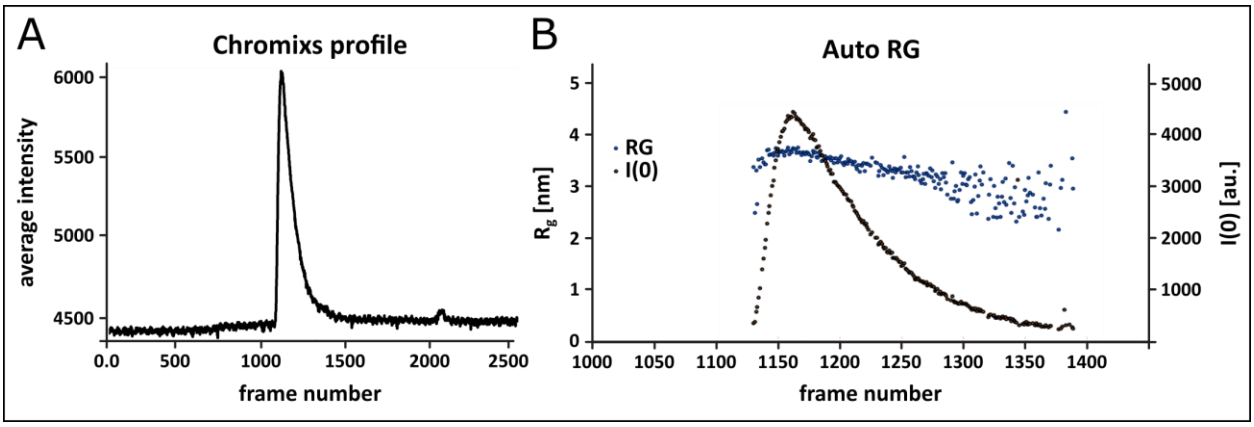

Figure 19 SEC-SAXS sample quality analysis. A) Chromixs profile plotting scattering intensity frames against the frame number. Slight increase in scattering at around frame 700 suggests larger components in the void volume, scattering intensity shift before and after the peak indicate capillary fouling. B) Auto RG plot showing I(0) and R_g of each frame. I(0) value shows a strong tailing, R_g values are not stable throughout the peak. Both indicates a mixture eluting from the column. Data analysis performed in collaboration with Melissa Graewert.

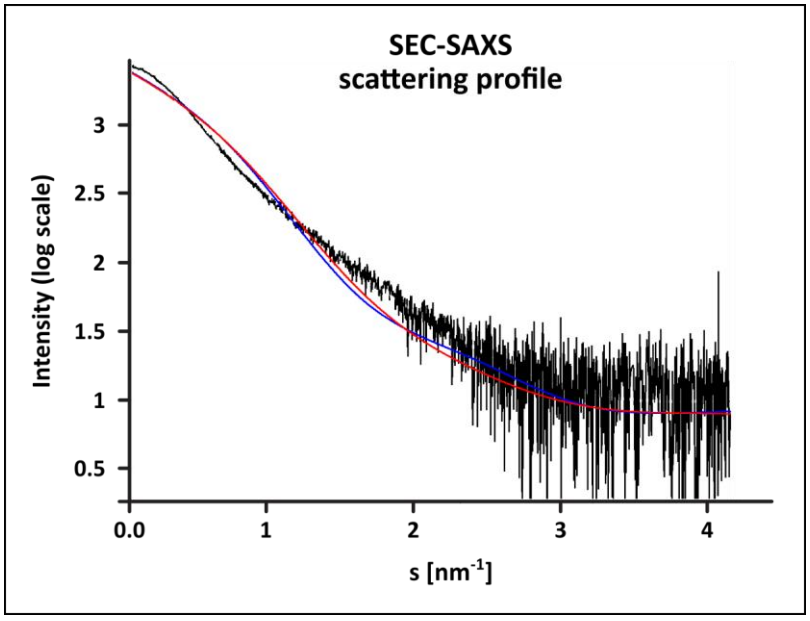

Figure 20 SEC-SAXS scattering profile of the titin obscurin complex. The theoretical scattering curves of both complex arrangement (blue – arrangement involving Z8-O58 interaction; red – arrangement involving Z8-O59 interaction) cannot be fitted with the experimental scattering curve. Experimental scattering curve suggest larger particle present. This might be a result of capillary fouling during the measurement. Data analysis performed in collaboration with Melissa Graewert.

5.1.9.4 SAXS Measurements of the complex formed by obscurin and the individual titin domains

To reduce the structural similarity of both possible complex arrangements of complex A and complex B the theoretical scattering curves of the two possible complex arrangements formed by obscurin O58O59 and titin Z8 while removing titin Z9 from the model were calculated. As the resulting 'sub-complex' is assumed to be smaller, minor differences between both possible Z8 positions are expected to be more pronounced. As shown in chapter 5.1.9.1 the angle of the linker between the two subunits of obscurin as seen in the crystal structure remains rigid in solution. This justifies the assumption that a titin Z8 - obscurin O58O59 and the Z9-O58O59 complexes in solution are comparable with the arrangement seen in the titin Z8Z9 - obscruin O58O59 crystal structure. As can be seen in Figure 21 A and B the scattering curve of complex B involving the Z8O59 interaction (red) shows a steeper descend at low s-ranges. This indicates a larger R_g value and a more elongated arrangement. Also, the pronounced shoulder at 1.5 nm^{-1} present for complex B indicates that Z8 is more exposed in complex B than in complex A, where Z8 is in the same plane as obscurin 58. The two curves differ already at lower s-ranges compared to the theoretical scattering curves of the complex involving Z9, thus, allowing to measure at lower complex concentrations. The scattering of the complex between obscurin O58O59 and Z8 was measured accordingly at concentrations between 0.34 mg ml^{-1} and 2.70 mg ml^{-1} using the batch approach. The fits between the theoretical scattering curves and the experimental curves of the two complexes were calculated (Table 35). Due to the noise-level, the chi^2 for both complex model measurements appear rather similar. However, the slight preference for complex A in all tested concentrations in combination with the visual inspection of fit, especially in the Kratky analysis suggests that the primary binding site of titin Z8 to obscurin is at the obscurin subunit O58.

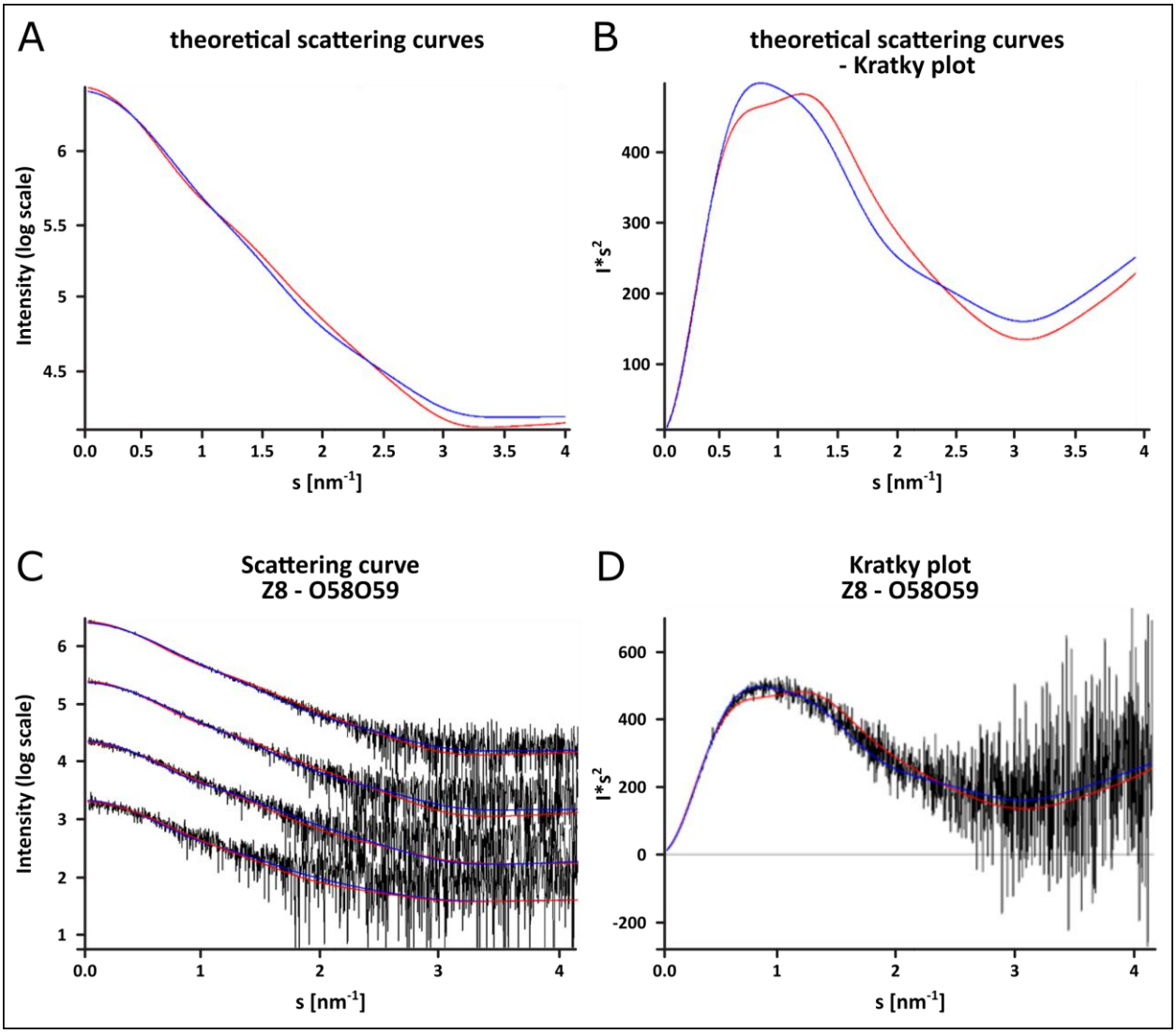

Figure 21 Theoretical and experimental scattering curves of the titin Z8 - obscurin O58O59 complex. A) Theoretical scattering curves calculated from the crystal structure for the two possible arrangements of titin Z8. Scattering curve of the complex containing the Z8O58 interaction is shown in blue, and the curve showing the weaker Z8O59 interaction is shown in red. Curve shown on log scale. B) Kratky analysis of both theoretical scattering curves for better illustration of differences between both curves shown in double log scale. C) Experimental scattering curves at different concentrations (2.7 mg/ml – 0.3 mg/ml, with the highest concentration at the top) fitted with the theoretical scattering curves (colouring see A and B). Fits for curves corresponding to the complex involving the Z8-O58 interaction (blue) are constantly better for all concentrations. D) Kratky analysis of the scattering curve with the highest concentration. A clear mismatch can be seen in the mid s-range between the theoretical curve corresponding to the complex involving the Z8-O59 interaction (red). Data analysis performed in collaboration with Melissa Graewert.

Table 35 Fit comparison of the scattering curves of both titin-obscurin complex arrangements. Calculated fits between the theoretical scattering curves and the experimental curves of the two possible titin-obscurin complex-arrangements. Slight preference for complex A in all tested concentrations.

Concentration [mg ml^{-1}]	**0.34**	**0.68**	**1.35**	**2.70**
Complex A	1.04	1.03	1.17	1.33
Complex B	1.07	1.03	1.26	1.50

Next, scattering data from the complex formed by obscurin O58O59 and titin Z9 were collected. First, the sample was measured in batch mode at concentrations between 0.4 and 3.2 mg ml^{-1}, showing large concentration dependencies (Figure 22). A comparison with the theoretical scattering curve of titin Z9 to obscurin O58O59 reveal larger particles at the highest concentration. At lower concentration, however, the larger particles did not contribute to the scattering signal and data could be used for further analysis at low resolution. The profiles indicate that the in-solution structure of titin Z9 to obscurin O58O59 measured in SAXS corresponds to the crystal structure. An additional analysis of the oligomeric mixture in the different concentrations measured in batch using 'Oligomer' (Konarev *et al.*, 2003) reveals that the complex dissociates at lower concentration. For this, the curves were fitted with different volume fractions of titin Z9, obscurin and the complex. To improve the data quality, SAXS data was also collected in SEC-SAXS mode. However, the complex dissociated into the subunits obscurin O58O59 and titin Z9 probably due to the high dilution during SEC run.

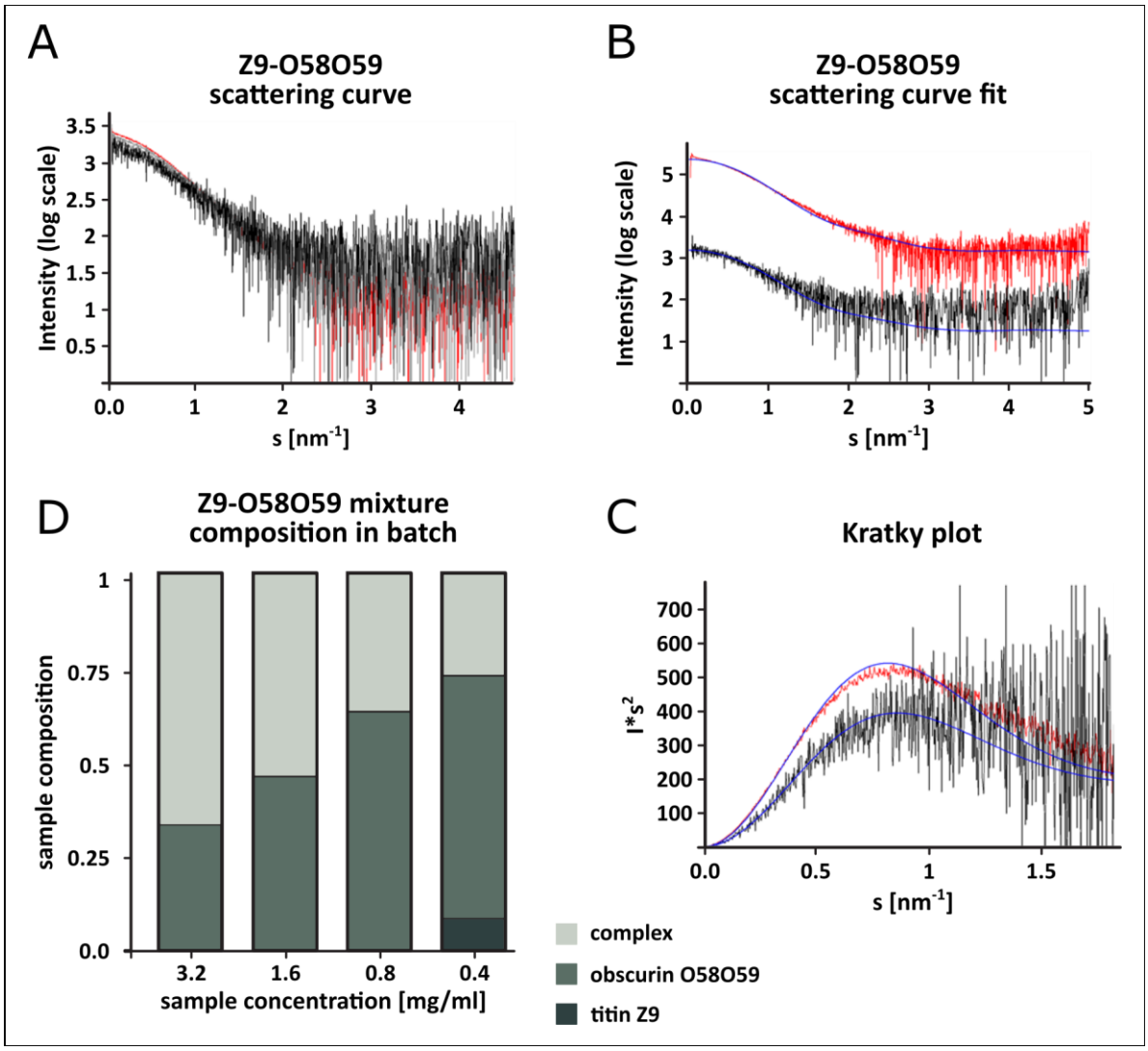

Figure 22 Batch SAXS measurements of titin Z9 and obscurin O58O59. A) SAXS data collected at four different concentrations and normalised against transmitted beam and concentration shown on log scale B) The data measured at 3.2 mg/ml (upper curve, red) and at 0.4 mg/ml (lower curve, black) are fitted with the theoretical scattering curve of the crystal structure of titin Z9 bound to obscurin O58O59 (blue curve). Note, the curves have been shifted along the Y-axis for presentation purposes. C) Kratky analysis of the curves shown in B). D) Composition of titin Z9 and obscurin measured in batch. The concentration at which the sample was measured is indicated (3.2 mg/ml down to 0.4 mg/ml). Data analysis performed in collaboration with Melissa Graewert.

Combining these data, SAXS measurements of the whole complex did not allow to confidently make assumptions about the correct arrangement of Z8 within the complex. On the one hand, the behaviour of the complex appeared to be concentration-dependent, and measurements without higher oligomeric states were only possible at lower concentrations. On the other hand, the structural differences between the two possible complexes are not pronounced enough to be measurable at such low concentrations. The reduction of complexity by reducing the involved

domains helped to improve the measurable difference between both possible arrangements. Here, the data suggest an interaction between Z8 and O58 supporting the interface evaluation by PISA and the mutation studies. Additionally, it was possible to receive a good low-resolution model of the individual obscurin O58O59 that reveals a high rigidity of the linker between both domains. The individual measurement of titin, however, did show tendencies to form larger particle even at low concentration that is in agreement with the observed general behaviour of titin during purification.

5.1.10 Interface Variants within titin Z8-Z9 and obscurin O58-O59

In this chapter, the focus is on analysing the potential impact of missense variants found in the involved Z-disk complex domains of titin and obscurin. The mean allele frequency (MAF) can be used as a first indicator of a likely benign or likely pathogenic variant. As proposed by Kobayashi *et al.* (Kobayashi *et al.*, 2017) an allele frequencies higher than 1.00E-4 should be considered "higher than expected" for diseases of Mendelian inheritance. Table 36 lists all variants found within the interface between titin Z8Z9 and obscurin O58O59. With a few exceptions mentioned below, all hereinafter described variants do fall below this threshold. A general analysis of the variant frequency within the domains reveals a general frequency of almost 0.5 indicating the presence of a known variant for almost every second amino acid (variant frequency: Z8: 0.44; Z9:0.47; O58: 0.32; O59: 0.42). Only obscurin O58 shows a slightly reduced frequency. When comparing the variant frequency within the interfaces with the overall frequency, the ratio is preserved indicating no selectivity for or against residues involved in interface. Only the interface between Z9 and O59 shows an increased frequency as seven out of eight involved residues have known variants. Except for Y1891C and A4484T which are both involved in the Z9-O59 interaction no clinical data are available for the interface variants.

Table 36 Variants located within the complex interfaces of titin Z8Z9 and obscurin O58O59. Interface, domains and residue number of the variant are shown. The type of single nucleotide variant (SNV), where applicable the 'reference SNP ID number' (rs), and the mean allele frequency (MAF) are given for each variant as well as the position within the structure and possible impact. Additional comments were made where necessary.

interface	domain	residue	SNV	rs	MAF	Structure	possible Structural/functional Impact	comment
Z8-O58	Z8	1711	S/F	None	1.09E-05	loop, exposed	interface, part of a serine/threonine kinase recognition motif	
		1741	D/G	None	4.07E-06	loop, exposed	interface	
		1788	T/I	None	4.13E-06	β-strand, hydrophobic core	not expected	
		1789	S/P	None	4.13E-06	β-strand, exposed	interface, fold destabilisation	
	O58	4387	E/K	None	3.22E-05	β-strand, exposed	interface	E4387R mutant showed only minor effect in ITC
		4396	R/C	None	5.97E-05	β-strand, exposed	interface, artifical disulfide formation	
Z9-O58	Z9	1879	Q/R	None	4.07E-06	loop, exposed	not expected	
		1905	Y/C	None	4.07E-06	loop, exposed	interface, artifical disulfide formation	
		1905	Y/H	None	4.07E-06	loop, exposed	not expected	
		1906	D/V	None	4.07E-06	loop, exposed	interface, fold destabilisation	D->R mutant aggregated strongly
		1906	D/Y	None	4.07E-06	loop, exposed	interface, fold destabilisation	D->R mutant aggregated strongly
	O58	4377	D/N	None	3.69E-05	loop, exposed	not expected	D4377R mutation had strongest effects in ITC
		4411	R/W	None	5.21E-05	β-strand, exposed	interface	
		4411	R/Q	None	1.05E-04	β-strand, exposed	not expected	R4411D mutant showed only minor effects in ITC

interface	domain	residue	SNV	rs	MAF	Structure	possible Structural/functional Impact	comment
Z9-O59	Z9	1882	R/H	None	6.50E-05	loop, exposed	interface	
		1882	R/C	None	2.17E-05	loop, exposed	interface, artificial disulfide formation	
		1889	V/F	None	4.70E-05	β-strand, partly exposed	interface	
		1891	Y/C	rs547305291	5.29E-05	β-strand, partly exposed	artificial disulfide formation	Dilated cardiomyopathy, Limb-girdle muscular dystrophy, uncertain clinical significance
	O59	4458	A/V	None	2.04E-05	loop, exposed	not expected	
		4484	A/T	rs116557268	2.64E-03	β-strand, partly exposed	not expected	very high MAF, no clinical significance
		4485	D/N	None	2.59E-04	β-strand, exposed	not expected	D4485N mutant showed only minor effects in ITC
		4486	G/S	None	1.11E-04	loop, exposed	not expected	high MAF

The titin variant Y1891C has been found in a patient with Dilated cardiomyopathy and Limb-girdle muscular dystrophy but has been reported with uncertain significance. It is part of the small β-sheet formed by titin Z9 and obscurin O59. As it introduces an exposed cysteine, this variant might facilitate false interactions with other molecules carrying exposed cysteines under oxidative stress. The obscurin variant A4484T is involved in the same β-sheet formation. Structurally it does not interfere with the β-sheet formation, thus, unlikely to largely interfere with the complex formation. No clinical significance has been assigned to this variant. A functional analyses by Arimura et al. showed no impact of A4484T on the titin Z8-Z9 interaction in M2H and IP assays or the incorporation into the Z-disk (Arimura *et al.*, 2007). These data are supported by the high mean allele frequency of 2.64E-3 that is largely above the threshold set for possibly pathogenic variants.

For the interface between Z8 and O58 five variants can be found: Z8: S1711F, D1741G, T1788I, S1789P; O58: E4387K, R4396C. S1711 is part of a serine/threonine kinase recognition motif KKKXTS that is thought to be phosphorylatable (presented in more detail in chapter 5.1.13.1). Hence, the serine to phenylalanine mutation would not just introduce a hydrophobic amino acid in the interaction interface but also cancel out a potential phosphorylation site. D1741 forms together with obscurin R4396 one of the three interaction clusters within the Z8-O58 interface and contributes to the interface with salt bridges. Due to the aspartate to glycine substitution, this interaction cluster would be removed from the interface. With the T1788I variant, a hydroxyl group is replaced by a more hydrophobic side chain. As the interface interaction is formed by the backbone atoms of T1788 and the side chain is pointing towards the hydrophobic core of the domain, no large impact on the interaction is expected. S1789 is located in the middle of the large β-strand G of titin Z8. A proline at this position could break the β-strand and destroy the structural arrangement of the main contributor to the Z8-O58 interaction and thus potentially interfere with the titin obscurin assembly. E4387K is a charge reversal mutation and thus expected to have a more severe impact on the interface. Nevertheless, data from ITC measurements show that even the harsher glutamate to arginine mutation was not able to have a large impact on the complex formation. R4396C again introduces a cysteine in an exposed position potentially facilitating disulphide bridges with other molecules.

The Z9-O58 interface reveals eight different variants within five residues: Z9: Q1879R, Y1905C/H, D1906V/Y: O58: D4377N, R4411W/Q. The impact of Q1879R variant is considered to be minor as both residues are polar residues and hydrogen bond formation to O58 L4346 and S4422 is still possible. The introduction of an exposed cysteine for Y1905C could again have implication on the interaction with other molecules (see above). Additionally, the loss of interaction with O58 D4377 is expected. For the Y1905H mutation, the interaction with O58 D4377 will probably not be affected. D1906 is highly involved in the interface interactions by the formation of multitude hydrogen bonds guided by the side chain. ITC demonstrated that the substitution of the aspartate with an arginine led to a destabilisation of the whole molecule making this residue not only important for the complex formation but also for the intramolecular stability. The two variants D1906V and D1906Y might have an impact on both the domain and the complex stability due to the loss of the side chain function. Aspartate to asparagine mutations as in D4377N are commonly found in databases. The loss of salt bridges though might weaken the interaction. Notably, the more drastic D4377R mutation that was tested in ITC experiments had the strongest effect of a single mutation on the *in vitro* complex formation. For residue R4411 the substitution with glutamate only had minor effects in the ITC experiments. While a glutamine variant is expected to be even less drastic, which is supported by the MAF of 1.05E-4 being above the threshold, tryptophan, as it is bulky and R4411 is very central in the interface, might slightly weaken the interface.

As already described above the Z9-O59 interface has a high abundance of known variants. In addition to the already mentioned Y1891C and A4484T the following variants can be found: Z9: R1882H/C, V1889F; O59: A4458V, D4485N, G4486S. The R1882 is part of the highly exposed loop region between β-strand C and D. The distance between R1882 and its interaction partner A4458 can only be covered by the long arginine, while a histidine mutation would not be able to form the interaction anymore. A Cysteine at this exposed position might have an influence on the domain's behaviour. The exchange of the valine to a phenylalanine in V1889F could, even though both are hydrophobic amino acids, impair the inter-domain β-sheet formation between Z9 and O59 as phenylalanine is significantly larger than valine. A4458V and G4486S both are not expected to have a huge impact on the interface. D4485N has been tested in ITC experiments

only showing a slight reduction of the interaction. For both variants D4485N and G4486S the MAF is with 2.59E-4 (D4485N) and 1.11E-4 (G4486S) relatively high and above the threshold. Considering the high MAF for A4484T, changes within this interface seem to be relatively common and without consequences for the Mendelian inheritance. The other possible interface of Z8 that is formed with O59 according to the crystal structure involved in total four residues from which two show variations that are located on domain O59. E4463K and A4507T are only involved in this interface with backbone interaction. Hence, the impact even of the relatively drastic glutamate to lysine variant might not be high.

Taking these data together, S1711F, S1789P, both located within titin Z8, and the cysteine variants R4396C (O58), R1882C (Z9), Y1891C (Z9), and Y1905C (Z9), have the potential to interfere with the complex formation. As outlined in chapter 5.1.6 the complex stability is relatively robust against single residue exchanges even if affecting the residues physico-chemical properties. The complex with its three interfaces seems to be able to compensate for local changes. Probably, only modifications that break structural motifs (e.g. proline substitutions) or facilitate the interaction with other molecules blocking the interaction interface (e.g. cysteine substitutions) might interfere with the complex formation.

5.1.11 Variants with connections to myopathies

Next, known variants within titin Z8-Z9 and obscurin O58-O59 which additionally can be found in disease databases like ClinVar (Landrum *et al.*, 2018) and are connected to patients with myopathies were analysed (Table 37).

Table 37 Variants that can be found in disease databases being assigned to a reference SNP ID number (rs). Domains and residue number of the variant are shown. The type of single nucleotide variant (SNV), 'rs' number, and the mean allele frequency (MAF) are given for each variant as well as the position within the structure and possible impact. The connection with a disease and their clinical significance are indicated. Additional comments were made where necessary.

domain	residue	SNV	RS (number of reports)	MAF	disease connection	clinical significance	Structure	possible Structural/functional Impact	comment
Z8	1744	P/L	rs75686037	6.69E-03	Dilated cardiomyopathy, Limb-girdle muscular dystrophy, Distal myopathy, Hypertrophic cardiomyopathy	favour benign	loop, exposed	not expected	high MAF
	1752	R/H	rs150737838	1.66E-04	NA	NA	β-strand, exposed	not expected	high MAF
	1755	N/S	rs201904897	3.98E-05	Dilated cardiomyopathy, Limb-girdle muscular dystrophy	Uncertain	loop, hydrophobic core	fold destabilisation	
	1755	N/H	rs559024675	4.07E-06	NA	NA	loop, hydrophobic core	fold destabilisation	
Z9	1854	E/D	rs571222694	4.06E-06	NA	NA	loop, exposed	not expected	
	1859	R/S	rs551538420	5.41E-05	sudden cardiac death	uncertain	β-strand, exposed	not expected	surrounded by conserved residues
	1861	R/C	rs532733393	6.14E-05	Dilated cardiomyopathy and Limb-girdle muscular dystrophy	uncertain	β-strand, exposed	artificial disulfide formation	surrounded by conserved residues
	1890	R/C	rs146496197	3.90E-04	Dilated cardiomyopathy, Limb-girdle muscular dystrophy, Distal myopathy, Hypertrophic cardiomyopathy	uncertain, favour benign	β-strand, exposed	artificial disulfide formation	close to Z9-O59 interface, high MAF
	1891	Y/C	rs547305291	5.29E-05	Dilated cardiomyopathy, Limb-girdle muscular dystrophy	uncertain	β-strand, partly exposed	artificial disulfide formation	Part of the Z9-O59 interface
	1914	A/V	rs374203813	3.25E-05	NA	uncertain, favour benign	β-strand, hydrophobic core		
	1914	A/T	rs118161093	1.34E-04	NA	NA	β-strand, hydrophobic core		high MAF

domain	residue	SNV	RS (number of reports)	MAF	disease connection	clinical significance	Structure	possible Structural/functional Impact	comment
O58	4344	R/Q	rs79023478	1.17E-02	Hypertrophic cardiomyopathy	uncertain	loop, exposed	not expected	impact discussed by (Arimura *et al.*, 2007; Hu *et al.*, 2017), high MAF
	4353	E/K	rs143021118	3.75E-04	NA	NA	hinge region of O58-59, loop, exposed	O58O59 arrangement destabilisation	high MAF
	4381	H/R	rs1150912	7.01E-01	NA	NA	loop, exposed	not expected	high MAF: natural variant
	4382	T/A	rs183406337	2.13E-03	NA	NA	loop, exposed	not expected	high MAF
	4392	E/D	rs116266256	7.18E-05	NA	NA	loop, exposed	not expected	
O59	4444	R/W	-	3.77E-05	distal muscular dystrophy	NA	loop, exposed	not expected	(Rossi *et al.*, 2017)
	4450	C/R	rs1188732	6.95E-01	NA	NA	β-strand, exposed	potential loss of a bond	high MAF: natural variant
	4471	A/T	rs183798858	1.66E-04	NA	NA	loop, exposed	not expected	high MAF
	4484	A/T	rs116557268	2.64E-03	NA	NA	β-strand, partly exposed	not expected	Part of the Z9-O59 interface, very high MAF
	4488	H/D	rs188287808	8.15E-03	NA	NA	β-strand, exposed	not expected	high MAF
	4489	H/Q	rs61825301	2.44E-01	NA	NA	β-strand, hydrophobic core		high MAF: natural variant
	4497	Q/L	rs116329268	1.81E-03	NA	NA	loop, exposed	not expected	high MAF
	4509	R/H	rs193004966	7.22E-03	NA	NA	loop, exposed	not expected	high MAF
	4511	A/T	rs139319906	8.15E-04	NA	NA	β-strand, exposed	not expected	high MAF
	4516	R/W	rs11810627	4.23E-01	NA	NA	β-strand, exposed	not expected	high MAF: natural variant

The first variant is P1744L (rs75686037). It has been found in various conditions like Dilated cardiomyopathy 1G, Limb-girdle muscular dystrophy, Distal myopathy or Hypertrophic cardiomyopathy, yet, interpreted to be likely benign. Looking at the structural data, its location within the loop region between β-strand C and D, a region not involved in any known interaction also supports the clinical interpretation. A proline to leucine mutation in this region is not expected to have an impact on the proteins behaviour. Furthermore, with a mean allele frequency (MAF) of 6.69E-3, it occurs too frequently to be connected to a rare genetic disease.

