Spring 2017

Obscurin acts as a variable force resistor

Aidan M. Willey
James Madison University

Obscurin acts as a variable force resistor

A Project Presented to

The Faculty of the Undergraduate

College of Chemistry and Biochemistry

James Madison University

In Partial Fulfillment of the Requirements

For the Degree of Bachelor of Sciences

By Aidan M. Willey

May 2017

Accepted by the faculty of the Department of Chemistry and Biochemistry, James Madison University, in partial fulfillment of the requirements for the Degree of Bachelor of Sciences

FACULTY COMMITTEE:

Faculty Project Advisor: Nathan T. Wright, Ph.D.,
Assistant Professor, Chemistry and Biochemistry

Reader: Isaiah C. Sumner, Ph.D.,
Assistant Professor, Chemistry and Biochemistry

Reader: Christopher E. Berndsen, Ph.D.,
Assistant Professor, Chemistry and Biochemistry

HONORS PROGRAM APPROVAL:

Bradley Newcomer, Ph.D.,
Dean, Honors College

This work is accepted for presentation, in part or in full, at JMU Chemistry Spring Symposium on March 24, 2017

Table of Contents

Figures

Introduction

Chapter 1

Chapter 2

Appendix

Acknowledgements

I would like to acknowledge Dr. Nathan T. Wright, for introducing me to his research laboratory. Along with Dr. Wright, I also would like to thank Dr. Isaiah Sumner, Dr. Christopher Berndsen, Dr. Gina Macdonald, and Dr. Paul Raston for their exemplary teaching abilities. In many ways, these professors have excited me about biophysical chemistry and influenced my desire to continue with research after graduation. I also owe thanks to the entire JMU Department of Chemistry and Biochemistry for giving me the privilege of using their instrumentation. Each and every professor in this department has also been a source of support and encouragement for me as well. This work was funded by the Tickle Fellowship (Aidan Willey) and Research Corporation (Nathan T. Wright) and Cottrell Scholar (Nathan T. Wright).

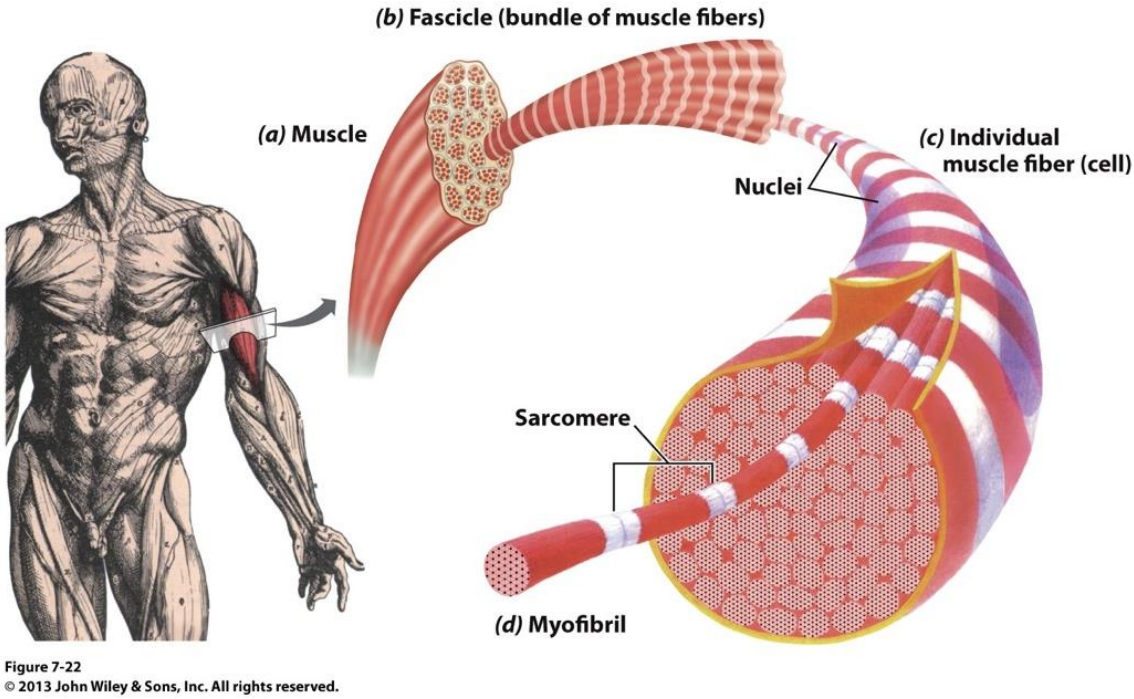
INTRODUCTION

The intricate structure of skeletal and cardiac muscle cells enables the generation of movement and force upon striated muscle contraction (Gautel, 2016). This complex architecture is essential in maintaining stability upon active muscle contraction and passive expansion. Connective tissues bind muscle fibers together to form bundles called fascicles, which bundle to form muscle tissue. This architecture makes the basic structure of skeletal and cardiac muscle. At a more molecular level, muscle fibers themselves contain thousands of myofibrils that are composed of numerous sarcomeres (Lieber, 2002). The sarcomere is the smallest functional unit of contraction and defines the area of interaction between thick and thin filaments. Thick and thin filaments in the sarcomere are mainly comprised of myosin and actin, which move past one another via a sliding filament mechanism upon muscle contraction (Figure 1A) (Huxley, 1973). The repetitive array of parallel actin and myosin proteins allows controlled force and movement to be generated throughout the entire muscle system (Figure 1B) (Gautel, 2016).

Actin and myosin rely on a multitude of other systems within myofibrils in order to function properly. Organizing actin and myosin into a functional sarcomeric unit is the job of the supporting cytoskeleton (Gautel, 2016). Besides simply bundling and positioning actin and myosin, the cytoskeleton also provides a connector between the contractile machinery and the rest of the cell (Kontragianni-Konstantopoulos, 2009). This attachment is necessary to relay contractile force, and can act as a biochemical signaling conduit. To achieve these diverse demands, the cytoskeleton relies, in part, on rope-like proteins that are extensively cross-linked with both the contractile apparatus and the surrounding cell components (Figure 1B). A family of giant modular signaling proteins serve this purpose and function as flexible scaffolds to the system, providing support to the overall muscle architecture (Young, 2001).

The two members of this class of cytoskeletal proteins are titin and obscurin (Kontragianni-Konstantopoulos, 2009).

A)



B)

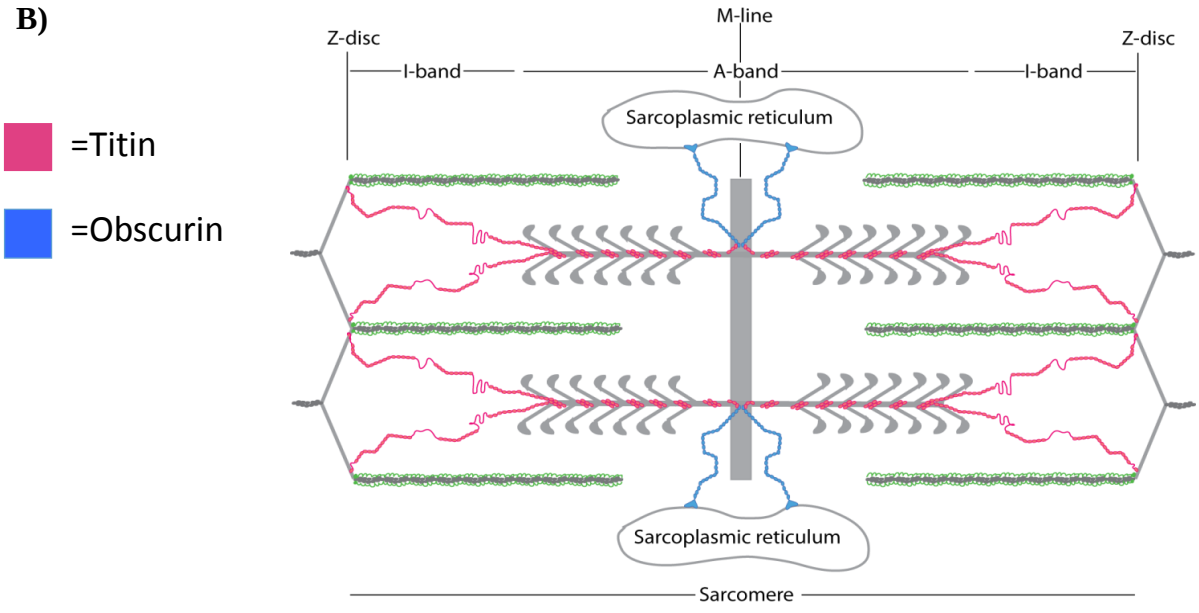


Figure 1. A) Diagram showing the structure and architecture in muscles. **B)** A blueprint showing how giant modular signaling proteins function within myofibrils to connect distal parts of the cell.

Obscurin is the most recently discovered giant muscle protein (Young, 2001). It is a very long, modular protein (800 kDa) and consists of around 70 independently-folded Ig-like and FnIII-like domains linked together (Meyer, 2013 and Gautel, 2011). Due to its modular architecture, individual obscurin domains can be excised from the rest of the molecule and still retain their native folds (Meyer, 2013). Obscurin links distal objects with cells; it is the only known connector between the cytoskeleton and the sarcoplasmic reticulum (Young, 2001). This link is achieved through interactions with titin (via obscurin domains Ig1 and Ig58/59) and with small anchorin (via the non-modular c-terminus) (Hu, 2013 and Gautel, 2011). Obscurin also plays a role in the lateral organization of myosin filaments, which is essential in maintaining global sarcomeric structure and alignment. (Gautel, 2016).

Obscurin interacts with many cell adhesion molecules, and seems to be involved in sensing whether a cell is fastened to the basal lamina (Kontragianni-Konstantopoulos, 2009). For example, obscurin domains phosphorylate cytoplasmic domains of N-cadherin, which is a major component of adherin junctions (Hu, 2013 and Kontragianni-Konstantopoulos, 2009). Thus, obscurin is involved in interactions between the contractile apparatus (via titin) and the basal lamina (Kontragianni-Konstantopoulos, 2009). Given obscurin's overall rope-like architecture and its stretch/adhesion function, it seems likely that obscurin acts as some kind force sensor/force resistor, acting to signal to cells whether they are positioned correctly within the larger tissue system (Caldwell, 2015). Aside from organizing cells, obscurin might also play a role in cellular signaling and homeostasis. Supporting this hypothesis, obscurin knockout cells acquire cancerous phenotypes and are also associated with both cardiomyopathies and muscular dystrophies (Wood, 2007). While unproven, this shows a place to look for a link between the cytoskeleton and how a cell senses its environment.

Myofibrils are simultaneously strong and flexible. While strength is derived through the well-understood mechanics behind the actin/myosin cross bridge formation, the molecular

mechanisms surrounding flexibility have proven more ambiguous (Gautel, 2011). In the past several years, multiple papers have demonstrated how long, modular, fiber-like proteins form a flexible web within muscle cells (Agarkova, 2005 and Guatel, 2011). Aside from organizing and anchoring the contractile apparatus, some of these long modular proteins also act as force resistors and force modulators to help control myocyte stretch. This is most well-documented in titin (Andres, 2016; Linke 2000; Lu, 1999). However, due to obscurin's architecture, binding partners, and physiological role, we hypothesize that obscurin also senses and modulates cell stretch.

Titin is hypothesized to be a force resistor due to -ray structural studies (von Castlemur, 2008). These studies concluded that linker length- the short peptide that connects tandem Ig-like titin domains- dictates the passive shape, or the persistence length, of the titin molecule. A longer length linker confers more flexibty while a shorter linker length creates a more rigid region (von Castlemur, 2008). A limitation in von Castlemur's experimentation involves purely experimental techniques using X-ay diffraction, which subjects the protein to significant crystal packing artifacts. In fact, later studies by Klau Schulten found very little difference in the stiffness between longer and shorter linkers (Murphy, 2007). However, these studies examined more course-grained effects of titin stretch; Schulten experimented on atomistic simulations. We circumvent these problems in our experiments by using the complementary techniques of solution state experimentation and atomic-scale computational simulations. Solution state experimentation will allow us to see how the protein functions in the presence of water, while computational techniques will allow us to watch the proteins response to force. By synthesizing these different techniques, we find that we obtain a more nuanced, deeper understanding of cytoskeletal protein motion.

Obscurin links distal objects in the cell. Due to its rope-like architecture, it is an intrinsic force resistor. Here, we try to determine whether obscurin behaves more like a spring or more like a rope (Figure 2).

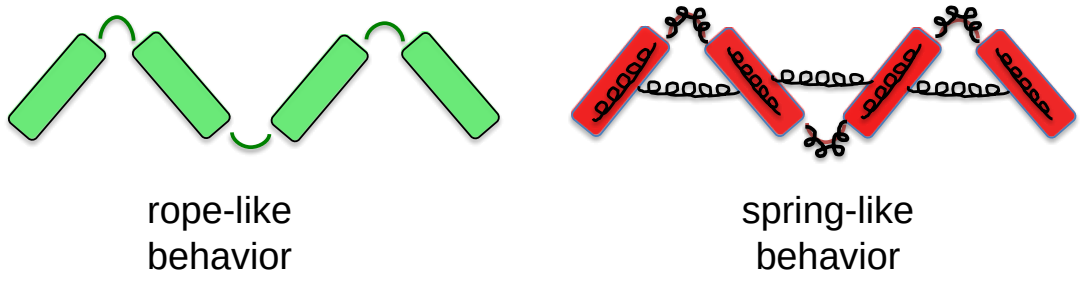


Figure 2. Two hypotheses predicting the different behaviors of dual domain systems

To do this, we use experimental methods to study high-resolution structures of tandem obscurin domains at rest and computation methods to study the structures when stretched. A ropelike behavior would suggest a more entropic model where the domains act independently and only resist force when fully stretched. A spring-like behavior characterizes an enthalpic model of resistance where there is more dependence between domains and resistance is present at many different lengths (Figure 2). Elucidating this nature of obscurin will help describe its function within the cell. In other words, understanding obscurin's stretch response allows us to classify obscurin as either a simple connector protein or as a stretch responder. Since obscurin is involved in cell stretch/adhesion signaling, we hypothesize that the mechanisms underlying force resistance play a large role in its function in the cell. This function may be important in muscles; obscurin knockouts animals display generalized muscle weakening (Kontrogianni-Konstantopoulos, 2009). Relatedly, the fact that obscurin knockout cells undergo an epithelial to mesenchymal transition suggests a role of obscurin signaling in normal cell homeostasis (Wood 2007). This, also may be related to

force resistance. However both of these phenotypes are of unknown etiology. Here, we begin to explore this gap in our knowledge.