The second variant in Z8 is N1755S (rs201904897) that was reported as a variant of uncertain significance in several individuals. While it has been connected to Dilated cardiomyopathy and Limb-girdle muscular dystrophy, it was also found in an individual that underwent exome sequencing not selected for cardiomyopathy, arrhythmia, or family history of sudden cardiac death (Ng *et al.*, 2013). The asparagine is pointing towards the domain core forming a polar interaction with Trp1738. Even though the variant is a conservative amino acid substitution the interaction will be lost which might have consequences for the domain stability, however, not expected to affect the domain largely.

The R1859S (rs551538420) variant in titin Z9 is of uncertain significance in the *TTN* gene. It has been reported in a three-year-old with sudden cardiac death. However, a second variant was found in the *TTN* gene, without providing segregation data (Campuzano *et al.*, 2014). The variant has a MAF of 5.41E-5, indicating it is not a common benign variant. The R1859S variant is a semi-conservative amino acid substitution. It is highly exposed and located on β-strand B in-between the three conserved residues motif AXFXC as described in chapter 5.1.4. The conserved residues are all oriented towards the hydrophobic core and are unlikely to be influenced by this variant. Additionally, R1859 is not involved or close to the interaction interfaces formed by titin and obscurin and therefore unlikely to interfere with the complex formation.

A second variant within this conserved triple residue motif with clinical indications has been found. R1861C (rs532733393) that was connected to Dilated cardiomyopathy and Limb-girdle muscular dystrophy with uncertain significance is directly neighbouring another cysteine (C1862). While C1862 is pointing toward the hydrophobic core, variant R1861C is highly exposed

potentially being able to form disulfide bridges with other molecules. It has a relatively low MAF of 6.14E-5 making it unclear whether this variant is pathogenic or rare benign.

Another residue substitution to a cysteine is R1890C (rs146496197). This variant was found in several patients with a wide range of pathological conditions like Dilated cardiomyopathy, Limb-girdle muscular dystrophy, Distal myopathy, Hypertrophic cardiomyopathy. It was classified as 'uncertain significance - favour benign'. This variant is located within the β-strand of titin Z9 that is involved in forming the Z9-O59 β-sheet interaction and is thus close to the interface forming residues. Here again, the introduced cysteine could under certain oxidising conditions form disulfide bridges with other molecules blocking the Z9-O59 interaction interface to obscurin. Notably, this variant has a world-population wide MAF of 3.90E-4 and has been identified in 0.2% (10/4406) of African American chromosome. These high frequencies suggest that it is likely benign.

Taking these data together, additional to the already mentioned variant Y1891C, three variants might have an impact on the protein's behaviour. While N1755S could influence the domain stability by cancelling out an intramolecular interaction in the hydrophobic core, the two variants R1861C and R1890C are introducing highly exposed cysteines to the domain surface potentially enabling disulfide bridge formation with other molecules. It is important to mention that Cysteine has a role that is very dependent on cellular location, making the effect of substitutions dangerous to interpret (e.g. (Barnes *et al.,* 1999)). All of the discussed variants have been of unknown/uncertain significance with tendencies to be benign and have not been clearly linked to a disease phenotype. They might coexist in the same patient with mutations in other well-established cardiomyopathy genes. These variants might have a role as disease modifiers when associated with other pathogenic variants or represent causal recessive mutations. In any case, while the structural data help to gain some insight of the potential role of these variants, it is necessary to evaluate their biological impact carefully.

5.1.12 Structural impact of functionally described variants

To date, all of the variants described above have only been connected to myopathies *via* the pathological background of the screened individuals. Further functional analyses of the variants

has only been carried out for three obscurin variants R4344Q (rs79023478) (Arimura *et al.*, 2007; Hu *et al.*, 2017), A4484T (rs116557268) (Arimura *et al.*, 2007) and R4444W (Rossi *et al.*, 2017). R4444W which is located on obscurin domain O59 has a low MAF of 3.77E-5 and thus cannot be excluded to be linked to a rare genetic disease. It has been found together with a frameshift mutation in the gene for protein filamin-C in two patients suffering from distal muscular dystrophy. The first two variants have large MAFs (1.17E-2 for R4344Q; 2.64E-3 for A4484T) that are considerably higher than the threshold suggesting common and therefore likely benign genomic variations. Nevertheless, it was described in both publications (Arimura *et al.*, 2007; Hu *et al.*, 2017) that variant R4344Q, which is located in obscurin domain O58, does have a physiological effect, while A4484T shows negative results (only addressed by Arimura et al.). Notably, both publications show different effects of the R4344Q variant. While Arimura et al. showed an impairment of the obscurin localisation to the Z-disk, Hu et al. demonstrate that under sedentary conditions, the R4344Q mutation results in Ca^{2+} deregulation by an increased binding of the monomeric Phospholamban (PLN). In the following chapter, a combined structural and biophysical approach will be used to investigate the impact of these variants on the titin-obscurin interaction and - for R4344Q - on the phospholamban binding.

5.1.12.1 Localising the disease-related obscurin residue R4444

To assess the possible impact of the variant R4444W, the high-resolution crystal structure of the titin-obscurin Z-disk complex was used. Analysis of the location reveals that the variant is not involved in or close to any known interface formed by obscurin O59 (Figure 23). It is highly exposed at the C-terminal tip of the domain. No further interactions of the residue within the structure were found.

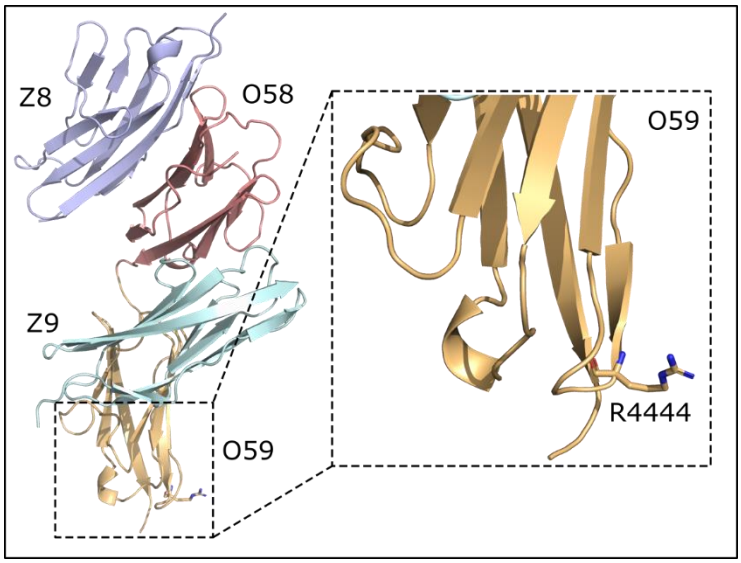

Figure 23 Localisation of the disease-related residue R4444 within the titin-obscurin Z-disk complex structure. R4444 is not involved in the interfaces forming the complex but located highly exposed on the C-terminal tip of the domain O59.

5.1.12.2 Localisation of disease-related residue R4344 and impact of the R4344Q mutation on the complex

The high-resolution crystal structure of the titin-obscurin Z-disk complex reveals that O58 R4344 is located at the N-terminal end of the O58 domain in a flexible region (Figure 24). Notably, this region is part of β-strand A in homologous structures. The residue is highly exposed and while forming backbone hydrogen bonds to β-strand B no other side chain specific interactions can be found. It is not located within any of the three interaction interfaces identified for Z-disk titin and obscurin. To validate this observation, the impact of the R4344Q variant on the complex formation was quantified by ITC. The K_d of 3.7 µM ± 0.3 µM is very close to the wild-type K_d (4.4 µM ± 0.1 µM) demonstrating that the interaction between titin and obscurin is not impaired by the variant. Next, it was investigated whether the protein folding and stability is affected. The correct folding was verified by CD measurements showing no change in the secondary structure composition compared to the wild-type. Thermal stability of both the wild-type and the variant was assessed by a thermal shift assay (Figure 24 C). The small thermal shift from 52.1 °C (wild-type) to 49.8 °C (R4344Q) (ΔT_m=-2.3°C) again demonstrates that no large changes in the behaviour of obscurin O58O59 are induced by the variant. Data presented here suggest that the variant R4344Q has no impact on the titin Z8Z9 – obscurin O58O59 interaction. Neither the

structural analysis of the variants position nor a number of different biophysical characterisations indicate an impairment of the titin-obscurin binding or a destabilisation of obscurin O58O59 itself. This is in good agreement with the high frequency this variant has been found in sequencing approaches.

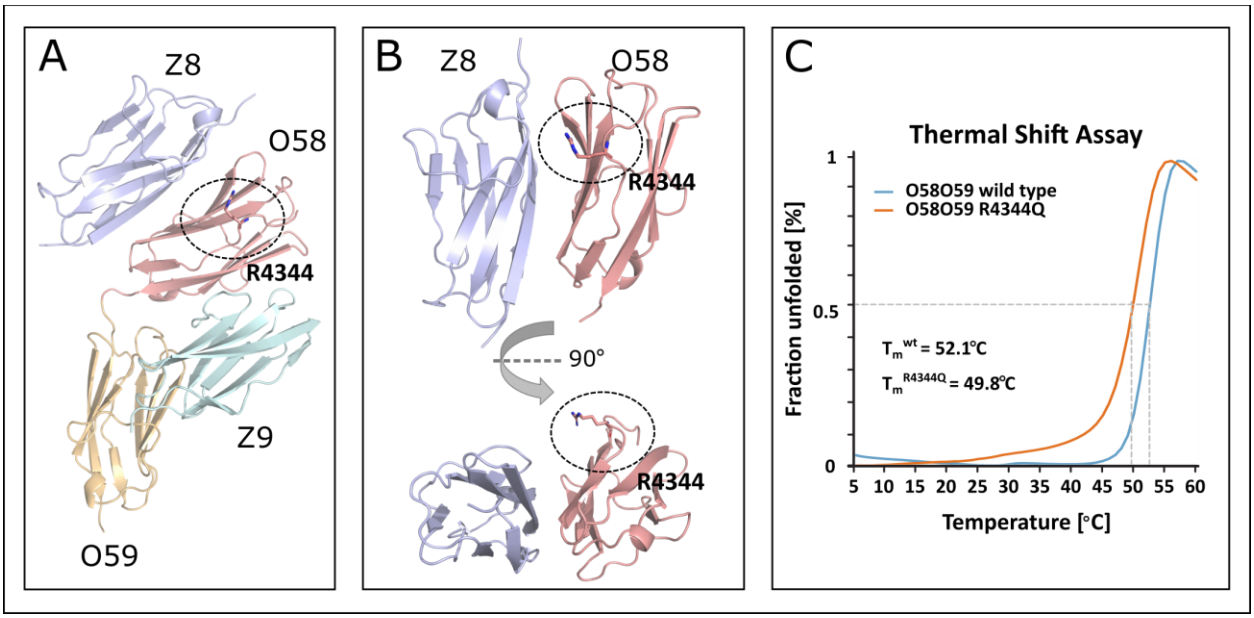

Figure 24 Structural investigation of the obscurin variant R4344Q location. A) Structural representation of the whole titin Z8Z9 – obscurin O58O59 complex and B) the closest interface between Z8 and O58 in two different perspectives. Residue R4344 (highlighted in the black circle) is not involved in any of the three interaction interfaces. C) Thermal shift assay performed on the obscurin O58O59 wild-type (blue) and HCM related mutation R4344Q (orange) described by Arimura et al. (2007). The thermal shift of ΔT_m=-2.3°C ($T_{m\,wt}$=52.1°C; $T_{mR4344Q}$=49.8°C) does not indicate a significant destabilisation of the mutated obscurin.

5.1.12.3 Impact of the R4344Q variant on the obscurin-phospholamban interaction

Next, the predicted binding of phospholamban (PLN) to obscurin O58 was approached. PLN is a 52 amino acid long integral membrane protein with an approximately 20 amino acid long cytosolic N-terminus. In its unphosphorylated state, it acts as an inhibitor of the sarcoplasmic reticulum ATPase dependent Ca^{2+}-pump SERCA2 (Rodriguez and Kranias, 2005) and therefore is part of the Ca^{2+} regulation machinery. The interaction between obscurin O58 and PLN was found in pull-down assays, and the binding mode was predicted by a molecular modelling approach (Hu *et al.*, 2017). The published data suggest a 10-fold increase of binding between O58 and PLN due to the R4344Q mutation. Figure 25 shows the arrangement of the cytosolic part of PLN as it binds to obscurin O58. A direct comparison of the Z8-O58 binding motif reveals that both interactions

would occupy the same site of obscurin, theoretically being able to compete for binding. To evaluate the proposed binding fluorescence anisotropy measurements with the N-terminally labelled cytosolic region of PLN (PLN1-19) was performed (Figure 26). The initial experiment showed a strong decrease of fluorescence anisotropy values at higher protein concentrations, especially for the mutant R4344Q (Figure 26 A). At the same time in the control measurement determining the total fluorescence intensities, the values increase at higher protein concentrations, again more drastically for the R4344Q mutant (Figure 26 B). While the anisotropy values should either increase upon binding or stay constant if no binding occurs, the total fluorescence intensities should be stable throughout the whole measurement, as only the labelled peptide that has a constant concentration should show fluorescence. To rule out systematic errors different buffer systems, incubation times, and concentration ranges were tested. Measurements could be improved by either extensive centrifugation or additional filtration of the protein sample prior to measurement, indicating that higher oligomers might interfere with the measurement. In an additional attempt, in which obscurin O58O59 alone was measured at different concentrations a decrease of anisotropy at higher concentrations could be observed. This points to a concentration-dependent intrinsic fluorescence of obscurin O58O59 that influences the anisotropy measurements. Notably, the effect seems to be larger for obscurin O58O59 R4344Q indicating that the mutant is more prone to form higher oligomers. Even though not valid for a final conclusion, as the behaviour of obscurin has a major impact on the measurement, the overall result points toward the absence of binding between obscurin O58O59 and PLN1-19.

As another approach to investigate whether phospholamban binds to obscurin O58, microscale thermophoresis (MST) was performed. Here, different concentrations of obscurin O58O59 wild-type or its mutant R4344Q were titrated to constant concentrations of labelled PLN1-19 (Figure 27). Even though the overall experimental conditions seemed to be optimal, as indicated by quality control measurements assessing the initial fluorescence, adsorption, ligand-induced fluorescence change, aggregation, and photobleaching, in both titrations the signal to noise ratio was very low. Even though a curve fitted on the data points, the signal for obscurin wild type was below the given threshold not allowing to calculate a reliable dissociation constant. The fit for

the measurement with obscurin R4344Q seemed to be better. Nevertheless, the signal to noise ratio was just above the threshold and thus still critically low making the calculated dissociation constant of K_d=239 µM not reliable. Consequently, it is not possible to confidently confirm or refute the binding of PLN to obscurin or an increased binding to the obscurin mutant R4344Q. Both approaches were not able to unambiguously validate the results of Hu et al.

Taking together, it is not possible to draw a conclusion about the impact of the R4344Q variant on physiological conditions. Neither the structural data nor biophysical analysis by SEC, ITC, thermal shift assay show an impairment of the complex formation as described by Arimura et al. Additionally, the increased binding of obscurin carrying the R4344Q variant to PLN could neither be confirmed nor ruled out. In particular, the large MAF of the R4344Q variant and its physiological impact remains elusive.

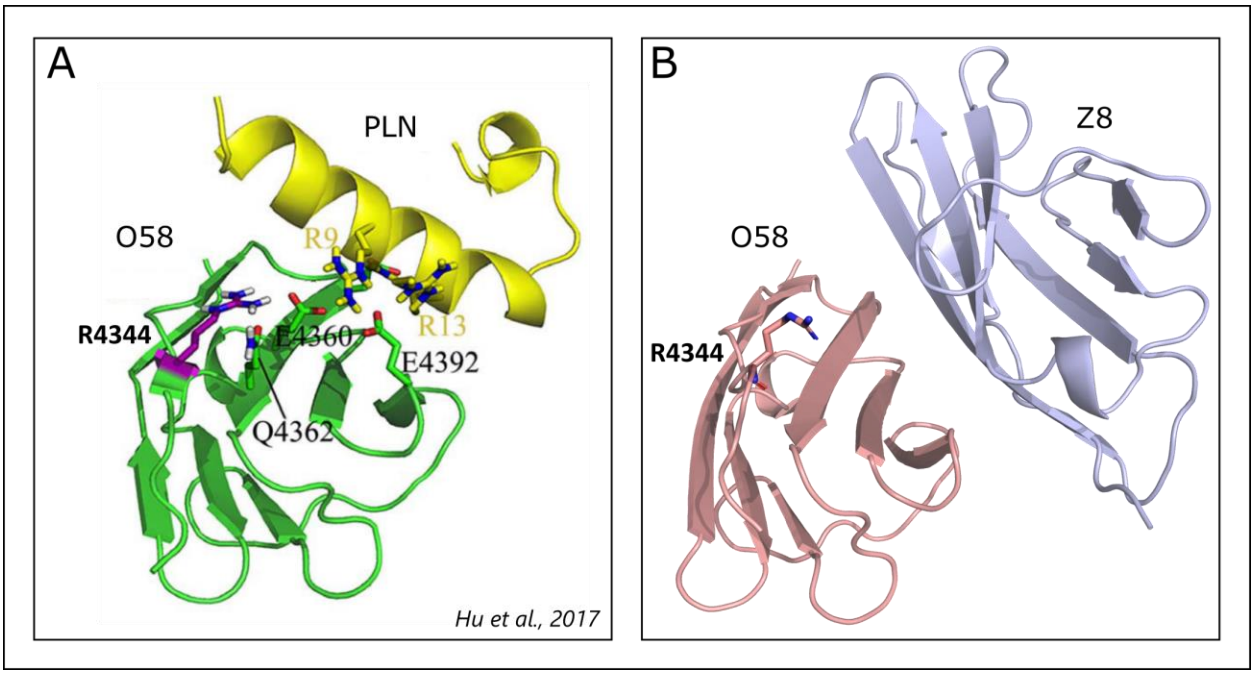

Figure 25 Phospholamban binding site on obscurin O58. Comparison of the obscurin O58 binding arrangement to A) the soluble part of Phospholamban (PLN) as predicted by (modified from (Hu et al., 2017)) in an modelling approach and B) titin Z8 as seen in the crystal structure. The binding position of both molecules overlap largely and theoretically could interfere with each other.

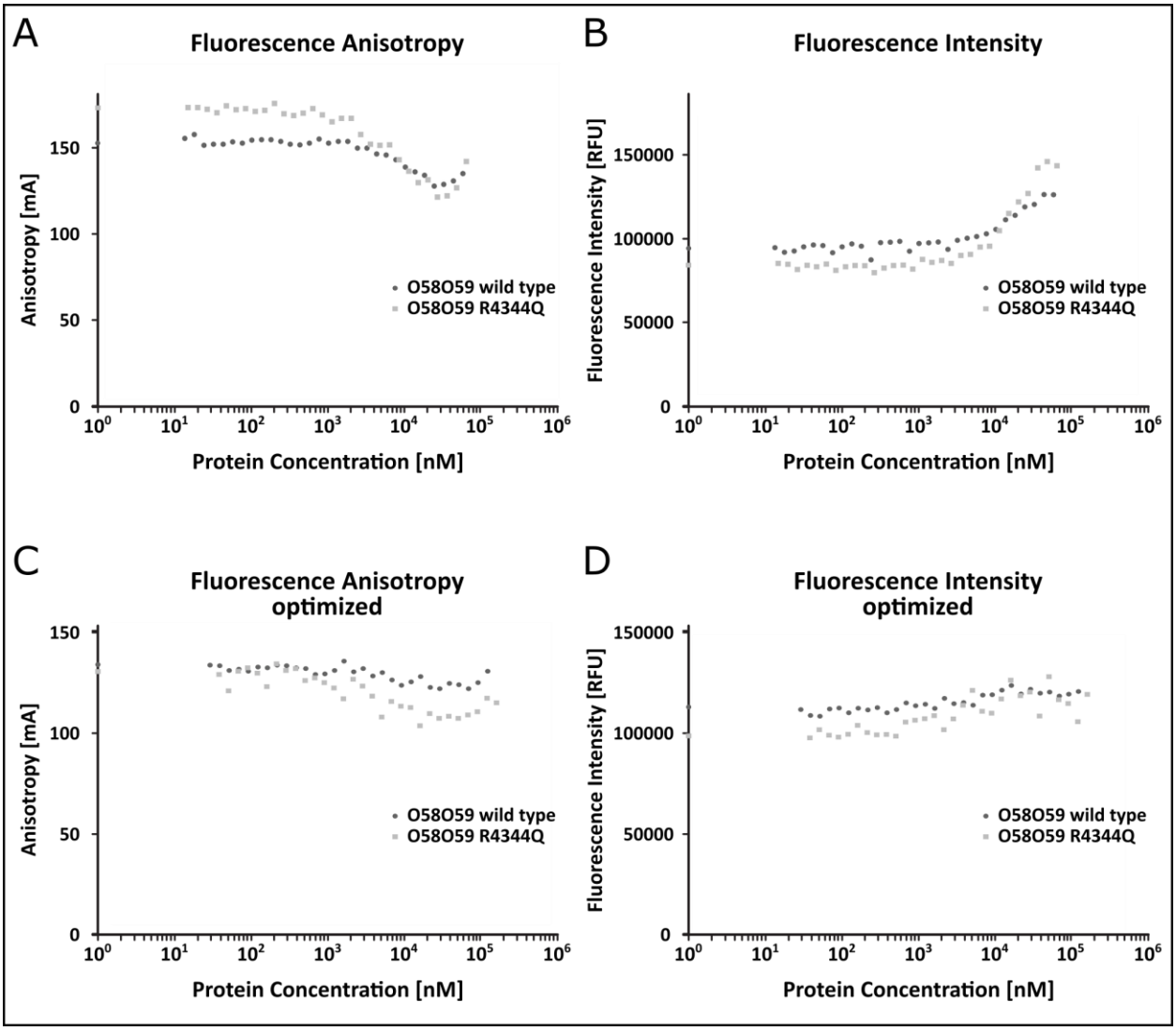

Figure 26 Fluorescence anisotropy based binding assay of PLN 1-19 and obscurin O58O59. A) Fluorescence anisotropy measurement of the N-terminally FITC-tagged PLN1-19 with obscurin O58O59 wild-type (dark grey) and R4344Q (light grey). Anisotropy values (millianisotropy - mA) decrease for wild-type and R4344Q mutant at higher protein concentrations. Expected behaviour for a binding of the peptide would be an increase of the fluorescence anisotropy values upon increasing protein concentration. B) Control measurements of the fluorescence intensities in relative fluorescence units (RFU). Slight increase can be observed for both wild-type and mutant. Values are expected to stay constant throughout the measurements as the concentration of labelled peptide stays constant. C) Optimised fluorescence anisotropy measurements. Anisotropy values are constant for the wild-type, while still slightly decreasing for R4344Q mutant. B) Control measurements of the optimised fluorescence intensities in relative fluorescence units (RFU). Slight increase can be observed for both wild-type and mutant.

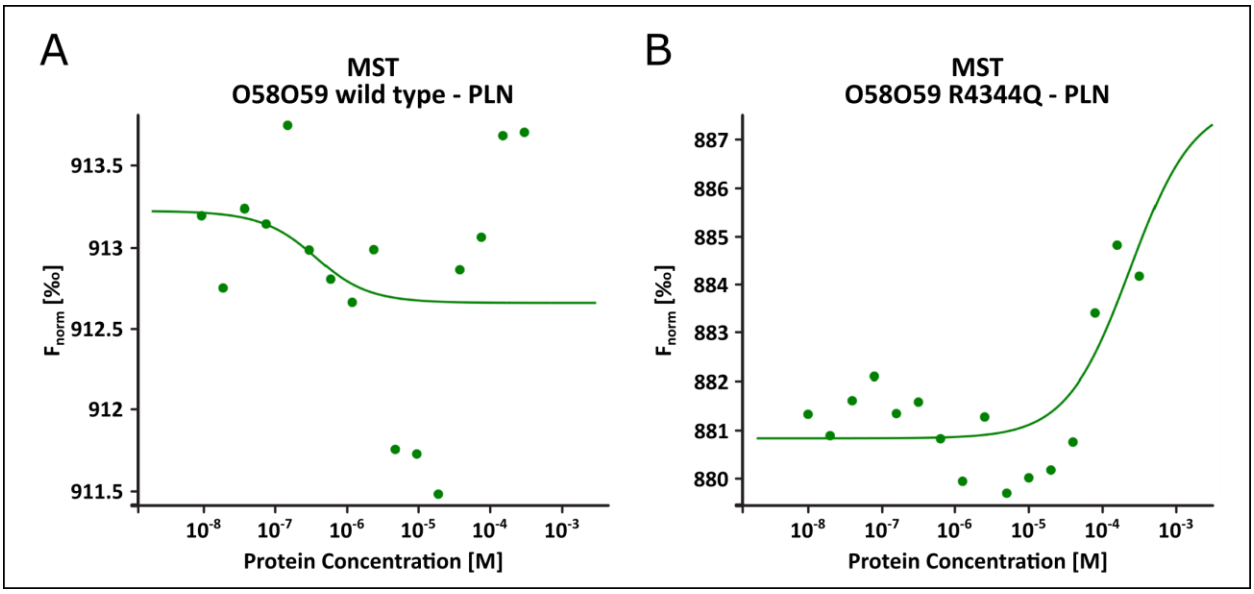

Figure 27 Microscale thermophoresis measurements of the binding of labelled PLN to obscurin O58O59 (A) and the R4344Q mutant (B). Normalised fluorescence is plotted against protein concentration. The signal for obscurin wild type was below the threshold not allowing to calculate a reliable dissociation constant. Even though the fit for the measurement with obscurin R4344Q seems to be better, the signal to noise ratio is just above the threshold still critically low making the calculated dissociation constant of K_d=239 µM not reliable.

5.1.13 In Vivo studies on the titin-obscurin complex formation

5.1.13.1 Phosphorylation of the titin-obscurin complex

In collaboration with Mathias Gautel at King's College London, it was possible to identify the location of three possible phosphorylation sites previously mapped on titin Z8Z9. In an *in vitro* screen of kinases able to phosphorylate Z8Z9, the mitogen-activated protein (MAP) kinase p38a was identified to specifically phosphorylate titin Z8Z9 at position T1727 (determined by mass spectrometry). Additionally, sequence analysis revealed the phosphorylation motif KKKXTS indicating a possible phosphorylation of position T1710 and/or S1711 in Z8Z9. All three phosphorylation sites are located within domain Z8. Residue 1727 is highly conserved as either threonine or serine in ten out of fourteen compared Ig-domains (sequence alignment performed on obscurin O58-O59, titin Z8-Z9 and the ten titin M-band domains M1-M10). A structure comparison (Figure 28) of the phosphorylation sites shows no involvement of residue T1727 in the interaction interface between titin and obscurin. In contrast, T1710 and S1711 are directly located at the binding interface between Z8 and O58, with S1711 forming hydrogen bonds to obscurin. ITC measurements of the phosphomimetic mutants support this finding (Table 38):

while T1727D did not show a difference to the wild-type interaction, T1710D and S1711D decreased the interaction. This effect is stronger for S1711D as it is directly involved in the intermolecular binding.

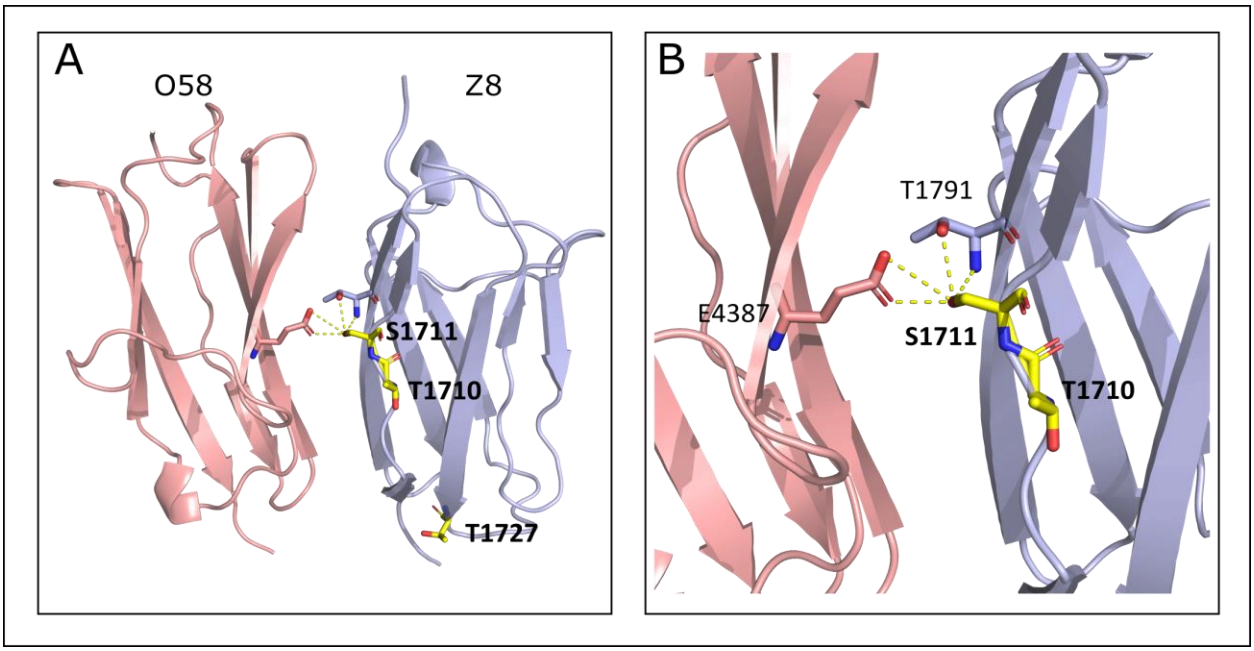

Figure 28 Phosphorylation sites of titin Z8. A) Position of the putative (T1710, S1711) and experimentally confirmed (T1727) phosphorylation site in the Z8-O58 interface shown in cartoon representation. B) Phosphorylation site at residues S1711 and T1710 including polar contacts. Initial discovery of the phosphorylation sties by Atsushi Fukuzawa (Gautel Laboratory).

Table 38 ITC measurements to analyse the impact of phosphomimetic mutations on the complex formation. Mean values and standard deviations derived from triplicates. With n (stoichiometry, Titin/Obscurin), H (enthalpy), T (temperature), S (entropy), K_d (dissociation constant).

Titin Z8Z9	Obscurin O58 O59	n	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})	K_d (μM)
wt	wt	0.93±0.09	-7.0±0.7	-0.4±0.7	4.4±0.1
		Phosphomimic mutations			
T1710D	wt	0.92±0.03	-6.6±0.4	-0.5±0.5	6.8±0.6
S1711D	wt	0.96±0.08	-5.9±0.7	-1.0±0.8	8.6±1.7
T1727D	wt	1.05±0.07	-7.1±0.1	-0.1±0.1	4.5±0.3

Nevertheless, a physiological phosphorylation would have a greater impact on the interaction. Since the kinase phosphorylating T1727 has been identified, experiments with the

phosphorylated titin Z8Z9 were performed by collaborators. They demonstrated that both pull-down assays and ITC with T1727p show lower affinity to obscurin O58O59 than the wild-type. At the same time, a non-phosphorylatable T1727A mutant shows higher affinity to obscurin O58O59 by *in vivo* FRET than wild-type (in COS1 cells). Surprisingly, the phosphorylation of T1727 by MAP kinase p38a was significantly lower in Z8Z9 carrying the phosphomimetic mutation S1711D compared to the wild-type. Taking together, the phosphorylation of titin at different residues might hint to a fine regulation of the titin-obscurin complex formation and lifetime. However, further investigation is necessary to understand the role of the individual phosphorylation sites and their interplay.

5.1.13.2 In vivo localisation experiments on proteins with interface mutations

Based on the structural analysis of the titin-obscurin complex and the *in vitro* experiments presented, the laboratory of Mathias Gautel performed *in vivo* experiments validating the accumulated data. The *in vivo* interaction between titin Z8Z9 and obscurin was assessed by localisation studies at the Z-disk. For this approach neonatal rat cardiomyocytes (NRC) that form fully functional sarcomeric structures were transfected with N-terminally eYFP-labelled obscurin O58O59 and its mutated versions. The Z-disk area was first determined by α-actinin base staining. α-actinin is an actin-binding protein that is known to localise to the Z-disk. Subsequently, the localisation of obscurin O58O59 or its mutants was assessed, and the ratio of fluorescence signal intensity between the Z-disk area and the non-Z-disk area in a defined part of the muscle cell evaluated (Figure 29 A+B). The localisation of obscurin O58O59 wild-type and eight different mutants were tested. The three mutants E4356R, E4387R and the combination E4356R/E4387R are involved in the Z8-O58 interface. Mutants D4377R and R4411E are part of the interface between titin Z9 and obscurin O58. The two mutants D4485R and D4485W are involved in the Z9-O59 interface. One mutant was designed to address the two interfaces Z9-O58 and Z9-O59 simultaneously involving D4377R, R4411E and D4485W. While all mutations in the interface between Z8 and O58 have a significant effect on the localisation of obscurin (adjusted p-value <0.0001), indicating an impairment of the titin-obscurin interaction, mutations within the other two interfaces were not able to reproduce the same effect. Even the triple mutation potentially affecting the O58-Z9-O59 interface did not have a significant effect. This underlines the

importance of the interface between Z8 and O58. At this point it remains to be elucidated whether i) the Z8-O58 interface is the only necessary interface to form the *in vivo* complex, ii) its proper formation is a prerequisite for the other interfaces to form or iii) the other interfaces have no physiological impact on the *in vivo* complex at all. However, *in vitro* data suggest that all three interfaces (reproducible K_d changes for all interface residue mutants in ITC measurements) and thus all four domains (Young, Ehler and Gautel, 2001) are involved in the complex formation. Notably, a doubling or tripling of the K_d is only expected to be around 10-20% of complex formation increase / decrease if the concentrations are assumed to be around the K_d. These minor changes would be difficult to detect in *in vivo* measurements.

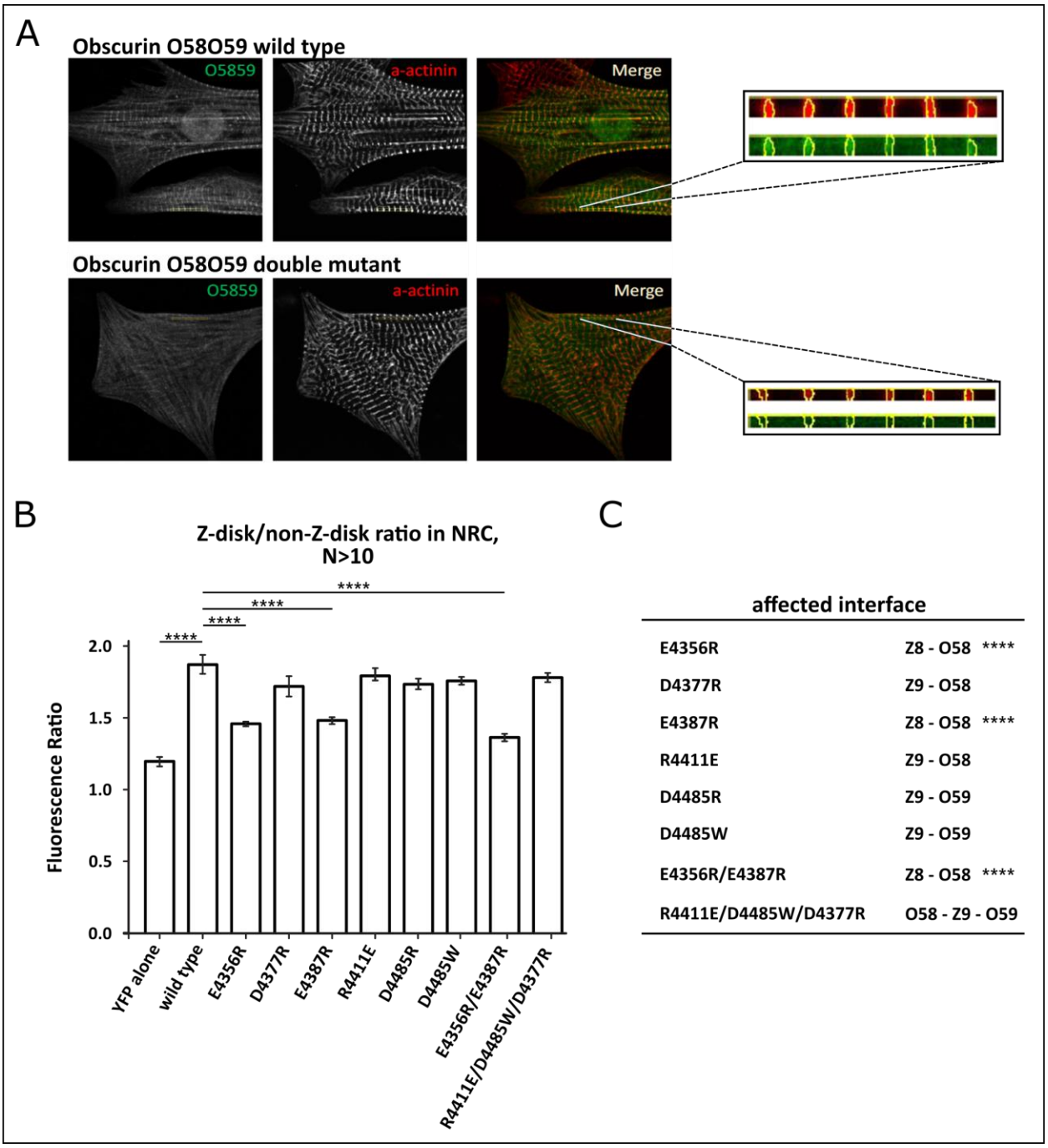

Figure 29 In vivo binding experiments in neonatal rat cardiomyocyte (NRC). A) Example fluorescence microscopy images of the location of N-terminally eYFP-labelled obscurin, α-actinin as an indication of the Z-disk localisation and an overlay of both signals. Ration between obscurin signal fluorescence intensity within the Z-disk and outside the Z-disk region was calculated and represented in B). Significant change in localisation (p-value <0.0001) is indicated, and only present in mutations affecting the Z8-O58 interface C). Data acquisition and microscopy performed by Atsushi Fukuzywa.

5.2 Myomesin 5 dimer and mutants

Previous attempt to solve the high-resolution crystal structure of myomesin FN-III domain 5 (My5) revealed a homodimer formation. Notably, a so-called domain swap in which both N-terminal β-strands exchange positions, each complementing the FN-III-fold of each other and forming a long β-strand. Biophysical bimolecular fluorescence complementation (BiFC) experiments performed by Chatziefthimiou and Temmerman (EMBL-Hamburg) prior to the work presented here, revealed three mutations located in the N-terminus of My5 able to deplete normal dimer formation.

5.2.1 Dimer purification

To investigate the in-solution behaviour of the myomesin My5 dimeric complex upon introduction of the mutants, single My5 domains of the wild-type and versions containing mutations P638R, P641R, or D643P were cloned and expressed. The preparative gel filtration (Figure 30 A) shows a significant change in the elution profile compared to the wild-type, but also when comparing the mutants among themselves. The wild-type shows four peaks: a prominent main peak at an elution volume of around 75 ml, one peak at higher elution volumes and two peaks at lower elution volumes. According to the molecular weights of the protein standard, the peaks could represent the My5 monomer (~85 ml), the dimer (~75 ml), the trimer (~70 ml), and the tetramer (~60 ml). The main peak has been shifted for all the tested mutants. For mutant P638R it is shifted to higher elution volumes indicating a reduced complex formation as it overlaps well with the wild type peak at around 85 ml probably representing the monomeric My5 construct explaining the decreased fluorescence signal in the BiFC experiments. For the other two mutants P641R and D643P the main peak is shifted towards smaller elution volumes around 70 ml. This hints to a higher molecular weight and, therefore, to an interaction of more than two My5 domains simultaneously. It overlaps well with the wild type peak assumed to be a trimer. A trimerisation, however, would not directly explain the decrease of the fluorescence signal in BiFC experiments unless the arrangement of the dimer is not part of the trimer. In addition, all mutants show a higher tendency for oligomerisation as can be seen on the large void volume peak. For a further analysis of the induction of higher oligomeric fractions, the wild-type and the

two mutants P641R and D643P were measured in SEC-MALLS experiments (Figure 30 B). For the wild-type, it was possible to confirm the presence of a tetrameric (~47 kDa), a trimeric (~36 kDa), a dimeric (~24 kDa), and a monomeric (~12 kDa) form of My5 by calculation of the corresponding molecular weights. The calculated molecular weight for the main peak of the two mutants also corresponds to a dimer and thus indicating a rearrangement of the two My5 domains in the dimer.

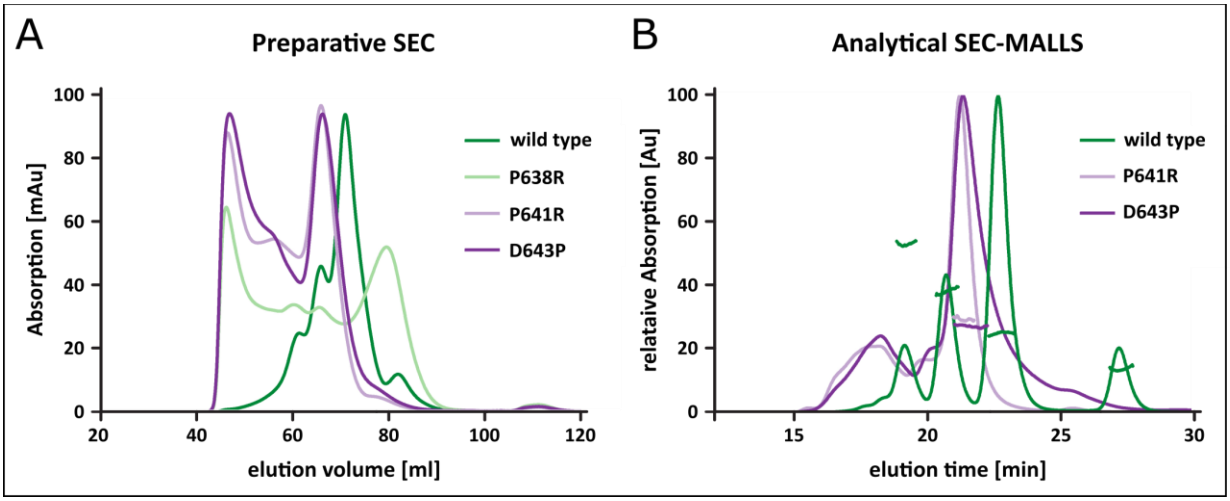

Figure 30 Purification profiles of Myomesin domain 5 wild-type and its mutants P638R, P641R, D643P. A) Preparative size exclusion chromatography. My5 wild-type shows four peaks of which the one at an elution volume around 75 ml is the most pronounced. P638R indicates a shift to the lower molecular peak at around >80 ml in combination of a decrease of the second peak. It also shows a higher tendency to form high molecular aggregates eluting in the void volume at around 45 ml. Same is true for P641R and D643P mutants. A peak shift of the most prominent peak to higher molecular weights and simultaneously an obliteration of the low molecular weight peak can be seen for both. B) Analytical size exclusion chromatography of the My5 wild-type and the two mutants P641R and D643P in combination with multi-angle laser light scattering (MALLS) measurements again indicates a shift of the main peak toward higher molecular weights. MALLS shows that all three main peaks have the molecular weight of the My5 dimer (~24 kDa) indicating a structural change of the dimer upon introduction of the mutants. Additional peaks for the wild-type correspond to the monomer (~12 kDa) at elution times around 27 min, the trimer (~36 kDa) at ~21 min, and the tetramer (~47 kDa) at ~19 min.

5.2.2 Analysis of the dimeric composition of the My5 dimer in solution using SAXS

SEC-SAXS measurements were performed on My5 wild type and the three My5 mutants P638R, P641R, and D643P to evaluate the differences between the constructs seen during purification and SEC-MALLS. These measurements were performed in collaboration with Melissa Graewert (Svergun group, EMBL-Hamburg).

5.2.2.1 Monomer shape and P638R comparison

First, the scattering profiles of the monomeric peak of My5 wild-type and the mutant P638R were measured. Different models were fitted to the experimental data: model 1 is the My5 monomer structure extracted from My5 dimer crystal structure without the 15 amino acid long 'swapped' N-terminus, model 2 is the My5 monomer extracted from My5 dimer crystal structure containing the complete N-terminus, model 3 is the structure of a My5 mutant (F700S, solved by Sauer et al.) that could be crystallised in a monomeric state representing a fully formed FN-III domain, and model 4 with a crystal structure of the monomeric My4 (PDB code: 5FM4). Chi^2 values of the model fits show (Figure 30) that for both scattering profiles model 2, representing the complete My5 domain with swapped N-terminus as seen in the structure, reaches the lowest and therefore the best values. This indicates that the N-terminus does not collapse into the domain structure forming a complete FN-III domain, when present as a monomer but rather stays exposed.

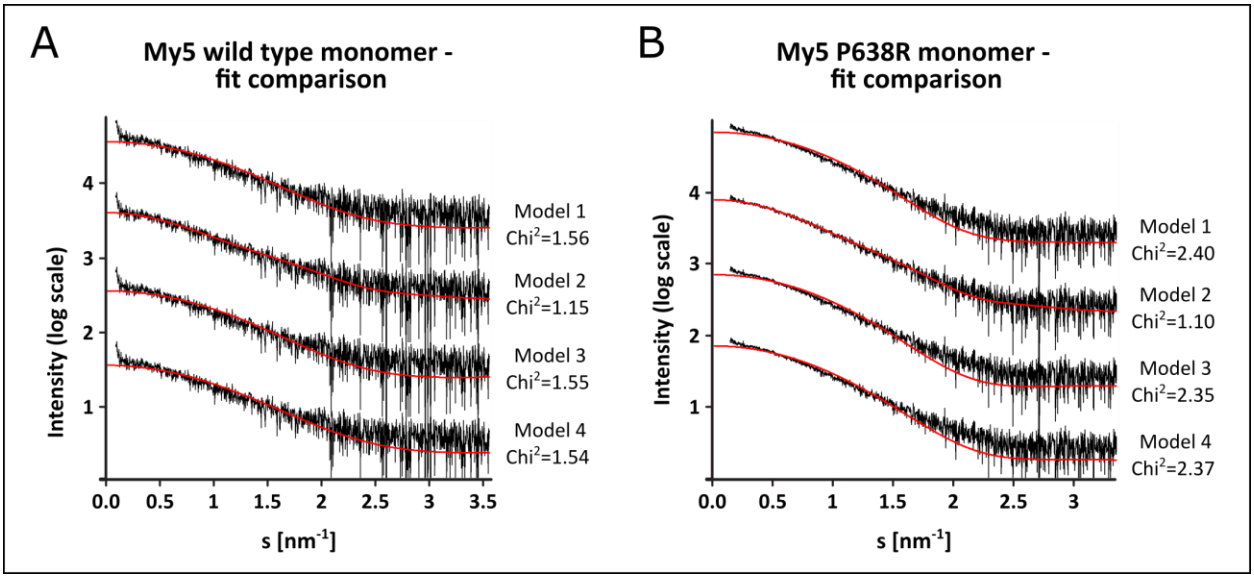

Figure 31 SAXS experiments on My5 wt and mutant P638R. Fits of different monomer models to the scattering curves of My5 wild-type (A) and My5 P638R (B). Model 1: monomer extracted from My5 dimer crystal structure without the 15 amino acid long 'swapped' N-terminus. Model 2: monomer extracted from My5 dimer crystal structure containing the whole N-terminus. Model 3: Structure a My5 mutant (F700S) that could be crystallised in a monomeric state representing a fully formed FN-III domain. Model 4: Crystal structure of the monomeric My4 (PDB code: 5FM4). For both, the wild-type and the mutant, Model 2 shows the best fit with a lower Chi^2 value compared to the other models indicating an 'opened' FN-III domain also in the monomeric state. Data analysis performed in collaboration with Melissa Graewert.

To further investigate the structural difference between the wild type and the mutant P638R leading to the loss of dimerisation, the Ensemble Optimisation Method (EOM) was applied. Here,

a pool of independent atomic models that are based on sequence and structural information is generated, and the averaged theoretical scattering intensity are compared to the experimental scattering. The calculation shows a higher R_g value for the wild-type (18.5 Å) than for the mutant (17.7 Å) (Figure 32 A+B). The comparison to the theoretical R_g values from the different crystal structures to which the data were fitted, also shows that the measured wild-type most probably has a largely exposed N-terminus as its R_g is closest to the R_g of the monomer extracted from the dimeric crystal. The mutant P638R monomer appears slightly more compact but is still larger than the completely folded monomeric models of My5 mutant structure (F700S) and My4. The visualisation of the two best fitting EOM models for wild-type (grey) and P638R (cyan) shows that the N-terminus of the wild type in both fits seems slightly more exposed and therefore more accessible for a potential domain swap between two My5 molecules.

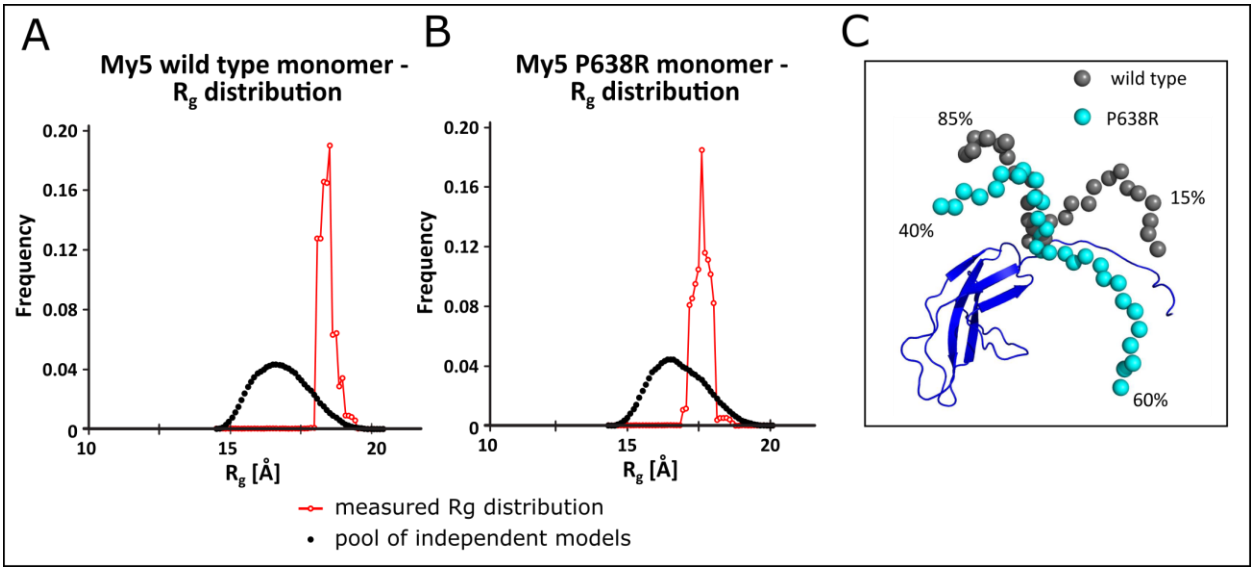

Figure 32 Comparison of R_g values for My5 wild-type monomer and P638R monomer derived from Ensemble Optimisation Method (EOM) calculations based on the experimental data of the respected monomer. A) R_g distribution for the My5 wild-type monomer shows a peak shifted to larger radii between 18 and 19 Å compared to a pool of independent models generated based on sequence and structural information. B) R_g distribution for the My5 P638R monomer also shows a shift, however, compared to the wild-type less shifted to larger radii. C) Visualisation of the two best fitting EOM models for wild-type (grey) and P638R (cyan). Frequencies are indicated. The N-terminus of My5 wild-type seems to be more exposed than for the mutant. Data analysis performed in collaboration with Melissa Graewert.

Table 39 Radii of gyration for the measured samples of My5 wild-type and My5 P638R as well as the calculated R_g values for different My5 models: Model1: monomer extracted from My5 dimer crystal structure without the 15 amino acid long 'swapped' N-terminus. Model 2: monomer extracted from My5 dimer crystal structure containing the whole N-terminus. Model 3: Structure a My5 mutant (F700S) that could be crystallised in a monomeric state representing a fully formed FN-III domain. Model 4: Crystal structure of the monomeric My4 (PDB code: 5FM4). Measured wild type is closest to R_g value of Model 2, indicating a largely exposed N-terminus. The mutant P638R monomer appears slightly more compact but is still larger than the completely folded models 3 and 4.

Domain	Measured wild-type monomer	Measured P638R monomer	Model 1: monomer from dimer without N-ter	Model 2: monomer from dimer with N-ter	Model 3: monomeric My5 mutant (F700S)	Model 4: monomeric My4
R_g [Å]	18.5	17.7	15.8	19.7	16.1	16.3

5.2.2.2 Dimer of P641R and D643P

In the next step, the scattering curves of the dimeric peaks of the wild-type and both mutants P641R and D643P were compared. The scattering curves reveal an additional feature, not present in the mutants, for the wild-type dimer that is even more pronounced in the Kratky plot at distances around 1.5-2 nm^{-1} (Figure 33). The additional shoulder is caused by the two globular parts that are connected by the β-sheet formed due to the domain swapping as seen in the crystal structure. The scattering curves of the mutants, therefore, indicate a more globular shape of the alternative dimer. A first model calculated on the SAXS data from My5 P641R (Figure 34) proposes a completely switched orientation of the two My5 molecules compared to the wild-type with the N-termini neither involved in the dimer formation nor buried in the FN-III-fold but exposed as seen for the monomers. The large distance between the two domains is probably due to a small fraction of slightly larger particles contributing to the SAXS curves that are still considered for modelling.

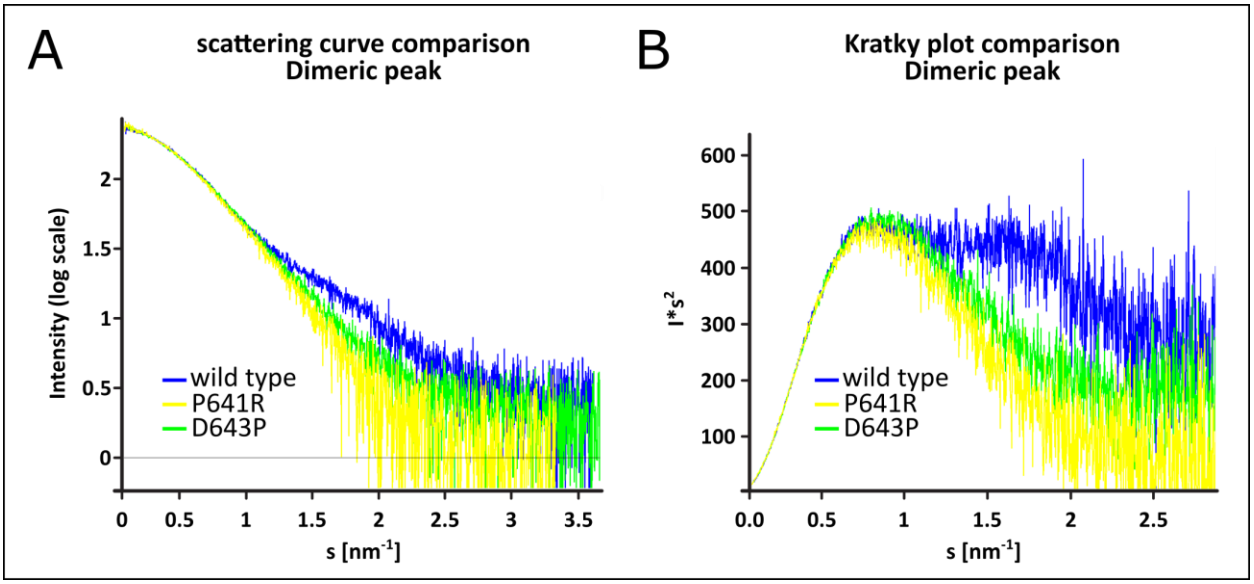

Figure 33 Scattering curve comparison of the dimeric peak of My5 wild-type, My5 P641R and My5 D643P. My5 wt (blue), My5 P641R (yellow) and My5 D643P (green) A) Scattering curve of the wild-type reveals an additional feature at distances around 1.5-2 nm^{-1} not present in the mutants. B) The Kratky plot enhances the difference between the wild type and the mutant scattering. While the mutants form an almost bell-shaped curve, indicating a globular shape of the dimer, the additional shoulder of the wild-type curve represents the two globular domains that are connected by the β-sheet formed by the two N-termini. Data analysis performed in collaboration with Melissa Graewert.

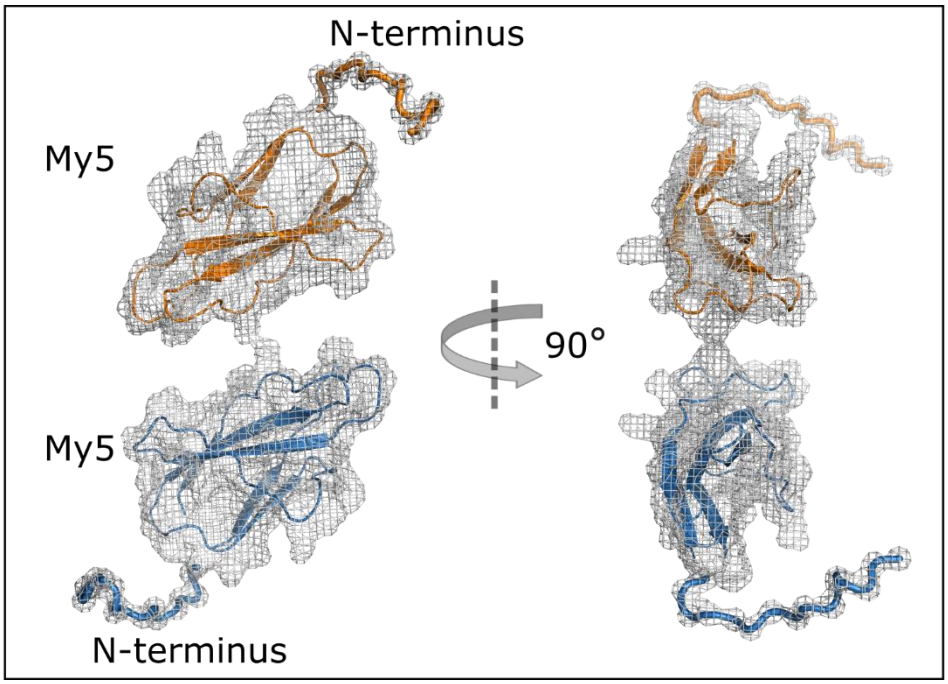

Figure 34 Preliminary SAXS model calculated from the scattering curve of My5 P641R (calculated by Dmitri Svergun (EMBL-Hamburg)). In this model, the N-termini of both molecules are still exposed and not involved in the dimer formation.

As a validation of the collected data, the fit between the experimental scattering data of My5 wild-type and My5 P641R, and the theoretical scattering curves of the dimeric crystal structure and the modelled structure was calculated. As the model structure was created on the scattering of P641R, it only served for a fit comparison with the wild-type scattering curve. For both experimental datasets, the modelled dimers show a slightly better fit (Figure 35). Nevertheless, the Kratky plot reveals that only the theoretical scattering curve derived from the crystal structure reflects the additional feature of the wild-type at distances around 1.5 nm^{-1} and does not fit properly with the experimental data of the mutant. The better fit is a result of the high noise level at higher distances. Taking these data together, it was possible to show that, differently than expected, in solution, the individual My5 monomer does not form a complete FN-III-fold. The N-terminal part of My5 does not only seem to flip out during domain swap upon dimerisation as seen in the crystal structure but is exposed the entire time. The mutation P638R is able to abolish dimer formation. The mutation seems to reduce the accessibility of the exposed N-terminus. Whether this is the reason for the deficient dimerisation or another mechanism during the domain swap that involves the residue P638 is not known. The other two mutation P641R and D543P decrease the monomeric fraction of My5 and form an alternative dimer, presumably not involving the N-terminus.