The high-resolution structure of roughly $1/10^{th}$ of obscurin has been previously determined, primarily through solution state nuclear magnetic resonance (NMR) analysis. We take advantage of this for the present studies. In particular, here we use two regions- Ig58/59 and Ig34/35/36/37 (Wakabayashi, 2005). These structures have been previously solved by Wakabayashi of Riken institute, but are not yet described in peer-reviewed journals (Wakabayashi, 2005). Starting our analysis on previously solved NMR-derived structures provides a major benefit, in that the sequence-specific assignments are known *a priori* (Wakabayashi, 2005). After we compare experimental NMR data to previously determined structures, we can then further analyze residual dipolar coupling (RDC) data in order to examine the interactions between the dual domain systems. Complementing this, small angle X-ray scattering (SAXS) analysis allows an independent method of assessing structural orientations of these domains. Mirroring this wet-lab work, we used steered molecular dynamics (SMD) to computationally simulate stretching between the dual domains. SMD was performed on the system in order to characterize overall force resistance behavior.

To determine NMR structure of the protein, we used labeled samples ^{15}N to increase resolution by adding extra dimensions. The 2D and 3D chemical shifts that we observe depend on both the frequency and environment of electric field. We use standard pulse programs developed to select for the detection of specific atoms, thus allowing us to sequence-specifically assign and study our proteins of interest. Overall, 2D ^{1}H-^{15}N HSQC and 3D ^{15}N TOCSY and ^{15}N NOESY NMR data was collected (See Figure 3).

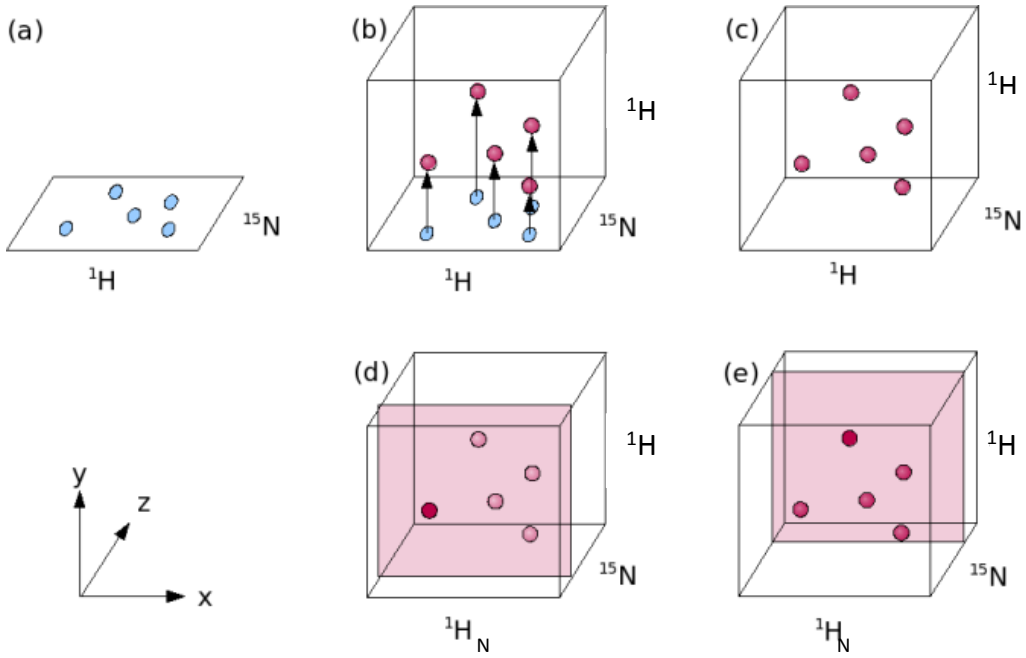


Figure 3. 2D and 3D NMR representation of how analysis is performed

2D ^{1}H-^{15}N HSQC is the first method of analysis and it serves as a fingerprint for subsequent experiments. Peaks in HSQC data collected represented each H-N spin system of a known amino acid in a given Ig domain (Figure 3A). These peaks are dependent on molecular environment. Since obscurin is a modular structure, the molecular environment of most residues in each domain are likely similar, regardless of whether the domain is expressed by itself or in the context of the larger obscurin molecule. We use this fact to assign the HSQC spectrum of dual-domain systems. To double-check our assignments, we also perform both 3D NMR experiments (Figures 3D,E). Peaks are assigned to their corresponding amino acids with the use of 3D NMR spectra. 3D-^{15}N TOCSY allows for sequence specific assignments, and it help to identify amino acid types. 3D-^{15}N NOESY data determines distance restraints for protons. Spectra using the nuclear overhauser effect (NOE) uses dipolar interactions between spins in order to correlate protons. The cross talk between spins determine the distances between atoms with unique chemical shifts. TOCSY spectra allow specific

assignments, while NOESY experiments map short-range distances within our domains of interest. We do not need to use HNCACB or CBCACONH spectra because the Ig35/36 domains have been previously solved (Wakabayashi, 2005).

While NOE and chemical shift perturbations provide short-range information about domain-domain orientation, there is a need for an independent measure of domain flexibility. One way to accomplish this is via residual dipolar couplings (RDC). NMR structures are further refined through residual dipolar coupling (RDC) data that generates angle restraints. RDC calculations provide information about the magnetic dipole-dipole interactions in a common frame of reference, and they will show the interdomain orientations of a dual domain system. RDC calculations are unique in that they can provide relative orientations between internuclei vectors that is independent of their distance separations (Chen, 2012). RDCs are collected from difference between isotropic and anisotropic samples. Isotropic samples can tumble rapidly because they are free in solution, but anisotropic solutions are partially aligned (polyacrylamide gels are used in this experiment). Thus these have an isotropic component. This dipolar coupling term can be visualized by subtracting out the scalar coupling term from the isotropic sample. This leaves a residual dipolar coupling term. Ultimately, dipolar coupling is the physical response of the relationship between the magnetic dipole of two atoms and the external magnetic field (Alba, 2002).

$$D_{PQ}(\theta, \square) = S\frac{\mu_o}{4\pi}\Upsilon_P\Upsilon_Q\,\hbar\left[A_a(3cos^2(\theta)-1)+\frac{3}{2}A_r sin^2(\phi)\right] \quad \textbf{(1)}$$

This equation shows how RDCs are calculated. D_{PQ} represents the vector connecting distances P and Q, which are parallel to the magnetic field, where the coupling is at its strongest. The values, Υ_P and Υ_Q, represent the gyromagnetic ratios between the spins, μ_o represents the vacuum permeability, and S accounts for the angular averaging due to internal

motion of the nuclei. All of these values are constants within the calculation. The dipolar coupling is mainly dependent on the magnitude of axial and rhombic components (A_a and A_r) of the alignment tensor. The rhombic component of the alignment tensor is the degree of alignment along two orthogonal axes in the plane perpendicular to the axial direction. By comparing the isotropic and anisotropic alignments, the values A_a and A_r, can be derived. This calculation of D_{PQ} can be converted into a degenerate angle, which provides information about the orientation of the internuclear bond vector relative to the protein's alignment tensor (see equation 1). A single RDC value is of little use. However, comparing values from multiple different bonds (in this case, N-H bonds) gives the relative orientation of one bond compared to another. From these values, a Quality (Q) factor can determined. The Q factor is a measure of the agreement between the predicted RDC values as back-calculated from the structure and the measured dipolar couplings. This estimates the quality of how well the RDC values agree with the structure (Chen, 2012). Overall a lower Q is ideal (<.40) is salvageable). The purpose of performing RDC on the dual domain system is to see if when we apply alignment tensors to one domain, we can get the other domain's RDC values to fit within the first alignment tensor. If we can, then the system behaves as a single body, and rotates together. This would suggest significant domain-domain cross-talk, if it doesn't, then there must be two alignment tensors and the system behaves more like a rope. Overall, RDCs aid in determining structure and orientation of the dual domain system, which can predict a ropelike versus spring like behavior.

Small angle X-ray scattering (SAXS) is another component that is able to determine structural composition. The NMR data is compared to the SAXS data in order to refine and improve the approximated structural models. An advantage with SAXS is that it offers an orthogonal description of biomolecular behavior, independent of NMR analysis. Specifically, we use these data to see how flexible our dual domain systems are, these conclusions can

then be compared or merged with NMR data for a more thorough, nuanced descriptions of the systems of interest. In SAXS, the protein sample is exposed to X-ray beams of a specified wavelength, and their scattering pattern is used to produce a degenerate, spatially averaged intensity profile. The scattering pattern of the X-ray depends on the proteins structure and size. SAXS gives low resolution imaging because it reduces the structural information down to two dimensions. Also, because the protein in solution is randomly oriented, the scattering pattern represents an average pattern from all possible orientations. The scattering pattern therefore does not depend on direction of the scattering vector, but only on the magnitude. SAXS measures the intensity as a function of scatter; which is because the scattering vector q results from a photon of wavelength λ scattering off of the sample at an angle 2θ (Skou, 2014). For a macromolecule in solution the intensity distribution must subtract the buffer profile from the sample profile (Skou, 2014).

$$I(q)=(2\pi \sin\theta)/\lambda \quad \textbf{(2)}$$

In Equation 2, 2θ is the scattering angle, and λ is the wavelength of the X-ray beam. Images of the sample can be integrated about the beam, and simple curves of scattering intensity profiles of I vs. q can be obtained, where q is normally measured in wavenumbers (Skou, 2014).

Overall, SAXS data is analyzed by using Fourier transform to in order to relate reciprocal space to real space, to determine the average protein intensity. SAXS data will be evaluated using Guinier plots, Porod plots, and P(r) vs. r distribution plots. Guinier plots will plot the logarithm of the scattering curve, I(q) against the q value; this relates the intensity of the particle to its size. The importance of the Guinier plot is to obtain the simple intensity profile, and the R_g value. The radius of gyration (R_g) shows the distribution of components of the protein around an axis, or the distance from an axis where the mass is assumed to be

concentrated. The next method of analysis of SAXS data is through Porod plots. These plots are utilized in order to describe and confirm a folded and globular protein; they measure log(I) versus log(q). These plots will show the relationship between the two domains in the system, and give insight about how the individual domains interact with each other. Finally, The P(r) vs. r plot is calculated by Fourier Transform to show the scattering curve in real space. This plot shows the probabilities that two points within an object (the P(r)) are a certain distance apart (r). Specifically regarding SAXS, this plot provides information about the paired set of distances in the macromolecular structure. It sums all of the possible pairs of points in the particle to create a histogram like function. The definition of this plot defines P(r)=0 at r=0, and at r =D_{max} at the maximum size of the particle (Murphy, 2007). This graph is useful for detecting conformational changes within the macromolecule. Overall, from this plot, D_{max} and R_g values can be extrapolated from the pair distribution plot. D_{max} represents the maximum size of the protein; it measures the diameter or rough length of the dual domain system. R_g represents the radius of gyration. The shape of the plot is also significant in this experiment because the shape of the line on the graph will signify whether the domains behave together as one, or as two separate systems (Figure 4).

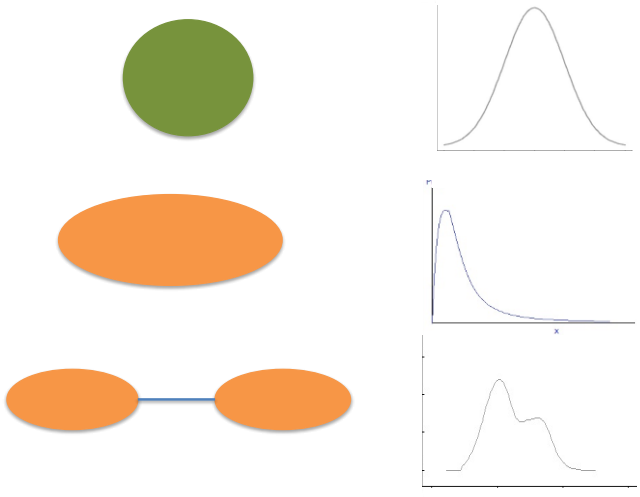


Figure 4. Analysis of the structure of protein domains using P(r) vs r distribution.

As Figure 4 shows, the pair distribution plot resembles a Gaussian distribution for a circular shape. Note that the probability of finding a point at a certain distance away increases proportionally as distance increases. For an elongated or elliptical protein, the probability curve is skewed to the left, and for a multidomain protein, the probability has two peaks. Our results will be evaluated using this rationale; a distribution graph showing one skewed peak will represent a system that behaves as a single body whereas a graph showing two peaks will represent a system that behaves as two separate bodies. A single body system will support a spring like model and a two separate body system will represent a rope like model. By integrating the results from these three solution based experiments, NMR, RDC, and SAXS, the structures and compositions of the dual domain systems will be thoroughly evaluated. After these have been analyzed, the force resistance of the systems can be simulated using computational methods.

The main component of how muscle cytoskeletal proteins function is how they respond to force. This mechanistic type of investigation is ideally suited to the use of steered molecular dynamics. SMD attempts to simulate the dynamics of time scaled events. It does this by applying a pulling force to impose a change in structure. The protein force field ($U_{ff}(R)$) is supplemented with the time dependent harmonic potential:

$$U_{tot}(R,t) = U_{ff}(R) + \frac{k}{2}\{x(R) - (x_o + vt)\}^2$$

In this equation, x(R) is a reaction coordinate, k is the harmonic spring constant, and v is the velocity (Caldwell, 2015). The velocity is fixed to a constant value prior to the run, and it pulls the dual domain system on one side, along a reaction coordinate. By using this type of simulation, we can computationally measure the stretch response of the dual domain system as it is pulled from one end. This is particularly useful for measuring small forces; AFM,

which is usually used for measuring forces in proteins, cannot generally resolve forces less than 10 pN, and the domain stretching we are looking at tends to be in the 3-10 pN range (Andres, 2016). Molecular dynamic simulations are also able to effectively imitate the molecular and biochemical environment that exists within and around obscurin (Israelewitz, 2001). Standardization of the model is set using minimization, heating, density, and equilibration calculation on the structure prior to the pull simulation in order to stabilize the molecule. After these standards have stabilized the model, it can be pulled from one end at a constant velocity. Following SMD, graphs will be generated from these simulations in order to visualize the force vs. distance, work vs. distance, angles, interactions, and twist between the dual domain system. Overall, the graphs will help to analyze the dual domain system's response to force resistance on an atomic level. It is hypothesized that the model will respond differently to these tests based on the flexibility vs rigidity of the system. By combining our experimental analysis with this kind of computational work, we will be able to characterize the overall behavior towards force that the Ig domains have (Israelewitz, 2001).

By studying these domains in tandem using multidimensional nuclear magnetic resonance (NMR) and small angle X-Ray scattering (SAXS) techniques, we gain insight into obscurin's shape and self-interactions. With the use of computational techniques, and residual dipolar coupling experiments (RDC), we gain understanding of how obscurin resists external force. These studies lead the way to more fully probe how obscurin uses its force resistance function to influence cell stretch, signaling, and organization. In this thesis, I concentrate on previously solved dual domain structures of obscurin, specifically Ig domains 34/35, 35/36 and 36/37. I then compare my results with the dual domain Ig58/59, which was analyzed previously (Caldwell, 2015).