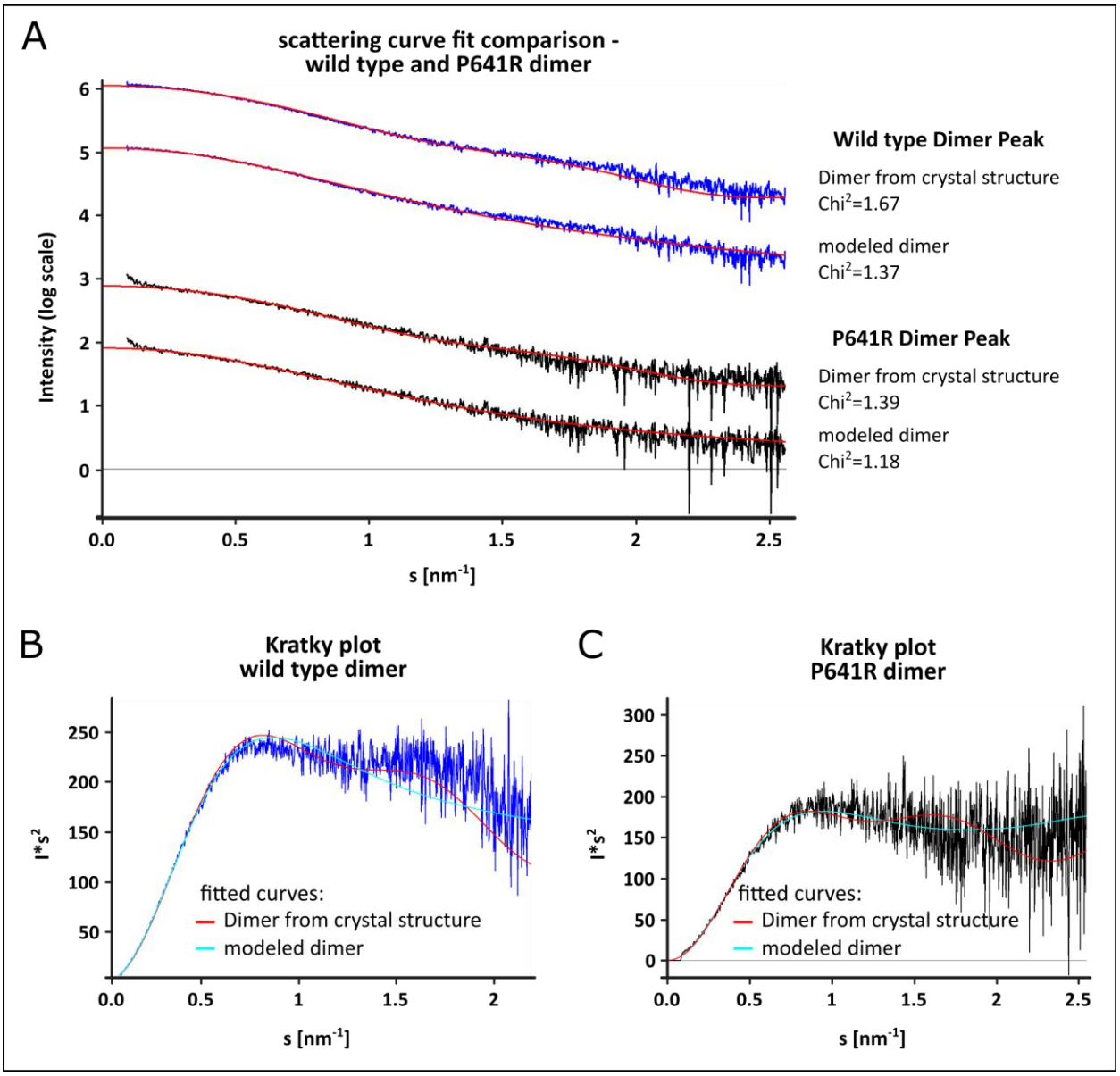

Figure 35 My5 wt and P641 mutant scattering curve comparison. Comparison of the theoretical scattering curve derived from the My5 crystal structure and the calculated model with the experimental scattering curves of My5 wild-type (blue) and P641R (black). A) Better fit of the modelled dimer for both the wild-type dimer peak and the mutant dimer peak. B) Kratky plot shows that the additional feature of the experimental wild-type scattering curve is much better represented by the theoretical scattering curve derived from the crystal structure. C) Kratky plot for the P641R mutant illustrates the lack of the characteristic feature at around 1.5 nm^{-1} and therefore a better fit for the calculated model dimer. Data analysis performed in collaboration with Melissa Graewert.

5.2.3 Sequence comparison of myomesin FN-III domains

To further investigate the unusual behaviour of the My5 FN-III domain even for the wild-type, a sequence alignment of the five FN-III domains of myomesin was performed. Sequence identities range between 41.6% for My4/My5 and 25.7% for My6/My8 (Table 40) and is thus considerably higher than what has been observed for the titin-obscurin domain comparison.

Table 40 Myomesin FN-III domain sequence identities in percent [%].

Myomesin FN-III sequence identities					
Domains	My4	My5	My6	My7	My8
My4	100	41.6	33.7	34.0	39.8
My5		100	33.3	31.7	39.6
My6			100	27.7	25.7
My7				100	36.2
My8					100

The sequence alignment reveals 13 highly conserved residues within all structures from which six are located within the C-terminal region (Figure 36). Notably, the two residues P638 and P641 that show a different dimerisation behaviour when mutated to arginine are among these conserved residues. Residue D643 is partly conserved in three out of five domain sequences. The completely folded FN-III domain structure of My4 indicates that all three residues are part of the conserved pattern within the structure and thus are likely weakening the packing of the N-terminus within the domain fold when mutated from a proline to an arginine. The swapped N-terminus of the My5 dimer positions itself in the same way within the complementary structure. At three distinct positions, My5 differs significantly from the otherwise conserved positions of the other domains. My5 has a methionine instead of a glycine at position 670, an additional threonine at position 688, and a cysteine instead of an alanine at position 714. While positions 670 and 688 are far apart from the N-terminal position of a completely folded FN-III domain, the cysteine at position 714 (A586 in My4 sequence) is directly located in the cavity where the N-terminus is integrated, as can be seen in the structure of My4 (Figure 37 A). The larger cysteine could potentially interfere and sterically prevent a tight packing of the N-terminus in the monomeric My5, opening the My5 monomer and facilitating the domain swap und thus

dimerisation. Upon mutation of P638 and P641 to an arginine, the steric hindrance could be high enough to prevent incorporation of the N-terminus in the dimer counterpart. In My4 and the My5 dimer, the homologous residue for D643 is forming a tight polar interaction (2.9-3.1 Å) with the lysine corresponding to My5 residue K659, a residue conserved in four out of five domains. This interaction might stabilise the N-terminus in the 'closed' domain conformation or the swapped dimer conformation. The mutation of D643 to a proline abolishes this interaction, weakening the N-terminal packing and potentially leading to a disruption of the My5 dimer.

As the interactions of the N-terminus of My5 in theoretical 'closed' state and in the dimeric state are assumed to be very similar an additional factor needs to play a role in favouring the domain swap and thus dimer formation over keeping the 'closed' monomeric structure. The crystal structure of the My5 dimer reveals a polar interaction between Arg651 and Asp701 (Figure 32 B). This interaction is directly located at the hinge region between the swapped β-strand A and β-strand B. While a negatively charged residue at the homologous position of Arg651 can be found in all but one compared myomesin structure, none of them have a positively charged residue at the corresponding positions of Asp701, making this interaction unique for My5. R651 and D701 might possibly stabilise the flipped out β-strand A, thus leading to favoured extended N-terminus even for the wild-type, as seen in SAXS. Additionally, the hydrophobic F700 might translocate residue D701 towards R651 and thus stabilise this interaction. Taking these data together, two major sequence derived factors might lead to the observed behaviour of My5. First, the exchange of a highly conserved alanine to a cysteine within the hydrophobic core might weaken the packing of the N-terminus, and second, the unique interaction between R651 and D701 might stabilise the flipped-out N-terminus.

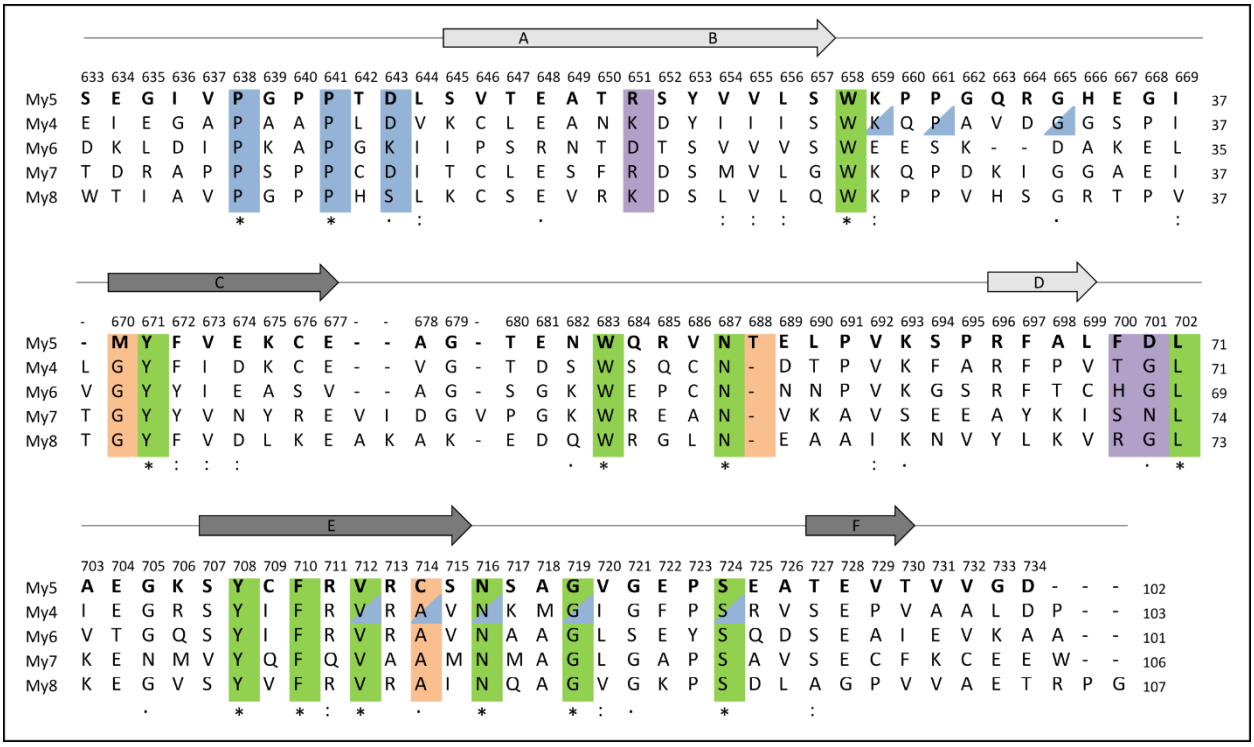

Figure 36 Sequence alignment of the myomesin FN-III domains My4-My8. Marked in green are the conserved residues. Residues marked in blue are the mutated residues that lead to an altered dimerisation behaviour, residues marked with a blue triangle are in close proximity to the mentioned residues, residues marked in orange are residues that are not conserved in My5 even though conserved in the other myomesin FN-III domains, residue marked in violet are involved in the interaction stabilising the hinge region. Asterisk indicates positions which have a single, fully conserved residue, colon indicates conservation between groups of strongly similar properties, period indicates conservation between groups of weakly similar properties. The positions of the secondary structural element shown on top of the alignment correspond to the My5 domain. Light grey and dark grey indicate the involvement of the β-strands in the different β-sheets.

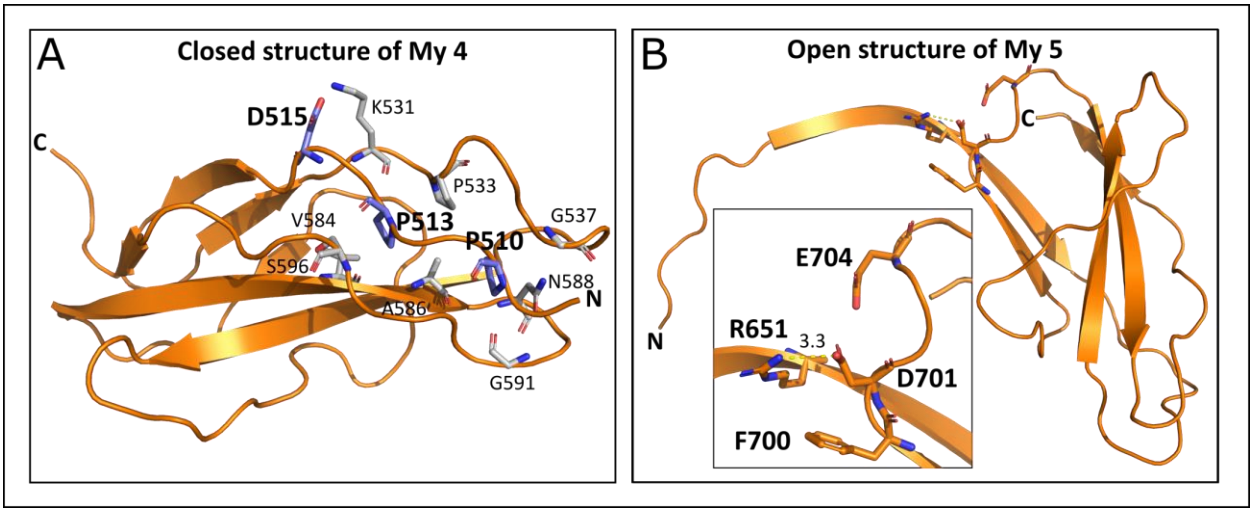

Figure 37 Comparison of the 'open' My5 structure and the 'closed' My4 structure (PDB code: 5FM4). A) Position of the My4 N-terminus. Residues in blue are the homologous residues of My5 P638, P641, and D643. Residues in grey are in close proximity to the N-terminus. All residues are highly conserved in all sequence aligned myomesin FN-III domains. B) Hinge region between the swapped β-strand A and β-strand B of My5. R651 and D701 show a polar interaction in the crystal structure possibly stabilising the flipped out β-strand A. The hydrophobic F700 might push residue D701 towards R651 and thus stabilise this interaction.

5.2.4 My5 Variants

Next, a search for variants located in My5 was performed. For 32 out of 102 residues, variants can be found from which most are affecting non-conserved residues. Notably, variants for all three N-terminal residues shown to have an effect on the dimerisation (P638, P641, and D643) can be found. Table 41 shows variants that are either affecting conserved residues or were deposited in a disease database. In both cases, six different variants were found. Variants that has been found in connection with a cardiovascular disease are all assigned benign or likely benign. All of the six disease-linked variants except for the one that also affects a conserved residue have a MAF above 1.00E-4 and are thus more frequent than expected for diseases of Mendelian inheritance. As expected, all six variants affecting conserved residues show MAF values lower than 1.00E-4. In total five variants might exert a direct effect on the behaviour of the N-terminal part of My5 and therefore could influence the dimerisation behaviour.

Table 41 Variants found in myomesin domain My5 that are either affecting conserved residues or were assigned with a 'rs' number.

residue	SNV	RS	MAF	disease connection	clinical significance	Structure	possible Structural/functional Impact	comment
638	P/S	-	8.34E-06	NA	NA	N-terminal loop, buried	Weakening of N-terminal packing in monomer and dimer	
641	P/L	rs138431527	2.50E-05	NA	NA	N-terminal loop, buried	not expected	
643	D/G	-	8.32E-06	NA	NA	N-terminal loop, exposed	Weakening of N-terminal packing in monomer and dimer	
651	R/Q	rs184721031	1.33E-04	Hypertrophic cardiomyopathy	Likely benign	β-strand, exposed in dimer; loop region, exposed in monomer	Destabilisation of the 'flipped-out' β-strand A due to loss of interaction with D701	High allele frequency; 0.13% (15/11558) of Latino chromosomes
688	T/S	rs188677538	2.71E-03	Hypertrophic cardiomyopathy, Cardiovascular phenotype	Likely benign	loop, exposed	not expected	High MAF
704	E/K	rs149528866	5.78E-03	Hypertrophic cardiomyopathy, Cardiovascular phenotype	Benign	loop, exposed	Destabilisation of the 'flipped-out' due to competition for interaction with D701	High MAF
708	Y/C	-	8.32E-06	NA	NA	β-strand, partly exposed	not expected	Conserved residue
710	F/C	-	1.66E-05	NA	NA	β-strand, buried	not expected	Conserved residue
710	F/L	-	8.32E-06	NA	NA	β-strand, buried	not expected	Conserved residue
713	R/H	rs183881662	6.68E-05	Arrhythmogenic right ventricular cardiomyopathy	uncertain	β-strand, exposed	not expected	
714	C/S	-	2.51E-05	NA	NA	β-strand, buried in cavity for N-terminus	Stabilisation of the closed conformation due to increased packing of N-terminus	Conserved residue
727	T/A	rs115382168	1.34E-03	Hypertrophic cardiomyopathy Cardiovascular phenotype	Benign	β-strand, partly exposed, pointing cavity for the N-terminus	not expected	High MAF

The variant P638S has a mean allele frequency (MAF) of 8.34E-6 and introduces a polar amino acid. In a completely folded FN-III domain this would interfere with the hydrophobic core which could lead to a weakening of the N-terminal packing. The second variant P641L introduces a hydrophobic amino acid. The residue points toward the hydrophobic core of the domain in both monomeric and dimeric state. Thus, this variant is not likely to interfere with the domain's behaviour. This is supported by an almost tenfold higher MAF of 2.50E-5 compared to the P641L variant. For residue D643G variant can be found that would abolish the previously mentioned interaction with K659 potentially destabilising the buried N-terminus in the monomer as well as in the dimer. Residue R651 is located in the loop region between β-strand A and B in the closed My5 conformation and part of the 'flipped-out' β-strand AB in the open/dimeric conformation. In the dimeric conformation, it interacts with D701 possibly stabilising the flipped β-strand.

An arginine to glutamine mutation as observed for residue 651 is normally not expected to have a large impact on the protein's behaviour. Nevertheless, it could abolish the interaction between residue R651 and D701 due to the reduced side chain length. Even though this variant has been connected to hypertrophic cardiomyopathy, the MAF of 1.33E-4 is too high to be the unique cause of disease development. Residue E704 has been found in connection to hypertrophic cardiomyopathy and cardiovascular phenotype. It is located in the loop region between β-strands D and E and in close proximity to the hinge region between β-strands A and B and thus close to Arg651. Even though not directly indicated in the crystal structure both residues are close enough to interact with each other and therefore stabilise the flipped-out β-strand. An exchange with a positively charged arginine at this position might compete with R651 for the interaction with D701 and at the same time repel R651 which might lead to a destabilisation of the flipped-out β-strand. The last variant that might have a structural or functional impact is C714S. This residue is buried in the cavity where the N-terminus is placed. A substitution with a smaller serine could lead to more space and better packing for the N-terminus and thus a stabilisation of the closed conformation.

5.2.5 My5 dimer force stability

The My5 crystal structure revealed a tight interaction of the two N-terminal parts of My5 forming a β-sheet upon domain swap. As myomesin is a muscle protein supposed to withstand high mechanical forces a strong interaction between both domains could serve as an additional junction within the M-band protein-protein interplay. To evaluate the stability of the My5 dimer upon mechanical stress, atomic force microscopy measurements on the dimer were aimed to be performed. Two constructs were designed, one with an yBBR tag for immobilisation on the surface (construct A) and another one carrying a sortase motif for coupling on the AFM cantilever (construct B) (Figure 38). To be able to purify dimers consisting of both constructs different purification tags were introduced. C-terminal of the My5 domains a linker including a very C-terminal cysteine was added. This is necessary as a reference fingerprint to identify the final complex disruption during AFM experiments.

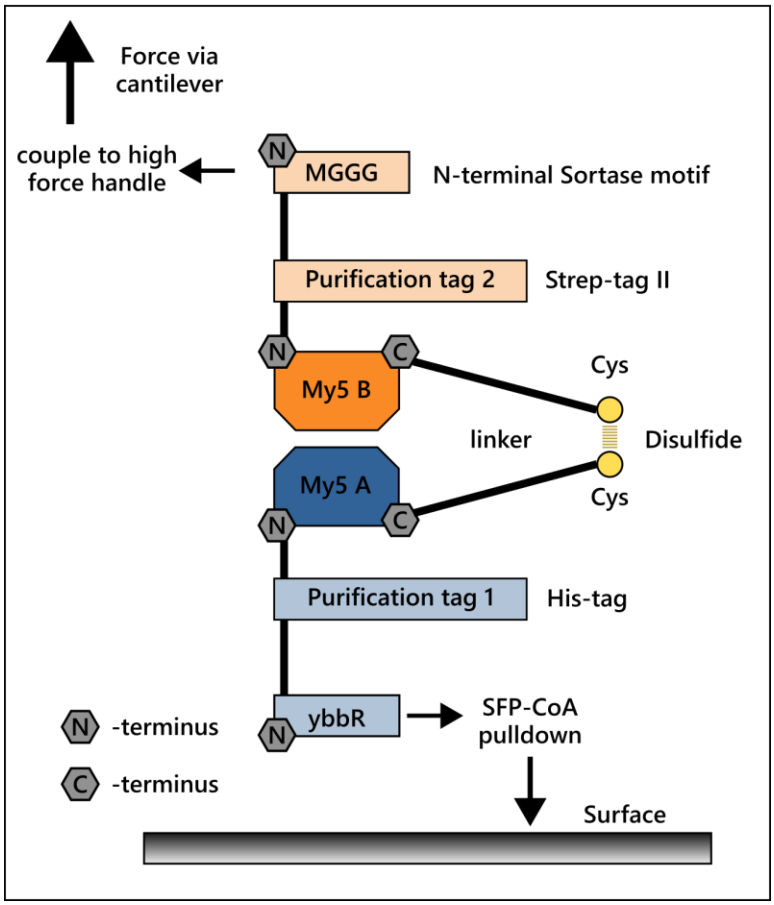

Figure 38 Scheme of the My5 construct architecture planned for the atomic force microscopy (AFM) measurements. Two constructs were designed each carrying two specific purification and AFM coupling tags. Construct A carries an N-terminal His-tag for Ni-NTA purification and an ybbR-tag for immobilisation on the AFM device surface. Construct B has an N-terminal Strep-tag II and a sortase motif to be coupled to the cantilever of the AFM device. Both constructs carry a C-terminal cysteine for disulphide formation necessary as a fingerprint force during measurement.

As a first attempt both constructs to form the complex were separately expressed in *E. coli* and subsequently combined during purification. While it was possible to purify both constructs *via*

their corresponding tag the second purification step did not purify any protein indicating that no AB-dimer was formed (Figure 39).

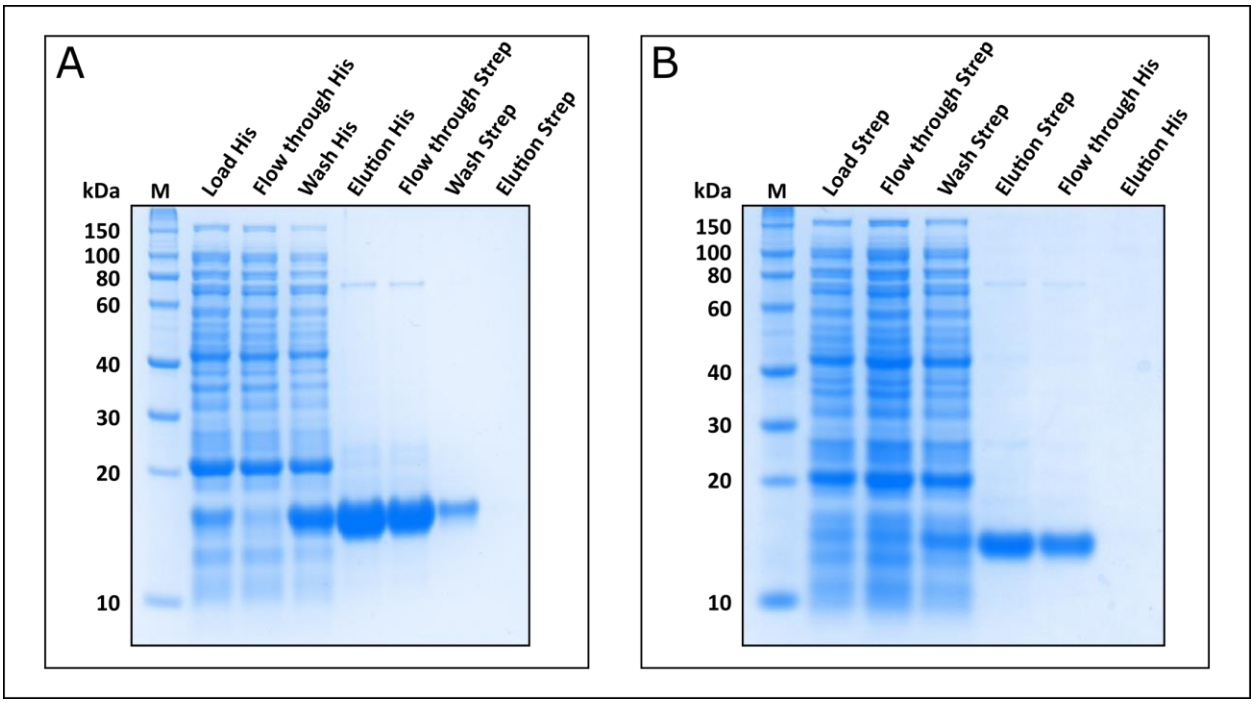

Figure 39 SDS-Page of the two-step purification of the My5 dimer for AFM. My5 construct A carrying a His-tag and My5 construct B carrying a strep-tag were separately expressed. Zell pellets were combined prior to protein purification. A) Dimer purification of the AFM dimer using the His-tag first and second the Strep-tag. The eluting fraction after His-tag purification step shows a high yield of protein. Wash His fraction also contains a large fraction of My5 indicating the presence of purely strep-tagged My5. Equal band intensity of the elution His and flow through strep fraction, and no protein after the step elution indicates no binding to the strep-tactin column and therefore the absence of step-tagged protein. B) Dimer purification of the AFM dimer using the Strep-tag first and second the His-tag. Protein can be purified via the first elution step, but no protein can be recovered after the second step.

Second, both constructs were transformed into the same expression host, expressed and purified. Gel filtration after the two-step purification indicates the presence of the dimer with a smaller fraction of monomeric My5 (Figure 40). The yield of dimeric protein (Figure 40 A) is very low compared to the yield after one purification step (Figure 40 B). This hints to two possibilities. Either the dimerisation/dissociation frequency is very high and formed AB dimers dissociate between the first and the second purification step resulting in the loss of a large fraction carrying the first purification tag or the dimer formation is very fast with a low on/off rate and happens directly after the translational process, thus, favouring the formation of AA and BB dimers. The monomeric fraction after the second purification step indicates that the dimer is able to dissociate and both the dimeric and the monomeric form are present in an equilibrium.

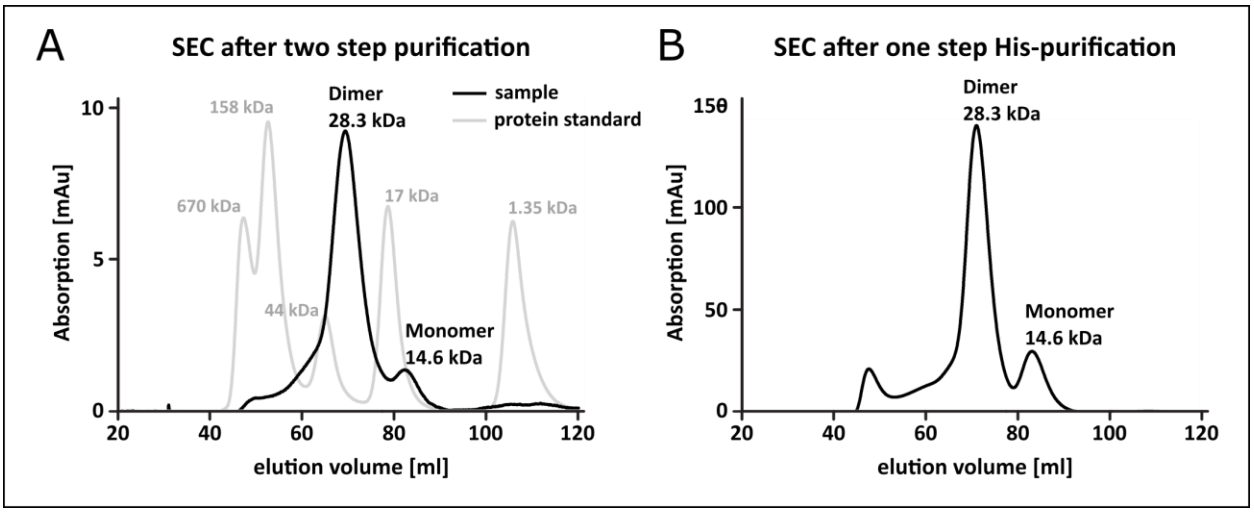

Figure 40 Preparative size exclusion chromatography of My5 after co-expression from different vectors. A) Elution profile after the two-step purification via His- and Strep-tag. Monomeric and dimeric peaks are clearly separated and verified by the comparison with the protein standard. Existence of the monomeric peak indicates an equilibrium between dimeric and monomeric form of My5, as only dimers could have been purified in the two-step purification. Protein yield is very low as indicated by the low absorption. B) As a comparison the elution profile after the first purification step via His-tag shows a substantially increased protein yield, indicating a preferred dimerisation of My5 constructs expressed on the same vector. The proportion between dimer and monomer did not change.

To increase the probability of AB-dimer formation directly after translation, a polycistronic vector was designed carrying both constructs next to each other for a consecutive translation. Additionally, DMSO was added during purification to facilitate the disulphide bridge formation and thus to stabilise the dimer (Tam *et al.*, 1991). The yield of dimeric protein was considerably higher after the second purification step and the dimer/monomer ratio increased (Figure 41). A large fraction of higher oligomeric proteins can be seen eluting from the gel filtration column which might be due to the addition of DMSO. However, the amount of purified dimer was sufficient for AFM experiments. These experiments are ongoing.

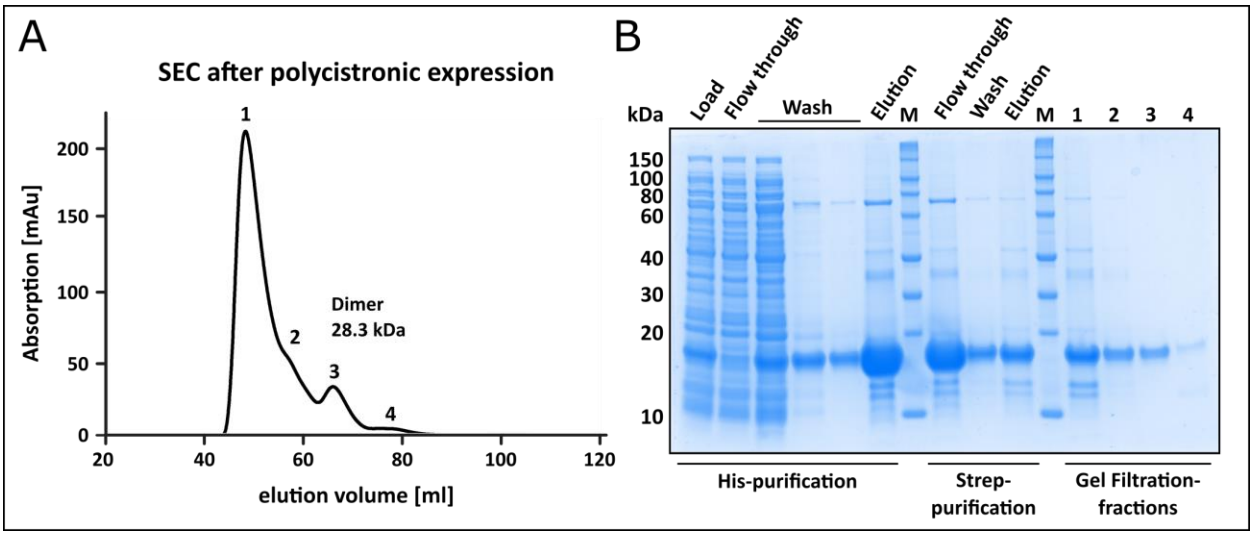

Figure 41 Preparative size exclusion chromatography of My5 after polycistronic co-expression. A) Elution profile after a two-step purification via His- and Strep-tag. Considerably higher void volume peak at ~45 ml (peak 1) and a pronounced shoulder at `55 ml (peak 2) indicate higher susceptibility of the My5 constructs to for higher oligomers. B) SDS-Page of the My5 dimer purification steps. Sample taken from peak three shows the high purity of the My5 dimer.

6 Discussion

6.1 The titin-obscurin Z-disk complex

The titin-obscurin complex located at the sarcomeric Z-disk and formed between titin domains Z8Z9 and obscurin domains O58O59 has been under investigation since its first report in 2001 (Young, Ehler and Gautel, 2001). However, not much is known about the functional role or its structural properties. Even though three of the four involved protein domains (obscurin O58, O59, and titin Z9) have been structurally characterised in several individual reports (Hu *et al.*, 2017; Rossi *et al.*, 2017; Wright and Letourneau, 2018), the high-resolution structure of the whole complex and thus the multifaceted domain interplay remained elusive.

Here, for the first time, it was possible to successfully obtain the atomic-resolution complex structure of the titin-obscurin Z-disk interaction, allowing a detailed analysis of the involved domain interfaces and the structural impact of variants linked to myopathies, and developing an attractive model for a complex regulation mechanism. In the following sections, first, the protein complex behaviour during purification and crystallisation will be discussed. Second, the general architecture of the Z-disk complex will be assessed under consideration of the different possible domain arrangements within the asymmetric unit of the crystal structure. Third, identified variants and their potential impact on the development of myopathy phenotypes will be evaluated. Finally, the general nature of the complex and its function will be explored addressing the questions, why the complex consists of four protein domains and how the localisation of obscurin at the Z-disk might be regulated.

6.1.1 Complex formation and its behaviour

The behaviour of the titin-obscurin complex before and during crystallisation was highly unusual compared to other proteins and protein complexes. Immediate precipitation was observed when mixing both proteins at high concentration. Furthermore, diluted protein also precipitated at low temperatures upon gradually increase of the concentration. As the concentration is a critical step for crystallisation, extensive process optimisation was performed. It was demonstrated that a

temperature increase during the concentration process was the key factor to obtain high complex concentrations. While both proteins did not show a sensitivity for low temperatures when stored separately on ice or in liquid nitrogen, the highly concentrated protein mixture was immediately precipitating when exposed to lower temperatures (<10 °C). Interestingly, crystallisation was only reproducible after increasing the temperature from 19 °C to 25 °C for 24 hours.

This shows that the temperature is a decisive factor upon complex formation leading to aggregation at low temperatures. Proteins unfold at high and low temperatures (Privalov, 1990; Somera, 1995). However, an unfolding event due to complex formation is unlikely when the individual components of the protein appear stable at the same temperature range. Nevertheless, it has been shown that protein interfaces are adapted to their native ambient temperature (He and Ma, 2016). Thus, a lower temperature might lead to an increase of a more undirected binding of the positively net charged titin domains and the negatively net charged obscurin domains. Simultaneously, the K_{off} rate of these interactions decreases at lower temperatures, possibly leading to false assemblies, providing an explanation for the observed behaviour of the mixed protein sample (Kim, Kwak and Kang, 2006).

Due to the long and flexible linker, the two titin domains might not necessarily interact with the same obscurin O58O59 molecule, forming a compact, energetically favoured complex. Instead, they might falsely interact with two distinct obscurin O58O59 molecules, crosslinking those and, therefore, allowing to form a vast interaction meshwork of several titin and obscurin molecules, ultimately leading to precipitation. At lower temperatures, this process is even more favoured as the temperature dependent entropy penalty for forming intermolecular interactions is reduced. However, at lower concentrations, false crosslinking by titin would be reduced as there are not enough obscurin molecules in close proximity, which could explain the observed behaviour.

6.1.2 Architecture of the complex

The asymmetric unit of the crystal revealed two copies of titin Z8Z9 and obscurin O58O59 organised in a ring-like structure. SEC and SEC-MALLS experiments performed on the complex showed a 1:1 dimer interaction of both proteins, ruling out a tetrameric complex as seen in the

asymmetric unit. The 47 amino acid long linker between both titin subunits is not visible in the crystal structure. Even though a degradation of the linker between the titin domains Z8 and Z9 was observable in solution when stored for several days at room temperature, the analysis of the crystal content revealed that the crystal is formed only by intact titin and obscurin, indicating that the linker is protected upon crystal formation. While it has been shown in literature that the linker between two domains can be an essential part of a complex structure (Pernigo *et al.*, 2017), the absence of structural information for the Z8-Z9 linker in the crystal structure, indicates its high flexibility and the lack of involvement in the complex interaction. This conclusion is supported by a recent study using NMR to assess the binding of obscurin and titin (Wright and Letourneau, 2018).

Due to the missing information of the titin Z8-Z9 linker region, two possible domain arrangements of the complex derive from the asymmetric unit. Titin domain Z8 can be assigned to two locations, one in which titin Z8 interacts with obscurin domain O58 and one in which Z8 interacts with obscurin O59. The remaining complex interfaces formed between obscurin domains O58O59 and titin domain Z9 are unambiguously defined showing Z9 positioning itself between both obscurin domains.

6.1.2.1 The Z8 domain of titin interacts with obscurin O58

To determine the correct positioning of titin Z8 and thus reveal the real interaction interfaces of the complex, computational analysis tools of the 'Protein interfaces, surfaces and assemblies' service (PISA) (Krissinel and Henrick, 2007) and biophysical methods such as ITC and SAXS were applied. It has been demonstrated that the interaction between Z8 and O59 is considerably smaller than the Z8-O58 interaction (324 $Å^2$ vs 497 $Å^2$). In addition, with hydrogen bond distances of >3.2 Å, it is also weaker compared with the Z8-O58 interaction (bond distances below 2.95 Å). Furthermore, the lower free energy gain ΔiG calculated for the interface formation and the high ΔiG p-value close to 0.5 indicate that the Z8-O59 interaction is likely an artefact of the crystal packing. ITC experiments, quantifying the interaction between titin Z8 and obscurin upon the introduction of mutations within both interfaces, showed no effect for the titin mutation K1782A affecting the Z8-O59 interface but revealed reproducible effects for all five mutated residues

located within the Z8-O58 interface. Even though the crystal structure indicated that the low-resolution structural differences between both Z8 arrangements are small, SAXS experiments on the complex formed by obscurin O58O59 and the individual titin Z8 domain once more suggested that the primary binding site of titin Z8 to obscurin is located at the subunit O58 of obscurin. Furthermore, the length of the linker between Z8 and Z9, and the fact that its length is highly conserved among different species indicates its pivotal role. The length of the linker is essential to allow the Z8-O58-Z9 arrangement, yet no density has been detected in the crystal structure. This indicates that the linker is not actively involved in the complex interaction and would thus not be essential for the alternative arrangement. Finally, phosphorylation sites found on titin domain Z8 are directly located within the interface of Z8 and O58 and have been shown to disrupt the complex formation in *in vivo* experiments. Taken together, it was possible to show that titin Z8 is interacting with the obscurin domain O58 which allows the conclusion of a Z8-O58-Z9-O59 arrangement of the titin-obscurin Z-disk complex.

6.1.2.2 The complex domain interaction has a hierarchical order

For the validation of the remaining complex interfaces, additional ITC experiments with interface mutants were performed. These assays demonstrated that all mutants involved in the interfaces lead to a decrease of interaction, whereas control mutations outside of any interface do not change the interaction strength compared to the wild-type. This indicates that the complex arrangement as seen in the crystal structure represents the in-solution domain-domain interaction and involves all four protein domains. While this is in agreement with most publications investigating the titin-obscurin Z-disk interaction (Young, Ehler and Gautel, 2001; Arimura *et al.*, 2007; Rossi *et al.*, 2017), it has also been shown that individual titin domains do only interact weakly with obscurin O58O59 (Young, Ehler and Gautel, 2001).

Interestingly, a recent publication presenting NMR data suggests that only titin Z8 is involved in the complex formation, while Z9 rather supports the structural stability of Z8 (Wright and Letourneau, 2018) which is directly contradictory to the previously mentioned publications and the work presented here. Further ITC experiments were performed either addressing multiple interfaces at once by introducing several mutations or using individual titin Z8 or Z9 domains to

evaluate the distinct contribution to the whole complex. It was possible to show that i) the double mutant addressing the Z9-O58 and Z9O59 interfaces shows the same reduction of interaction as the single mutant which only addresses the Z9-O58 interface, indicating a minor role of the Z9O59 interaction within the interaction of the whole complex, ii) the double mutant within the Z8-O58 interface only disrupts the complex formation completely when combined with an additional Z9-O58 interface mutation, showing that these two interfaces are mainly responsible for the complex stability, and iii) even though Z8 appears less stable without Z9, it seems to interact more strongly with obscurin O58O59 than Z9 alone.

These results might indicate a hierarchical relevance of the interfaces for the complex stability. The two interfaces between obscurin O58 and both titin domains appear to be the main interface, synergistically stabilising the complex. The smaller interface between Z9 and O59 seems less important. This is supported by the finding that the sequence identity of obscurin domain O59 in all five species investigated is the lowest of all complex domains, ranging between 77% and 39% and is completely missing in the obscurin sequence in horses even though all other domains are highly conserved (>94%) (chapter 5.1.8).

Additionally, the cross-species sequence comparison revealed that the Z9-O59 interface is the only interface in which amino acid substitutions are located that could disrupt the interface. However, these substitutions are not affecting the native behaviour of the complex, indicating a minor physiological role of this interface. Furthermore, it was possible to show in SAXS experiments that the linker between both domains of obscurin O58O59 remains rigid in solution, which is supported by NMR experiments performed by Hu et al. (2017). This indicates that the involvement of O59 is not essential but rather a side product of the arrangement between titin Z8Z9 and obscurin O58.

6.1.3 Variants within the titin-obscurin complex interface

Due to the dramatic decrease of costs to generate whole genome sequencing data, nowadays, large databases are able to provide enormous amounts of information, allowing for specific search for variants in precisely defined chromosomal regions which are coding for specific proteins. While population databases collect genomic data from broad population cohorts,

disease databases provide information about disease-associated genetic variants. For the titin-obscurin complex, these databases were used to i) identify variants located within the interaction interfaces (Table 36) and ii) assess the impact of variants linked to disease phenotypes (Table 37).

The variant analysis revealed an overall frequency for the domains between 32% (O58) and 47% (Z9), indicating the presence of a known variant for every third to almost every second amino acid, interestingly without a selectivity for or against residues involved in the three complex interfaces. Altogether, 22 variants located within the three interfaces were found. Considering the relatively mild impact of the introduced mutants for ITC measurements (Table 32), most of the variants found are not expected to have a significant impact on the complex formation as they either represent mild residue exchanges or have a mean allele frequency (MAF) too high to affect the interaction. The stability of the complex appears rather robust against single residue substitution even if they fundamentally change the physico-chemical nature of the residue. The complex with its three separate and independent interfaces seems to compensate these local changes. In most cases, the identified variants alone would presumably not be sufficient to disrupt the complex formation. This is supported by the finding that only two of the 22 identified interface variants are also assigned in the disease databases ClinVar (Landrum *et al.*, 2018), namely Y1891C (rs547305291) (Z9) and A4484T (rs116557268) (O59). While Y1891C has only been submitted to the database without a published report, a functional analysis of the obscurin variant A4484T showed neither an impact on the titin Z8-Z9 interaction in mammalian two-hybrid (M2H) and co-immunoprecipitation (co-IP) assays nor an incorporation into the Z-disk (Arimura *et al.*, 2007). This is in agreement with the high MAF of 2.64E-03 that is largely above the threshold set for possibly pathogenic variants.

Possibly, only single modifications that have broader effects, such as disrupting structural motifs (e.g. proline substitutions), weakening of the overall fold stability of domains, or facilitating the interaction with other molecules and thus blocking the interaction interface (e.g. cysteine substitutions) (Betts and Russell, 2007), might interfere with the complex formation. In the complex interface only five residues where found to have the potential to interfere with the complex formation: S1789P, located in titin Z8 and possibly disrupting the β-strand G, which is the main interaction site for the Z8-O58 interface, and the cysteine variants R4396C (O58),

R1882C (Z9), Y1891C (Z9), and Y1905C (Z9). Except for Y1891C, no clinical indication has been found in the ClinVar database (Landrum *et al.*, 2018) for these variants. Y1891C is connected to dilated cardiomyopathy and limb-girdle muscular dystrophy, however, with uncertain clinical significance. Y1891C is involved in the interface between Z9 and O59, being part of the small β-sheet formed by both β-strands D of Z9 and O59. It has already been discussed in the previous section that the Z9-O59 interaction probably does not play a pivotal role for the titin-obscurin complex formation. However, the introduced cysteine in the Y1891C variant could, under oxidising conditions, form disulfide bridges with other molecules and thus not only affect the Z9-O59 binding site but, due to the precise fitting of Z9 between O58 and O59, also affect the interaction between Z9 and O58.

Although it has also been shown that Z8 alone can bind to O58O59 (chapter 5.1.7), this variant might drastically destabilise the *in vivo* complex and lead to the described phenotypes. Interestingly, dilated cardiomyopathy may be induced by different factors like a heart attack or substances such as alcohol or heavy metals (Gavazzi *et al.*, 2000; Hazebroek, Dennert and Heymans, 2012). These toxins could induce oxidative stress in the myocytes, triggering the disulphide bridge formation (Cumming *et al.*, 2004), which in turn can lead to protein aggregation. Even though this mechanism would also apply for the other mentioned interface variants that substitute for a cysteine, these have not been linked to any disease phenotype. However, considering the low MAF (>6.00E-5) of these variants and the mentioned potential involvement of other factors, one cannot exclude a similar effect.

6.1.4 Disease-related variants not involved in the titin-obscurin interaction interface

As could be demonstrated for titin variant Y1891C, structural information can be used to draw conclusions about the impact of disease-related variants found within these structures. Genomic disease databases such as ClinVar (Landrum *et al.*, 2018) enable the structural analyses of clinically described variants linked to specific disease phenotypes. Apart from Y1891C, 25 different variants, affecting the domains of the titin-obscurin Z-disk complex, were found in disease databases, and one (R4444W) in a recently published report (Rossi *et al.*, 2017). Seven of

these variants were directly linked to myopathies, however, with varying clinical significances (Landrum *et al.*, 2018). Interestingly, five of these seven disease-related variants are located within the gene sequence coding for titin Z8 (two variants) and titin Z9 (three variants). While this might indicate the relevance of titin in the development of myopathies, it is important to mention that the titin gene (*TTN*) is, in contrast to obscurin, known to be a major cause for human-inherited muscle diseases and thus has been the target of extensive research over the last years, implicitly increasing the amount of clinical data that can be mapped to known variants (Chauveau, Rowell and Ferreiro, 2014).

This thesis demonstrates that the variant N1755S, located in a loop region of Z8 and pointing to the domains core, might destabilise the domain structure. This could have an influence on the development of dilated cardiomyopathy and limb-girdle muscular dystrophy. However, the destabilising effect would not be expected as severe, supporting the uncertain significance of this variant and a report in which the variant was assigned to an individual without cardiomyopathy, arrhythmia, or any family history of sudden cardiac death (Ng *et al.*, 2013).

Another variant that has been assigned with uncertain significance is the semi-conservative amino acid substitution R1859S, exposed in the β-strand B in between the conserved residue motif AXFXC and not involved or close to known interaction interfaces of titin Z9. The conserved residues are all oriented towards the hydrophobic core and are unlikely to be influenced by this variant. The variant has been found together with a second titin variant (not located within Z8Z9) in connection with sudden cardiac death (Campuzano *et al.*, 2014). Even though, the MAF of 5.41E-05 shows that it is not a common benign variant, from the structural point of view this variant is not expected to have vital impact on the structural stability or the titin-obscurin interaction, nevertheless other yet to be described interactions of this domain could be affected.

As discussed earlier, cysteine substitutions, which introduce exposed cysteines on the surface of domains, might exert pathologically relevant effects on the ability of the domain to interact with other molecules. The variant R1861C, which is associated with dilated cardiomyopathy and limb-girdle muscular dystrophy with uncertain significance (ClinVar database (Landrum *et al.*, 2018)), is not involved or close to the interaction interfaces of Z9. However, it might still affect the

binding between titin and obscurin when the interfaces are blocked due to a disulfide bridge formation with a larger molecule under oxidising conditions (Zhou and Freed, 2004; Betts and Russell, 2007).

Interestingly, the two variants P1744L (Z8) and R1890C (Z9), even though showing a MAF ranging between 6.69E-03 for P1744L and 3.90E-04 for R1890C that are considered as too high for a rare genetic disease (Kobayashi *et al.,* 2017), have been found in numerous patients suffering from conditions such as dilated cardiomyopathy, limb-girdle muscular dystrophy, distal myopathy, or hypertrophic cardiomyopathy (ClinVar database (Landrum *et al.*, 2018)). P1744L is exposed on the loop region between β-strand C and D of the Ig domain structure of titin Z8, a region not involved in any known interaction, and not expected to have a structural impact. R1890C is located within the β-strand of titin Z9 that is involved in forming the Z9-O59 β-sheet interaction, however, not directly involved in the titin-obscurin interaction. The high MAF, in combination with the large amount of different conditions found in connections with these two variants, allows for two scenarios: either the variant is a natural, benign variant, not related to the disease, and was coincidently found together with the disease phenotype, or these variants are causative only in combination with other, yet unknown, variants or factors, as discussed for Y1891C in the previous section.

Most of the variants discussed would not explain the development of a pathogenic phenotype on their own by, for example, disrupting the complex structure or severely destabilising domain integrities. However, these variants might have a role as disease modifier when synergising with other specific variants.

Exemplary for this concept is the variant R4444W which is located in obscurin O59 and has been found in patients affected by distal myopathy. These patients also carry a frameshift mutation in the FLNC gene encoding filamin-C, an actin-binding-like protein located at the Z-disk (Razinia *et al.*, 2012; Rossi *et al.*, 2017). Because the FLNC mutation was also present in one healthy individual not carrying the R4444W variant, the authors concluded that the obscurin mutation might be necessary for the full expression of the myopathic phenotype. These studies additionally report that the obscurin variant is able to decrease the binding between titin Z8Z9 and obscurin.

In contrast, the structural analysis performed in this thesis does not indicate such an effect, as the affected residue is not located near the O59 interaction interface. However, it is possible that the mutation influences the domain stability and thus impairs binding, as has been observed for other mutants created for ITC experiments (chapter 5.1.6). Alternatively, the variant could impair other, yet unknown, interactions of obscurin O59 in the sarcomere.

6.1.4.1 Obscurin variant R4344Q does not show an impact on the titin-obscurin complex formation

Obscurin carrying the variant R4344Q, located in the domain O58, has been studied in several reports investigating the functional impact on the proteins behaviour (Arimura *et al.*, 2007; Hu *et al.*, 2017). R4344Q is one of two obscurin O58O59 variants that have been linked to myopathies. Arimura et al. (2007) demonstrated a conformational change on the surface of O58 due to the R4344Q variant using a molecular modelling approach. Additionally, it was shown in M2H and co-IP assays that the binding between obscurin O58O59 and titin Z8Z9 was decreased. Consequently, Arimura et al. concluded that the R4344Q variant is affecting the recruitment of obscurin at the Z-disk which ultimately results in an impaired localisation of the 'signalling-end' of obscurin at the Z-disk-I-Band junction and the development of hypertrophic cardiomyopathy.

The crystal structure obtained in this thesis shows that, indeed, the residue R4344 is exposed at the domain surface. However, it is not involved in the protein complex interfaces between titin and obscurin. Furthermore, it was neither possible to reproduce the negative effect of the R4344Q variant on the complex interaction using ITC nor to explain the reported decrease of interaction by a general destabilisation of the domains fold determined by thermal shift assays. Additionally, the mean allele frequency of 1.17E-02 for this variant is far too high to be considered a rare variant, indicating that the variant is most probably not involved in the development of the hypertrophic cardiomyopathy phenotype.

Surprisingly, Hu et al. (2017) could also not support the finding that the variant R4344Q is involved in the titin-obscurin binding. Instead, they presented evidence for a gain-of-function mechanism in which obscurin-binding to the Ca^{2+} homeostasis-regulating molecule phospholamban (PLN) is increased by ten. Interestingly, the structure presented in this work

revealed a close proximity between the obscurin interfaces formed with titin Z8 and PLN. Looking at the theoretical localisation of both interactions, however, it is unlikely that they would compete with each other. The titin-obscurin complex forms at the Z-disk-I-band junction and the PLN-obscurin interaction can only take place near the sarcoplasmic reticulum (SR) membrane, as PLN is a transmembrane protein of the SR and would thus only be relevant for obscurin bound at the M-band. However, even though different biophysical methods such as fluorescence anisotropy and microscale thermophoresis were applied, it was not possible to reproduce the reported increased binding of the obscurin variant R4344Q compared with the wild type. This discrepancy could arguably result from the use of different experimental setups, such as *in vitro* versus *in vivo*, and exogenous short fragments versus endogenous full-length proteins. Nevertheless, it does not explain how the high MAF reported for this variant can meaningfully be connected to a rare genetic disease like hypertrophic cardiomyopathy. Even though this work presents strong evidence that the R4344Q variant has no direct effect, neither on the titin nor the PLN interaction, the ambiguous results from other reports make it difficult to draw a final conclusion.

6.1.4.2 Genetic origin of myopathies

In contrast to complex diseases, which are influenced by a combination of multiple genes and environmental factors, it is a common concept that myopathies are mostly mendelian diseases in which a single gene can be causative of a pathogenic phenotype (Arbustini *et al.*, 2000; Maron, 2002; Guglieri *et al.*, 2008; Piluso *et al.*, 2010).

This concept holds true for single genetic variants causing either the abolishment of an entire pivotal interaction (Sauer *et al.*, 2010; Pernigo *et al.*, 2015) or the truncation of a protein, as for example known for many variants in titin (Chauveau, Bonnemann, *et al.*, 2014; Chauveau, Rowell and Ferreiro, 2014). However, in this thesis, it was possible to show that many variants which are linked to myopathies would probably not be able to be causative of a disease phenotype on their own and thus could qualify to be part of a complex disease mechanism. To date, it is debated whether complex diseases are caused by either a huge number of common variants with small

individual effects (infinitesimal model) (Visscher, Hill and Wray, 2008) or 'major-effect' rare variants (rare allele model) (Cirulli and Goldstein, 2010).

The truth may lie between these lines. The sarcomere is one of the most complex structures known in nature, highly regulated with thousands of proteins assembling and interacting with each other. As reviewed for complex diseases in general by Gibson (2012), a combination of rare and common variants, each more or less dysregulating, misaligning or weakening the sarcomeric structure, might only be causative of a disease phenotype, if the individual effects reach a certain threshold. However, identification of these individual contributors and their combinations would be challenging, as correlation does not imply causation, and diseases are often categorical divided into datasets of 'cases' and 'controls', and thus one is prone to falsely assign variants which contribute to the threshold as being benign (Gibson, 2012).

6.1.5 The function and regulation of the titin-obscurin complex

Obscurin and titin interact with each other at different sarcomeric locations, more than 1 µm apart from one another. Interestingly, both complexes seem to be regulated, as they show preferential localisation of obscurin molecules during different phases of myofibrillogenesis and in mature myofibrils (Young, Ehler and Gautel, 2001; Borisov *et al.*, 2004; Kontrogianni-Konstantopoulos and Bloch, 2005; Borisov, Martynova and Russell, 2008; Carlsson, Yu and Thornell, 2008). Notably, the M-band interaction seems to be dominant in mature sarcomeres. Being a giant muscle protein with thousands of amino acids, the regulation of obscurin is highly essential for muscle cells, following the principle: 'use it or degrade and recycle its components'. The localisation pattern and the underlying regulatory mechanism, however, is not yet fully understood (Young, Ehler and Gautel, 2001; Borisov *et al.*, 2004; Kontrogianni-Konstantopoulos and Bloch, 2005; Borisov, Martynova and Russell, 2008; Carlsson, Yu and Thornell, 2008). Interestingly, the comparison of both titin-obscurin complexes reveals a specific difference. The titin-obscurin complex at the Z-disk consists of four domains, forming three defined interaction interfaces. The M-band complex forms by the interaction of one domain of each protein, raising the question of why the Z-disk complex requires four domains.

Remarkably, even though the Z-disk complex has a much larger interface area, both interactions appear rather similar in terms of affinity/avidity, with dissociation constants in the low micromolar range and a Gibbs free energy of 7-8 kcal mol^{-1} (Pernigo *et al.*, 2010; Sauer *et al.*, 2010). It is likely that the combined affinity of all three interfaces is necessary to reach the critical stability for the Z-disk complex to guide the signalling-end of obscurin close to the Z-disk as has been reported (Arimura *et al.*, 2007; Kontrogianni-Konstantopoulos *et al.*, 2009).

The interplay of the four domains also brings the two titin domains, which are separated by a 47 amino acid long linker, close together, shortening the titin molecule up to 14 nm. While the Z-like arrangement of the titin Z8-linker-Z9 construct (see Figure 42.1) would allow such a condensing effect, the proximity to the neighbouring elastic PEVK region of titin (Gautel *et al.*, 1996; Yajima *et al.*, 1996; Linke *et al.*, 2002) and the overall length of more than 1000 nm of titin makes the physiological effect arguable. However, this comparatively small effect might play a role during myofibrillogenesis where titin has been shown to extend from the Z-disk to the M-band successively and is not yet finally aligned (van der Loop *et al.*, 1996; Yang, Obinata and Shimada, 2000).

It is not known whether the two titin domains are located closely together or whether the linker is further extended in mature sarcomeres. If both domains are separated, the tetrameric complex might also function as a sensor for failed sarcomeric integrity: as a result of e.g. overextension, titin could loosen, bringing the two titin domains Z8Z9 in close proximity, which in turn could be recognised by obscurin. This would bring the signalling end again close to the Z-disk and help reorganise the structure, a role of obscurin that is well established (Bowman *et al.*, 2008; Kontrogianni-Konstantopoulos *et al.*, 2009).

Another possible function of the titin-obscurin complex may lie in the subdivision of the complex into interfaces with different affinities. This could potentially allow the complex to be modulated in its overall avidity by regulating individual interfaces. Thus, the localisation or copy number of obscurin molecules at the Z-disk can be finely tuned - a regulation that goes beyond binary ‘on’ and ‘off’. This could explain the observation that in mature myofibrils obscurin localises primarily at the M-band and to a lesser extent at the Z-disk (Borisov *et al.*, 2004; Borisov, Martynova and

Russell, 2008), even though for every M-band interaction there should be a corresponding binding site on titin for obscurin at the Z-disk.

Nevertheless, a model for Z-disk complex regulation has not yet been described in the literature. Interestingly, the localisation of two phosphorylation sites that have been found by collaborators (laboratory of Mathias Gautel (King's College London)) were identified using the crystal structure obtained in this thesis. Both sites are located in the titin domain Z8, one within the interface formed with O58 (T1710/S1711), and the other not involved in the complex interfaces (T1727). The phosphomimetic mutation of residue S1711, as well as the phosphorylated T1727, showed a weakening but no disruption of the complex in ITC and pull-down experiments, indicating that the other two interfaces are able to keep the complex intact.

Due to the lack of other regulatory mechanisms for the Z9-O58 and Z9-O59 interfaces, this could demonstrate the necessity of a composite regulation mechanism, potentially involving i) the separation of the two titin domains, which would disrupt the weaker Z9-O58 and Z9-O59 interaction, and ii) the phosphorylation of the stronger Z8-O58 interface (Figure 42). The separation of the two titin domains might be the result of either a tightening of titin during myofibrillogenesis, as a consequence of the final alignment of titin along the assembling thick filament, or, the lateral alignment and fusion of myofibril (Borisov *et al.*, 2004; Borisov, Martynova and Russell, 2008). In a second step, the remaining Z8-O58 complex would first be further weakened by phosphorylation of the accessible T1727 residue, making residue S1711, which is involved in the interface, more accessible for subsequent phosphorylation and final disruption of the complex. The phosphorylation of residue S1711 would subsequently block the interface and protect Z8 against a new complex formation with obscurin. Interestingly, *in vivo* interaction data with interface mutants, acquired by collaborators (laboratory of Mathias Gautel (King's College London)), also support the idea of separated titin domains in mature sarcomeres, as only mutants that were located within the Z8-O58 interface were affecting the localisation of obscurin.

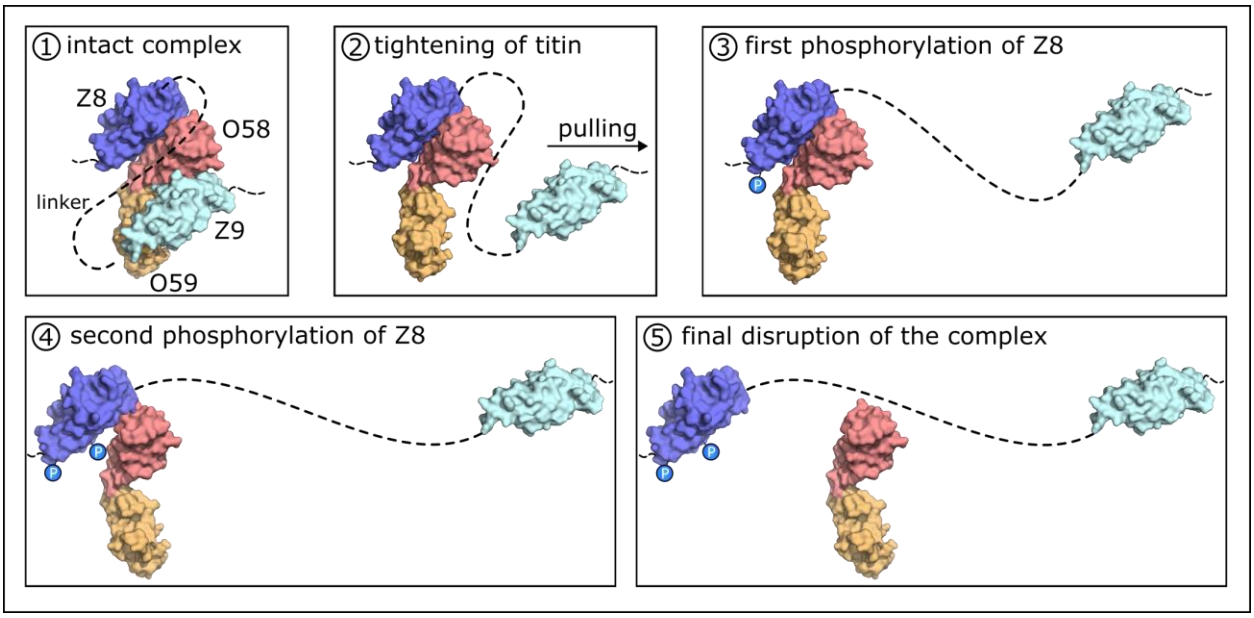

Figure 42 Titin-Obscurin Z-disk complex regulation model. The intact complex (1) gets weakened by the separation of the two titin domains and disruption of the lower affinity interfaces between Z9 and obscurin O58O59 (2) due to a tightening of the titin molecule e.g. during myofibrillogenesis. Weakening of the remaining Z8-O58 interface by phosphorylation of Z8 residue T1727 (3). This allows eventually the phosphorylation of Z8 residue S1711 (or T1710), located within the Z8-O58 interface (4), to ultimately disrupt the titin-obscurin complex (5).