A RESTRICTING FEATURE OF THIS STUDY

A technique known as Atomic Force Microscopy (AFM) is commonly used to study the mechanics and molecular stretching of proteins in muscle cells (Andres, 2016 and Manibog, 2017). In particular, AFM has recently been used to study the flexibility and elasticity in giant modular proteins (Andres, 2016 and Pernigo, 2017). It is also used to study the network of protein interactions, which is a large factor in how proteins are able to withstand such force (Manibog, 2017).

Regarding force resistance of proteins, AFM is able to characterize the domain-domain bursting and unfolding of titin and obscurin. AFM studies on the folding of titin show that domain bursting occurs at approximately 200-300 pN and unfolding unfolds below 10 pN (Andres, 2016). Other AFM studies show that even the smaller amounts of force, around 30 pN, cause the titin/obscurin interaction to break (Pernigo, 2010, 2017).

In this study we attempt to understand the interactions between titin and obscurin and therefore focus on how the individual domain-domain system of obscurin behaves upon stretch (Rohs, 1999). AFM cannot be used in this study due to the scale at which we focus. When studying dual domain systems in solution, we are generally studying forces less than 10 pN. AFM ultimately is too harsh on the obscurin molecule with regard to studying the interactions between domains. Instead, in this study molecular dynamics will be used to investigate how dual domain systems respond to force (Ho, 2010).

Chapter 1 – Structural and Flexibility Studies of Ig58/59 and Ig35/36

METHODS

Transformation:

BL21 cells were thawed in ice and 2μL of various human obscurin DNA fragments in a pet24a vector were added via standard transformational procedures. Resulting colonies were picked, and these bacterial stocks were then either stored on the LB-Kan plate at 4 °C or as glycerol stocks at -80 °C

Protein Purification:

Cells were grown in ^{15}N solution until the OD_{600} = 0.50-0.60. Cells were induced with IPTG (0.20 g/L) for 2.5-5 hrs at at 37 ºC. Cells were centrifuged (10 min, 5000 RPM, JLA 9.100) and the pellet was kept and frozen overnight. The pellet was thawed and sonicated with 100 μL PMSF. After clarification via centrifugation (40 min, 12000 RPM, JA 25.50), the supernatant was applied to a nickel column followed by a G-75 sizing column. Purified protein was confirmed via SDS PAGE gel and concentrated to 0.5-2 mM (Rudloff, 2015). Protein concentration was determined using a UV-VIS spectrometer and the extinction coefficient (Ig58/59-24520 L mol^{-1} cm^{-1}, Ig35/36-22960 L mol^{-1} cm^{-1}).

Nuclear Magnetic Resonance.

All NMR experiments were performed on the 600 MHz Bruker Avance II magnet, equipped with a 5 mm triple resonance probe and z axis pulse field gradient coils. HSQC and TOCSY experiments were performed on Ig dual domain samples in 20 mM Tris, 50 mM NaCl, pH 7.5 and 10% D_2O at 20 °C. The experimental NMR peaks were compared to single domain peaks found in previous literature (Wakabayashi, 2005).

HSQC experiments were collected with 2048 points for the H dimension and 200 points for the NH dimension. TOCSY experiments were collected with 2048x128x64 points in the aliphatic dimension with a spin lock of 70 μsec (Rudloff, 2015). NMR data was processed using NMRPipe and visualized with SPARKY (Pinotsis, 2012).

RDC

Anisotropic IPAP experiments for RDC determination were performed using the same conditions as for the HSQC and TOCSY with the exception of using a stretched polyacrylamide gel (Rudloff, 2015). The gel was prepared using 4% acrylamide, and soaked with buffer prior to soaking with protein. RDC values were calculated using PALES software (Hummer, 2003).

SAXS

1.0, 3.0, and 5.0 mg/mL samples of various obscurin samples were prepared in the G75 buffer and evaluated on the Argonne beamline12-id-B. Porod plots and Gunier Plots were created of the samples using PRIMUS. Multifoxs was used to analyze the fit of SAXS and RDC data together (Schneidman, 2016).

RESULTS:

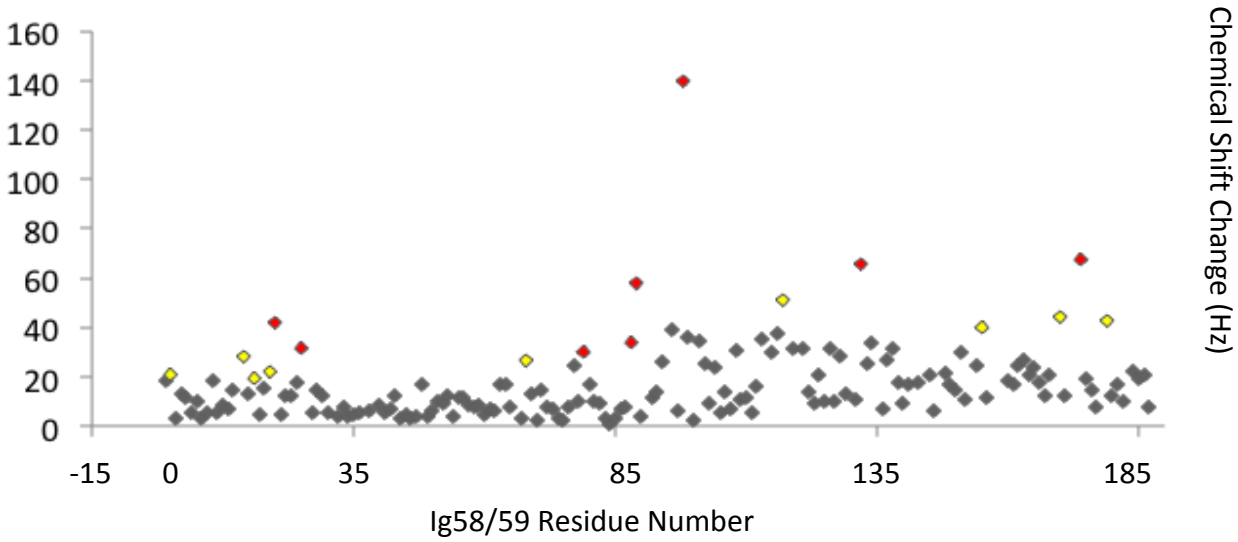
The compositional structure of Ig58/59 using NMR, RDC and SAXS analysis

Obscurin is a highly modular protein, linked together by individual Ig domains. This characteristic of obscurin makes it possible to isolate individual domains and study them without factoring in the rest of the molecule (Young, 2001). Using NMR, RDC and SAXS analysis, the interactions between neighboring domains and structural characteristics of the dual domain system were defined. From these data we could obtain information about domain-linker-domain models. Previous NMR work suggested that Ig58/59 was flexible. Thus we began this current project by re-examining the Ig58/59 data in greater detail. Ultimately by understanding the structure of the dual domain systems, their resistance to force can be elucidated. This region of obscurin is of interest due to its binding to titin Ig9/10 domains (Rohs, 1999). Overall, this study could potentially provide a basis to characterize the interaction between titin and obscurin.

Domain-domain independence of Ig58/59 found through NMR analysis

NMR results for Ig58/59 were previously analyzed in our lab and were used in order to more fully compare Ig35/36 results to Ig58/59 results (Caldwell, 2015). The HSQC of the dual domain Ig58/59 construct is well dispersed. By overlaying the dual domain construct with the individual solved Ig58 and Ig59 HSQCs, we could compare the chemical shift changes brought about by the domains being in tandem (Caldwell, 2015). The chemical shift changes were mapped onto a model of Ig58/59, and showed no significant localization of shifts (Figure 5A). Overall, this NMR data analysis of Ig58/59 strongly suggests that the domains do not significantly interact with one another. This result suggests that the domains are independent of each other (Caldwell, 2015).

A.



B.

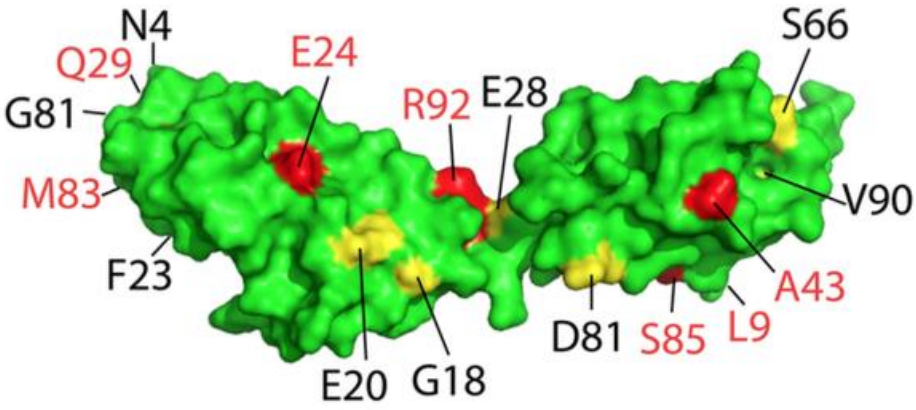


Figure 5. A) A graphical representation of the residues that had significant shifts (two standard deviations shown in yellow; three standard deviations in red) between individual HSQCs and the dual domain HSQC. **B)** Model of Ig58/59 that shows the residues with significant chemical shifts. (Caldwell Thesis, 2015).

The domains showed little interaction with each other as depicted in Figures 5 A and B. This result was unexpected. Due to the short linker region between Ig58/59, it was expected that

this would have an influence on the interactions between the domains, especially in areas near the linker (to the right side of Ig58 and the left side of Ig59 in Figure 5).

RDC shows that Ig58/59 tumbles freely in solution

Residual dipolar couplings were performed in order to extend and refine the NMR analysis. NMR data suggested independence between domain interactions. We used residual dipolar couplings to detect whether the linker might be involved, rather than the domains, in the dual domain system's response to force. RDC compares anisotropic and isotropic samples of Ig58/59 in order to better understand the behavior of the dual domain system in solution. RDCs measure the angles of the different samples relative to one another (since the anisotropic sample is directionally dependent), in order to ultimately find the orientation between them (Figure 6A). The determined Q-factor for the Ig58/59 model was 0.48, meaning that the data did not fit well with the model. Overall, the RDC calculations suggest that the Ig58/59 model tumbles freely in solution and cannot fit under one alignment tensor.

A.

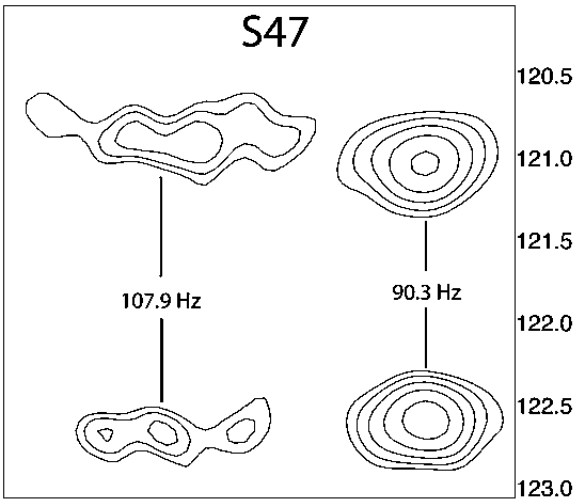


Figure 6. (A) Analysis of RDC values, showing isotropic (right) and anisotropic (left) splittings. **(B)** Table shows Ig58 and Ig59 RDC fitted data. The data was not able to fit the dual-domain model to the individual domains with the same alignment tensor. since the model could not fit under one alignment tensor, this suggests the two domains tumble freely in solution.

B.

Model	Q factor	Number of RDCs
Ig58	0.26	15
Ig59	0.33	17
Ig58/59 model	0.48	32

SAXS shows Ig58/59 exists in an extended conformation

SAXS was also used in order to further the structural analysis on Ig 58/59. Using a Guinier plot to analyze SAXS data, we confirmed the presence of a globular protein. Analysis of SAXS using GNOM was utilized; GNOM is an indirect transform program used for SAXS data processing. GNOM outputted a P(r) vs r distribution plot for the dual domain system, which characterizes the amount of extension between the domains. This analysis was degrees of aggregation (Tassone, n.d.).

The distribution suggested that the two domains exist in a moderately extended configuration. The left skewed peak with a smaller hump protruding from the downward slope indicates the presence of 2 separated domains (as explained in Figure 7). The R_g value of 27.2 Å also indicated that the model existed in two separated domains.

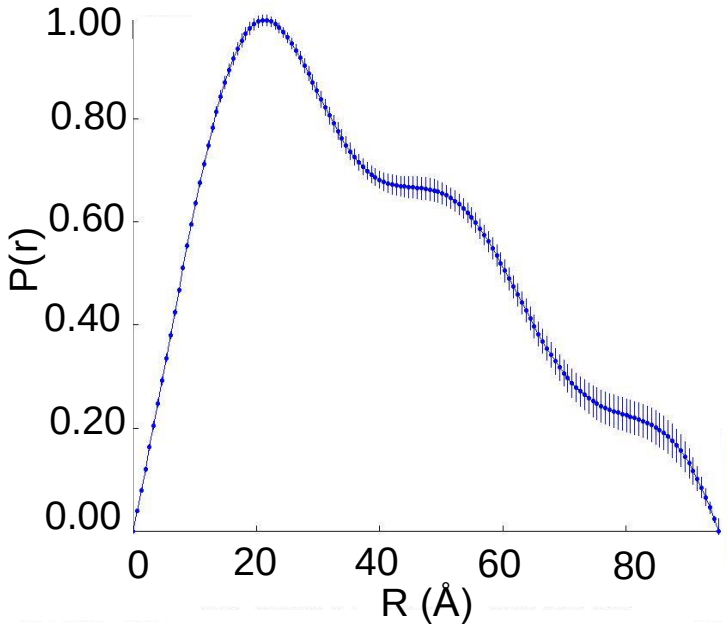


Figure 7. (A) SAXS analysis for Ig58/59; the P(r) vs. R analysis suggests two separated domains (R_g=27.3 Å, D_{max}=95 Å, α=0.81).

Agreement of experimental techniques

Following the previous three different analyses, data from individual RDCs and SAXS analyses were compiled together. This was done in order to create a model that assembled the results together to more thoroughly characterize the behavior of the dual domain structure. The program MultiFoXS was used in order to perform the compilation of SAXS models (Schneidman, 2016). The RDC values were then compared to these SAXS models using PALES, and the four structures that fit best to the data (top 20%) were visualized in pymol (Figure 8). This ensemble clearly shows a dual domain system that rotates freely in solution and has a flexible and more mobile behavior. This is perhaps the best visualization of the remarkable agreement of our experimental data.