6.2 The myomesin 5 dimer

The sarcomeric M-band cytoskeleton protein myomesin-1 is known to be part of the triangular interaction with titin and obscurin and considered highly important for the cross-linking of the thick filament myosin. By forming a homodimer at the very C-terminal My13 domain (Lange, Himmel, *et al.*, 2005; Pinotsis *et al.*, 2008, 2012), myomesin is able to span between the M4-lines of two horizontally aligned thick filaments, each entering two adjacent sarcomeres (Obermann *et al.*, 1997). Furthermore, it has been shown that the myomesin domain My4 forms a complex with the third domain of obscurin (Pernigo *et al.*, 2017). This complex also shows tendencies to dimerise. Interestingly, an unpublished crystal structure acquired by Sauer *et al.* (EMBL-Hamburg) shows an additional dimerisation side for myomesin domain My5, possibly providing yet another junction within the M-band cross-linking architecture. Remarkably, the dimer formation is driven by an unusual mechanism called three-dimensional domain swapping (Bennett, Schlunegger and Eisenberg, 1995; Heringa and Taylor, 1997). In the following sections, a potential mechanism for the dimerisation, the physiological relevance of the dimer formation and its possible role for the sarcomeric stability will be discussed.

6.2.1 The mechanism of My5 dimerisation

In the process of three-dimensional domain swapping two identical proteins exchange a structural component while keeping the native-like contacts with the complementary portion of the other protein (Rousseau, Schymkowitz and Itzhaki, 2003; Gronenborn, 2009). In addition to the native-like ‘primary’ interface formed upon domain swapping, a ‘secondary’ interface between the hinge regions of the crossing swapped domains can form. The transition between the closed monomer and the domain-swapped dimer/oligomer has been shown to depend on three factors affecting the free energy difference between both states (Kuhlman *et al.*, 2001; Liu *et al.*, 2001; Rousseau *et al.*, 2001; Schymkowitz *et al.*, 2001): i) the monomer is thermodynamically favoured due to the higher entropy of two single versus one complexed molecule, ii) the formation of the ‘secondary’ interface, which is only found in the swapped dimer/oligomer and therefore introduces additional interactions, favours the dimer, and iii) the conformational change within the hinge region can introduce or release strain tension in the

molecule. The hinge region has been described to play a central role in domain swapping mechanisms: its length, stability and residue composition can either hinder or promote domain swapping in molecules (Rousseau, Schymkowitz and Itzhaki, 2012).

Experiments prior to this thesis performed by Sauer, Chatziefthimiou and Temmerman, that aimed to disrupt the complex formation, showed that a disruption was only possible by mutating residues located within the 'primary' interface of the swapped dimer. Mutants within the 'secondary' interface were not affecting the dimer formation. Remarkably, two of the 'primary-interface'-mutants are part of a conserved pattern and buried in the domain fold. The introduction of the charged aspartate (P638D and P641D), therefore, could destabilise the N-terminal insertion and prevent the domain swap, as was observed in SEC and SAXS measurements. The other mutant (D643P) in the N-terminus is exposed and might form an interaction with residue K659, thus, stabilising the N-terminus. An Asp to Pro mutation will abolish this electrostatic interaction and possibly change the N-terminal conformation, hindering the incorporation into the domain fold. This indicates that either the formation of the 'primary' and not the 'secondary' interface is the driving force of the dimer formation, or, that the 'primary' interface needs to form prior to the 'secondary' interface in a 'zipper'-like mechanism (Ozkan *et al.*, 2007), starting at the very N-terminal end and proceeding along the peptide chain in the direction of the C-terminus. At the same time, this suggests that the hinge region favours a more extended conformation, exceeding the thermodynamic contribution of the monomeric form.

Interestingly, as has been shown in chapter 5.2.2.1, SAXS measurements on the monomeric form of My5 wild-type indicate that the 'open' conformation, with the N-terminal part not collapsed into a complete FN-III-fold but being highly exposed, is – unexpectedly – favoured in solution. This 'open' conformation, resulting in an unprotected hydrophobic domain-core and a mostly accessible N-terminus, is thermodynamically unfavourable and thus provides an additional factor that could drive the dimerisation. In contrast to the literature that is considering these exposed conformations as intermediate states between monomers and swapped dimers (Wolynes, Luthey-Schulten and Onuchic, 1996; Yang *et al.*, 2004), the individual My5 wild-type domain does not seem to be able to stably form a completely folded FN-III domain.

The crystal structure of the dimer reveals an interesting and non-conserved, thus unique, residue interplay at the hinge region of the My5 domain. Residue R651 forms a salt bridge with residue D701. Additionally, the hydrophobic residue F700 points towards R651 and could push this residue in the direction of E704, which is also close to the R651-D701 interaction, potentially being able to establish a D701-R651-E704 interaction cluster (Figure 37). Those interactions might stabilise the exposed N-terminus, thus, destabilising the completely folded FN-III domain. Notably, the crystallisation of a monomeric form of My5 was only possible after the mutation of the above-mentioned residue F700 to a serine (Sauer *et al.,* unpublished), underlining the importance of these interactions from non-conserved residues for the distinct behaviour of My5.

The sequence alignment furthermore reveals a conserved residue pattern, mainly contributing to the hydrophobic core of the domain. Interestingly, a substitution of an alanine to cysteine (position 714 in My5), which is unique to My5, makes the sulfur side chain directly face the N-terminus in the completely folded FN-III domain. The cysteine might weaken the packing of the N-terminus within the domain fold. In combination with the induced strain at the hinge region, this leads to a thermodynamic shift towards an open conformation in the monomeric state. The dimerisation *via* domain swapping, therefore, would overcome the unfavourable 'open' conformation comprising a solvent accessible hydrophobic core by re-establishing a native-like interface, while simultaneously keeping the hinge region in a favourable conformation.

6.2.2 The physiological relevance of the My5 dimer

Domain swapping has been described to have functional significance in many cellular processes, such as receptor binding (Louie *et al.,* 1997) or DNA cleavage (Newcomer, 2002). Furthermore, it has been proposed that domain swapping might be an important factor for the evolution of functional dimers (Bennett, Schlunegger and Eisenberg, 1995; Lynch, 2013). However, several known structures show domain swapping only in truncated domains of a complete molecule, whereas the complete molecule itself does not show this behaviour and is only present as a monomer (Ultsch *et al.*, 1999). Therefore, domain swapping might also occur as an artefact and a consequence of the truncation, suggesting that smaller domains undergo domain swapping more easily than larger domains (Ultsch *et al.*, 1999; Liu and Eisenberg, 2002). In this regard it is

important that larger constructs of myomesin, that were purified for electron microscopy experiments (Chatziefthimiou *et al.,* unpublished) do not show My5 induced dimerisation, raising the question of a physiological relevance of this dimer.

Interestingly, the analysis of disease-related variants located in My5 shows that the variants of the two residues R651 and E704 are connected to the development of hypertrophic cardiomyopathy. As shown in the last section, these two residues are potentially involved in the stabilisation of the exposed conformation of the hinge region. The two variants R651Q and E704K could each lead to a loss of this interaction, reducing the strain in the hinge region and thus stabilising the monomeric form. This could, indeed, indicate a physiological significance for the ability of the N-terminus to get exposed and potentially undergo domain swapping with a second My5 domain. However, as literature is proposing (Ultsch *et al.*, 1999; Liu and Eisenberg, 2002), it is possible that the N-terminus of My5 is only permanently exposed in the truncated version of the individual My5 domain and, due to stabilising effects of the surrounding polypeptide chain, not accessible in the whole myomesin structure. However, myomesin is a muscle scaffolding protein and thus subjected to additional forces not present in solution. Consequently, an absence of My5 mediated dimerisation in solution for larger myomesin constructs is not necessarily indicative of an absence in endogenous proteins. It has been demonstrated earlier that the incorporation of the N-terminus in the FN-III-fold is potentially weaker compared with other FN-III domains. This implies that the N-terminus detaches from the domain core more easily upon application of forces along the myomesin filament under physiological conditions, transforming the My5 domain to an open conformation. This can explain why a dimerisation could not be observed in long myomesin constructs that are in-solution and thus relaxed.

Even though domain swapping can require the complete unfolding of the protein (Landrieu *et al.,* 2001), an opening of the domains structure into an incompletely folded state has also been described to be sufficient (Miller and Marqusee, 2011; Kang *et al.,* 2012). Therefore, the domain swap of My5 could occur spontaneously during the incorporation/spanning of myomesin into the M-band or under the condition of external tension and is not restricted to the timepoint of initial protein folding. This could also involve a theoretical force-induced regulation of the My5 homo-

dimerisation, analogous to what has been described for the titin kinase which shows a tension-induced opening of its catalytic site (Mayans *et al.,* 1998; Gräter *et al.*, 2005; Puchner *et al.*, 2008).

6.2.3 Putative structural role of the My5 dimer

The current cross-linking model of myomesin involves not only the N-terminal interaction with myosin (Obermann *et al.*, 1997) and the C-terminal homodimerisation (Lange, Himmel, *et al.*, 2005; Pinotsis *et al.*, 2008, 2012) but also the possibility of a homodimerisation of two My4-Obs3 complexes (Pernigo *et al.*, 2017). Interestingly, the later complex introduces a new geometric constraint and a relatively sharp bending of myomesin to the previous M-band model (Lange, Himmel, *et al.*, 2005) which in turn proposed an almost diagonal spanning of myomesin. This additional interaction could, besides the anchoring of obscurin(-like-1) to myomesin, also be important for the regulation of the vertical distance between parallel myosin filaments. However, little is known about the structural stability of the homodimer formed by two My4-Obs3 complexes. The complex not only would have to withstand forces originating from the obscurin molecule but would also be exposed to forces coming from longitudinal and transversal movement of the adjacent and parallel myosin filaments (Figure 43). Notably, the stability of a complex can be highly dependent on the direction of the applied forces. This has impressively been demonstrated for the interaction between the staphylococcal adhesin SdrG and the human fibrinogen β, which form a complex that can withstand forces of almost 2.5 nN, whereas non-native pulling resulted in a 40-fold stability decrease to 60 pN (Milles *et al.*, 2018).

The high complexity of forces applied to the homodimer formed by two My4-Obs3 complexes could favour the involvement of an additional interaction, such as the dimerisation of My5, supporting the stability and preserving the integrity of the myomesin junction. Additionally, the homodimer of the My4-Obs3 complex only forms at high concentrations (Pernigo *et al.*, 2017). The My5 self-assembly-point being close to My4 could, hence, also bring two My4-Obs3 complexes together and thus facilitate the My4 dimerisation in the first place.

Structurally, both complexes could coexist without interfering with each other, preserving the overall polypeptide chain direction (Figure 43 B). The two interaction points might each stabilise the junction upon different directional forces. This additional junction would ultimately connect

two different myomesin dimers, which are C-terminally linked, and lead to a stiffer arrangement of the myomesin molecules and thus a better force distribution between the two M4 M-lines and a more rigid M-band stabilisation (Figure 43 A). Due to the theoretical length of a myomesin C-terminally connected dimer of approximately 100 nm (Berkemeier *et al.*, 2011) and the M4-M4 distance of only 44 nm, the absence of this junction could lead to vertical and horizontal misalignment of the M-band and the myosin filaments (Figure 43 C). Forces originating from myosin filaments, which antiparallelly extend into two adjacent sarcomere-halves, would not disperse but concentrate at the symmetry interface located at the M1 M-line.

It is not known how elastic the thick filament within the M-band region is. The elastic property of myomesin that has been postulated to be able to extend by approximately 50 nm allowing reversible deformation of the sarcomeric M-band (Berkemeier *et al.,* 2011; Pinotsis *et al.,* 2012), however, would only be able to exploit its whole potential in the absence of the My4-My5 induced junction. This supports the idea that the My5 complex formation is induced upon stretching forces, potentially also facilitating the My4 homodimerisation, and subsequently realigning the M-band organisation after extensive muscle contraction and M-band misalignment.

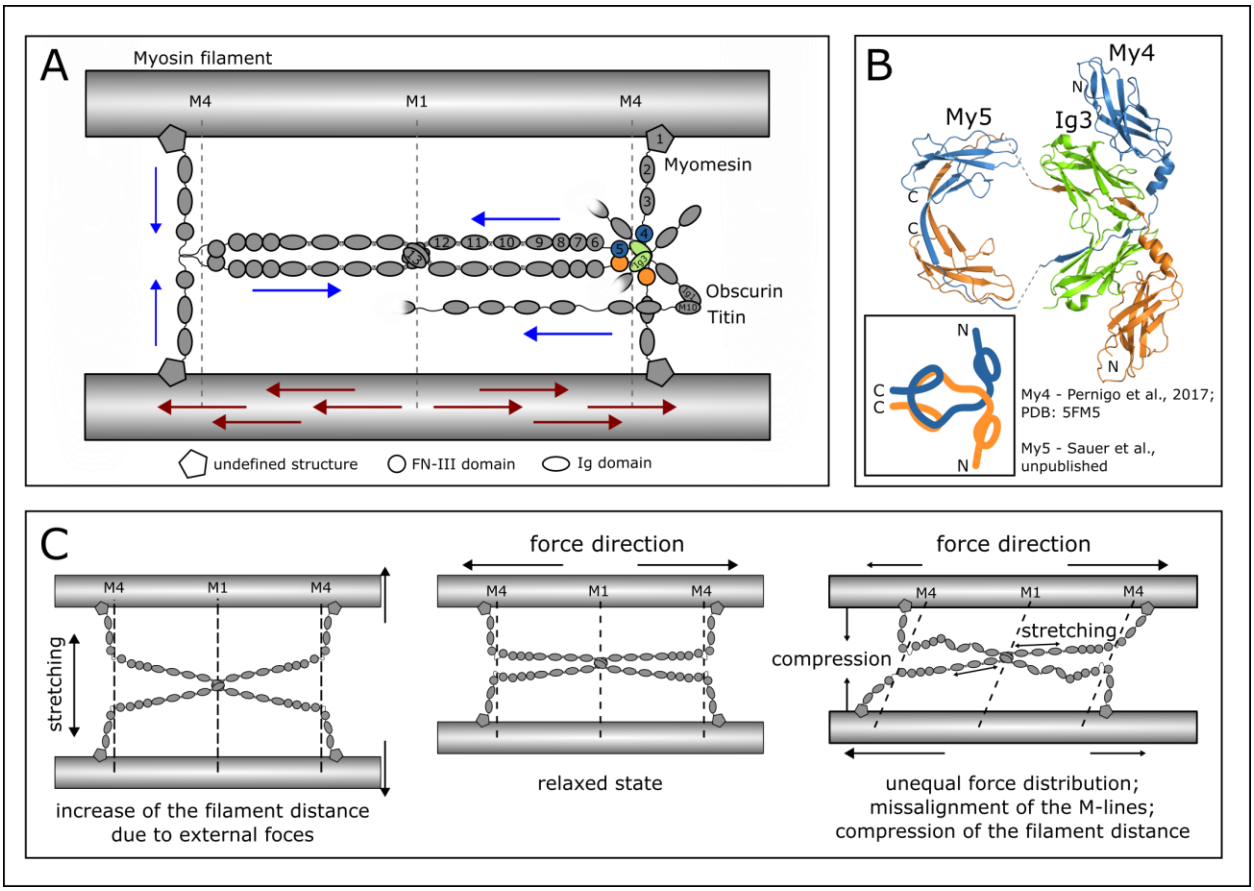

Figure 43 Proposed Myomesin architecture at the M-band including the My5 dimer. A) Arrangement of myomesin upon My5 dimer formation, introducing a second tightly packed junction close to the (My4-Obs3)$_2$ dimer. The additional My4-My5 junction induces stable arrangement of all four involved myomesin filaments. Active forces (red arrows) and passive forces (blue arrows) are distributed along the whole M4-M1-M4 region. B) Arrangement of the My5 homodimer (Sauer et al., unpublished) and the (My4-Obs3)$_2$ dimer (Pernigo *et al.*, 2017). *A simplified illustration of the protein sequence (box) shows the entanglement between both myomesin chains. C) Theoretical consequences of the M-band arrangement in the absence of the additional My4-My5 junction (arrows indicate directional forces). A vertical widening (left) or a horizontal misalignment and vertical compression (right) could destabilise the M-band structure.*

It was possible to show that the My5 domain-swapped dimer not only is an artefact of the truncational approach of My5 crystallisation but can also form under physiological conditions and even play a meaningful role for the M-band stability. Two possible models illustrating the function of the My5 dimer within the myomesin filament and the M-band structure can therefore be proposed.

In the first model, the My5 dimer forms upon incorporation of myomesin into the M-band structure due to induced strain on the N-terminus. Thus, it facilitates the dimerisation of two My4-Obs3 complexes by bringing them in close proximity and forms a tightly packed myomesin

dimer junction in the middle of two parallel myosin filaments. In this model, the interaction and thus the arrangement, introducing additional stiffness and stability to the M-band region, are permanent. In the second model, the N-terminus of My5 functions as a strain sensor as it is only exposed and able to undergo domain swapping upon pulling forces along the myomesin filament during muscle contraction, providing an additional structural support for the My4-Obs3 junction. However, whether the My4-obscurin interaction and the self-assembly of My5 are cross-regulating each other and able to coexist remains elusive. In any case, a stable junction around the My4-My5 domains would be required to keep the M-band perfectly aligned in all three dimensions.

Combining all published structures with unpublished work performed by Sauer and Chatziefthimiou et al. (EMBL Hamburg) it was possible to create the most complete model to date of the myomesin structure (Figure 44). The model incorporates the crystal packing from the crystallisation of My11-My13 (Pinotsis *et al.*, 2012) indicating a possible dimer of dimers around the domain My13. It very well represents and supports the myomesin architecture and M-band arrangement that was proposed in Figure 43. Notably, both the $(My4\text{-}Obs3)_2$ dimer and the My5 dimer simultaneously fit into the model. The missing structural information of domains My2, My3 and My6 was substituted by dummy domains with the expected measurements. The distances measured in this model are slightly higher than described by Obermann et al. (1996), with a distance between M1 M-line and My4-My5 of approximately 25 nm (opposed to approx. 20 nm (Obermann *et al.*, 1996)), probably due to the crystal packing or an imperfectly placed My6 arrangement.

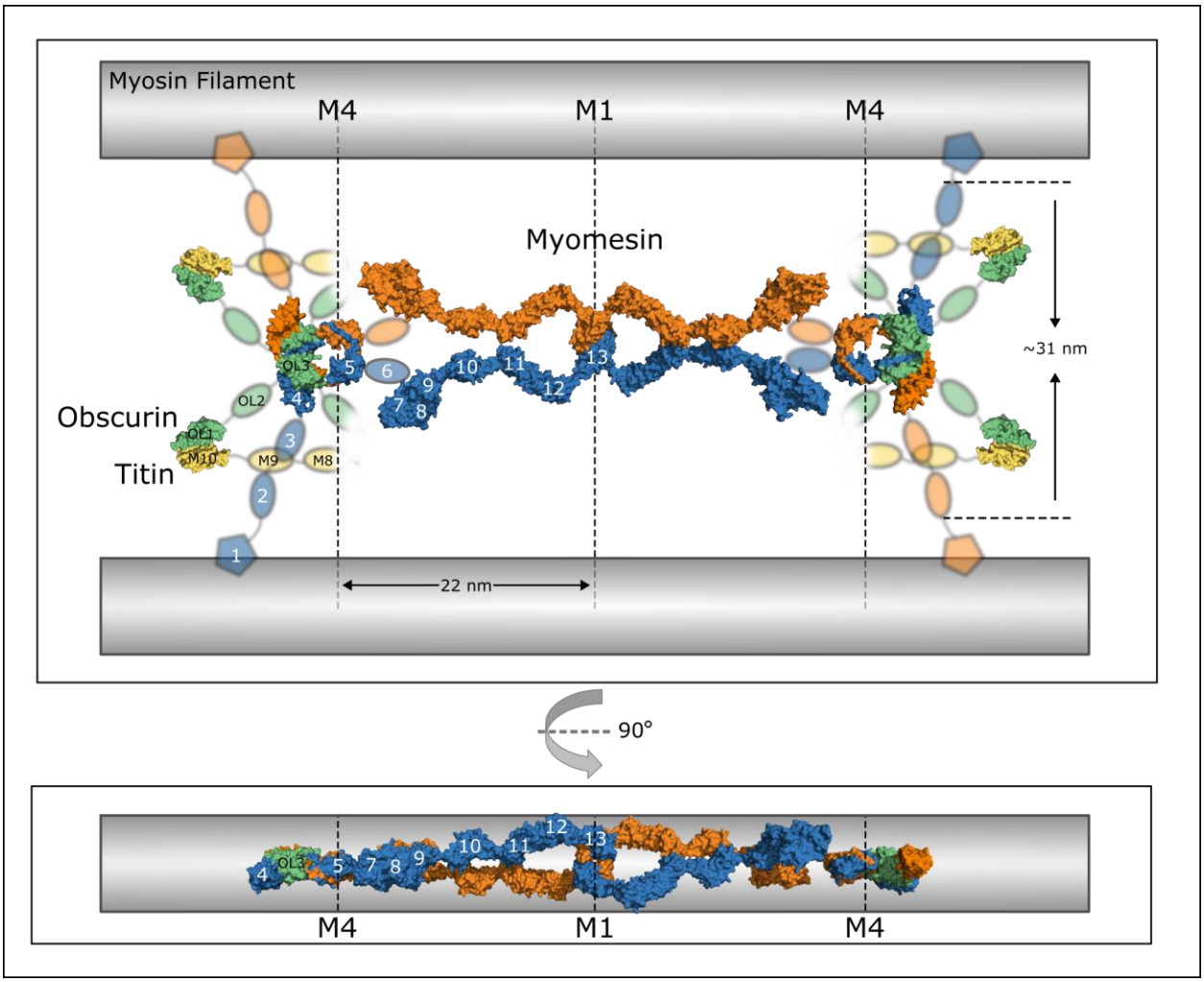

Figure 44 Composite model of all available high-resolution structures of myomesin. Crystal packing of My13 indicates a tetrameric arrangement at the M1 M-line formed by two dimers. Arrangement of domains My7 to My13 shows that these domains are not aligned in a straight way. My7-My9 are closely packed. The missing structural information of domains My2, My3 and My6 was substituted by dummy domains with the expected dimensions. Titin-obscurin complex (PDB: 3knb) (Sauer et al., 2010), My4-Ol3 (PDB: 5fm5) (Pernigo et al., 2017), My9-My13 (PDB: 2y23 and 2y25) (Pinotsis et al., 2012), My7My8 (Chatziefthimiou et al., unpublished), My5 (Sauer et al., unpublished).

7 Summary and future perspectives

The fundamental functions of muscles are enabling flexibility and dynamics on the one hand while providing stability on the other. This contradictory principle of motion and rigidity needs to be finely regulated, precisely maintained, and reliably stabilised and thus requires highly ordered sub-cellular structures. Numerous proteins, such as titin, obscurin and myomesin, are part of this multitasking network contributing to the muscles function.

This thesis provides the very first comprehensive model of the titin-obscurin Z-disk interaction, presenting for the first time a complete high-resolution structure of this complex, proposing attractive hypotheses for the involvement of four instead of two domains in the complex formation, and offering a novel model for the regulatory mechanism of the complex. In the second part, it is demonstrated that the domain swapping dimerisation of myomesin domain My5 has a physiological relevance and can play an important role for the sarcomeric stability. Furthermore, this thesis presents a composite model of all available high-resolution structures of myomesin, representing the most complete structural model of myomesin to date, and further supporting the proposed myomesin M-band arrangement.

The giant muscle protein titin spans half of the sarcomere guiding numerous proteins, such as the sarcomeric filament protein obscurin, to their individual location on the one hand, and functioning as a molecular spring, stabilising the sarcomeric structure, on the other. Previous studies revealed two separate interaction sites between titin and obscurin: one at the M-band between the very N-terminal Ig domain Ig1 of obscurin and the very C-terminal domain M10 of titin (Fukuzawa *et al.*, 2008), and one at the Z-disk between obscurin O58-O59 and titin Z8-Z9 (Young, Ehler and Gautel, 2001). The interaction at the M-band has already been structurally determined and functionally described (Pernigo *et al.*, 2010, 2015; Sauer *et al.*, 2010). In contrast, the complete molecular structure of the titin-obscurin interaction at the Z-disk remains elusive. The localisation of obscurin within the sarcomere changes during maturation of the myofibril indicating a regulation of the Z-disk complex, however, without any proposed mechanism reported so far. Moreover, several reports about variants located within the complex domains

and linked myopathies indicate an involvement of the Z-disk complex development pathogenic phenotypes.

The aim of the first part of this thesis was to crystallise and structurally describe the titin-obscurin Z-disk complex in order to allow a detailed analysis of the involved domain interfaces and reported variants, and draw conclusions about the complex behaviour, function and regulation.

It was possible to solve the high-resolution structure of the complete titin-obscurin complex that is located at the Z-disk. The structure reveals a coordinated domain interplay of titin domains Z8 and Z9 and the obscurin domains O58 and O59 and thus confirms previous reports proposing a tetrameric complex formation (Young, Ehler and Gautel, 2001; Arimura *et al.*, 2007; Rossi *et al.*, 2017). The obtained structural information, ITC and SAXS experiments demonstrate that the domains form contacts between titin Z8 and obscurin O58, between obscurin O58 and titin Z9, and between titin Z9 and obscurin O59, whereas the 47 amino acid long linker between titin Z8 and Z9 was not involved in the binding between these two molecules. Furthermore, it was possible to reveal a hierarchical order of the individual interfaces for the complex stability, indicating that the interaction formed between obscurin O58 and both titin domains is of major importance – primarily due to the more stable Z8-O58 interaction, whereas the interaction induced by obscurin domain O59 seems to play a minor role.

The existence of three complex interfaces opposed to one interface, which is found in the second interaction between titin and obscurin located at the sarcomeric M-band, with, however, a similar complex stability, led to four proposed hypotheses of the functional meaning of this extended interface:

(1) **Complex stability:** The combined affinity of all three interfaces is needed to reach the critical stability for the Z-disk complex, necessary to stably guide the signalling end of obscurin close to the Z-disk.

(2) **Condensation of the length of titin:** The tetrameric complex brings the two titin domains close together, shortening the titin molecule by 14 nm potentially being important for the alignment of titin in the sarcomere.

(3) **Sensor for the sarcomeric integrity of titin:** It is not known whether the distance between the two titin domains is changing throughout the sarcomeric lifetime, extending the linker and disabling the formation of the tetrameric complex seen in the crystal structure. In case a separation of both domains is of physiological relevance, for example, due to a straightening of titin during myofibrillogenesis, a loosening of the titin molecule could bring both domains back into close proximity. This could re-establish the complex formation with obscurin and imply that the tetrameric complex might function as a strain-sensor for failed sarcomeric integrity of titin.

(4) **Multi-step regulation of the complex:** *In vivo* experiments indicated the absence of the interfaces created by titin domain Z9 in mature muscle cells. In combination with the presence of phosphorylation sites only affecting the Z8-O58 interface and the modular organisation of the titin-obscurin complex being subdivided into interfaces with different affinities, an attractive model for the complex regulation was proposed. In a first step, the two titin domains are separated by tightening of the titin molecule, disrupting the interfaces between obscurin and titin Z9. In a second step, the remaining interface between titin Z8 and obscurin O58 is initially weakened by the phosphorylation of the phosphorylation-site not involved in the interface, which consequently enables the phosphorylation of the second site located within the remaining interface, ultimately leading to a disruption of the titin-obscurin interaction (Figure 42). This novel regulatory mechanism for the titin-obscurin Z-disk complex facilitates a finely tuned obscurin concentration at the Z-disk, going beyond a mere on or off state. Thus, it is an attractive model to explain how the cell is able to effectively locate and spacio-temporally direct the giant obscurin molecule, which is due to its sheer size highly expensive for the cell to maintain.

To test these hypotheses, further experiments are necessary. This includes investigating the distance between both titin domains during the sarcomeric lifetime, identifying the kinase responsible for the second phosphorylation step and subsequently evaluating the localisation of obscurin in kinase knockout mutants.

The ability of the complex to compensate for single amino acid substitutions within the interfaces, as has been demonstrated in ITC experiments, was also reflected in the analysis of natural variants located within the domain structures. Several variants that are linked to the

development of myopathies were structurally analysed. None of the analysed variants was expected to have a significant impact on the domains behaviour. This included the disease-related variant R4344Q, located on obscurin O58, that has previously been connected to the development of hypertrophic cardiomyopathy and investigated in several studies. One exception were variants that entail an amino acid substitution to a cysteine, like the variant R1861C, that could potentially form unexpected interfaces with other molecules under oxidising conditions and therefore block native interfaces. Additionally, many of these variants did show a mean allele frequency too high to be solely causative for rare genetic diseases of mendelian origin. Consequently, it was questioned whether the common concept that most myopathies are Mendelian diseases can still hold on. With the sarcomere being one of the most complex structures known, finely regulated with thousands of proteins interacting, a combination of rare and common variants, each more or less dysregulating, misaligning or weakening the sarcomeric construct, might also be causative for a particular disease phenotype. This demonstrates, that even though structural data help to gain insight into the potential role of variants, it is necessary to carefully evaluate their biological impact by complementary biophysical, biochemical and functional approaches.

Myomesin is located at the M-band and a major contributor to the cross-linking of the thick filament by connecting two myosin filaments with each other forming a C-terminal homo-dimer. Elastic properties at the C-terminal region of myomesin enable the molecule to extend, possibly allowing reversible deformation of the sarcomeric M-band. An unpublished high-resolution structure suggests a homo-dimerisation of the fifth domain of myomesin, with an unusual mechanism called three-dimensional domain swapping.

In the second part of the thesis, the previously observed homodimerisation of the fifth domain of the M-band scaffolding protein myomesin was investigated regarding the domain swapping dimerisation mechanism and its physiological relevance. It was possible to show that the native monomeric fold of the My5 domain was most certainly destabilised by two My5 specific structural features, leading to an exposed N-terminus. First, the hinge region, which connects the central part of the domain with the exchanged N-terminal region, contains specific residue interactions that are stabilising an open domain conformation. Variants linked to myopathies for

some of these residues further support their essential role for the My5 domain. Second, the incorporation of the N-terminus might be further destabilised by a unique cysteine located in the hydrophobic core able to reduce the packing of the N-terminus. This induces a thermodynamic shift towards an open conformation with an exposed hydrophobic core. By the formation of a domain-swapped dimer, the hinge region can be kept in a favourable conformation while simultaneously 'closing' the domain fold, re-establishing a native-like interface, and additionally further stabilising the dimer with a new secondary interface between the hinge regions of both swapped domains.

Even though the permanently exposed N-terminus, that was also validated by SAXS measurements, and thus the domain-swapped dimer might be an artefact of the truncated protein, the mentioned features of My5 indicate a My5 specific destabilisation of the domain region. A model for the endogenous protein was proposed in which the N-terminus is exposed, and thus domain swapping is induced upon strain application on the myomesin molecule. This could happen during the incorporation of the molecule into the M-band structure, or as a strain sensing mechanism upon muscle contraction. In both cases, a new self-assembly point in the myomesin structure would be created, close to the My4-Obs3 dimer that has also been shown to additionally homodimerise in previous studies. Both complexes could potentially support each other by forming a tight junction and increasing the stability of the myomesin induced myosin cross-linking to withstand multi-directional forces (Figure 43).