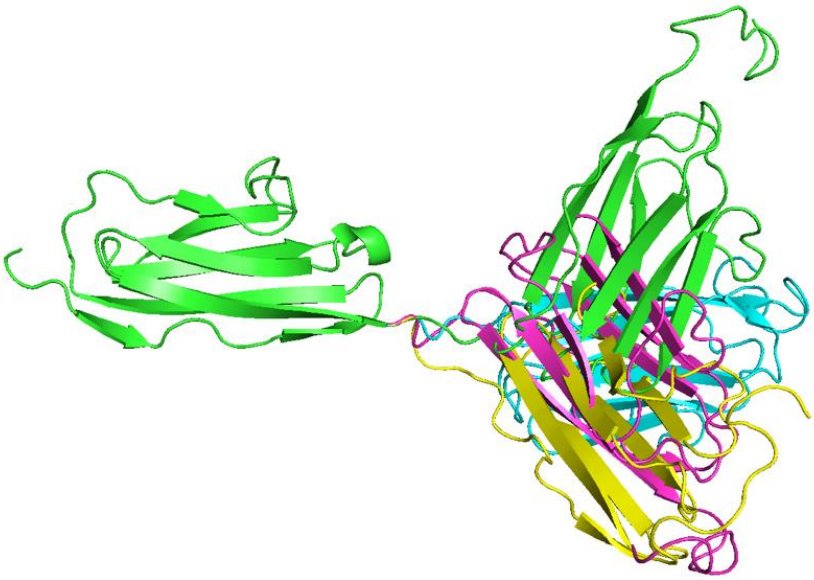


Figure 8. Compiled models of Ig58/59 structures (Q value= 0.41). Models were calculated based on SAXS and RDC data and the top 4 models (representing the top 20% of the data) are shown. The top 4 models show that Ig58/59 has a flexible domain-domain interaction.

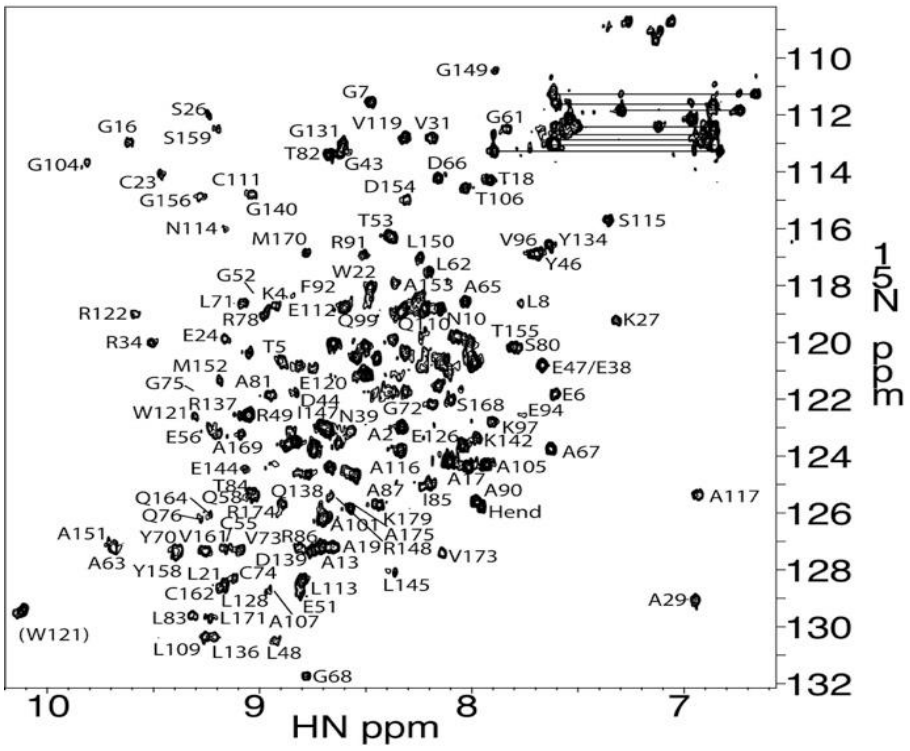
The compositional structure of Ig35/36 using NMR, RDC and SAXS analysis

From the Ig58/59 data, we next reasoned that all of obscurin was very flexible. If this were true, it would be a dramatic departure from titin, which is spring-like. To test this hypothesis, we next applied these same methods to a different obscurin Ig region, Ig35/36.

NMR analysis of Ig35/36 shows domain domain independence

The NMR results for Ig35/36 were compared to previous literature values for individual domains (Wakabayashi, 2005). Since we did not have to first sequence-specifically assign the HSQC, this process was much accelerated for this dual domain structure. Previous values of individual domains of Ig 35 and Ig36 were plotted together on an excel graph (Figure 9 A). The experimental HSQC spectrum of Ig35/36 was analyzed and annotated (Figure 9 B). 155 of the 176 residues were identified and labeled. The spectrum showed a well-dispersed Ig 35/36 construct (Figure 10 B). The peaks were compared to the individual HSQC peaks, and they showed similar compositions with little chemical shift changes (Figures 9 A, B). Overall the NMR analysis suggested that there were little to no domain domain interactions, similar to our Ig58/59 findings. The random distribution of significant chemical shifts throughout the structure suggested little to no correlation between the movement of the amino acids and the dual domain system.

(A)



(B)

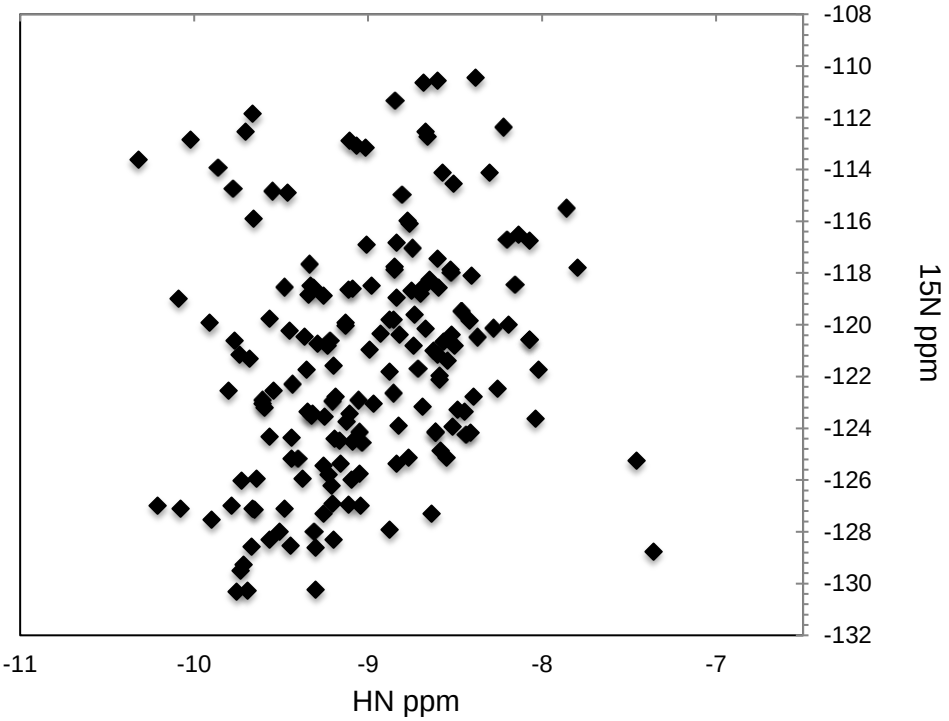
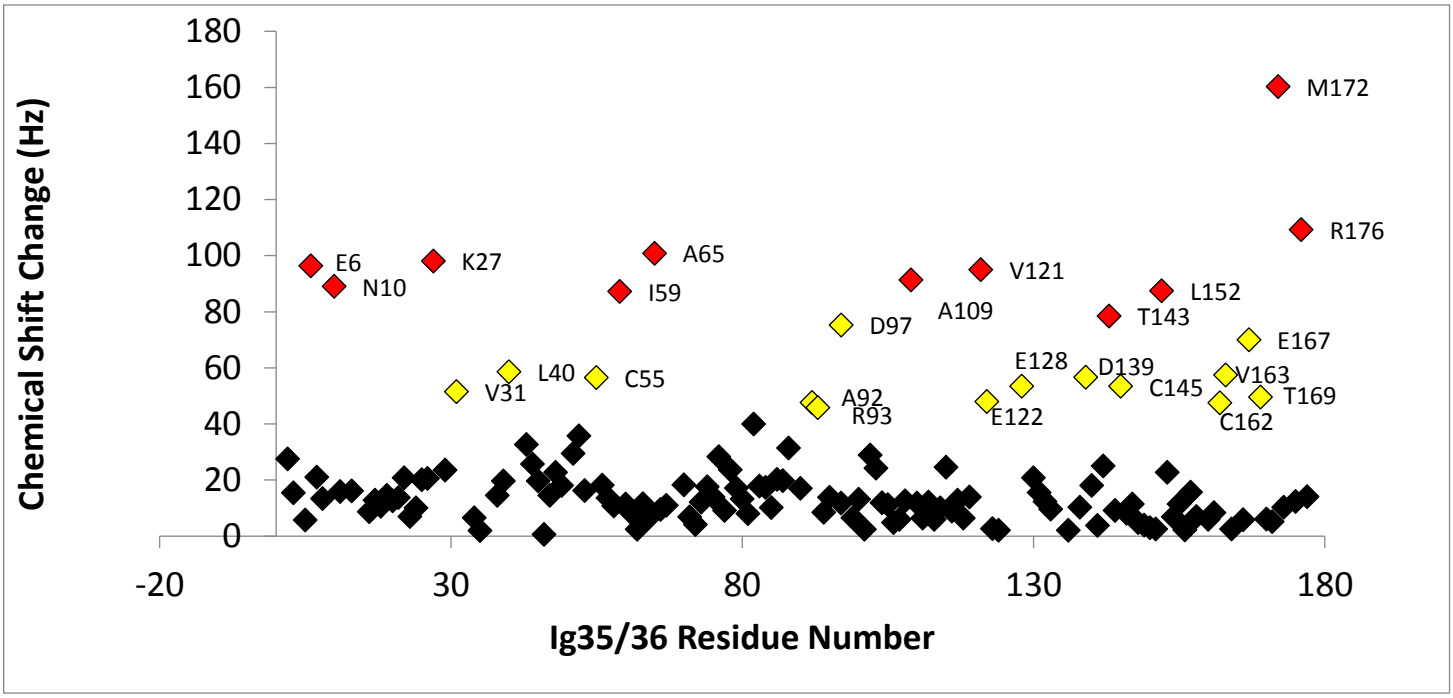


Figure 9. **(A)** Superimposed HSQC shifts, derived from literature values of individual Ig35 and Ig36 NMR shifts and plotted using excel. **(B)** Experimental HSQC of the tandem Ig35/36 construct. Note, the dual-domain system has similar chemical shifts when compared to (A).

(A)



(B)

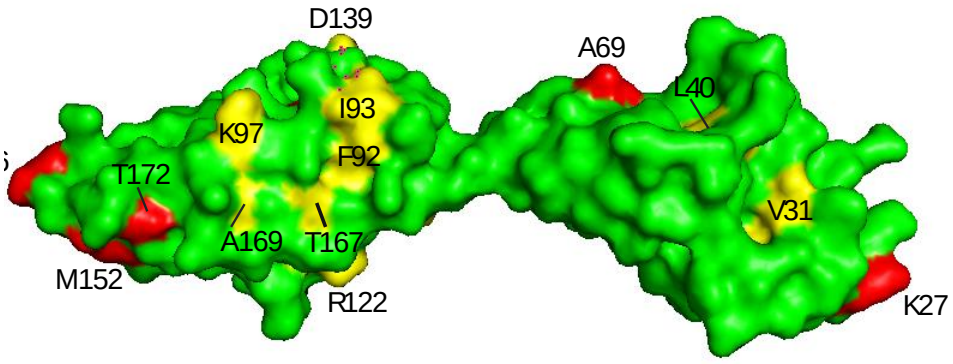


Figure 10. (A) Chemical shift changes of the Ig35/36 construct, as compared to the individual Ig35 and Ig36 constructs. **(B)** Model of Ig35/36 with significant chemical shifts (two standard deviations shown in yellow; three standard deviations in red). Note, the random distribution on the surface of the model suggests that the domains do not significantly interact with each other.

RDC calculations show that Ig35/36 tumbles as a unit in solution

Through further analysis, we attempted to determine whether the linker region facilitated interactions between the domains. Residual dipolar couplings expanded the NMR data and provided us with long-range angle restraints. The residual dipolar coupling values compared anisotropic and isotropic samples of Ig35/36 to determine the orientation of the system and understand its behavior in solution (Figure 11A). Like with Ig58/59, here individual models were calculated from the RDC data. However unlike Ig58/59, where the combined Q-factor was substantively worse than those calculated from the individual domains, here the Q-factor was roughly the average between these two domains (Figure 11B). This signifies that a single alignment tensor can be reasonably fit to Ig35/36. In contrast, NMR analysis shows that the domains of Ig35/36 are independent of one another. Combining the analyses together suggests that while the domains do not form a tight interaction with each other, but the domains tumble in a fixed orientation relative to one another in solution.

(A)

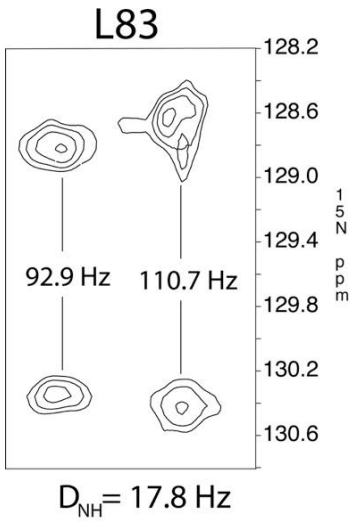


Figure 11. (A) Analysis of RDC values, showing isotropic (right) and anisotropic (left) splittings. **(B)** Table shows Ig35 and Ig36 RDC data fit the published structures reasonably well. The data also fit the dual-domain model roughly as well as the individual domains, with the same alignment tensor. This suggests the two domains tumble in solution as one unit.

(B)

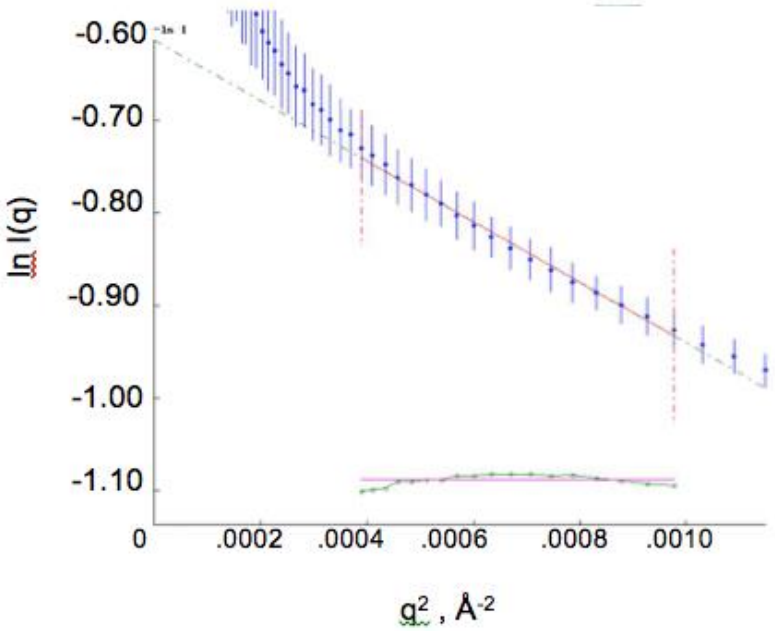
Model	Q factor	Number of RDCs
Ig35 literature	0.5	17
Ig36 literature	0.17	22
Ig35/36 model	0.17-0.38	39

data analyses of Ig58/59 and Ig35/36 together show that even though obscurin linker regions are uniformly short, they behave differently. This result disproves our original hypothesis that obscurin behaves flexibly. Instead, some linkers create stiff links between domains (like Ig35/36), while others create flexible linkers (like Ig58/59). From here we reasoned that the linker region composition plays a significant role in modulating obscurin stiffness, while the domains themselves seem to not interact with their neighboring domains.

SAXS analysis suggests Ig35/36 exists in a semi-extended conformation

To independently verify our NMR data, we next turned to SAXS. SAXS analysis also helped to evaluate the role of the linker on the behavior of the dual domain system. A Guiner plot shows the presence of a globular protein (Figure 12A). The first 55 points were removed from the analysis due to aggregation. The P(r) vs r distribution plot showed a left skewed graph, which suggested that the domains were in close proximity with each other (Figure 12B). The R_g value of 22.7 Å also suggested the domains existing close together, when compared to Ig58/59 (27.3 Å). Overall, The SAXS data demonstrates that Ig 35/36 exists in a semi-extended conformation and looks like a single ovoid molecule, instead of two independently rotating domains connected by a linker like Ig58/59.

(A)



(B)

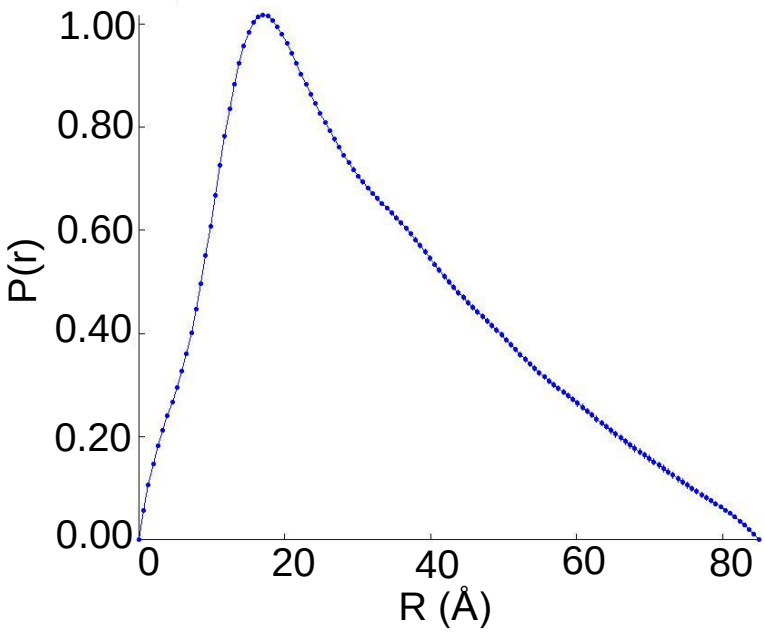


Figure 12. (A) Guinier plot of Ig35/36 **(B)** SAXS analysis for Ig35/36; the P(r) vs. R analysis suggests an elongated two-domain structure, with the domains closely associated (R_g=22.7, D_{max}=85, α=0.64).

Agreement of experimental techniques

Like with Ig58/59, we next conducted a joint SAXS/RDC analysis of Ig35/36 using MultiFoXS (Schneidman-Duhovny, 2016).The four structures that best fit to the data (top 20%) were visualized in pymol (Figure 13). This ensemble shows a dual domain system that has a fixed and semi-rigid characterization.

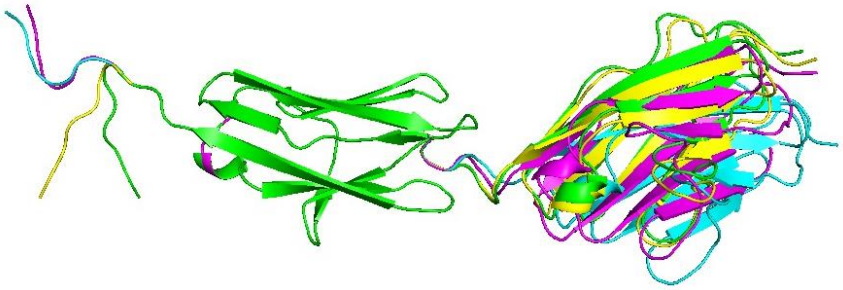


Figure 13. Compiled models of Ig35/36 structures (Q value= 0.38). Models were calculated based on SAXS and RDC data and the top 4 models (representing the top 20% of the data) are shown. The top 4 models show that Ig35/36 has a fixed domain-domain orientation.

DISCUSSION

By using NMR, RDC and SAXS analysis in tandem to study Ig dual domains, we were able to better understand and characterize the domain-linker-domain models in obscurin. NMR analysis showed similar results regarding domain/domain independency for both Ig58/59 and Ig35/36. Based on previous studies on titin, these NMR results showing little interdomain interactions were unexpected (Andres, 2016 and von Castlemur, 2008). Domains within titin are found to have more interactions and, if the linker is short, have a significant dependency on each other (von Castlemur, 2008). In obscurin however, we see no such domain-domain interactions (Figures 4, 9). Through further analysis with RDC and SAXS, we found different behavioral characteristics for Ig35/36 domains and Ig 58/59 domains. Regarding RDC, Ig58/59 could not be solved using one alignment tensor, suggesting that the molecule tumbled freely in solution. In contrast, RDC calculations on Ig35/36 were solved using one alignment tensor, suggesting that this system tumbles as a unit in solution. SAXS analysis further supported these differing behaviors Ig58/59 exists as two separate domains, while Ig35/36 behaves more like a single domain.

This data suggests that Ig58/59 behaves as a freely rotating, mobile system while Ig35/36 exists as a more fixed and rigid structure. Because NMR data suggests that the domains are independent of each other, we speculate that the linker between the domains might dictate the behavior of these domains. The linker of Ig58/59 contains a glycine residue and a tryptophan residue (GW), while Ig35/36 contains a leucine residue and a proline residue (LP). Substituting a glycine to a proline is likely to restrict the motion; we suggest the linker in Ig58/59 allows more flexibility while the linker in Ig35/36 causes a more rigid characteristic. Further analysis of the linker hypothesis will be discussed in the discussion of Chapter 2.

CHAPTER 2 COMPUTATIONAL SIMULATIONS AND COMPARISONS OF FORCE RESISTANCE

METHODS

SMD

Steered molecular dynamics was performed using pdb files of Ig domains (Ig3536-v3.pdb, Ig3536-GW.pdb., Ig5859-2.pdb, Ig5859-LP.pdb). SMD simulations were performed with the PMEMD module of the Amber 12 MD software package, using AMBERff12SB force field and in explicit solvent. SMD simulations were run in a manner similar to (Caldwell, 2015). Mutations were included in domains with YASARA using the 'mutate' function. Structures were optimized, heated, accounted for density, and equilibrated prior to pulling simulations. the cutoff distances were set to 8.0 Å.

For equilibrium simulations, a constant temperature of 300 K was imposed using a Langevin thermostat with a collision frequency of 1 ps^{-1}. A constant velocity of 1.0 Å/ns (0.1 m/s) was used in order to simulate biologically relevant pulling forces. The SMD spring constant (rk2) was set to 0.2 and the temperature used was 300.0 K (Caldwell, 2015). (Mcgee, n.d.)

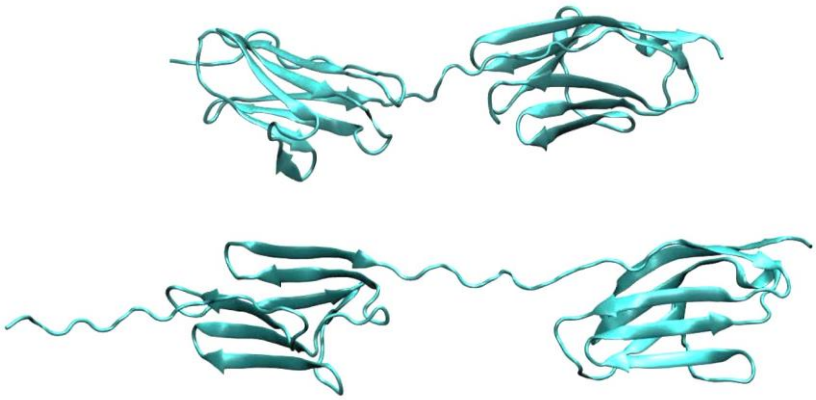
Analysis was visualized by creating force vs. distance plots and angle vs. distance plots using Gnuplot. VMD was utilized in order to create movie files. For further details, see appendix for detailed instructions done in explicit solvent.

RESULTS

SMD simulations on the force resistance of Ig58/39 and Ig35/36

Obscurin exists predominantly in muscle cells, and thus is subjected to physical extension and contraction. Our experimental data suggests that different parts of obscurin have different flexibilities. To more closely examine this facet of obscurin dynamics we used steered molecular dynamics (SMD) (Figure 14). In chapter 1, Ig35/36 showed a rigid and fixed behavior while Ig58/59 showed a more flexible and mobile behavior. A limitation to this technique is that it only looks at the models in solution and not in motion. In physiological environments, muscle proteins respond to different magnitudes of force. For example, in muscles, obscurin will go through periods of compression and stretch. Ideally, we want to evaluate obscurin as it exists in motion in its natural state within the intact sarcomere. SMD is a decent substitute for this *in vivo* testing, since stretch and compression mirror the physical action of muscle. We used SMD tests to simulate force and observe the models responses to stretch. Explicit solvent is used, which is advantageous because it accounts for the surrounding water molecules in the system and can simulate for phenomena such as hydrodynamic drag. In the simulations, one domain was pulled at the end at a force of 1.0 Å/ns (0.1 m/s) in order to represent a biologically realistic stretching movement (Figure 14).

Before pulling



After pulling

Figure 14.
Schematic of the pulling simulations performed on SMD. This is a representation of Ig35/36 being pulled at 1.0 Å /ns. This shows the change in linker length across simulation time.

Pushing simulations were also performed on both Ig dual domains in order to validate that the structures did not have an unrealistic conformation. Between -20.0 to 0.0 Å

compressions, both Ig58/59 and Ig35/36 systems showed no domain-domain resistance and ultimately gives a better sense of the energy well of the structure. For the pulling simulations, from 0.0 to 45.0 Å the domain stretching showed increase work as distance increased. Pulling was performed until the force caused the domains to unfold, which occurred beyond approximately 50.0 Å.

A comparison of Ig58/59 and Ig35/36 shows significant differences to stretch. (Figure 15 A, B). Ig58/59 begins resisting force slightly from 10.0-20.0 Å stretch, and after 20.0 Å begins to resist force more strongly until becoming fully stretched at ~50 Å (Figure 15 A, B). For Ig58/59 stretching, about 60 kcal/mol of work is put into the system in order to fully extend the model. In contrast, Ig35/36 begins resisting force strongly starting at 5.0 Å of stretch and continues to require a large amount of work until becoming fully extended at ~45.0 Å. For Ig35/36 stretching, about 180 kcal/mol of work is put into the system in order to fully extend the model.

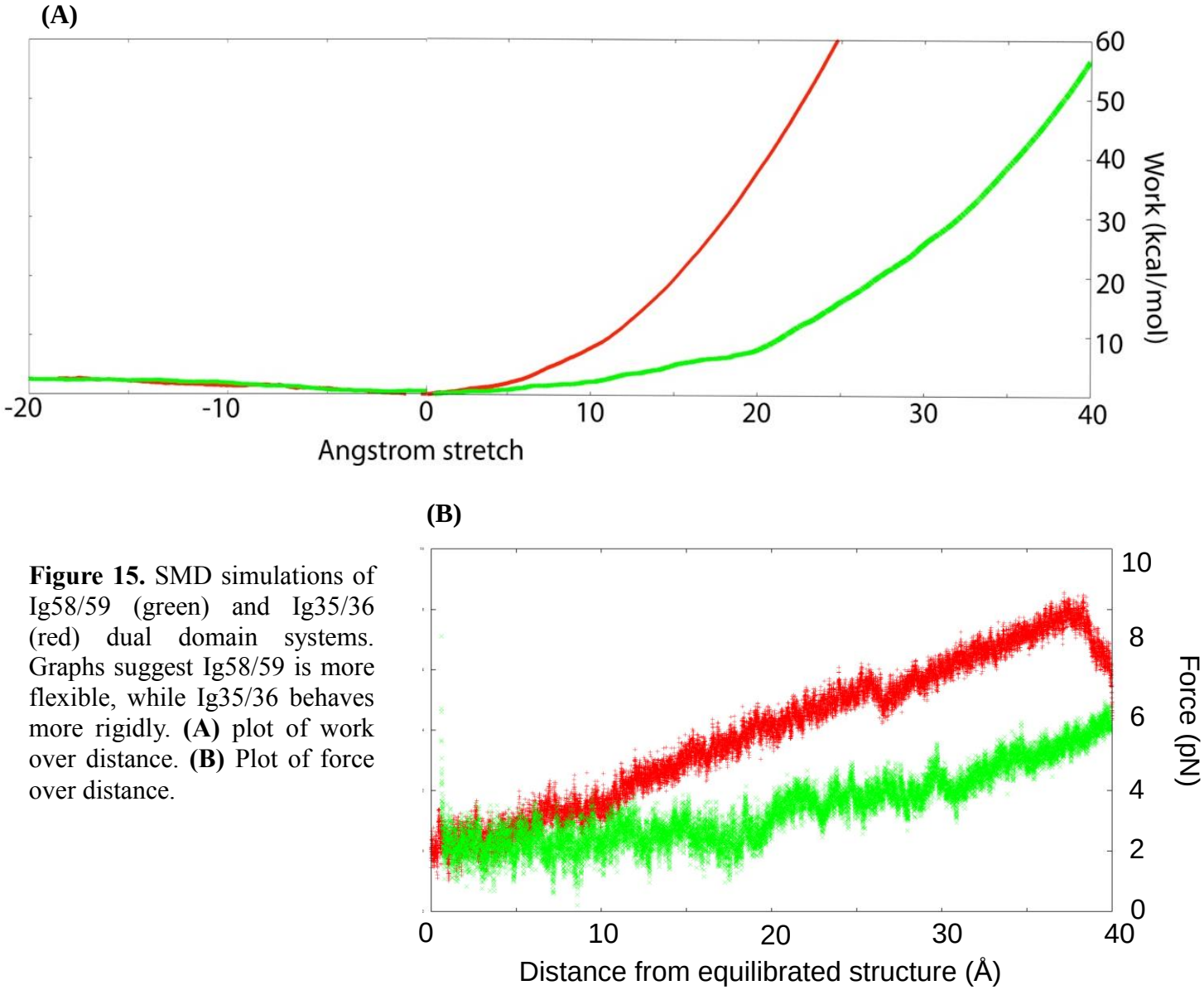


Figure 15. SMD simulations of Ig58/59 (green) and Ig35/36 (red) dual domain systems. Graphs suggest Ig58/59 is more flexible, while Ig35/36 behaves more rigidly. **(A)** plot of work over distance. **(B)** Plot of force over distance.