Notably, the elastic region of the myomesin molecule located at the C-terminal section of myomesin, however, would only be able to exploit its whole potential in the absence of the My4-My5 induced junction. This indicates that the My5 dimerisation is only induced by stretching forces during myomesin incorporation in the M-band or extensive muscle contraction which might help to realign the M-band organisation. Furthermore, a structural overview, representing the most complete structural model of myomesin to date by combining all published and yet to be published myomesin structures, illustrates that the proposed tetrameric myomesin structure is in agreement with the expected myomesin domain distances (Figure 44).

To complete the structural overview, it is necessary to obtain the missing structures preferably as multidomain structures to receive information about their respective orientation. Furthermore, it is essential to evaluate the role of the My5 homodimer further. This can be achieved *via in vivo* experiments, overexpressing the My5 domain to compete with endogenous dimerisation. A disruption of the M-band or even sarcomeric structure would further indicate a physiological role of this dimerisation. Additionally, AFM experiments will provide insight into the structural stability of the complex upon force application. Finally, the relationship between the My5 dimerisation and the My4-Obs3 dimerisation needs to be evaluated, to address the questions whether both complexes can exist simultaneously, or whether they are cross-regulating each other.

Overall, this work provides strong evidence that muscle related diseases are not necessarily of Mendelian origin but could, in some cases, be considered complex diseases. This concept questions the prevalent stance to search for 'the one' mutant responsible for the development of a myopathy and might change the general accepted view on myopathies. Moreover, this thesis demonstrates that the behaviour of muscle proteins needs to be carefully interpreted under the consideration of their endogenous environment, being exposed to external forces that are not present in solution. In both presented cases, external forces applied to the proteins could have an impact on their fundamental behaviour and regulating their function. It was shown that the structural information is key to further understand the proteins nature and to shed light on regulatory mechanisms. However, only the combination of structural, biophysical, clinical, and statistical methods, allows a comprehensive assessment of the physiological role of proteins and their involvement in pathogenic mechanisms.

Bibliography

Ackermann, M. A. *et al.* (2014) 'Obscurins: Goliaths and Davids take over non-muscle tissues', *PLoS ONE*. Edited by F. Cappello. Public Library of Science, 9(2), p. e88162. doi: 10.1371/journal.pone.0088162.

Adams, P. D. *et al.* (2010) 'PHENIX: A comprehensive Python-based system for macromolecular structure solution', *Acta Crystallographica Section D: Biological Crystallography*. International Union of Crystallography, 66(2), pp. 213–221. doi: 10.1107/S0907444909052925.

Adams, R. J. and Pollard, T. D. (1986) 'Propulsion of organelles isolated from Acanthamoeba along actin filaments by myosin-I', *Nature*. Nature Publishing Group, 322(6081), pp. 754–756. doi: 10.1038/322754a0.

Agarkova, I. *et al.* (2003) 'M-band: a safeguard for sarcomere stability?', *Journal of muscle research and cell motility*, 24(2–3), pp. 191–203. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14609030 (Accessed: 2 July 2018).

Agarkova, I. *et al.* (2004) 'The molecular composition of the sarcomeric M-band correlates with muscle fiber type', *European Journal of Cell Biology*. Urban & Fischer, 83(5), pp. 193–204. doi: 10.1078/0171-9335-00383.

Agarkova, I. and Perriard, J.-C. (2005) 'The M-band: an elastic web that crosslinks thick filaments in the center of the sarcomere', *Trends in Cell Biology*. Elsevier Current Trends, 15(9), pp. 477–485. doi: 10.1016/J.TCB.2005.07.001.

Akiyama, N. *et al.* (2006) 'Transverse Stiffness of Myofibrils of Skeletal and Cardiac Muscles Studied by Atomic Force Microscopy', *The Journal of Physiological Sciences*. PHYSIOLOGICAL SOCIETY OF JAPAN, 56(2), pp. 145–151. doi: 10.2170/physiolsci.RP003205.

Arbustini, E. *et al.* (2000) 'Familial dilated cardiomyopathy: from clinical presentation to molecular genetics', *European Heart Journal*. Oxford University Press, 21(22), pp. 1825–1832. doi: 10.1053/euhj.2000.2173.

Arimura, T. *et al.* (2007) 'Structural analysis of obscurin gene in hypertrophic cardiomyopathy', *Biochemical and Biophysical Research Communications*, 362(2), pp. 281–287. doi: 10.1016/j.bbrc.2007.07.183.

Bagnato, P. *et al.* (2003) 'Binding of an ankyrin-1 isoform to obscurin suggests a molecular link between the sarcoplasmic reticulum and myofibrils in striated muscles.', *The Journal of cell biology*. Rockefeller University Press, 160(2), pp. 245–53. doi: 10.1083/jcb.200208109.

Bang, M. L. *et al.* (2001) 'The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system.', *Circulation research*, 89(11), pp. 1065–72. doi: 10.1161/hh2301.100981.

Barnes, M. R. *et al.* (1999) 'A lipid-binding domain in Wnt: a case of mistaken identity?', *Current*

biology : CB. Elsevier, 9(19), pp. R717-9. doi: 10.1016/S0960-9822(99)80465-5.

Benian, G. M. *et al.* (1989) 'Sequence of an unusually large protein implicated in regulation of myosin activity in C. elegans', *Nature*. Nature Publishing Group, 342(6245), pp. 45–50. doi: 10.1038/342045a0.

Benian, G. M. *et al.* (1996) 'The Caenorhabditis elegans gene unc-89, required fpr muscle M-line assembly, encodes a giant modular protein composed of Ig and signal transduction domains.', *The Journal of cell biology*. Rockefeller University Press, 132(5), pp. 835–48. doi: 10.1083/JCB.132.5.835.

Benian, G. M. and Mayans, O. (2015) 'Titin and obscurin: Giants holding hands and discovery of a new Ig domain subset', *Journal of Molecular Biology*. Elsevier Ltd, 427(4), pp. 707–714. doi: 10.1016/j.jmb.2014.12.017.

Bennett, M. J., Schlunegger, M. P. and Eisenberg, D. (1995) *3D domain swapping: A mechanism for oligomer assembly*, *Protein Science*. Wiley-Blackwell. doi: 10.1002/pro.5560041202.

Berkemeier, F. *et al.* (2011) 'Fast-folding alpha-helices as reversible strain absorbers in the muscle protein myomesin.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 108(34), pp. 14139–44. doi: 10.1073/pnas.1105734108.

Betts, M. J. and Russell, R. B. (2007) 'Amino-Acid Properties and Consequences of Substitutions', in *Bioinformatics for Geneticists: A Bioinformatics Primer for the Analysis of Genetic Data: Second Edition*. Chichester, UK: John Wiley & Sons, Ltd, pp. 311–342. doi: 10.1002/9780470059180.ch13.

Blanchet, C. E. *et al.* (2015) 'Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY)', *Journal of Applied Crystallography*, 48(2), pp. 431–443. doi: 10.1107/S160057671500254X.

Blausen.com staff (2014) 'Medical gallery of Blausen Medical 2014', *WikiJournal of Medicine 1 (2):10*. doi: 10.15347/wjm/2014.010. ISSN 2002-4436.

Borisov, A. B. *et al.* (2004) 'Dynamics of Obscurin Localization During Differentiation and Remodeling of Cardiac Myocytes: Obscurin as an Integrator of Myofibrillar Structure', *The Journal of Histochemistry & Journal of Histochemistry & Cytochemistry*, 52(9), pp. 1117–1127. doi: 10.1369/jhc.3A6183.2004.

Borisov, A. B. *et al.* (2006) 'Essential role of obscurin in cardiac myofibrillogenesis and hypertrophic response: Evidence from small interfering RNA-mediated gene silencing', *Histochemistry and Cell Biology*. Springer-Verlag, 125(3), pp. 227–238. doi: 10.1007/s00418-005-0069-x.

Borisov, A. B., Martynova, M. G. and Russell, M. W. (2008) 'Early incorporation of obscurin into nascent sarcomeres: implication for myofibril assembly during cardiac myogenesis', *Histochemistry and Cell Biology*. Springer-Verlag, 129(4), pp. 463–478. doi: 10.1007/s00418-008-0378-y.

Bork, P., Holm, L. and Sander, C. (1994) 'The Immunoglobulin Fold: Structural Classification,

Sequence Patterns and Common Core', *Journal of Molecular Biology*. Academic Press, 242(4), pp. 309–320. doi: 10.1006/JMBI.1994.1582.

Bowman, A. L. *et al.* (2007) 'Different obscurin isoforms localize to distinct sites at sarcomeres', *FEBS Letters*. Wiley-Blackwell, 581(8), pp. 1549–1554. doi: 10.1016/j.febslet.2007.03.011.

Bowman, A. L. *et al.* (2008) 'The Rho-Guanine Nucleotide Exchange Factor Domain of Obscurin Regulates Assembly of Titin at the Z-Disk through Interactions with Ran Binding Protein 9', *Molecular Biology of the Cell*. Edited by M. B. Omary, 19(9), pp. 3782–3792. doi: 10.1091/mbc.e08-03-0237.

Burkart, C. *et al.* (2007) 'Modular Proteins from the Drosophila sallimus (sls) Gene and their Expression in Muscles with Different Extensibility', *Journal of Molecular Biology*. Academic Press, 367(4), pp. 953–969. doi: 10.1016/J.JMB.2007.01.059.

Campuzano, O. *et al.* (2014) 'The role of clinical, genetic and segregation evaluation in sudden infant death', *Forensic Science International*. Elsevier, 242, pp. 9–15. doi: 10.1016/J.FORSCIINT.2014.06.007.

Carlsson, L., Yu, J.-G. and Thornell, L.-E. (2008) 'New aspects of obscurin in human striated muscles', *Histochemistry and Cell Biology*. Springer-Verlag, 130(1), pp. 91–103. doi: 10.1007/s00418-008-0413-z.

Chauveau, C., Bönnemann, C. G., *et al.* (2014) 'G.P.34', *Neuromuscular Disorders*. Elsevier, 24(9–10), pp. 804–805. doi: 10.1016/j.nmd.2014.06.048.

Chauveau, C., Bonnemann, C. G., *et al.* (2014) 'Recessive TTN truncating mutations define novel forms of core myopathy with heart disease', *Human Molecular Genetics*, 23(4), pp. 980–991. doi: 10.1093/hmg/ddt494.

Chauveau, C., Rowell, J. and Ferreiro, A. (2014) 'A rising titan: TTN review and mutation update', *Human Mutation*, 35(9), pp. 1046–1059. doi: 10.1002/humu.22611.

Chen, V. B. *et al.* (2010) 'MolProbity : all-atom structure validation for macromolecular crystallography', *Acta Crystallographica Section D Biological Crystallography*. International Union of Crystallography, 66(1), pp. 12–21. doi: 10.1107/S0907444909042073.

Cirulli, E. T. and Goldstein, D. B. (2010) 'Uncovering the roles of rare variants in common disease through whole-genome sequencing', *Nature Reviews Genetics*. Nature Publishing Group, 11(6), pp. 415–425. doi: 10.1038/nrg2779.

Clark, K. A. *et al.* (2002) 'Striated Muscle Cytoarchitecture: An Intricate Web of Form and Function', *Annual Review of Cell and Developmental Biology*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA , 18(1), pp. 637–706. doi: 10.1146/annurev.cellbio.18.012502.105840.

Coronel, R., de Groot, J. . and van Lieshout, J. . (2001) 'Defining heart failure', *Cardiovascular Research*. Oxford University Press, 50(3), pp. 419–422. doi: 10.1016/S0008-6363(01)00284-X.

Cumming, R. C. *et al.* (2004) 'Protein Disulfide Bond Formation in the Cytoplasm during Oxidative

Stress', *Journal of Biological Chemistry*, 279(21), pp. 21749–21758. doi: 10.1074/jbc.M312267200.

Danowski, B. A. *et al.* (1992) 'Costameres are sites of force transmission to the substratum in adult rat cardiomyocytes.', *The Journal of cell biology*, 118(6), pp. 1411–20. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1522115 (Accessed: 29 October 2018).

Dlugosz, A. A. *et al.* (1984) 'The relationship between stress fiber-like structures and nascent myofibrils in cultured cardiac myocytes.', *The Journal of cell biology*, 99(6), pp. 2268–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6438115 (Accessed: 28 October 2018).

Du, A. *et al.* (2003) 'Myofibrillogenesis in the first cardiomyocytes formed from isolated quail precardiac mesoderm'. doi: 10.1016/S0012-1606(03)00104-0.

Ebashi, S., Endo, M. and Ohtsuki, I. (1969) 'Control of muscle contraction', *Quarterly Reviews of Biophysics*. Cambridge University Press, 2(04), p. 351. doi: 10.1017/S0033583500001190.

Ehler, E. *et al.* (1999) 'Myofibrillogenesis in the developing chicken heart: assembly of Z-disk, M-line and the thick filaments.', *Journal of cell science*, 112 (Pt 10), pp. 1529–39. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10212147 (Accessed: 28 October 2018).

Ehler, E., Fowler, V. M. and Perriard, J.-C. (2004) 'Myofibrillogenesis in the developing chicken heart: Role of actin isoforms and of the pointed end actin capping protein tropomodulin during thin filament assembly', *Developmental Dynamics*. Wiley-Blackwell, 229(4), pp. 745–755. doi: 10.1002/dvdy.10482.

Ehler, E. and Gautel, M. (2008) 'The sarcomere and sarcomerogenesis', *Advances in Experimental Medicine and Biology*, 642, pp. 1–14. doi: 10.1007/978-0-387-84847-1_1.

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

Ervasti, J. M. (2003) 'Costameres: the Achilles' Heel of Herculean Muscle', *Journal of Biological Chemistry*, 278(16), pp. 13591–13594. doi: 10.1074/jbc.R200021200.

Evilä, A. *et al.* (2016) 'Targeted next-generation sequencing assay for detection of mutations in primary myopathies', *Neuromuscular Disorders*. Elsevier, 26(1), pp. 7–15. doi: 10.1016/J.NMD.2015.10.003.

Exome Aggregation Consortium *et al.* (2015) 'Analysis of protein-coding genetic variation in 60,706 humans', *bioRxiv*. Cold Spring Harbor Laboratory, p. 030338. doi: 10.1101/030338.

Fattorusso, R. *et al.* (1999) 'NMR structure of the human oncofoetal fibronectin ED-B domain, a specific marker for angiogenesis.', *Structure (London, England : 1993)*, 7(4), pp. 381–90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10196121 (Accessed: 4 November 2018).

Frank, D. *et al.* (2006) 'The sarcomeric Z-disc: a nodal point in signalling and disease', *Journal of Molecular Medicine*, 84(6), pp. 446–468. doi: 10.1007/s00109-005-0033-1.

Franke, D. *et al.* (2017) 'ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering

from macromolecular solutions', *Journal of Applied Crystallography*. International Union of Crystallography, 50(4), pp. 1212–1225. doi: 10.1107/S1600576717007786.

Franke, D. and Svergun, D. I. (2009) 'DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering', *Journal of Applied Crystallography*. International Union of Crystallography, 42(2), pp. 342–346. doi: 10.1107/S0021889809000338.

Freiburg, A. *et al.* (2000) 'Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity', *Circulation Research*. American Heart Association, Inc., 86(11), pp. 1114–1121. doi: 10.1161/01.RES.86.11.1114.

Freiburg, A. and Gautel, M. (1996) 'A Molecular Map of the Interactions between Titin and Myosin-Binding Protein C. Implications for Sarcomeric Assembly in Familial Hypertrophic Cardiomyopathy', *European Journal of Biochemistry*. Wiley/Blackwell (10.1111), 235(1–2), pp. 317–323. doi: 10.1111/j.1432-1033.1996.00317.x.

Fukuzawa, A. *et al.* (2008) 'Interactions with titin and myomesin target obscurin and obscurin-like 1 to the M-band: implications for hereditary myopathies', *Journal of cell science*, 121(Pt 11), pp. 1841–1851. doi: 10.1242/jcs.028019.

Fukuzawa, A., Idowu, S. and Gautel, M. (2005) 'Complete human gene structure of obscurin: Implications for isoform generation by differential splicing', in *Journal of Muscle Research and Cell Motility*, pp. 427–434. doi: 10.1007/s10974-005-9025-6.

Fürst, D. O. *et al.* (1988) 'The organization of titin filaments in the half-sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: a map of ten nonrepetitive epitopes starting at the Z line extends close to the M line.', *The Journal of cell biology*. Rockefeller University Press, 106(5), pp. 1563–72. doi: 10.1083/JCB.106.5.1563.

Fürst, D. O. and Gautel, M. (1995) 'The anatomy of a molecular giant: how the sarcomere cytoskeleton is assembled from immunoglobulin superfamily molecules.', *Journal of molecular and cellular cardiology*. Elsevier, 27(4), pp. 951–9. doi: 10.1016/0022-2828(95)90064-0.

Gautel, M. *et al.* (1996) 'The central Z-disk region of titin is assembled from a novel repeat in variable copy numbers.', *Journal of cell science*, 109 (Pt 1, pp. 2747–54. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8937992 (Accessed: 31 October 2018).

Gavazzi, A. *et al.* (2000) 'Alcohol abuse and dilated cardiomyopathy in men', *The American Journal of Cardiology*. Excerpta Medica, 85(9), pp. 1114–1118. doi: 10.1016/S0002-9149(00)00706-2.

Gerull, B. *et al.* (2002) 'Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy.', *Nature genetics*. Nature Publishing Group, 30(2), pp. 201–4. doi: 10.1038/ng815.

Gibson, G. (2012) 'Rare and common variants: twenty arguments', *Nature Reviews Genetics*. Nature Publishing Group, 13(2), pp. 135–145. doi: 10.1038/nrg3118.

Graewert, M. A. *et al.* (2015) 'Automated pipeline for purification, biophysical and X-ray analysis

of biomacromolecular solutions', *Scientific Reports*. Nature Publishing Group, 5(1), p. 10734. doi: 10.1038/srep10734.

Gräter, F. *et al.* (2005) 'Mechanically Induced Titin Kinase Activation Studied by Force-Probe Molecular Dynamics Simulations', *Biophysical Journal*. Cell Press, 88(2), pp. 790–804. doi: 10.1529/BIOPHYSJ.104.052423.

Greenfield, N. J. (2007) 'Using circular dichroism spectra to estimate protein secondary structure', *Nature Protocols*. NIH Public Access, 1(6), pp. 2876–2890. doi: 10.1038/nprot.2006.202.

Gronenborn, A. M. (2009) 'Protein acrobatics in pairs--dimerization via domain swapping.', *Current opinion in structural biology*. NIH Public Access, 19(1), pp. 39–49. doi: 10.1016/j.sbi.2008.12.002.

Grove, B. K. *et al.* (1984) 'A new 185,000-dalton skeletal muscle protein detected by monoclonal antibodies.', *The Journal of cell biology*. Rockefeller University Press, 98(2), pp. 518–24. doi: 10.1083/JCB.98.2.518.

Guergueltcheva, V. *et al.* (2011) *Distal myopathy with upper limb predominance caused by filamin C haploinsufficiency*. Available at: http://www.hgvs.org/mutnomen (Accessed: 24 November 2018).

Guglieri, M. *et al.* (2008) 'Limb–girdle muscular dystrophies', *Current Opinion in Neurology*, 21(5), pp. 576–584. doi: 10.1097/WCO.0b013e32830efdc2.

Hackman, P. *et al.* (2008) 'Truncating mutations in C-terminal titin may cause more severe tibial muscular dystrophy (TMD).', *Neuromuscular disorders : NMD*. Elsevier, 18(12), pp. 922–8. doi: 10.1016/j.nmd.2008.07.010.

Harpaz, Y. and Chothia, C. (1994) 'Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains.', *Journal of molecular biology*, pp. 528–539. doi: 10.1006/jmbi.1994.1312.

Hazebroek, M., Dennert, R. and Heymans, S. (2012) 'Idiopathic dilated cardiomyopathy: possible triggers and treatment strategies.', *Netherlands heart journal : monthly journal of the Netherlands Society of Cardiology and the Netherlands Heart Foundation*. Springer, 20(7–8), pp. 332–5. doi: 10.1007/s12471-012-0285-7.

He, Y.-M. and Ma, B.-G. (2016) 'Abundance and Temperature Dependency of Protein-Protein Interaction Revealed by Interface Structure Analysis and Stability Evolution', *Scientific Reports*. Nature Publishing Group, 6(1), p. 26737. doi: 10.1038/srep26737.

Heringa, J. and Taylor, W. R. (1997) 'Three-dimensional domain duplication, swapping and stealing', *Current Opinion in Structural Biology*. doi: 10.1016/S0959-440X(97)80060-7.

Higgins, D. G. *et al.* (1994) 'The evolution of titin and related giant muscle proteins', *Journal of Molecular Evolution*. Springer-Verlag, 38(4), pp. 395–404. doi: 10.1007/BF00163156.

Hinits, Y. and Hughes, S. M. (2007) 'Mef2s are required for thick filament formation in nascent

muscle fibres', *Development*, 134(13), pp. 2511–2519. doi: 10.1242/dev.007088.

Holtzer, H. *et al.* (1997) 'Independent Assembly of 1.6 .MU.m Long Bipolar MHC Filaments and I-Z-I Bodies.', *Cell Structure and Function*. Japan Society for Cell Biology, 22(1), pp. 83–93. doi: 10.1247/csf.22.83.

Horowits, R. *et al.* (1986) 'A physiological role for titin and nebulin in skeletal muscle', *Nature*. Nature Publishing Group, 323(6084), pp. 160–164. doi: 10.1038/323160a0.

Hu, L. R. *et al.* (2017) 'Deregulated Ca2+ cycling underlies the development of hypertrophic cardiomyopathy due to mutant obscurins', *Science Advances*, 3(6), p. e1603081. doi: DOI: 10.1126/sciadv.1603081.

Hu, L. Y. R. and Kontrogianni-Konstantopoulos, A. (2013) 'The kinase domains of obscurin interact with intercellular adhesion proteins', *FASEB Journal*, 27(5), pp. 2001–2012. doi: 10.1096/fj.12-221317.

Huebsch, K. A. *et al.* (2005) 'Mdm muscular dystrophy: Interactions with calpain 3 and a novel functional role for titin's N2A domain', *Human Molecular Genetics*. Oxford University Press, 14(19), pp. 2801–2811. doi: 10.1093/hmg/ddi313.

Huxley, A. F. and Niedergerke, R. (1954) 'Structural Changes in Muscle During Contraction: Interference Microscopy of Living Muscle Fibres', *Nature*. Nature Publishing Group, 173(4412), pp. 971–973. doi: 10.1038/173971a0.

Huxley, H. and Hanson, J. (1954) 'Changes in the Cross-Striations of Muscle during Contraction and Stretch and their Structural Interpretation', *Nature*. Nature Publishing Group, 173(4412), pp. 973–976. doi: 10.1038/173973a0.

Jeffrey, G. A. (1997) *An introduction to hydrogen bonding*. Oxford University Press. Available at: https://books.google.de/books/about/An_introduction_to_hydrogen_bonding.html?id=ZRAFifo37QsC&redir_esc=y (Accessed: 12 October 2018).

Kabsch, W. (2010) 'Integration, scaling, space-group assignment and post-refinement', *Acta Crystallographica Section D: Biological Crystallography*, 66(2), pp. 133–144. doi: 10.1107/S0907444909047374.

Kang, X. *et al.* (2012) 'Foldon unfolding mediates the interconversion between Mpro-C monomer and 3D domain-swapped dimer', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1205241109.

Karplus, P. A. and Diederichs, K. (2012) 'Linking crystallographic model and data quality', *Science*. American Association for the Advancement of Science, 336(6084), pp. 1030–1033. doi: 10.1126/science.1218231.

Karplus, P. A. and Diederichs, K. (2015) 'Assessing and maximizing data quality in macromolecular crystallography', *Current Opinion in Structural Biology*. Elsevier Current Trends, pp. 60–68. doi: 10.1016/j.sbi.2015.07.003.

Kim, D., Kwak, Y.-G. and Kang, S. H. (2006) 'Real-time observation of temperature-dependent

protein–protein interactions using real-time dual-color detection system', *Analytica Chimica Acta*, 577(2), pp. 163–170. doi: 10.1016/j.aca.2006.06.047.

Kobayashi, Y. *et al.* (2017) 'Pathogenic variant burden in the ExAC database: an empirical approach to evaluating population data for clinical variant interpretation.', *Genome medicine*. BioMed Central, 9(1), p. 13. doi: 10.1186/s13073-017-0403-7.

Konarev, P. V. *et al.* (2003) 'PRIMUS: A Windows PC-based system for small-angle scattering data analysis', *Journal of Applied Crystallography*. International Union of Crystallography, 36(5), pp. 1277–1282. doi: 10.1107/S0021889803012779.

Konhilas, J. P., Irving, T. C. and de Tombe, P. P. (2002) 'Length-dependent activation in three striated muscle types of the rat.', *The Journal of physiology*. Wiley-Blackwell, 544(Pt 1), pp. 225–36. doi: 10.1113/JPHYSIOL.2002.024505.

Kontrogianni-Konstantopoulos, A. *et al.* (2003) 'Obscurin Is a Ligand for Small Ankyrin 1 in Skeletal Muscle', *Molecular Biology of the Cell*. Edited by G. Guidotti, 14(3), pp. 1138–1148. doi: 10.1091/mbc.e02-07-0411.

Kontrogianni-Konstantopoulos, A. (2006) 'Obscurin modulates the assembly and organization of sarcomeres and the sarcoplasmic reticulum', *The FASEB Journal*, 20(12), pp. 2102–2111. doi: 10.1096/fj.06-5761com.

Kontrogianni-Konstantopoulos, A. *et al.* (2009) 'Muscle giants: Molecular Scaffolds in Sarcomerogenesis', *Physiological Reviews*, 89(4), pp. 1217–67. doi: 10.1152/physrev.00017.2009.

Kontrogianni-Konstantopoulos, A. and Bloch, R. J. (2005) 'Obscurin: A multitasking muscle giant', in *Journal of Muscle Research and Cell Motility*. Kluwer Academic Publishers, pp. 419–426. doi: 10.1007/s10974-005-9024-7.

Kozin, M. B. and Svergun, D. I. (2001) 'Automated matching of high- and low-resolution structural models', *Journal of Applied Crystallography*. International Union of Crystallography, 34(1), pp. 33–41. doi: 10.1107/S0021889800014126.

Krissinel, E. and Henrick, K. (2007) 'Inference of Macromolecular Assemblies from Crystalline State', *Journal of Molecular Biology*. Academic Press, 372(3), pp. 774–797. doi: 10.1016/J.JMB.2007.05.022.

Kuhlman, B. *et al.* (2001) 'Conversion of monomeric protein L to an obligate dimer by computational protein design.', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.181354398.

Labeit, S. *et al.* (1992) 'Towards a molecular understanding of titin.', *The EMBO journal*. European Molecular Biology Organization, 11(5), pp. 1711–6. doi: S0022283696904416 [pii].

Labeit, S. and Kolmerer, B. (1995) 'Titins: Giant Proteins in Charge of Muscle Ultrastructure and Elasticity', *Science*, 270(5234), pp. 293–296. doi: 10.1126/science.270.5234.293.

Labeit, S., Kolmerer, B. and Linke, W. A. (1997) 'The giant protein titin: Emerging roles in

physiology and pathophysiology', *Circulation Research*, pp. 290–294. doi: 10.1161/01.RES.80.2.290.

Laddach, A. *et al.* (2017) 'TITINdb—a computational tool to assess titin's role as a disease gene', *Bioinformatics*. Edited by A. Valencia. Oxford University Press, 33(21), pp. 3482–3485. doi: 10.1093/bioinformatics/btx424.

Landrieu, I. *et al.* (2001) 'p13SUC1and the WW domain of PIN1 bind to the same phosphothreonine-proline epitope', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M006420200.

Landrum, M. J. *et al.* (2018) 'ClinVar: improving access to variant interpretations and supporting evidence', *Nucleic Acids Research*, 46(D1), pp. D1062–D1067. doi: 10.1093/nar/gkx1153.

Lange, S., Xiang, F., *et al.* (2005) 'Cell biology: The kinase domain of titin controls muscle gene expression and protein turnover', *Science*. American Association for the Advancement of Science, 308(5728), pp. 1599–1603. doi: 10.1126/science.1110463.

Lange, S., Himmel, M., *et al.* (2005) 'Dimerisation of Myomesin: Implications for the Structure of the Sarcomeric M-band', *Journal of Molecular Biology*. Academic Press, 345(2), pp. 289–298. doi: 10.1016/J.JMB.2004.10.040.

Lange, S., Ehler, E. and Gautel, M. (2006) 'From A to Z and back? Multicompartment proteins in the sarcomere', *Trends in Cell Biology*, pp. 11–18. doi: 10.1016/j.tcb.2005.11.007.

Leahy, D. J., Aukhil, I. and Erickson, H. P. (1996) '2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region.', *Cell*, 84(1), pp. 155–64. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8548820 (Accessed: 4 November 2018).

Lin, Z. *et al.* (1994) 'Sequential appearance of muscle-specific proteins in myoblasts as a function of time after cell division: Evidence for a conserved myoblast differentiation program in skeletal muscle', *Cell Motility and the Cytoskeleton*, 29(1), pp. 1–19. doi: 10.1002/cm.970290102.

Linke, W. A. *et al.* (1997) 'Actin-titin interaction in cardiac myofibrils: probing a physiological role.', *Biophysical journal*. Elsevier, 73(2), pp. 905–19. doi: 10.1016/S0006-3495(97)78123-2.

Linke, W. A. *et al.* (2002) 'PEVK domain of titin: An entropic spring with actin-binding properties', in *Journal of Structural Biology*. Academic Press, pp. 194–205. doi: 10.1006/jsbi.2002.4468.

Liu, Y. *et al.* (2001) 'A domain-swapped RNase A dimer with implications for amyloid formation', *Nature Structural Biology*. doi: 10.1038/84941.

Liu, Y. and Eisenberg, D. (2002) '3D domain swapping: As domains continue to swap', *Protein Science*. Wiley-Blackwell, 11(6), pp. 1285–1299. doi: 10.1110/ps.0201402.

van der Loop, F. T. *et al.* (1996) *Integration of titin into the sarcomeres of cultured differentiating human skeletal muscle cells.*, *European journal of cell biology*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8741211 (Accessed: 28 September 2017).

Louie, G. V. *et al.* (1997) 'Crystal structure of the complex of diphtheria toxin with an extracellular

fragment of its receptor', *Molecular Cell*. doi: 10.1016/S1097-2765(00)80008-8.

Luther, P. K. (2009) 'The vertebrate muscle Z-disc: sarcomere anchor for structure and signalling', *Journal of Muscle Research and Cell Motility*. Springer Netherlands, 30(5–6), pp. 171–185. doi: 10.1007/s10974-009-9189-6.

Luther, P. and Squire, J. (1978) 'Three-dimensional structure of the vertebrate muscle M-region', *Journal of Molecular Biology*. Academic Press, 125(3), pp. 313–324. doi: 10.1016/0022-2836(78)90405-9.