While the slopes of the of the graphs for Ig58/59 and Ig35/36 likely vary due to the domain composition, the angstrom distance where the amount of work starts to increase is primarily due to the linker region between the domain systems. This result suggests that the linker in Ig domains might be involved in the characterizing force resistance.

SMD offers an advantage of high-resolution visualization of these pulling events. To gain a better understanding of how these domains behaved when stretched, we measured the domain/domain angles and the domain twist (Figure 16, Figure 17). The domain/domain angle measurement characterizes the relative ease of extension the model has when being

stretched. The plot can be correlated with the work vs. distance plot show how much the angle is increasing between the domains with how work is being done on the model. The dihedral angles characterized the twisting of the dual domain system. Twist can characterize how rigid and resistant the molecule is to movement, and how the domains behave relative to one another.

The domain/domain angles in Ig58/59 show a wide-angle range when little to no work is involved, occurring from 0.0-30.0 Å (Figures 16 A, C). The twist dihedral angle shows the domains twisting frequently upon stretching, occurring between 30.0-45.0 Å (Figures 16 B, C). This suggests that Ig58/59 can accommodate a lot of stretch and resembles a rope like structure.

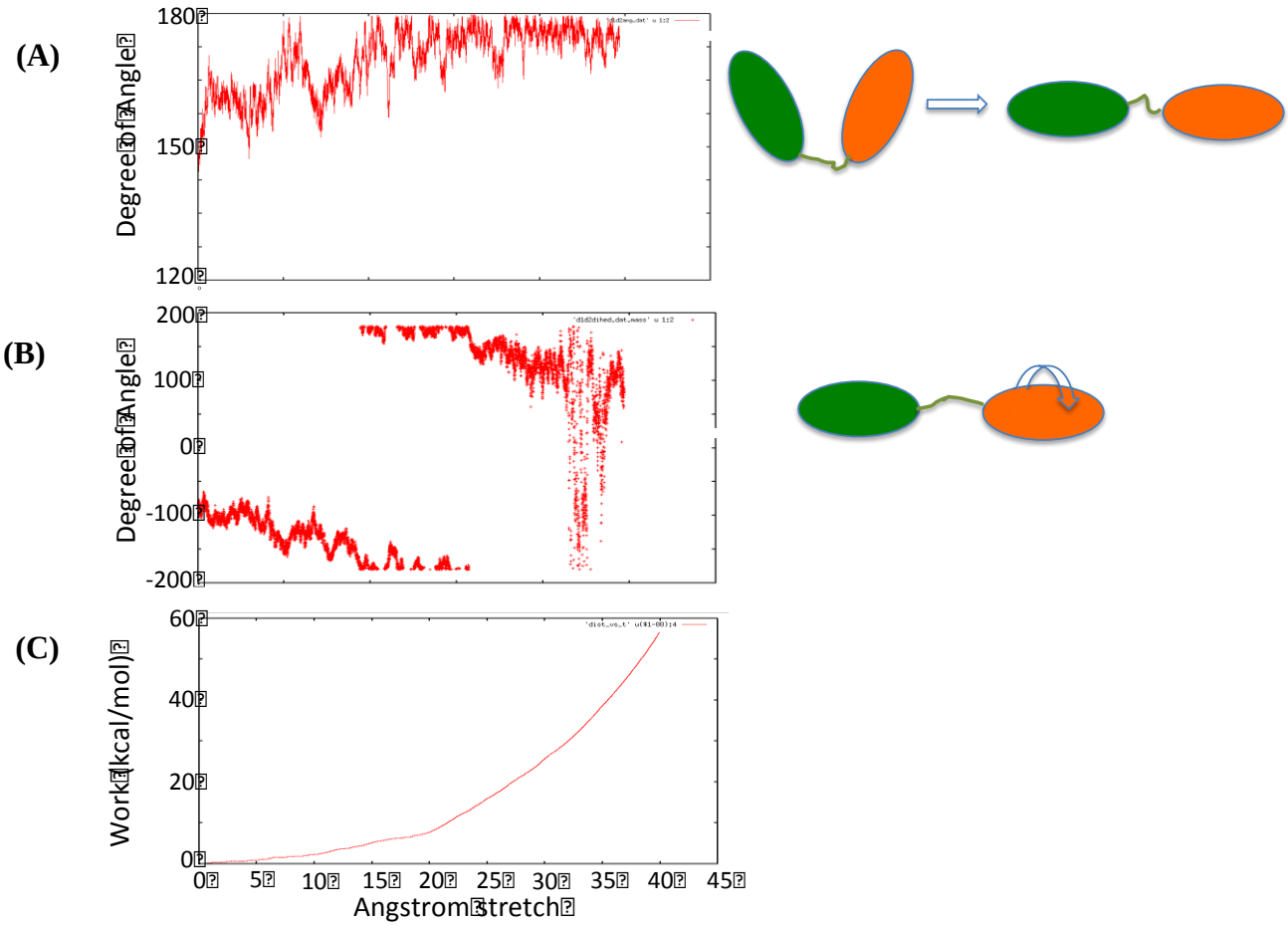


Figure 16. (A) Domain-domain angles of Ig58/59 **(B)** Twist (dihedral) angles of Ig58/59 **(C)** Work vs. Angstrom stretch of Ig 58/59 ; it takes 60 kcal/mol of work to fully stretch Ig58/59.

In contrast, the domain/domain angles in Ig35/36 show the angle increasing when work is introduced into the system, occurring from 5.0-35.0 Å (Figures 17 A, C). In other words, Ig35/36 angle increase correlates strongly with the work involved in stretching the domains apart. The twist (dihedral) angle in Ig35/36 shows how one domain can twist relative to the other. This twisting event in Ig35/36 occurs just prior to the domains being fully stretched, as seen from 35.0-45.0 Å (Figures 17 B, C). This suggests that Ig35/36 cannot accommodate much stretch in its structure, and behaves more rigidly.

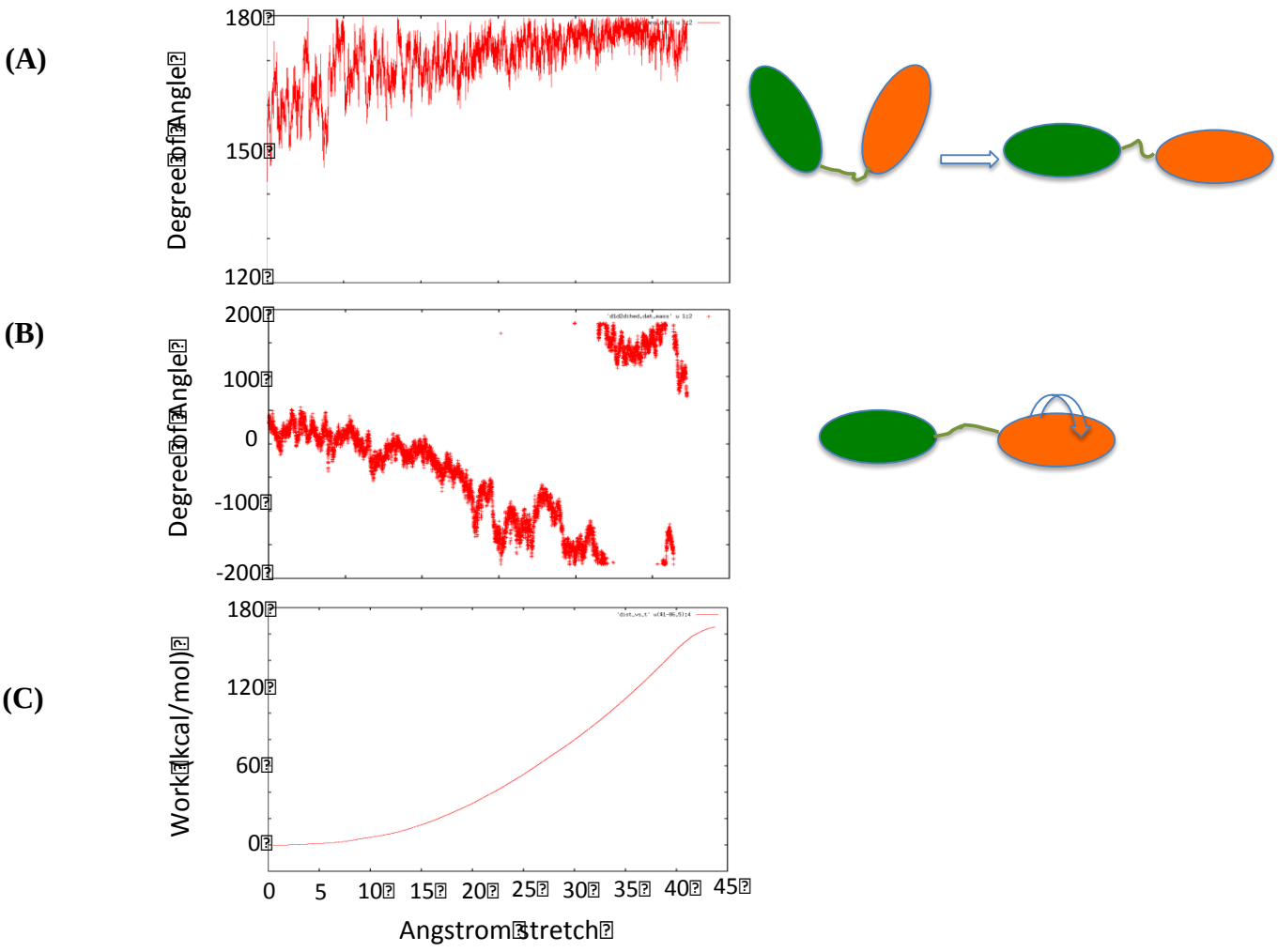


Figure 17. (A) Domain-domain angles of Ig35/36 **(B)** Twist (dihedral) angles of Ig35/36 **(C)** Work vs. Angstrom stretch of Ig35/36; it takes 180 kcal/mol of work to fully stretch Ig35/36.

SMD simulations of mutated linker sequences in Ig58/59 and Ig35/36

SMD simulations in Results Part I support the experimental findings in Chapter 1. Results showed in order to pull the domains apart from each other, Ig58/59 required less work, while Ig35/36 required more work. This shows that Ig35/36 resists force more strongly compared to Ig58/59. Following the initial simulations on the domain systems we hypothesized the linker sequence had a role in characterizing the different responses to stretch. We next evaluated Ig58/59 and Ig35/36 domains with mutated linkers. To do this, we mutated the residues of interest in the linker region for Ig58/59 and Ig35/36 and ran the same simulations.

In Ig35/36, we mutated the linker region RALPAR to RAGWAR, in order to resemble the linker sequence of Ig58/59. We compared the mutated sequence to the wild type sequences to show how the mutation alters the behavioral response. Mutated Ig35/36 shows a significantly different response compared to the wild type model (Figure 18). The response to force shifts to the right; the amount of work starts to increase in between 15.0-30.0 Å compared to 5.0-15.0 Å. This indicates that the mutation causes the model to resemble the behavior of Ig58/59 and suggests that the mutated Ig35/36 resists force less strongly.

In Ig58/59, we mutated the linker region RGWRLE to RLPRLE in order to resemble the linker sequence of Ig35/36. We once again compared the mutated and wildtype sequences. Mutated Ig58/59 shows less change compared to mutated Ig35/36 (Figure 18). The amount of work is almost identical to the original Ig58/59; the increase starts at approximately 20.0 Å which is in the range of the original SMD simulation of I58/59 (15.0-30.0 Å). This indicates that the residue swaps in Ig58/59 have little influence on the behavior than we predicted.

This mutation experiment has fewer variables than comparing completely different domains. Comparing different domains introduces extra variables that we may not recognize or be able to account for. By making mutations in SMD, we work with very similar systems having minor changes, and can thus obtain more information about how the linker specifically affects stretch. This should reduce conflicting variables.

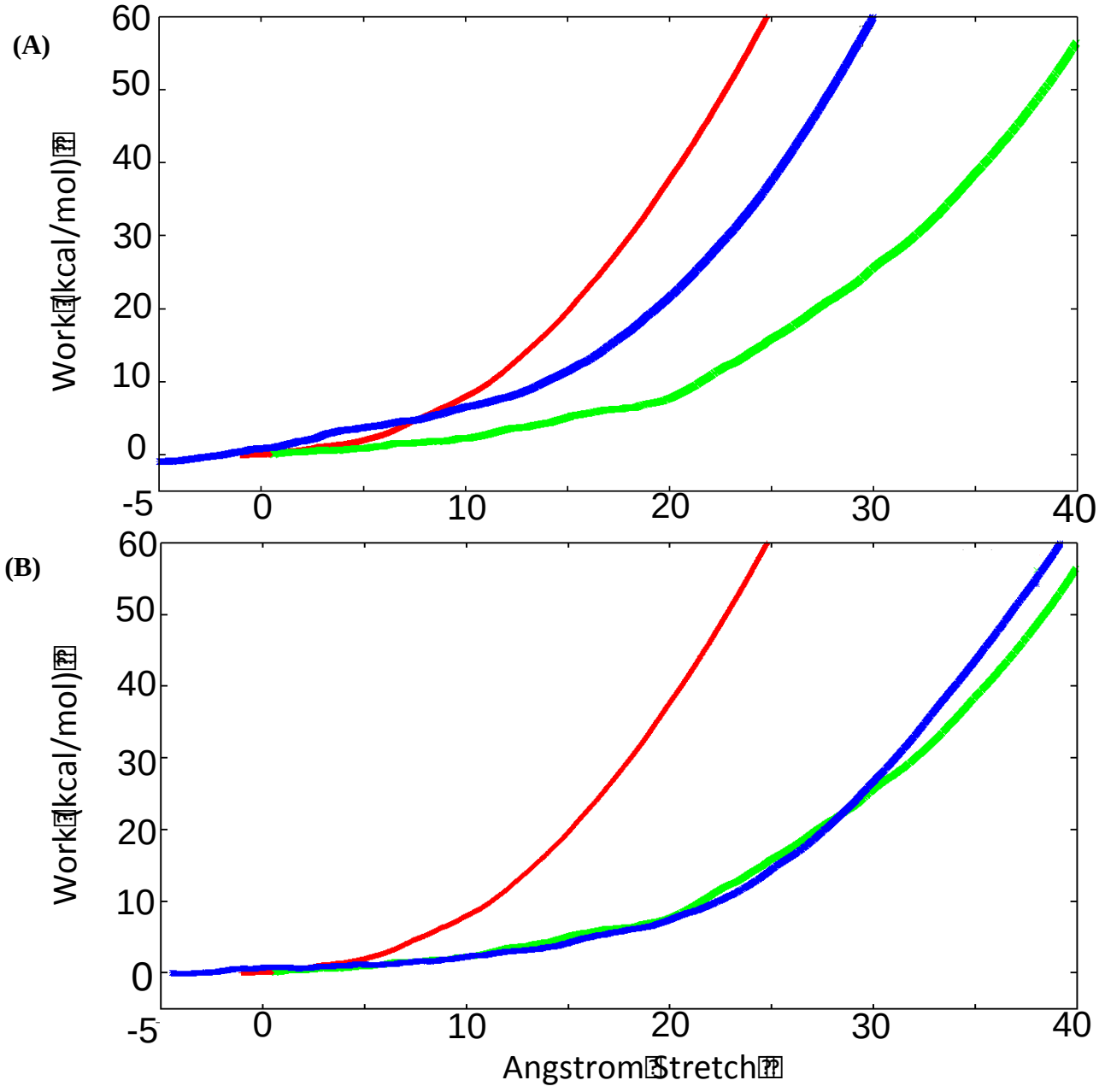


Figure 18. **(A)** SMD simulation of mutated Ig35/36 model (blue). The leucine and proline residues in the linker region were substituted with glycine and tryptophan residues. **(B)** SMD simulation of mutated Ig58/59 model (blue). The glycine and tryptophan residues were substituted for leucine and proline

DISCUSSION

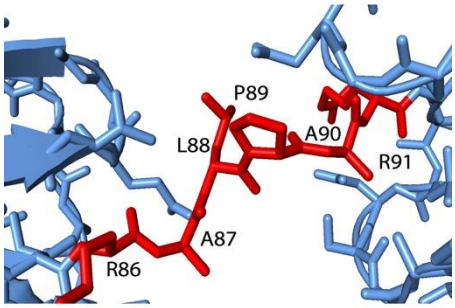
From the results in Chapters 1 and 2, we observe flexibility and mobility in the Ig58/59 structure, and rigidity in the Ig35/36 structure. Chapter 1 used experimental techniques to evaluate Ig58/59 to find domain-domain independence, a freely tumbling model in solution, and an extended conformation. In contrast, Ig35/36 exhibited domain-domain independence, and an overall system in a semi-extended conformation that tumbled as a unit in solution. Supporting these results, SMD showed Ig58/59 resisting force less strongly compared to Ig35/36. Thus, obscurin behaves as a variable force resistor- some parts of the molecule are spring-like and some parts are rope-like. This distinction seems to come from the linker region's composition. This is different from titin, where the flexibility is mostly based on linker length (von Castlemur, 2008)

By comparing the simulations to the experimental data in Chapter 1, we found that more rigid systems have higher resistance to force while more flexible systems resist force more readily. In correlation to RDC and SAXS experimental data, the computational data supported the different behaviors of the systems. The linker region plays a significant role in the protein's response to stretch (Figure 18). We suggest the increase in work is caused by the linker beginning to stretch. SMD shows the work increasing in Ig58/59 from 15.0-30.0 Å while in Ig35/36, from 5.0-15.0 Å (Figure 15). Thus the graph of Ig58/59 and Ig35/36 being pulled shows that the linker sequence in Ig58/59 begins to stretch at a larger angstrom distance compared to Ig35/36.

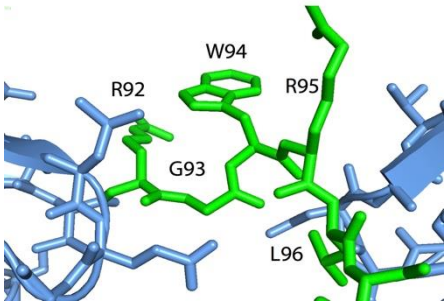
Following this hypothesis, we evaluate the sequences of the linker regions (Figure 19 A, B). The linker sequence of Ig58/59 contains a glycine and a tryptophan, which could potentially introduce the flexibility into the system. The tryptophan residue could cause the linker to have a kink, and might lead to increased rotation. Additionally, the glycine residue might also facilitate rotation/ increase flexibility. In contrast, the linker sequence of Ig35/36

contains a proline and a leucine which could cause rigidity. The structure of the bonds in the proline residue, might cause the linker to adopt a more anchored and fixed structure. In other words, the structure of the proline residue might enforce a specific geometry in the linker region.

(A)



(B)



(C)

Domain	Linker Sequence	Behavior
Ig35/36	RALPAR	Fixed
Ig34/35	KALPAK	TBD
Ig58/59	RGWRLE	Mobile
Ig57/58	KVTAK	TBD

Figure 19. (A) Linker region of modeled Ig35/36. **(B)** Linker region of modeled Ig58/59. **(C)** Table of the linker regions for four dual domain systems of interest.

We mutated the linker regions of the structures, as shown in Figures 18 A and B, and observed a significant change in behavior in the mutated Ig35/36 model. This supports the hypothesis that the linker region dictates the behavior and flexibility of the domain-domain structure. The mutated Ig58/59 model did not produce as significant of a change, and needs further evaluation (Figure 18 B).

Overall, the conclusions gathered from this study indicate that the linker sequence plays a significant role in the behavior of the domain. For future directions we propose evaluating the

linker region further by mutating the entire linker sequence. We will add in the entire linker sequence and do SMD simulations. Also, we have only done the linker mutations one time. These results need to be repeated, especially since adding LP to the Ig58/59 linker only slightly increased the stiffness of this molecule. Finally, we also plan to evaluate other dual domain structures within obscurin (Figure 19 C). All of these projects are ongoing, and are being worked on in the Wright Lab.

APPENDIX

Molecular Dynamics Simulation instructions

log in: ssh –X (ip address)
(ip address for the supercomputer that we are using)
:then enter password

make a folder: mkdir (name of folder)
enter into a folder: cd
list folders: ls
take contents of file and print to screen: cat
to move up a folder: ../
to remove a folder: rm

copy a pdb from the computer to a supercomputer:
Open the desktop where the .pdb file is located on X11
to copy this .pdb file into the supercomputer type:
scp 3435.pdb wrightnt@134.126.158.145:/home/wrightnt/aidan/Ig34-39

In order to set the conditions for the molecule we must use the program XLeap which will neutralize, set parameters, and put in either explicit/implicit solvent.
This program will basically tell the atoms how to interact:

AMBER:
EXPLICIT SOLVENT
to open amber (this will open AMBER 12SB forcefield):
type:
xleap –x –f $AMBERHOME/dat/leap/cmd/leaprc.ff12SB
type:
model
click:
> File> Load PDB File (click file and load pdb project that you want to load)

in order to neutralize, type:
charge model (this will tell you the charge)
addions model K+ (how ever many to neutralize)
addions model Cl-

in order to put in explicit solvent, type:
solvateOct model TIP3PBOX 30.0

in order to set parameters, type:
loadamberparams frcmod.ionsjc_tip3p

to save files that you created, type:
saveamberparm model modelname.prmtop modelname.inpcrd

inpcrd- input coordinates
input prmtop => Amberfile
inpcrd => Amber7 restart file

close AMBER by typing quit

IMPLICIT SOLVENT
to open amber (this will open AMBER 12SB forcefield):
type:
xleap –x –f $AMBERHOME/dat/leap/cmd/leaprc.ff12SB
type:
model
click:
> File> Load PDB File (click file and load pdb project that you want to load)

to set the default PBRadii mbondi3
saveamberparm model model name.prmtop modelname.inpcrd

close AMBER by typing quit

Running the Simulation
IMPLICIT SOLVENT:
to run the stretch simulation requires 3 steps:
1. Minimize structure
2. Equilibrate
3. Pull

*these files will be created in emacs:
some tips:
make sure to hit enter after the backslah
ntb=0 implies implicit solvent
igb=8 gives implicit solvation model, designed for ff12SB
cut=99 which tells the program how many angstroms away each atom can see
to save in emacs: control x, control s
to exit emacs: control x, control c

Minimization:
type:
emacs modelname.in (this is an input file)
in emacs, type:
modelname-equilibration
&cntrl
imin=1,
maxcyc=5000,
ncyc=500,
ntb =0,
igb =8,
cut =99,
/

exit emacs

the next step is to start the minimization, this requires a command:
pmemd.cuda –O –i modelname.in –o modelname.out –c modelname.inpcrd –p modelname.prmtop –r modelname-min.rst

If you want to see the molecule in VMD:
type:
vmd modelname.prmtop
in VMD: file- new molecule-
load prmtop file
file rst. file
file type: Amber 7 restart

Equilibration:
emacs modelname-equil.in
model name-equil
&cntrl
irest = 0, ntx = 1, ig=-1,
imin=0, ntb=0,
igb=8, ntpr= 1000, ntwx=1000
ntt = 3, gamma_ln = 1.0,
tempi = 0.0, temp0 = 300.0,
nstlim = 10000000, dt = 0.002,
cut =100,
ntwr = 2000,
ntc = 2, ntf = 2,
ioutfm=1,ntxo=2,
/
Save and quit
** If you are starting equil. for the first time (for that model), irest=0 and ntx=1. If restarting equil. irest=1 and ntx=5

emacs submit-script.sh (which has been copied into folder)
at bottom:
the next step is to start the equilibration, this requires a command:
pmemd.cuda –O –i modelname-equil.in –o modelname-equil.out –c modelname-min.rst –p modelname.prmtop –r modelname-equil.rst –x modelname-equil.mdcrd

to see if equilibration is done:
vmd modelname.prmtop
file-new molecule
prmtop.load .mdcrd file file type: NetCDF, load all at once
Extensions- Analysis- RMSD trajectory tool
protein (whole thing) or resid # to # (certain residues)
check backbone and plot
align RMSD

Pulling or compression:
emacs modelname-equil.in

```
        model name-equil
        &cntrl
        irest=1,ntx=5, ig=-1
        imin=0, ntb=0
        igb=8, ntpr=1000, ntwx=1000
        ntt=3, gamma_ln=1.0
        tempi=300.0, temp0=300.0
        nstlim=100000000, dt=0.001
        cut=999
     jar=1
     ntwr=1000
    NTC=2, NtF=2
    ioutfm=1, ntxo=2
    /
   &wt type='DUMPFREQ',  istep=1000, /
   &wt type='END', /
  DISANG= dist.RST
  DUMPAVE = dist_vs_t
  LISTIN= POUT
  LISTOUT=POUT
```

write dist.RST file:

For constant velocity:

&rst iat=___, r2=___, r2a=___, rk2=____

iat is the two atoms which are being pulled- choose alpha carbon in VMD
r2 is the starting distance (vmd-hold 2 while clicking both atoms)
r2a is the final distane
rk2 is the spring constant

to run the simulation:
pmemd.cuda –O –i modelname-equil.in –o modelname-pull.out –c modelname-equil.rst –p modelname.prmtop –r modelname-pull.rst –x modelname-pull.nc

EXPLICIT SOLVENT:
To run stretch simulation:
5 steps-

1. Optimise structure
2. Heating calculation
3. Density- NPT
4. Equilibration (2days)
5. Pulling (3-5 days)

3435minimization example

```
 &cntrl
  imin   = 1,
  maxcyc = 1000,
  ncyc   = 500,
  ntb    = 1,
  cut    = 10.0
/
```

pmemd.cuda –O –i modelname.in –o model name.out –c modelname.inpcrd –p modelname.prmtop –r modelname-min.rst

3435Heating example
3435heating
 &cntrl
  imin   = 0,
  irest = 0, ig=-1,
  ntx   = 1,
  ntb   = 1,
  cut   = 10.0,
ntr =1,
  ntc   = 2,
  ntf   = 2,
  tempi = 0.0,
  temp0 = 300.0,
  ntt   = 3,
  gamma_ln = 1.0,
  nstlim = 10000, dt = 0.002
  ntpr = 100, ntwx = 100, ntwr = -500
ioutfm=1, ntxo=2
 /

pmemd.cuda -0 –I 3536heat.in –o Ig3536heat.out –c Ig3536-small3.inpcrd-p Ig3536-small3.prmtop –r Ig3536heat.rst

3435density
 &cntrl
  imin = 0, irest = 1, ntx = 5,
  ntb = 2, pres0 = 1.0, ntp = 1,
  taup = 2.0,
  cut = 10.0, ntr = 0,
  ntc = 2, ntf = 2,
  tempi = 310.0, temp0 = 310.0,
  ntt = 3, gamma_ln = 1.0,
  nstlim = 50000, dt = 0.002,
  ntpr = 500, ntwx = 500, ntwr = 1000,
ioutfm=1,ntxo=2
ig=-1
/

pmemd.cuda -O -i Ig3536density.in -o Ig3536density.out -c Ig3536heat.rst_500 -p Ig3536-small3.prmtop -r Ig3536density.rst -x Ig3536density.nc

Ig3536-equil.in
&cntrl
  imin = 0, irest = 1, ntx = 5,
  ntb = 1,
  cut = 8.0, ntr = 0,

```
  ntc = 2, ntf = 2,
  tempi = 310.0, temp0 = 310.0,
  ntt = 3, gamma_ln = 1.0,
  nstlim = 10000000, dt = 0.002,
  ntpr = 2000, ntwx = 2000, ntwr = 2000,
ioutfm=1,ntxo=2,
ig=-1
/
```

pmemd.cuda -O -i ig3536-equil.in -o Ig3536-equil.out -c Ig3536density.rst -p Ig3536-small3.prmtop -r Ig3536-equil.rst -x Ig3536-equil.mdcrd

Pulling or compression:

```
emacs modelname-equil.in
        model name-equil
        &cntrl
        irest=1,ntx=5, ig=-1
        imin=0, ntb=1,
        igb=0, ntpr=1000, ntwx=1000
        ntt=3, gamma_ln=1.0
        tempi=300.0, temp0=300.0
        nstlim=100000000, dt=0.002
        cut=8
     jar=1
     ntwr=1000
    NTC=2, NtF=2
    ioutfm=1, ntxo=2
    /
   &wt type='DUMPFREQ',  istep=1000, /
   &wt type='END', /
  DISANG= dist.RST
  DUMPAVE = dist_vs_t
  LISTIN= POUT
  LISTOUT=POUT
```

To run pulling:

pmemd.cuda –O –i Ig3435equilibration –o 3435.pull.out –c 3435-small-equil.rst –p 3435-small.prmtop –r 3435-small-pull.rst –x 3435-small-pull.nc &

To run ptraj scripts:

dist.RST

```
&rst iat=13,2690,
r2=84.36,
r2a=184.36,
rk2=10.,/
```

ptraj runs:

ptraj.script.distance
trajin 3435-small-pull.nc
distance l1dist :86 :91 out l1dist.dat
distance d1dist :5 :85 out d1dist.dat
distance d2dist :92 :174 out d2dist.dat

cpptraj modelname.prmtop <ptraj.script.distance> distance.out

ptraj.script.angles-
trajin 3435-small-pull.nc
angle d1d2ang :5@C,CA,N :86-92@CA,C,N :174@CA,C,N out d1d2ang.dat mass
dihedral d1d2dihed :5@C,CA,N :77@CA,C,N :106@CA,C,N :163@CA,C,N out d1d2dihed.dat mass

cpptraj modelname.prmtop <ptraj.script.angles> angles.out

ptraj.script.interactions
trajin 3435-small-pull.nc
#Domain-linker contacts
nativecontacts :5-85 :86-91 writecontacts native-d1l1.dat resout nativeres-d1l1.dat distance 8.0 out native-d1l1.out first name native-d1l1 byresidue
nativecontacts :86-91 :92-174 writecontacts native-d2l1.dat resout nativeres-d2l1.dat distance 8.0 out native-d2l1.out first name native-d2l1 byresidue

#Domain-Domain Contacts
#quit
nativecontacts :5-85 :92-174 writecontacts native-d1d2.dat resout nativeres-d1d2.dat distance 8.0 out native-d1d2.out first name native-d1d2 byresidue

cpptraj modelname.prmtop <ptraj.script.interactions> interactions.out

EMACS
Save: control x, control s
Exit: control x, control c

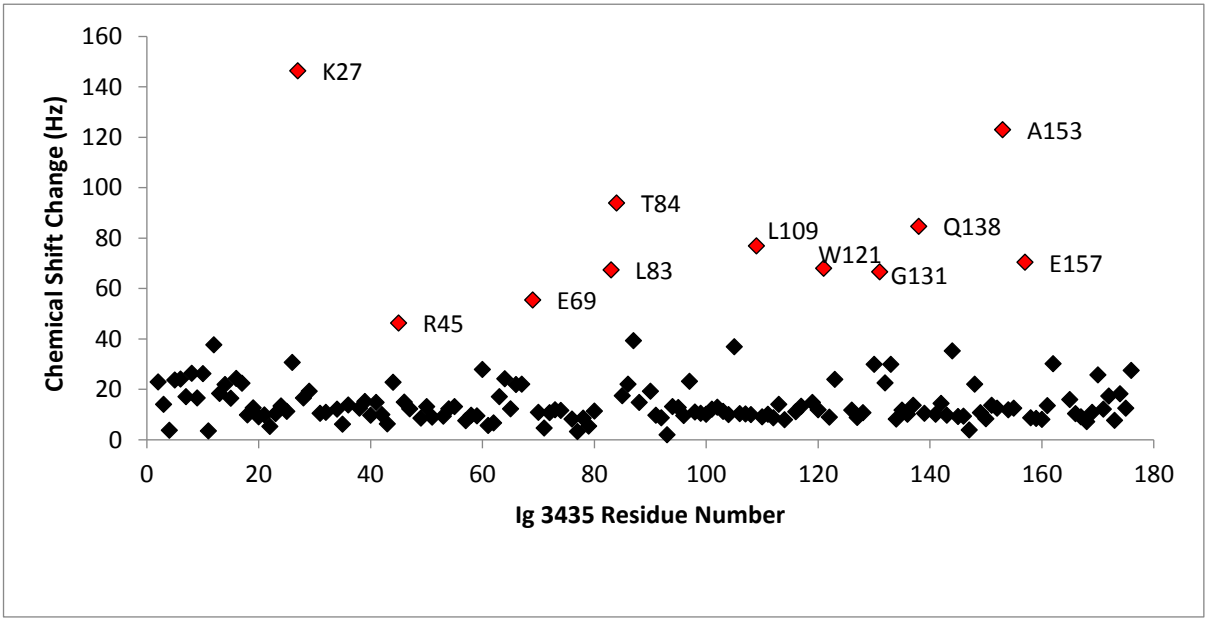
to get plots
gnuplot
plot ‘dist_vs_t’ u 0:4 w l
plot 'd1d2dihed.dat.mass' u 1:2 w l
etc.

Ig34/35 Evaluation using NMR and SMD

NMR analysis suggests domain-domain independence

The NMR results for Ig34/35 correlate with the results seen in Ig58/59 and Ig35/36. The analysis suggests domain-domain independence. The random distribution of significant chemical shifts suggests that there is no interaction between the domains in the dual domain system

A



B

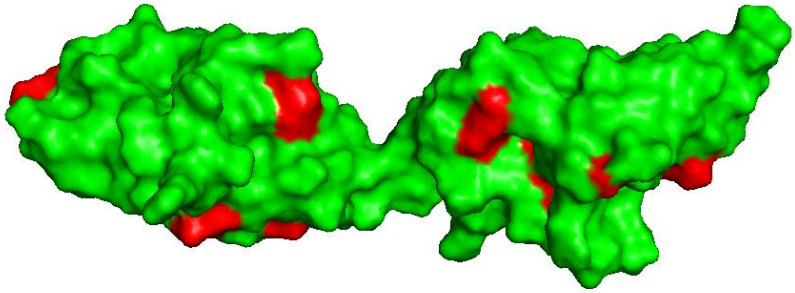


Figure 20. A) A graphical representation of the residues that had significant shifts (two standard deviations shown in yellow; three standard deviations in red) between individual HSQCs and the dual domain HSQC. **B)** Model of Ig58/59 that shows the residues with significant chemical shifts. (Caldwell Thesis, 2015).

SMD suggests Ig34/35 behaves similarly to Ig35/36

Ig34/35 SMD simulations are more similar to Ig35/36 simulations. The two simulations show small differences in the amount of work used on the systems, but overall have the same trend. The Ig34/35 model requires more work; the work input begins at approximately 10.0 Å. This data suggests that like Ig35/36, Ig34/35 behaves as a rigid and fixed domain system. This supports our final hypothesis: the linker sequence has a defining role in domain-domain behaviors.

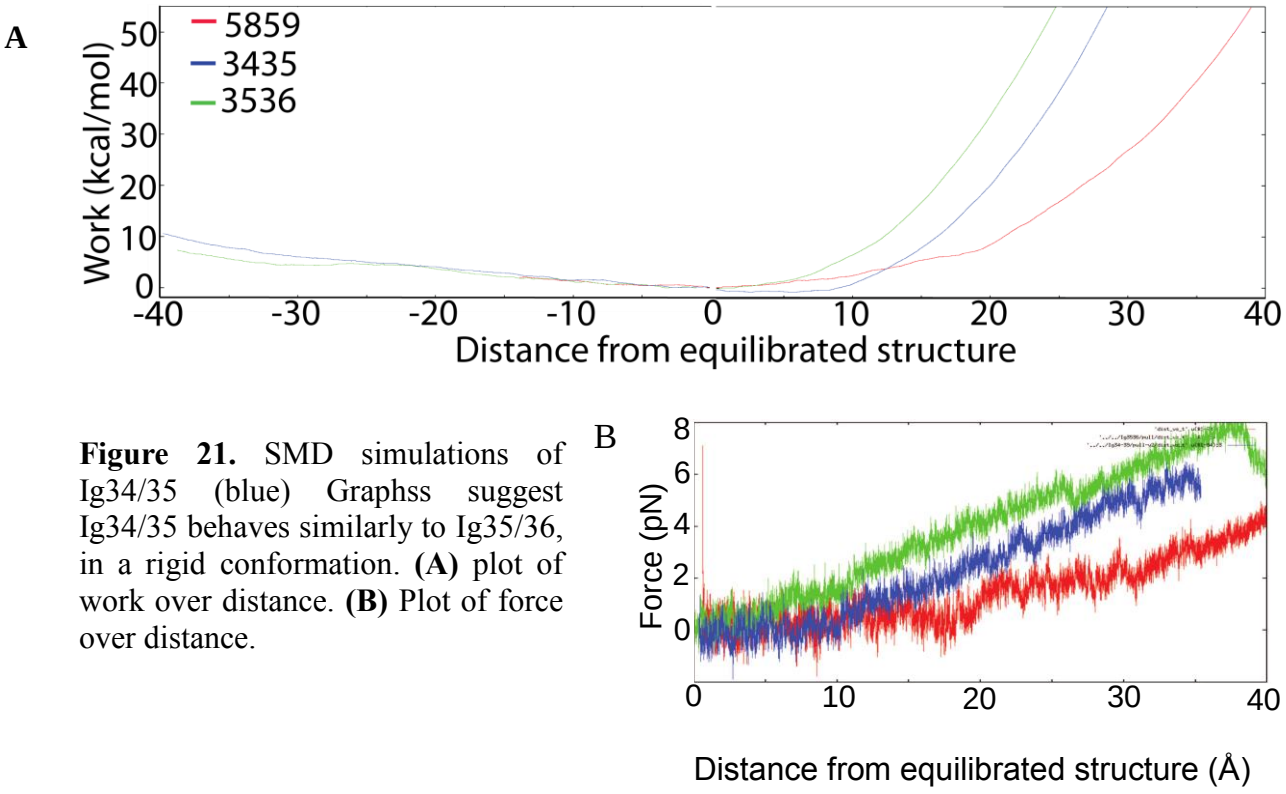


Figure 21. SMD simulations of Ig34/35 (blue) Graphss suggest Ig34/35 behaves similarly to Ig35/36, in a rigid conformation. **(A)** plot of work over distance. **(B)** Plot of force over distance.

Ig34/35 is a very unstable protein that is sensitive to changes in temperature and pH. RDC and SAXS data could not be evaluated for this reason. Anisotropic samples for RDC calculations had denatured protein and were not interpretable. SAXS data contained too much aggregated sample.

Literature

Ackermann, M., King, B., Perry N., Bobbili P., Rudloff, M., Berndsen C., Wright N., Hecker, P., and KontrogianniKonstantopoulos, A. (2017). Small Obscurins at the Intercalated Disc Mediate Cardiomyocyte Adhesion and Size via the PI3K/AKT/mTOR Pathway.Uner Review, Human Physiology

Agarkova, I,, Perriard, J.C. (2005) The M-band: an elastic web that crosslinks thick filaments in the center of the sarcomere. Trends in cell biology 15:477-485.

Alba, E., Tjandra, N. (2002) NMR dipolar couplings for the structure determination of biopolymers in solution. Progress in Nuclear Magnetic Resonance Spectroscopy, 40, 175-197.

Andres, J. (2016) Work done by titin protein folding assists muscle contraction. Cell Reports *14*, 1339-1347

Arimura T., Matsumoto, Y. Okazaki, O. Hayashi, T., Takahashi, M., Inagaki, N., Hinohara, K., Ashizawa, N., Yano, K., Kimura, A. (2007) Structural analysis of obscurin gene in hypertrophic cardiomyopathy. Biochemical and Biophysical Research Communications 362 (2), 281-287

Caldwell, T. A., Sumner, I. C, Wright, N. T. (2015) Mechanical dissociation of the M-band titin/obscurin complex is directionally dependent FEBS Letters 589, 1735-1739

Caldwell, Tracy. (2015) Examination of the structure, force resistance, and elasticity of muscle proteins. James Madison University Senior Honors Projects 110

Chen, K., Tjandra N. (2012). The use of Residual dipolar couplings in studying proteins by NMR. Top Curr. Chem 2012; 326 47-67.

Gautel M (2011) The sarcomeric cytoskeleton: who picks up the strain? Current opinion in cell biology 23, 39-46.

Gautel, M., Djinovic, Carugo K. (2016) The sarcomeric reticulum: from molecules to motion. Journal of Experimental Biology, 219 135-145.

Ho B.K., Agard D.A. (2010) an improved strategy for generating forces in steered molecular dynamics PLoS ONE 5(9)

Hu, L., Kontrogianni-Konstantopoulos, A. (2013) The Kinase domains of obscurin interact with intercellular adhesion proteins. The FASEB Journal (5) 2001-2012.

Hummer, G., Szabo, A. (2003) Kinetics from nonequilibrium single-molecule pulling experiments. Biophysical journal 85:5-15.

Huxley, H.E. (1973) Structural basis of muscular contraction Journal dePhysique Colloques 34, C10-13-C10-25

Isralewitz B, Gao M, Schulten K (2001) Steered molecular dynamics and mechanical functions of proteins. Current opinion in structural biology 11, 224-230.

Kontrogianni-Konstantopoulos A, Ackermann MA, Bowman AL, Yap SV, Bloch RJ (2009) Muscle giants: molecular scaffolds in sarcomerogenesis. Physiological reviews, 89, 1217-1267.

Lee, E., Hsin, J., von Castlemur, E., Mayans, O., Schulten, K., (2010) Tertiary and Secondary Structure Elasticity of a Six-Ig Titin Chain, Biophysical Journal 98 (6), 1085-1095

Lieber, R. L. (2002) Skeletal Muscle Structure, Function, and Plasticity, 2nd ed Lippincott Williams and Wilkins: Philidelphia

Linke WA (2000) Stretching molecular springs: elasticity of titin filaments in vertebrate striated muscle. Histology and histopathology 15, 799-811

Lu H, Isralewitz B, Krammer A, Vogel V, Schulten K (1998) Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. Biophysical journal *75*, 662-671

Manibog, K., Yen, C. F., Sivasankar, S. (2017) Measuring Force-Induced Dissociation Kinetics of Protein Complexes Using Single-Molecule Atomic Force Microscopy. Methods in Enzymology 582, 297-320

McGee Jr TD, Bureau H, Allen C, Hernandez R. AMBER Advanced Tutorial 26.

Meyer LC, Wright NT (2013) Structure of giant muscle proteins. Frontiers in physiology, 4, 368.

Murphy, R., Tsai, A. (2007) Misbehaving proteins: Protein (Mis)folding, aggregation, and Stability, Illustrated Edition Springer Science and Business Media: New York

Pernigo et al. (2017) Binding of Myomesin to Obscurin-Like-1 at the Muscle M-Band Provides a Strategy for Isoform-Specific Mechanical Protection. Structure 25, 107-120

Pernigo S, Fukuzawa A, Bertz M, Holt M, Rief M, Steiner RA, Gautel M (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), 2908-2913

Pinotsis, N., Chatziefthimiou, S.D., Berkemeier, F., Beuron, F., Mavridis, I.M., Konarev, P.V., Svergun, D.I., Morris, E., Rief, M., Wilmanns, M. (2012) Superhelical architecture of the myosin filament-linking protein myomesin with unusual elastic properties. PLoS biology 10(2) e 1001261

Rohs R, Etchebest C, Lavery R, (1999) Unraveling proteins: a molecular mechanics study. Biophysical Journal 76, 2760-2768

Rudloff, M. W., Woosley, A. N., Wright, N. T. (2015) Biophysical characterization of naturally occurring titin M10 mutations Protein Society 24, 946-955

Schneidman-Duhovny, D., Hammel, M., Tainer, JA., and Sali. A., (2016) FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. NAR

Skou, S., Gillilan, R., Ando, N. (2014). Synchotron based small angle Xray scattering (SAXS) of proteins in solution. Nat Protec. 9(7) 1727-1739

Tassone, C. (n.d) Small angle x-ray scattering: an overview. Stanford Synchotron Radiation Lightsource

Von castlemur, E. et al. (2008). A regular pattern of Ig super-motifs defines segmental flexibility as the elastic mechanism of the titin chain. PNAS 105(4) 1186, 1191.

Wakabayashi, M. , Yoshida, M. , Hayashi, F. , Yokoyama, S. (2005)
Solution structure of the second ig-like domain from human Obscurin. Pdb, Structure 13: 473-482

Wood et al. (2007). The genomic landscapes of human breast and colorectal cancers. Science 318 (5853) 1108-1113.

Young P, Ehler E, Gautel M. (2001) Obscurin, a giant sarcomeric Rho guanine nucleotide exhchange factor protein involved in sarcomere assembly. J Cell Biology 154, 123-136