Lynch, M. (2013) 'Evolutionary diversification of the multimeric states of proteins', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1310980110.

Lyon, R. C. *et al.* (2015) 'Mechanotransduction in Cardiac Hypertrophy and Failure', *Circulation Research*, 116(8), pp. 1462–1476. doi: 10.1161/CIRCRESAHA.116.304937.

Maron, B. J. (2002) 'Hypertrophic Cardiomyopathy', *JAMA*. American Medical Association, 287(10), pp. 1308–1320. doi: 10.1001/jama.287.10.1308.

Marston, S. *et al.* (2015) 'OBSCN mutations associated with dilated cardiomyopathy and haploinsufficiency', *PLoS ONE*, 10(9), pp. 1–10. doi: 10.1371/journal.pone.0138568.

Maruyama, K. *et al.* (1987) 'Binding of Actin Filaments to Connectin', *J.Biochem.* Oxford University Press, pp. 1339–1346. doi: 10.1093/oxfordjournals.jbchem.a122001.

Mayans, O. *et al.* (1998) 'Structural basis for activation of the titin kinase domain during myofibrillogenesis', *Nature*. Nature Publishing Group, 395(6705), pp. 863–869. doi: 10.1038/27603.

McCoy, A. J. *et al.* (2007) 'Phaser crystallographic software', *Journal of Applied Crystallography*. International Union of Crystallography, 40(4), pp. 658–674. doi: 10.1107/S0021889807021206.

Micsonai, A. *et al.* (2018) 'BeStSel: A web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra', *Nucleic Acids Research*. Oxford University Press, 46(W1), pp. W315–W322. doi: 10.1093/nar/gky497.

Miller, G. *et al.* (2003) 'A targeted deletion of the C-terminal end of titin, including the titin kinase domain, impairs myofibrillogenesis.', *Journal of cell science*. The Company of Biologists Ltd, 116(Pt 23), pp. 4811–9. doi: 10.1242/jcs.00768.

Miller, K. H. and Marqusee, S. (2011) 'Propensity for C-terminal domain swapping correlates with increased regional flexibility in the C-terminus of RNase A', *Protein Science*. doi: 10.1002/pro.708.

Milles, L. F. *et al.* (2018) 'Molecular mechanism of extreme mechanostability in a pathogen adhesin.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 359(6383), pp. 1527–1533. doi: 10.1126/science.aar2094.

Mozaffarian, D. *et al.* (2016) 'Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association.', *Circulation*. Lippincott Williams and Wilkins, 133(4), pp. e38-360. doi: 10.1161/CIR.0000000000000350.

Newcomer, M. E. (2002) 'Protein folding and three-dimensional domain swapping: A strained relationship?', *Current Opinion in Structural Biology*. doi: 10.1016/S0959-440X(02)00288-9.

Ng, D. *et al.* (2013) 'Interpreting Secondary Cardiac Disease Variants in an Exome Cohort', *Circulation: Cardiovascular Genetics*, 6(4), pp. 337–346. doi: 10.1161/CIRCGENETICS.113.000039.

Obermann, W. M. *et al.* (1996) 'The structure of the sarcomeric M band: localization of defined domains of myomesin, M-protein, and the 250-kD carboxy-terminal region of titin by immunoelectron microscopy.', *The Journal of cell biology*. Rockefeller University Press, 134(6), pp. 1441–53. doi: 10.1083/JCB.134.6.1441.

Obermann, W. M. *et al.* (1997) 'Molecular structure of the sarcomeric M band: mapping of titin and myosin binding domains in myomesin and the identification of a potential regulatory phosphorylation site in myomesin.', *The EMBO journal*. EMBO Press, 16(2), pp. 211–20. doi: 10.1093/emboj/16.2.211.

Ojima, K. *et al.* (1999) 'Initiation and maturation of I-Z-I bodies in the growth tips of transfected myotubes.', *Journal of cell science*, 112 (Pt 22), pp. 4101–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10547369 (Accessed: 5 November 2018).

Okagaki, T. *et al.* (1993) 'The major myosin-binding domain of skeletal muscle MyBP-C (C protein) resides in the COOH-terminal, immunoglobulin C2 motif.', *The Journal of cell biology*. Rockefeller University Press, 123(3), pp. 619–26. doi: 10.1083/JCB.123.3.619.

Olivetti, G. *et al.* (1996) 'Aging, cardiac hypertrophy and ischemic cardiomyopathy do not affect the proportion of mononucleated and multinucleated myocytes in the human heart', *Journal of Molecular and Cellular Cardiology*. Elsevier, 28(7), pp. 1463–1477. doi: 10.1006/jmcc.1996.0137.

Ozkan, S. B. *et al.* (2007) *Protein folding by zipping and assembly*, *PNAS July*. Available at: http://cacs.usc.edu/education/cs653/Okzan-ZipAssemble-PNAS07.pdf (Accessed: 5 December 2018).

Pantoliano, M. W. *et al.* (2001) 'High-Density Miniaturized Thermal Shift Assays as a General Strategy for Drug Discovery', *Journal of Biomolecular Screening*. Sage PublicationsSage CA: Thousand Oaks, CA, 6(6), pp. 429–440. doi: 10.1177/108705710100600609.

Pask, H. T. *et al.* (1994) 'M-band structure, M-bridge interactions and contraction speed in vertebrate cardiac muscles', *Journal of Muscle Research and Cell Motility*. Kluwer Academic Publishers, 15(6), pp. 633–645. doi: 10.1007/BF00121071.

Pernigo, S. *et al.* (2010) 'Structural insight into M-band assembly and mechanics from the titin-obscurin-like-1 complex.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(7), pp. 2908–13. doi: 10.1073/pnas.0913736107.

Pernigo, S. *et al.* (2015) 'The crystal structure of the human titin:obscurin complex reveals a conserved yet specific muscle M-band zipper module', *Journal of Molecular Biology*. Academic Press, 427(4), pp. 718–736. doi: 10.1016/j.jmb.2014.11.019.

Pernigo, S. *et al.* (2017) 'Binding of Myomesin to Obscurin-Like-1 at the Muscle M-Band Provides

a Strategy for Isoform-Specific Mechanical Protection', *Structure*. Elsevier Ltd., 25(1), pp. 107–120. doi: 10.1016/j.str.2016.11.015.

Perry, N. A. *et al.* (2013) 'Obscurins: Unassuming giants enter the spotlight', *IUBMB Life*. Wiley-Blackwell, 65(6), pp. 479–486. doi: 10.1002/iub.1157.

Person, V. *et al.* (2000) 'Antisense oligonucleotide experiments elucidate the essential role of titin in sarcomerogenesis in adult rat cardiomyocytes in long-term culture', *Journal of cell science*, 113 Pt 21, pp. 3851–3859. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11034912 (Accessed: 28 October 2018).

Piluso, G. *et al.* (2010) 'Mendelian bases of myopathies, cardiomyopathies, and neuromyopathies.', *Acta myologica : myopathies and cardiomyopathies : official journal of the Mediterranean Society of Myology*. Pacini Editore, 29(1), pp. 1–20. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22029103 (Accessed: 26 November 2018).

Pinotsis, N. *et al.* (2006) 'Evidence for a dimeric assembly of two titin/telethonin complexes induced by the telethonin C-terminus', *Journal of Structural Biology*. Academic Press, 155(2), pp. 239–250. doi: 10.1016/j.jsb.2006.03.028.

Pinotsis, N. *et al.* (2008) 'Molecular basis of the C-terminal tail-to-tail assembly of the sarcomeric filament protein myomesin', *EMBO Journal*. Wiley-Blackwell, 27(1), pp. 253–264. doi: 10.1038/sj.emboj.7601944.

Pinotsis, N. *et al.* (2012) 'Superhelical Architecture of the Myosin Filament-Linking Protein Myomesin with Unusual Elastic Properties', *PLoS Biology*. Edited by G. A. Petsko. Public Library of Science, 10(2), p. e1001261. doi: 10.1371/journal.pbio.1001261.

Pollazzon, M. *et al.* (2010) 'The first Italian family with tibial muscular dystrophy caused by a novel titin mutation', *Journal of Neurology*. Springer-Verlag, 257(4), pp. 575–579. doi: 10.1007/s00415-009-5372-3.

Privalov, P. L. (1990) 'Cold Denaturation of Protein', *Critical Reviews in Biochemistry and Molecular Biology*, 25(5), pp. 281–306. doi: 10.3109/10409239009090613.

Puchner, E. M. *et al.* (2008) 'Mechanoenzymatics of titin kinase', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 105(36), pp. 13385–13390. doi: 10.1073/pnas.0805034105.

Raeker, M. Ö. *et al.* (2006) 'Obscurin is required for the lateral alignment of striated myofibrils in zebrafish', *Developmental Dynamics*. Wiley-Blackwell, 235(8), pp. 2018–2029. doi: 10.1002/dvdy.20812.

Randazzo, D. *et al.* (2013) 'Obscurin is required for ankyrinB-dependent dystrophin localization and sarcolemma integrity.', *The Journal of cell biology*. The Rockefeller University Press, 200(4), pp. 523–36. doi: 10.1083/jcb.201205118.

Raynaud, F. *et al.* (2005) 'Calpain 1-titin interactions concentrate calpain 1 in the Z-band edges and in the N2-line region within the skeletal myofibril', *FEBS Journal*. Wiley/Blackwell (10.1111),

272(10), pp. 2578–2590. doi: 10.1111/j.1742-4658.2005.04683.x.

Raynaud, F., Astier, C. and Benyamin, Y. (2004) 'Evidence for a direct but sequential binding of titin to tropomyosin and actin filaments', *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. Elsevier, 1700(2), pp. 171–178. doi: 10.1016/J.BBAPAP.2004.05.001.

Razinia, Z. *et al.* (2012) 'Filamins in Mechanosensing and Signaling', *Annual Review of Biophysics*, 41(1), pp. 227–246. doi: 10.1146/annurev-biophys-050511-102252.

Reddy, A S., Day, I. S. (2001) 'Analysis of the myosins encoded in the recently completed Arabidopsis thaliana genome sequence.', *Genome biology*, 2(7), p. RESEARCH0024. doi: 10.1046/j.1440-0952.2003.01004.x.

Reedy, M. K., Holmes, K. C. and Tregear, R. T. (1965) 'Induced changes in orientation of the cross-bridges of glycerinated insect flight muscle', *Nature*. Nature Publishing Group, 207(5003), pp. 1276–1280. doi: 10.1038/2071276a0.

Remaut, H. and Waksman, G. (2006) 'Protein-protein interaction through beta-strand addition', *Trends in Biochemical Sciences*. Elsevier Current Trends, 31(8), pp. 436–444. doi: 10.1016/j.tibs.2006.06.007.

Rhee, D., Sanger, J. M. and Sanger, J. W. (1994) 'The premyofibril: Evidence for its role in myofibrillogenesis', *Cell Motility and the Cytoskeleton*, 28(1), pp. 1–24. doi: 10.1002/cm.970280102.

Ribeiro, E. de A. *et al.* (2014) 'The structure and regulation of human muscle α-actinin.', *Cell*. Elsevier, 159(6), pp. 1447–60. doi: 10.1016/j.cell.2014.10.056.

Rodriguez, P. and Kranias, E. G. (2005) 'Phospholamban: a key determinant of cardiac function and dysfunction.', *Archives des maladies du coeur et des vaisseaux*, 98(12), pp. 1239–43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16435604 (Accessed: 11 September 2018).

Rossi, D. *et al.* (2017) 'A novel FLNC frameshift and an OBSCN variant in a family with distal muscular dystrophy', *PLoS ONE*, 12(10). doi: 10.1371/journal.pone.0186642.

Rousseau, F. *et al.* (2001) 'Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.101542098.

Rousseau, F., Schymkowitz, J. and Itzhaki, L. S. (2012) 'Implications of 3D domain swapping for protein folding, misfolding and function', *Advances in Experimental Medicine and Biology*. doi: 10.1007/978-1-4614-3229-6_9.

Rousseau, F., Schymkowitz, J. W. H. and Itzhaki, L. S. (2003) 'The unfolding story of three-dimensional domain swapping', *Structure*. doi: 10.1016/S0969-2126(03)00029-7.

Rowland, T. J. *et al.* (2016) 'Obscurin Variants in Patients With Left Ventricular Noncompaction', *Journal of the American College of Cardiology*. Elsevier, 68(20), pp. 2237–2238. doi: 10.1016/J.JACC.2016.08.052.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) 'DNA sequencing with chain-terminating inhibitors.', *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp. 5463–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/271968 (Accessed: 28 October 2018).

Sanger, J. M. J. W. *et al.* (2005) 'How to build a myofibril', in *Journal of Muscle Research and Cell Motility*. Kluwer Academic Publishers, pp. 343–354. doi: 10.1007/s10974-005-9016-7.

Sanger, J. W. *et al.* (2002) 'Myofibrillogenesis in skeletal muscle cells.', *Clinical orthopaedics and related research*, (403 Suppl), pp. S153-62. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12394464 (Accessed: 28 October 2018).

Sanger, J. W. *et al.* (2010) 'Assembly and Dynamics of Myofibrils', *Journal of Biomedicine and Biotechnology*. Hindawi Publishing Corporation, 2010. doi: 10.1155/2010/858606.

Sauer, F. *et al.* (2010) 'Molecular basis of the head-to-tail assembly of giant muscle proteins obscurin-like 1 and titin.', *EMBO reports*. Nature Publishing Group, 11(7), pp. 534–540. doi: 10.1038/embor.2010.65.

Sauer, F. (2011) *Structural studies on the association of filamentous proteins in the human M-Bands*. EMBL and University Würzburg.

Savarese, G. and Lund, L. H. (2017) 'Global Public Health Burden of Heart Failure.', *Cardiac failure review*. Radcliffe Cardiology, 3(1), pp. 7–11. doi: 10.15420/cfr.2016:25:2.

Schoenauer, R. *et al.* (2005) 'Myomesin is a Molecular Spring with Adaptable Elasticity', *Journal of Molecular Biology*. Academic Press, 349(2), pp. 367–379. doi: 10.1016/J.JMB.2005.03.055.

Schoenauer, R. *et al.* (2008) 'Myomesin 3, a Novel Structural Component of the M-band in Striated Muscle', *Journal of Molecular Biology*. Academic Press, 376(2), pp. 338–351. doi: 10.1016/J.JMB.2007.11.048.

Schoenauer, R. *et al.* (2011) 'EH-myomesin splice isoform is a novel marker for dilated cardiomyopathy', *Basic Research in Cardiology*. Springer-Verlag, 106(2), pp. 233–247. doi: 10.1007/s00395-010-0131-2.

Schultheiss, T. *et al.* (1990) 'Differential distribution of subsets of myofibrillar proteins in cardiac nonstriated and striated myofibrils', *Journal of Cell Biology*, 110(4), pp. 1159–1172. doi: 10.1083/jcb.110.4.1159.

Schymkowitz, J. W. H. *et al.* (2001) 'Observation of signal transduction in three-dimensional domain swapping', *Nature Structural Biology*. doi: 10.1038/nsb1001-888.

Seidel, S. A. I. *et al.* (2013) 'Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions', *Methods*, pp. 301–315. doi: 10.1016/j.ymeth.2012.12.005.

Shriver, M. *et al.* (2015) 'Loss of giant obscurins from breast epithelium promotes epithelial-to-mesenchymal transition, tumorigenicity and metastasis', *Oncogene*. Nature Publishing Group, 34(32), pp. 4248–4259. doi: 10.1038/onc.2014.358.

Sievers, F. *et al.* (2011) 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.', *Molecular systems biology*. EMBO Press, 7(1), p. 539. doi: 10.1038/msb.2011.75.

Sjöström, M. and Squire, J. M. (1977) 'Fine structure of the A-band in cryo-sections. The structure of the A-band of human skeletal muscle fibres from ultra-thin cryo-sections negatively stained', *Journal of Molecular Biology*. Academic Press, 109(1), pp. 49–68. doi: 10.1016/S0022-2836(77)80045-4.

Somera, G. N. (1995) *PROTEINS AND TEMPERATURE*, *Anllu. Rev. Ph)'siol*. Available at: www.annualreviews.org (Accessed: 21 November 2018).

Sorimachi, H. *et al.* (1997) 'Tissue-specific expression and α-actinin binding properties of the Z-disc titin: Implications for the nature of vertebrate Z-discs', *Journal of Molecular Biology*, 270(5), pp. 688–695. doi: 10.1006/jmbi.1997.1145.

Spooner, P. M. *et al.* (2012) 'Large Isoforms of UNC-89 (Obscurin) Are Required for Muscle Cell Architecture and Optimal Calcium Release in Caenorhabditis elegans', *PLoS ONE*. Edited by T. Lamitina. Public Library of Science, 7(7), p. e40182. doi: 10.1371/journal.pone.0040182.

Squire, J. M. *et al.* (2005) 'Molecular Architecture in Muscle Contractile Assemblies', *Advances in Protein Chemistry*. Academic Press, 71, pp. 17–87. doi: 10.1016/S0065-3233(04)71002-5.

Stout, A. L. *et al.* (2008) 'Tracking changes in Z-band organization during myofibrillogenesis with FRET imaging', *Cell Motility and the Cytoskeleton*, 65(5), pp. 353–367. doi: 10.1002/cm.20265.

Svergun, D., Barberato, C. and Koch, M. H. (1995) 'CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates', *Journal of Applied Crystallography*. International Union of Crystallography (IUCr), 28(6), pp. 768–773. doi: 10.1107/S0021889895007047.

Svergun, D. I. (1992) 'Determination of the regularization parameter in indirect-transform methods using perceptual criteria', *Journal of Applied Crystallography*. International Union of Crystallography (IUCr), 25(pt 4), pp. 495–503. doi: 10.1107/S0021889892001663.

Svergun, D. I. (1999) 'Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing', *Biophysical Journal*. Cell Press, 76(6), pp. 2879–2886. doi: 10.1016/S0006-3495(99)77443-6.

Svergun, D. I., Semenyuk, A. V. and Feigin, L. A. (1988) 'Small-angle-scattering-data treatment by the regularization method', *Acta Crystallographica Section A*. International Union of Crystallography, 44(3), pp. 244–251. doi: 10.1107/S0108767387011255.

Swords, W. E. (2003) 'Chemical Transformation of E. coli', in *E. coli Plasmid Vectors*. New Jersey: Humana Press, pp. 49–54. doi: 10.1385/1-59259-409-3:49.

Tam, J. P. *et al.* (1991) 'Disulfide Bond Formation in Peptides by Dimethyl Sulfoxide. Scope and Applications', *Journal of the American Chemical Society*, 113(17), pp. 6657–6662. doi: 10.1021/ja00017a044.

Tan, L.-B. *et al.* (2010) 'So many definitions of heart failure: are they all universally valid? A critical appraisal', *Expert Review of Cardiovascular Therapy*. Expert Reviews Ltd., 8(2), pp. 217–229. Available at: http://go.galegroup.com/ps/anonymous?id=GALE%7CA251006351&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=14779072&p=AONE&sw=w (Accessed: 7 November 2018).

Tardiff, J. C. *et al.* (2015) 'Targets for therapy in sarcomeric cardiomyopathies', *Cardiovascular Research*. Oxford University Press, 105(4), pp. 457–470. doi: 10.1093/cvr/cvv023.

Tokuyasu, K. T. and Maher, P. A. (1987) 'Immunocytochemical studies of cardiac myofibrillogenesis in early chick embryos. II. Generation of α-actinin dots within titin spots at the time of the first myofibril formation', *Journal of Cell Biology*. Rockefeller University Press, 105(6 I), pp. 2795–2801. doi: 10.1083/jcb.105.6.2795.

Tria, G. *et al.* (2015) 'Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering', *IUCrJ*. International Union of Crystallography, 2(2), pp. 207–217. doi: 10.1107/S205225251500202X.

Trinick, J. and Tskhovrebova, L. (1999) 'Titin: a molecular control freak', *Trends in Cell Biology*. Elsevier Current Trends, 9(10), pp. 377–380. doi: 10.1016/S0962-8924(99)01641-4.

Trombitás, K., Greaser, M. L. and Pollack, G. H. (1997) 'Interaction between titin and thin filaments in intact cardiac muscle', *Journal of Muscle Research and Cell Motility*. Kluwer Academic Publishers, 18(3), pp. 345–351. doi: 10.1023/A:1018626210300.

Tskhovrebova, L. and Trinick, J. (2003) 'Titin: properties and family relationships', *Nature Reviews Molecular Cell Biology*, 4(9), pp. 679–689. doi: 10.1038/nrm1198.

Turnacioglu, K. K. *et al.* (1996) 'Partial characterization of zeugmatin indicates that it is part of the z-band region of titin', *Cell Motility and the Cytoskeleton*, 34(2), pp. 108–121. doi: 10.1002/(SICI)1097-0169(1996)34:2<108::AID-CM3>3.0.CO;2-7.

Udd, B. *et al.* (2005) 'Titinopathies and extension of the M-line mutation phenotype beyond distal myopathy and LGMD2J', *Neurology*, 64(4), pp. 636–642. doi: 10.1212/01.WNL.0000151853.50144.82.

Udd, B. (2012) 'Distal myopathies – New genetic entities expand diagnostic challenge', *Neuromuscular Disorders*. Elsevier, 22(1), pp. 5–12. doi: 10.1016/J.NMD.2011.10.003.

Udd, B. and Griggs, R. (2001) 'Distal myopathies.', *Current opinion in neurology*, 14(5), pp. 561–6. doi: 10.1097/00019052-200110000-00003.

Ultsch, M. H. *et al.* (1999) 'Crystal structures of the neurotrophin-binding domain of TrkA, TrkB and TrkC', *Journal of Molecular Biology*. Academic Press, 290(1), pp. 149–159. doi: 10.1006/JMBI.1999.2816.

van der Ven, P. F. *et al.* (2000) 'A functional knock-out of titin results in defective myofibril assembly.', *Journal of cell science*, 113 (Pt 8), pp. 1405–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10725223 (Accessed: 28 October 2018).

Van der Ven, P. F. *et al.* (1999) 'Thick filament assembly occurs after the formation of a cytoskeletal scaffold.', *Journal of muscle research and cell motility*, 20(5–6), pp. 569–79. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10555075 (Accessed: 28 October 2018).

van der Ven, P. F. and Fürst, D. O. (1997) 'Assembly of titin, myomesin and M-protein into the sarcomeric M band in differentiating human skeletal muscle cells in vitro.', *Cell structure and function*, 22(1), pp. 163–71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9113403 (Accessed: 28 October 2018).

Visscher, P. M., Hill, W. G. and Wray, N. R. (2008) 'Heritability in the genomics era — concepts and misconceptions', *Nature Reviews Genetics*. Nature Publishing Group, 9(4), pp. 255–266. doi: 10.1038/nrg2322.

Volkov, V. V. and Svergun, D. I. (2003) 'Uniqueness of ab initio shape determination in small-angle scattering', in *Journal of Applied Crystallography*. International Union of Crystallography (IUCr), pp. 860–864. doi: 10.1107/S0021889803000268.

Wakayama, J. *et al.* (2000) 'Atomic Force Microscopic Evidence for Z-band as a Rigid Disc Fixing the Sarcomere Structure of Skeletal Muscle.', *Cell Structure and Function*. Japan Society for Cell Biology, 25(6), pp. 361–365. doi: 10.1247/csf.25.361.

Wang, J. *et al.* (2005) 'Dynamics of Z-band based proteins in developing skeletal muscle cells', *Cell Motility and the Cytoskeleton*. Wiley-Blackwell, 61(1), pp. 34–48. doi: 10.1002/cm.20063.

Wolynes, P. G., Luthey-Schulten, Z. and Onuchic, J. N. (1996) 'Fast-folding eriments and the topography of protein folding energy landscapes', *Chemistry & Biology*. Cell Press, 3(6), pp. 425–432. doi: 10.1016/S1074-5521(96)90090-3.

Wright, N. T. and Letourneau, A. G. (2018) 'Structural Insights on the Obscurin-Binding Domains in Titin', *Protein & Peptide Letters*, 25. doi: 10.2174/0929866525666181004102031.

Xiao, S. and Gräter, F. (2014) 'Molecular basis of the mechanical hierarchy in myomesin dimers for sarcomere integrity.', *Biophysical journal*. Elsevier, 107(4), pp. 965–73. doi: 10.1016/j.bpj.2014.06.043.

Xu, J. *et al.* (2015) 'Investigation of Pathogenic Genes in Chinese sporadic Hypertrophic Cardiomyopathy Patients by Whole Exome Sequencing', *Scientific Reports*. Nature Publishing Group, 5(1), p. 16609. doi: 10.1038/srep16609.

Yajima, H. *et al.* (1996) 'A 11.5-kb 5'-terminal cDNA sequence of chicken breast muscle connectin/titin reveals its Z line binding region', *Biochemical and Biophysical Research Communications*. Academic Press, 223(1), pp. 160–164. doi: 10.1006/bbrc.1996.0862.

Yamada, T. *et al.* (2003) 'Molecular Organizations of Myofibrils of Skeletal Muscle Studied by Atomic Force Microscopy', in. Springer, Boston, MA, pp. 285–294. doi: 10.1007/978-1-4419-9029-7_27.

Yang, S. *et al.* (2004) 'Domain swapping is a consequence of minimal frustration', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0403724101.

Yang, Y. G., Obinata, T. and Shimada, Y. (2000) 'Developmental relationship of myosin binding proteins (myomesin, connectin and C-protein) to myosin in chicken somites as studied by immunofluorescence microscopy.', *Cell structure and function*, 25(3), pp. 177–85. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10984101 (Accessed: 28 October 2018).

Young, P. *et al.* (1998) 'Molecular structure of the sarcomeric Z-disk: Two types of titin interactions lead to an asymmetrical sorting of α-actinin', *EMBO Journal*. European Molecular Biology Organization, 17(6), pp. 1614–1624. doi: 10.1093/emboj/17.6.1614.

Young, P., Ehler, E. and Gautel, M. (2001) 'Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor protein involved in sarcomere assembly', *Journal of Cell Biology*, 154(1), pp. 123–136. doi: 10.1083/jcb.200102110.

Zhang, Y., Werling, U. and Edelmann, W. (2012) 'SLiCE: A novel bacterial cell extract-based DNA cloning method', *Nucleic Acids Research*. Oxford University Press, 40(8), pp. 1–10. doi: 10.1093/nar/gkr1288.

Zhang, Y., Werling, U. and Edelmann, W. (2014) 'Seamless Ligation Cloning Extract (SLiCE) cloning method.', *Methods in molecular biology (Clifton, N.J.)*. NIH Public Access, 1116, pp. 235–44. doi: 10.1007/978-1-62703-764-8_16.

Zhou, W. and Freed, C. R. (2004) 'Tyrosine-to-cysteine modification of human alpha-synuclein enhances protein aggregation and cellular toxicity.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 279(11), pp. 10128–35. doi: 10.1074/jbc.M307563200.

Abbreviations

APS	ammonium persulfate
ASU	asymmetric unit
AMP	adenosine monophosphate
ATP	adenosine triphosphate
BiFC	bimolecular fluorescence complementation
BSA	bovine serum albumin
CD	circular dichroism
co-IP	co-immunoprecipitation
CV	column volume
DCM	dilated cardiomyopathy
DNA	deoxyribonucleic acid
E. coli	Escherichia coli
EDTA	ethylenediaminetetraacetate
EM	electron microscopy
ESRF	European Synchrotron Radiation Facility
ExAC	Exome Aggregation Consortium
FITC	fluorescein isothiocyanate
FN-III	fibronectin-type-III
HCM	hypertrophic cardiomyopathy
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Ig	immunoglobulin
IPTG	isopropyl-b-thiogalactopyranoside
LGMD2J	limb-grindle muscular dystrophy type 2J
LB	lysogeny broth
MAF	mean allele frequency
MALLS	multiple angle laser light scattering
MES	2-(N-morpholino) ethanesulfonic acid
MW	molecular weight
MyBPC	myosin-binding protein C
M2H	mammalian two-hybrid
NHS	N-hydroxysuccinimide
NTA	nitrilotriacetic acid
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
P_i	inorganic phosphate
PLN	phospholamban
Rmsd	root mean square deviation
ROS	reactive oxygen species
Rpm	revolutions per minute

SAXS	small angle x-ray scattering
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SERCA2	Sarco-/endoplasmatic reticulum Ca^{2+}-ATPase
SH3	Src homology 3
SPR	surface plasmon resoncne spectroscopy
SR	sarcoplasmic reticulum
TCEP	tris-(2-carboxyethyl)phosphine
TEMED	tetramethylethylenediamine
TEV	tobacco etch virus
TMD	tibial muscular dystrophy
Tris	tris-(hydroxymethyl)aminomethane
UV	ultraviolet

Contributions

The in this thesis presented work was entirely performed by myself with the following exceptions:

(1) The model refinement of the titin-obscurin complex was performed by Spyros Chatziefthimiou, EMBL-Hamburg
(2) The collection and processing of the SAXS data was performed in collaboration with Melissa Graewert and Dmitri Svergun.
(3) In vivo binding experiments in NRC were performed by Atsushi Fukuzawa (Gautel Laboratory, King's College London)
(4) The unpublished dimeric structure of My5 was solved by Florian Sauer, EMBL-Hamburg (currently: University of Würzburg)

Acknowledgement

Firstly, I would like to express my sincere gratitude to my advisor Prof. Matthias Wilmanns for his trust in my abilities and his support during my PhD study and related research. He made it possible for me to pursue my ideas, to thrive my professional skills and to present my research all over the world. For this, I am deeply thankful!

Besides my advisor, I would like to thank the rest of my thesis advisory committee: Prof. Reinhard Schneppenheim, Prof. Maya Köhn, and Dr. Dmitri Svergun, for their insightful comments and encouragement, but also for the critical questions which pushed me to widen my research from various perspectives.

It is a pleasure to thank Dr. Melissa Gräwert for all her support and work for the SAXS experiments, helping to explore the nature of the proteins I worked with, and Prof. Mathias Gautel and Dr. Atushi Fukuzawa for their exciting collaboration and *in vivo* studies.

My sincere thanks also go to Dr. Spyros Chatziefthimiou, Dr. Inga Pfeffer, Dr. Annabel Parret, and Dr. Vivian Pogenberg for all the helpful discussions, great technical and personal support and open ears throughout the years.

Additionally, I'm grateful to thank my fellow office- and lab-mates Thomas, Maria, Yonca, Riccardo, Karen, Tom, Nishit, Emma, Diana, Morlin, Sonja, and last but not least Lisa for the stimulating discussions and for all the fun we have had in the last four years. It was really a pleasure meeting and becoming friends with you!

I owe my deepest gratitude to my older brother Daniel. Without you, I would not be where and who I am now. Thank you very much for introducing me to science and always being an inspiration as a scientist and a person! I am so proud of you!

To my family, thank you for your love, support, and unwavering belief in me. You have made everything possible for me, and I am deeply grateful for this! No matter how things go, I know that you will always be there for me.

And finally, I would like to thank my dear Stéphanie for her love and endless support, for all the late nights and early mornings, and for keeping me sane over the past months. Thank you for being my source of inspiration, proof-reader, editor, and strong shoulder. But most of all, thank you for being my best friend. I owe you everything.

Curriculum Vitae

Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt.

Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift: