Examination of the Structure, Force Resistance, and Elasticity of Muscle Proteins

_______________________

An Honors Program Project Presented to
the Faculty of the Undergraduate
College of Science and Mathematics
James Madison University

by Tracy Anne Caldwell

May 2015

Accepted by the faculty of the Department of Chemistry and Biochemistry, James Madison University, in partial fulfillment of the requirements for the Honors Program.

FACULTY COMMITTEE:

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

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

Reader: Gina MacDonald, Ph.D.,
Professor, Chemistry and Biochemistry

HONORS PROGRAM APPROVAL:

Philip Frana, Ph.D.,
Interim Director, Honors Program

PUBLIC PRESENTATION

This work is accepted for presentation, in part or in full, at Chemistry Department Spring Symposium on 3/26/15.
Table of Contents

List of Figures 3
Acknowledgments 5
Preface 6
Obscurin segmental flexibility defines a role in force resistance 14
Obscurin and titin display a directional preference to force resistance 35
Re-examination of the Structure and Elasticity of the Ig65-Ig70 Segment of Titin 51
References 62
Appendix I: Transformation 78
Appendix II: Protein Growth 79
Appendix III: Protein Purification 80
Appendix IV: MD and SMD 81
# List of Figures

**Figures**

<table>
<thead>
<tr>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schematic of human muscle and the individual sarcomere</td>
<td>6</td>
</tr>
<tr>
<td>HSQC example</td>
<td>8</td>
</tr>
<tr>
<td>2D vs. 3D NMR spectra</td>
<td>9</td>
</tr>
<tr>
<td>NMR spectra and beta sheet interactions</td>
<td>16</td>
</tr>
<tr>
<td>NMR and crystal structure of Ig59</td>
<td>17</td>
</tr>
<tr>
<td>Ig58/59 HSQC and surface representation</td>
<td>21</td>
</tr>
<tr>
<td>Titin multi-domain structure schematic</td>
<td>22</td>
</tr>
<tr>
<td>SASSIE plot</td>
<td>23</td>
</tr>
<tr>
<td>RDC derived best Ig58/59 structures</td>
<td>23</td>
</tr>
<tr>
<td>Ig58/59 pulling and compression work plots</td>
<td>25</td>
</tr>
<tr>
<td>Ig58/59 pulling time point structures</td>
<td>26</td>
</tr>
<tr>
<td>Peeling and sheering schematic of M10-Ig1, force and work plots</td>
<td>38</td>
</tr>
<tr>
<td>Structures of peeled and sheered structures, hydrogen bonds,</td>
<td>40</td>
</tr>
<tr>
<td>and hydrophobic contacts</td>
<td></td>
</tr>
<tr>
<td>Mutation force and work plots</td>
<td>43</td>
</tr>
<tr>
<td>RMSD of models</td>
<td>48</td>
</tr>
<tr>
<td>Force, work, hydrogen bonds, and hydrophobic contacts of slow</td>
<td>49</td>
</tr>
<tr>
<td>pull peeling and sheering simulations</td>
<td></td>
</tr>
<tr>
<td>RMSD of individual domains during pulling</td>
<td>50</td>
</tr>
<tr>
<td>Ig6 structure, B-factors, crystal packing</td>
<td>54</td>
</tr>
</tbody>
</table>
Work plot of compression, time point structures 57
Alignment of two structures of Ig67-69 59
Crystal packing example 60
Linker regions 60
Time points of SMD of I6 61

Tables
NMR statistics 18
X-ray crystallography statistics 19
I6 crystallography statistics 53
Acknowledgements

I would like to thank Dr. Nathan Wright, Dr. Christopher Berndsen, and Dr. Gina MacDonald for serving on my senior project committee as well as for all of their help and guidance. I would also like to thank Dr. Isaiah Sumner for help with simulations and the entire James Madison University Department of Chemistry and Biochemistry for their support and use of instrumentation. This work was supported by the Jeffress Memorial Trust J-1041 (to NTW), Research Corporation 22450 (to NTW), James Madison University startup funds (to NTW, CEB, and IS), the Second Century Scholarship (to TAC), the Tickle Fellowship (to TAC), and the Polocsay Award in Undergraduate Research (to TAC).
Preface

Muscles are very highly ordered structures: from bundles of muscle fibers to individual muscle cells (myocytes) to individual sarcomeres (Figure 1A). The sarcomere, composed primarily of the proteins actin and myosin, is the smallest unit of contraction. Thousands can be found in a single muscle fiber (termed myofibril). When muscles contract or relax, long filaments of actin and myosin slide past one another (Figure 1B).

![Figure 1. A) Schematic of muscle structure and organization. B) Schematic of an individual sarcomere showing the giant muscle proteins obscurin (blue), titin (pink), and nebulin (green).](image-url)
If actin and myosin alone slid past each other, with no other proteins involved, the muscle cell would not contract. Instead, the contractile apparatus must be attached to the rest of the cell. These connections are facilitated through a complex web of specialized cytoskeletal proteins. The giant muscle protein, obscurin, is one of these elements, and works, in part, by connecting the M-line of the sarcomere, the sarcoplasmic reticulum, and T-tubule structure (blue, Figure 1B)\(^3\). On a molecular level, obscurin performs this connection by binding with small ankyrin at the sarcoplasmic reticulum and titin at the sarcomere\(^4,5\). Titin is the largest known polypeptide, which regulates the length of the sarcomere (pink, Figure 1B)\(^3\). Both obscurin and titin are binding sites to other proteins such as myosin-binding protein-C slow (sMyBP-C variant 1), calmodulin, and myomesin to obscurin, and telethonin, small ankyrin 1, filaminC, nebulin, tropomyosin, myomesin, and calmodulin to titin \(^3,4,6-7\). These interactions contribute further to cytoskeletal crosslinking, as well as provide opportunities for cellular regulation.

Obscurin and titin are made up of independently folded domains that can be studied individually. Both are comprised of mostly Ig (immunoglobulin) or FnIII (Fibronectin type III)-like domains, which are made of two beta sheets held together by a hydrophobic core\(^3\). High resolution structures of a limited number of both titin and obscurin domains have been determined using both nuclear magnetic resonance (NMR) and X-ray crystallography. These structures have been complemented by low resolution methods such as small angle X-ray scattering (SAXS) and cryo-electron microscopy (cryo-EM)\(^3\). Here, other high and low resolution structures not previously published will be presented in order to investigate how their response to force, elasticity, flexibility, and orientation of domains aids in their function.
NMR structure determination utilizes the natural spin of isotope labeled samples ($^{13}\text{C}$, $^{15}\text{N}$). These spin states will give rise to a nuclear magnetic resonance which depends on both the frequency of the electric field and the environment that the atoms are subject to. Pulse programs have been developed by many scientists. Each pulse program allows for the acquisition of certain information about the structure of the protein. For example, the first step in solving an NMR protein structure is to assign a heteronuclear single quantum coherence spectra (HSQC). The pulse program for this experiment allows for the resonance of only NH groups to be detected. Since each amino acid in a protein contains one NH in the backbone, each peak in this spectra corresponds to one amino acid in a unique environment (see “Obscurin segmental flexibility defines a role in force resistance” Figure 3A) (Figure 2). The next step is to assign each of these peaks to its specific amino acid. This is completed through the use of 3-dimentional NMR

![Figure 2](image_url). Example of an HSQC spectra where every peak corresponds to one amino acid of the protein.
spectra (Figure 3). Finally, spectra based on the nuclear Overhauser effect (NOE) are used which utilize cross talk between different spins to determine the distances between atoms. This cross-talk allows for the mapping of inter-protein interactions that leads to a 3-dimensional high resolution structure. NMR structures are supplemented by residual dipolar coupling (RDC) data that generates angle restraints. RDCs are collected from the difference between isotropic and anisotropic samples. Isotropic samples, which are free in solution, can tumble rapidly, but anisotropic solutions, which are prepared in various alignment media (like polyacrylamide gels), have restricted rotation and therefore are more likely to be in a single orientation. The differences between the isotropic and anisotropic sample is the RDC value which can then be converted into a degenerate angle that provides information about the orientation of the internuclear bond vector relative to the protein’s alignment tensor (see “Obscurin segmental flexibility defines a role in force resistance” Figure 1D). RDCs are especially useful in

**Figure 3.** Two dimensional spectra are like a page (left), while three dimensional spectra are like a page in a book (right).
providing long-range structural data which aids in structure determination as well as the determination of the orientation between two domains.

X-ray crystallography is the second method for high resolution structure determination. To solve a crystal structure, the protein first must be crystallized. This is completed by placing vapor diffusion methods. The solution will diffuse into the protein drop and crystals consisting of a regular packing of protein molecules will form due to an increase in precipitant concentration. The crystals are then exposed to an X-ray source. The waves from the X-ray beam scatter and the resulting electron diffraction is detected. These diffractions can then be Fourier transformed into a map of the electron density of the protein. When the diffraction is detected, only the amplitude or intensity is recorded, but the phase of the diffracted wave is not. Molecular replacement using a homologous structure or isomorphous replacement using a heavy atom is necessary to solve the phases to create an electron map that accurately reflects the contents of the crystal. The protein sequence is then fit to the electron density map and refined to determine the high-resolution 3D structure.

NMR and X-ray crystallography, have solved the structures of many of the domains of the muscle proteins obscurin and titin. While the structure of these proteins has been studied by several scientists, the mechanism by which they, and other muscle proteins function in the muscle is only partially understood, and will be investigated here. A main component of how muscle proteins function is how they respond to force. This type of investigation is ideally suited to the use of steered molecular dynamics (SMD). SMD is part of a large body of literature that, in part, attempts to simulate the dynamics of rare (long-time scale) events. In SMD, the protein force field \( U_{FF}(R) \) is augmented by the addition of a time-dependent,
harmonic potential, i.e.,

\[ U_{\text{tot}}(R, t) = U_{\text{FF}}(R) + \frac{K}{2} (\mathbf{x}(R) - (\mathbf{x}_0 + \nu t))^2 \]  

where \( R \) represents the atomic coordinates, \( \mathbf{x}(R) \) is a reaction coordinate, \( K \) is the harmonic spring constant and \( \nu \) is velocity. The constant velocity potential terms pulls the protein along a pre-selected reaction coordinate. In addition to the sampling of rare events, SMD can simulate atomic force microscopy (AFM) in which a molecule is tethered to the end of a cantilever. Force is then applied along the cantilever and the protein can be pulled apart. SMD has been used to successfully replicate experimental AFM data and has been used to investigate how proteins rupture under force.

Studies of force in muscle proteins is important because myocytes are intrinsically both strong and flexible. While strength is derived through the well understood mechanics surrounding the \( \text{Ca}^{2+} \) and ATP-dependent actin/myosin cross bridge formation, the molecular mechanisms governing flexibility have proven more elusive. In the past several years, multiple papers have demonstrated how long, modular, fiber-like proteins form a flexible web within muscle cells. The M-band, Z-disk, and the structural lattice holding the contractile apparatus in place are all comprised, in part, from components of this web. Other than anchoring the macromolecular contractile machinery in place, some of these long modular proteins also act as force resistors and force modulators to help control myocyte stretch. Studies on titin have shown that its modular nature affords a large degree of conformational flexibility, and that the protein resists increasing stretch in a modified entropic-spring-like model. In this model, stretch resistance is accomplished through a combination of an entropically-derived increase in energy required to straighten out multiple domains, along with inter-domain enthalpic interactions (see “Obscurin segmental flexibility defines a role in force resistance” Figure 4).
This enthalpic-entropic spring design allows titin to provide appropriate stretch resistance at both low and high force loads, which in turn protects the myocyte from undue mechanical stretch damage\textsuperscript{33-34}. One question that remains is whether titin’s reaction to stretch is unique, or if a modified entropic spring-like mechanism can explain the behavior of other long modular proteins as well\textsuperscript{33-34}. While titin is the only protein to span the entire half-sarcomere, other structural proteins with similar architecture such as myomesin, M-protein, obscurin, and MyBP-C also contribute to the filamentous cytoskeleton of myocytes\textsuperscript{42-43}. Although these proteins have different orientations relative to the sarcomere than titin, they are subjected to similar mechanical forces, and thus may also act as stretch resistors. The first question we will consider is if obscurin could also act as a stretch resistor.

Titin’s flexibility and elasticity have been extensively studied experimentally by techniques such as atomic force microscopy (AFM)\textsuperscript{25,44-46}. However, the forces seen to break titin’s domains using these techniques are less than the forces exerted by the muscle (reviewed in Tskhovrebova)\textsuperscript{47}. How then could the muscle stand up to normal everyday stretch? These experiments are simplistic and do not take into account the intricate web of protein interactions which stabilize the sarcomere. One of these important interactions is between the two giant muscle proteins titin and obscurin\textsuperscript{48}. The second question we will consider is whether the orientation of this interaction plays a role in its resistance to force.

An important aspect of a muscle protein’s elasticity, flexibility, and therefore force resistance is the orientation of its domains. These orientations are due to both inter- and intra-protein interactions, which lead to several distinct spring mechanisms\textsuperscript{49}. The third and last
question we will consider is the discrepancy between computational and crystal structure data for a six domain section of titin.
Obscurin segmental flexibility defines a role in force resistance

Obscurin is the most recently discovered giant muscle protein. This protein is vital to muscle cell organization and maintenance. Knockout and knockdown studies in mice show myocytes with poorly organized M and A bands of the sarcomere, poor sarcomere organization around the sarcomere, altered muscle development, and generalized muscle weakening. Obscurin is organized in a modular fashion and many of its more than 70 domains bind to a specific cytoskeletal, signaling, or membrane-associated proteins. Obscurin is currently the only known connection between the contractile apparatus and the sarcoplasmic reticulum, through interactions with titin and small ankyrin. Additional obscurin targets include other parts of the contractile apparatus such as sarcomeric myosin and myosin binding protein-C (slow variant), as well as signaling molecules like calmodulin and Rho-A. Genetic analysis and mouse modeling show obscurin to have three main functions in myocytes: it is a key member of the sarcomeric cytoskeleton, it connects the sarcomere to surrounding membranes, and it plays a role in stretch signaling.

Given that obscurin and titin have a similar global architecture, it is tempting to speculate that tandem obscurin Ig-like domains behave similarly to titin upon force application. However as no multi-domain obscurin structures have been solved, no direct comparison is possible. Another complicating factor is the fact that obscurin and titin are oriented differently within the sarcomere and thus subjected to differing levels and kinds of mechanical force. To better define obscurin’s role in muscle stretch mechanics, here we report the structure of obscurin Ig59. This data, combined with reported NMR data on Ig58, allows a high-resolution model of the obscurin Ig58/59 dual domain in solution (in press). Through NMR experiments and SMD simulations
we find that, despite the short linker between Ig58 and Ig59, the two domains do not significantly interact with each other. These data also show that this region of obscurin behaves more like a classical entropic spring than tandem titin domains. However, increased force application begins to stretch out individual domains and leads eventually to ‘domain bursting’, a phenomenon seen in titin to protect against extreme stretch. This unique obscurin stretch response suggests that obscurin can acts as a passive connector at low forces or at rest, but can ‘turn on’ and become a force resistor when strong mechanical stretch is applied to the myocyte.

Results

To better understand the molecular mechanism of obscurin’s stretch response, we first solved the high-resolution structure of Ig59. This domain was studied for several reasons. First, the high-resolution structure of Ig58 was recently solved with NMR and X-ray crystallography. Thus elucidating the Ig59 solution structure allows for the eventual study of the Ig58/59 duel domain system. As Ig58/59 are both necessary for obscurin to bind to titin, solving the structure of the duel domain system provides a base for eventually characterizing how obscurin and titin interact.

The solution structure of Ig59 contains a total of 1257 restraints, including 228 intraresidue NOE restraints, 382 sequential restraints, 95 medium-range restraints, 292 long-range restraints 64 H-bonds, 144 dihedral restraints, and 53 H-N residual dipolar coupling (RDC) restraints (Figure 1, Table 1). Together, this gives an average of greater than 12 restraints per residue. NOE correlations were assigned on the basis of the $^{1}$H, $^{13}$C, and $^{15}$N backbone and sidechain chemical shifts (Figure 2A). With the exception of Met 0, Arg 2, Arg 80, and most of
Figure 1. A) Observed beta sheet interactions of Ig59 as seen in the $^{15}\text{N}$-edited NOESY. B) Example of TOCSY showing neighboring and cross-strand interactions. C) Example of NOESY backbone walking. D) Example of RDC spectra showing isotropic (right) and anisotropic (left) samples.
the His tag at the C-terminus, every backbone H-N bond and most sidechain C-H shifts are visible in these NMR experiments. The resulting models show beta sheets extending from Glu 7 to Lys 10 (strand A’), Ala 12 to Arg 15 (strand A), Ala 18 to Thr 25 (strand B), Ser 36 to Ile 40 (strand C), Trp 51 to Asp 56 (strand D), His 59 to Leu 65 (strand E), Gly 73 to Ala 78 (strand F), and Ala 84 to Leu 91 (strand G). Overall, strands A, B, E and D form one b-sheet and A’, G, F, and C form another sheet (Figure 2B), folding into a characteristic Ig-like fold (Meyer 2014) (Figure 2B). H-N RDC values independently verify the validity of this structure.

Figure 2. A) NMR structure of Ig59 comprised of NOE restraints and RDCs. Colored according to figure 1B B) Ig59 has a characteristic Ig-like fold with strands A-G. C) Overlay of Ig59 NMR (green) and crystal (blue) structures.
Table 1. NMR-derived restraints and statistics of 20 NMR structures\(^1\)

<table>
<thead>
<tr>
<th></th>
<th>&lt;20&gt;</th>
<th>best</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>rmsd from distance constraints (Å)</strong>(^2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total (1061)</td>
<td>0.028 ± 0.002</td>
<td>0.029</td>
</tr>
<tr>
<td>intraresidue (228)</td>
<td>0.005 ± 0.004</td>
<td>0.003</td>
</tr>
<tr>
<td>sequential (</td>
<td>(</td>
<td>i−j</td>
</tr>
<tr>
<td>medium range (1 &lt; (</td>
<td>i−j</td>
<td>≤ 1) (95)</td>
</tr>
<tr>
<td>long range (</td>
<td>(</td>
<td>i−j</td>
</tr>
<tr>
<td>hydrogen bonds (64)</td>
<td>0.056 ± 0.006</td>
<td>0.050</td>
</tr>
<tr>
<td><strong>rmsd from exptl dihedral constraints (°)</strong> (\Phi, \Psi) (144)</td>
<td>0.687 ± 0.116</td>
<td>0.684</td>
</tr>
<tr>
<td><strong>rmsd from dipolar coupling restraints (Hz)</strong> (D_{NH}) (52)</td>
<td>0.89 ± 0.09</td>
<td>0.93</td>
</tr>
<tr>
<td><strong>rmsd from exptl (^{13})C chemical shifts</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(^{13})Ca (ppm)</td>
<td>1.59 ± 0.17</td>
<td>1.31</td>
</tr>
<tr>
<td>(^{13})Cb (ppm)</td>
<td>1.46 ± 0.04</td>
<td>1.47</td>
</tr>
<tr>
<td><strong>rmsd from idealized geometry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bonds (Å)</td>
<td>0.004 ± 0.001</td>
<td>0.004</td>
</tr>
<tr>
<td>angles (°)</td>
<td>0.633 ± 0.020</td>
<td>0.640</td>
</tr>
<tr>
<td>impropers (°)</td>
<td>0.442 ± 0.036</td>
<td>0.431</td>
</tr>
<tr>
<td><strong>Lennard-Jones potential energy (kcal/mol)</strong>(^3)</td>
<td>-363 ± 10</td>
<td>-359</td>
</tr>
<tr>
<td>Q-value (^4)</td>
<td>0.27 ± 0.04</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>% most favorable region in the Ramachandran plot</strong>(^5)</td>
<td>74.1 ± 3.0</td>
<td>72.8</td>
</tr>
<tr>
<td><strong>rmsd of the mean structure (Å)</strong>(^6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>all backbone atoms (3-91)</td>
<td>0.609 ± 0.063</td>
<td>0.494</td>
</tr>
<tr>
<td>all heavy atoms (3-91)</td>
<td>1.2128 ± 0.080</td>
<td>1.157</td>
</tr>
</tbody>
</table>

---

\(^1\) The 20 ensemble structures, \(<20>\), are the results of simulated annealing calculations. The best structure is the closest to the average structure. The values shown for the \(<20>\) are the mean ± standard deviation.

\(^2\) None of the 20 structures has a distance violation > 0.35 Å or a dihedral angle violation of > 5°. The force constants used in the SA calculations are as follows: 1000 kcal mol\(^{-1}\) Å\(^2\) for bond length, 500 kcal mol\(^{-1}\) rad\(^{-2}\) for angles and improper torsions, 4 kcal mol\(^{-1}\) Å\(^{-4}\) for the quartic van der Waals (vdw) repulsion term (hard-sphere effective vdw set to 0.8 times their values in CHARMM parameters), 50 kcal mol\(^{-1}\) Å\(^{-2}\) for experimental distance constraints, 1 kcal mol\(^{-1}\) Å\(^{-2}\) for distance symmetry constraints, 0.5 kcal mol\(^{-1}\) ppm\(^{-2}\) for the \(^{13}\)C chemical shift constraints, and 1.0 for the conformational database potential. The force constants (in kcal Hz\(^{-2}\)) used for dipolar coupling restraints is 0.50.

\(^3\) Lennard-Jones van der Waals energies were calculated using CHARMM parameters and were not used in any stage of the structure determination.

\(^4\) Q-values were determined by randomly removing 10% of all RDC values. To ensure accuracy, an ensemble of structures with a second randomly removed subset of RDCs was also run. The Q-value of this second set was similar to the first.

\(^5\) PROCHECK was utilized to generate the Ramachandran plot.

\(^6\) Backbone calculations include C\(^\alpha\), N, and C\(^\prime\) atoms. Only residues 3–91 are included since no long-range NOE correlations were observed for residues 1–2 and 92–104.
Independently, Ig59 crystal structure was solved to a resolution of 1.18 Å (Figure 2C, Table 2). Of all known Ig-like domains, Ig59 is closest in sequence homology to Ig58 (30 percent), and the backbone RMSD between the high-resolution structures of these two sequences is 1.8 Å (Figure 2C). Most of these differences occur in the loop regions where the protein is more flexible, which results in a higher b-factor and lower number of NOEs. Thus, the Ig-like fold is similar. However, the surface of the molecules have different topologies, with Ig58 containing more exposed charged moieties and Ig59 having more solvent-exposed hydrophobic patches.

**Table 2.** X-Ray crystallography statistics of obscurin Ig59.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.97918</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>30.49-1.18</td>
</tr>
<tr>
<td>Space group</td>
<td>P 3 1 2 1</td>
</tr>
<tr>
<td>Unit cell (Å)</td>
<td>60.98 60.98 47.56</td>
</tr>
<tr>
<td>Unit cell (°)</td>
<td>90 90 120</td>
</tr>
<tr>
<td>Total reflections</td>
<td>662701 (35265)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>33764 (3303)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>19.5 (10.7)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.82 (92.53)</td>
</tr>
<tr>
<td>Mean I/sigma(I)</td>
<td>22.98 (5.07)</td>
</tr>
<tr>
<td>Wilson B-factor</td>
<td>14.83</td>
</tr>
<tr>
<td>R-merge</td>
<td>0.1345 (0.5394)</td>
</tr>
<tr>
<td>R-meas</td>
<td>0.1389</td>
</tr>
<tr>
<td>CC1/2</td>
<td>0.988 (0.912)</td>
</tr>
<tr>
<td>CC*</td>
<td>0.997 (0.977)</td>
</tr>
<tr>
<td>R-work</td>
<td>0.1642 (0.2634)</td>
</tr>
<tr>
<td>R-free</td>
<td>0.1851 (0.2974)</td>
</tr>
<tr>
<td>Number of non-H atoms</td>
<td>775</td>
</tr>
<tr>
<td>macromolecules</td>
<td>678</td>
</tr>
<tr>
<td>water</td>
<td>97</td>
</tr>
<tr>
<td>Protein residues</td>
<td>90</td>
</tr>
<tr>
<td>RMS (bonds)</td>
<td>0.033</td>
</tr>
<tr>
<td>RMS (angles)</td>
<td>1.46</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
<td>99</td>
</tr>
<tr>
<td>Ramachandran outliers (%)</td>
<td>0</td>
</tr>
<tr>
<td>Clashscore</td>
<td>2.23</td>
</tr>
<tr>
<td>Average B-factor</td>
<td>20.5</td>
</tr>
</tbody>
</table>
Obscurin is a highly modular protein; individual domains can be easily excised and studied without the context of the rest of the molecule. However, there is a possibility that while each domain is independently folded, it could interact with neighboring domains to create an ordered superstructure. This kind of domain-domain interaction occurs extensively in titin. The superstructural motif is driven through interactions of Ig domains to either the Ig-Ig linker region or to neighboring Ig domains. To test whether obscurin also has a defined superstructural organization, we created an Ig58/59 construct. The resulting HSQC spectrum of this dual-Ig construct is well dispersed (Figure 2A). While the individual Ig58 and Ig59 HSQCs closely overlay with Ig58/59, 3D $^{15}$N-edited CBCA(CO)NH and an HNCO NMR experiments were also completed to verify the sequence-specific assignments. The resulting chemical shift changes, when mapped onto a model of Ig58/59, are shown in Figure 2B. The most significant chemical shift occurs at G92 of Ig58. This chemical shift change is expected; in the individual Ig58 structure, G92 neighbors the His6 tail while in the Ig58/59 construct it neighbors the Trp-Arg linker region between the two domains. The fact that there are no chemical shift changes greater than ~0.1 ppm, combined with there being no localized chemical shift changes, or ‘hot spots’, strongly suggests that the two domains do not significantly interact with each other.
Figure 3. A) HSQC of Ig58/59. B) Residues with a chemical shift difference of two (yellow) or three (red) standard deviations between Ig58 or Ig59 and Ig58/59.
This result was surprising; obscurin only has a 2-3 residue linker between Ig58 and Ig59, and we expected this short linker to facilitate domain-domain interactions as observed in titin (Figure 4). The lack of large chemical shift changes, even near residues close to the Ig-Ig interface, suggest that Ig58/59 domains do not interact with one another. As additional evidence of a semi-elongated system, there are no observed inter-domain NOEs between Ig58 and Ig59. To further explore the overall shape of Ig58/59, SAXS data demonstrates that the two domains exist primarily in a moderately extended conformation. Traditional rigid-body analysis showed that Ig58/59 had an Rg value of 26.2Å, and models of this structure indicate two domains. The program SASSIE was used to better visualize the likely motion between Ig59 and Ig58. Analysis of these data reveal that Ig59 is generally oriented in a moderately extended orientation relative to Ig58 (Figure 5).

**Figure 4.** Schematic of inter-domain interactions as seen in titin. These inter-domain interactions are not seen between obscurin Ig58 and Ig59. Taken from Lee et al.\textsuperscript{37}
Figure 5. SASSIE derived plot showing that moderately extended Ig58/59 models fit experimental SAXS data best.

Figure 6. A) The best Ig58/59 structure as determined by RDC values (Q factor= 0.27). B) The 10 best RDC-derived Ig58/59 structures with Q-factors below 0.40.

In order to determine a higher-resolution model of Ig58/59, a set of H-N RDC constraints were added to the merged Ig58 and Ig59 NOE data. This was justified because of the lack of chemical shift changes between the single and double domain, and due to the aforementioned
lack of any noticeable differences in NOE spectra. Due to the size of Ig58/59 and spectral crowding, fewer RDCs were collected than in either of the individual domains. However, the 15 RDC measurements in Ig58 and 17 in Ig 59 were sufficient to orient the two domains relative to each other (Figure 6A). The resulting model agrees with our other NMR and SAXS data and shows both domains partially (145°) extended away from each other. Given that there are only local physical restraints holding the domains in this orientation, and given the degeneracy of RDC data, this lowest NMR structure should be considered one possible structure. A more realistic model of Ig58/59 is likely the cluster of structures calculated from SAXS data, which shows multiple different semi-extended orientations of Ig59 when oriented to Ig58. In fact, the 10 best NMR RDC structures are oriented in a wide range of angles, once again showing the flexibility of the linker region (Figure 6B).

While the above experiments describe how obscurin Ig58/59 behaves in solution, the ultimate goal is to understand how obscurin acts in the context of the intact sarcomere. Obscurin links the contractile apparatus to the surrounding SR membrane structures. It is reasonable to expect obscurin to experience a significant amount of both pulling and compression forces, depending on the state of the muscle and the orientation of the protein. To simulate these forces, we performed steered molecular dynamics (SMD) simulations on the entire high-resolution Ig58/59 model. While NMR and SAXS experiments show Ig58/59 to be semi-extended, SMD compression of Ig58/59 show that these two domains can be brought together by roughly 6 Å and still be in the realm of thermal noise (taken to be 0.7 kcal/mol) (Figure 7A). Further compression progressively requires more energy as more domain/domain contacts are formed. Likewise, extension models show that the dual domains can only be extended by 2 Å under
thermal noise (Figure 7B). After this point, Ig58/59 can still be extended with minimal force application until the linker region is fully extended (Figure 8). One domain must then partially unfold to attain any further extension. Together, these data suggest that obscurin behaves like a true entropic spring under low force loads, and can expand and contract via Brownian motion by roughly 8 Å per two domains. Additional stretching breaks stabilizing bonds within individual obscurin Ig-like domains (Figure 8). Tandem obscurin domains thus behave more like an enthalpic-entropic spring with the application of significant stretch.

**Figure 7.** A) Work plot of compression shows that the duel domain system can bend to about 145° under thermal noise. B) Work plot of extension shows straightening of domains followed by the unfolding of Ig59 with labeled time points corresponding to Figure 8. Gray lines are representative of the threshold for thermal noise (0.7 kcal/mol).
Figure 8. Constant velocity pulling of obscurin Ig58/59 showing straightening of domains, straightening of linker, and subsequent unfolding of Ig59.
When Ig58/59 is stretched, Ig 59 broke always down first with an initial slipping of beta strand 1 followed by detachment of this strand. Upon further stretching hydrogen bonds between beta strand 2 and 3 are peeled apart. Once this first beta sheet is unraveled, the rest of the domain rapidly unfolds. In order to probe the strength of each individual domain, constant velocity SMD was performed on each domain separately. When Ig59 was pulled at a constant velocity of 6A/ns it broke down in a similar fashion to the constant velocity pulling of Ig5859 with a slippage of strand 1 followed by a detachment of this strand marking the beginning of the domain collapse. Strands 8, 2, 5, and 6 separate sequentially leaving 3 and 4 the only strands still bonded. A similar constant velocity pull of Ig58 suggests a somewhat similar pattern of unfolding with beta strand 1 detaching followed by 8, 6 and 5 and leaving 3 and 4 bound. Ig58 does not seem to have the same slippage as is seen in the unfolding of Ig59. The unfolding of Ig58 also requires has a slightly higher work value.

Constant force simulations were also performed in order to investigate the maximum force the complex could withstand. A constant force of 100 pN straightened the two domains but was not enough to pull apart either domain (similar to 4ns in figure 8). A force of 300 pN causes Ig59 to unfold but Ig58 remains folded (similar to 15ns in figure 8). This agrees with the constant velocity simulations, as 300 pN corresponds to a work of 4.3 kcal/mol which greater than the work required to pull apart Ig59 (3.105 kcal/mol) but around the same as the work required to break down Ig58 (4.069 kcal/mol). A force of 500 pN caused the domains to rapidly unfold, with Ig59 unraveling before Ig58. Ig59 was also seen to unfold first in all constant velocity simulations.
Discussion

The structure and motion of a dual-domain obscurin segment provides insight into how obscurin may function in a live myocyte. Unlike titin, obscurin does not have extensive domain-domain interactions between domains. Assuming all 60 obscurin Ig-like domains act like Ig58/59, and simplifying the system to discount the effect of hydrodynamic drag, this lack of interdomain contacts should allow obscurin to expand and contract up to 24 nm via thermal fluctuations. The entire obscurin chain could also extend 140 nm further with gradually increasing force so that all domains are straightened. Any additional extension of the obscurin Ig-like region would result in the partial or complete unfolding of individual Ig domains, and would likely require forces significantly greater than those in physiological environments. Such ‘domain bursting’, at similar forces, is common among poly-Ig domain proteins and has been proposed as a defense mechanism for overstretching\textsuperscript{54}. In general, obscurin’s reaction to force is reminiscent but not identical to titin; obscurin is predicted to have a much shallower stretch response in the early part of extension. Therefore, obscurin also behaves as a modified entropic spring but the enthalpic contribution, especially under small amounts of stretch, seems to be small.

These differences in how obscurin and titin react to stretch underlie each protein’s putative role in the myocyte. Titin plays a central role in preventing muscle overstrecthing, and thus needs to be able to respond appropriately to a broad range of forces. In contrast, obscurin primarily links the sarcolemma with the cytoskeleton. In this role, obscurin may need, at times, to simply act as a rope that loosely connects the two structures. However, the sarcoplasmic reticulum and the contractile apparatus need to remain near each other to ensure efficient calcium
binding and muscle contraction. To facilitate this, obscurin’s domain organization allows the protein to passively resist increasing SR/sarcomere separation above a certain distance threshold. In this model, obscurin is situated to help mitigate and resist strong mechanical forces. While titin also performs this task, titin only can act as a resistor to forces that are parallel with the thick and thin filaments. Since obscurin appears to be oriented randomly relative to the sarcomere, it can in turn react to mechanical stretching emanating from multiple directions, especially if those forces are moderately strong. Given that obscurin also contains several signaling moieties such as a RhoGEF/PH domain and two kinase domains, there is a possibility that obscurin behaves as both a passive force resistor and an active mechanosensor, where it reacts to certain force loads. However before obscurin can be classified fundamental questions, such as whether obscurin signaling domains are actually influenced by force, need to first be addressed.

In comparison, titin domains are typically separated by roughly 4-6 residues, and these domains often have extensive interdomain contacts (Von Castelmur). Additionally, computer simulations of longer titin chains suggest that the domains need to have a significant level of interaction for titin to behave as it does within the muscle (Schulten). Given that obscurin domains have very short interdomain linkers (1-3 residues), it is surprising that obscurin does not have more extensive interdomain interactions. However, there are very few sidechains that could be involved in interactions, given the population of residues in Ig58 and Ig59 that face towards the linker. Close examination of additional multi-domain obscurin fragments is necessary to determine whether the extended domain architecture is characteristic of obscurin, or if this is unique to this particular sequence.
While the structural effects of the obscurin mutation itself have been well characterized, the molecular mechanism of the obscurin/titin interaction are still unknown. However, with the completion of this work, there are several clues as to how titin and obscurin likely interact. Since Ig58/59 exists in a largely extended orientation, the overall obscurin/titin binding shape likely requires this shape. Also, the obscurin/titin binding event is unusual in that both proteins require two Ig domains for efficient binding. Elucidating the particulars of this interaction may provide insight into how obscurin, titin, and other poly-Ig containing proteins regulate target specificity.

Two mutations, one in Ig58 and one in Ig59, have been recently studied by others and can lend further insight into the obscurin-titin binding interaction. A mutation of obscurin Ig58 (R4344Q) leads to hypertrophic cardiomyopathy in humans. A recent knock-in study of obscurin Ig58/59 to mice leads to sarcomeric dysregulation, changes in myogenesis, and signaling changes (unpublished data). Presumably, these physiological effects are the result of the obscurin Ig58/59 being unable to bind to its molecular target, titin ZIg9/10. The R4344Q mutation does not disrupt the overall structure but does disrupt a highly charged surface. Therefore, the obscurin/titin interaction probably depends on electrostatic interactions. A mutation of obscurin Ig59 (A4445W) leads to a family mutation. While this mutation has yet to be studied structurally, it has been shown to abolish binding to titin. Structural analysis of this mutation will lend further insight to how obscurin and titin bind.
Materials and Methods

Protein preparation and NMR collection  All chemicals were ACS grade or higher and were typically purchased from Fisher Scientific unless otherwise specified. Recombinant $^{15}$N, $^{15}$N-$^{13}$C, and unlabeled protein were purified after overexpression in *Escherichia coli* (BL21(DE3)) using a pET24a vector system (Novagen, San Diego CA) in a manner similar to (Rudloff 2015). All NMR experiments were collected on a 600 MHz Bruker Avance II spectrometer equipped with a TXI room temperature 5 mm probe with z-axis pulse field gradient coils. All NMR samples were collected at 25°C in 20 mM Tris pH 7.5, 20 mM NaCl, 0.35 mM NaN$_3$, and 0.5-2.5 mM protein with 10% D$_2$O. We collected a 2D HSQC, and standard triple resonance experiments including HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, C(CO)NH, H(CCCO)NH, $^{15}$N-edited TOCSY, $^{15}$N-edited NOESY, $^{13}$C-edited NOESY and pseudo-3D IPAP experiment for H-N residual dipolar couplings, as previously described (Rudloff 2015). Both NOESY experiments used 110 ms mixing time. Most experiments were collected with 128, 64 and 1024 points in the $T_1$, $T_2$, and $T_3$ dimensions, respectively. NMR data were processed with NMRPipe$^{55}$, extended in the indirect dimension via linear prediction, and the resulting spectra were analyzed via Sparky$^{56}$.

Standard Bruker IPAP experiments using 256 pts for each $T_1$ dimension were used to collect RDC data in isotropic and axially-compressed 5.5% acrylamide gel samples, as previously described$^{57}$. We used the program PALES for RDC alignment tensor fitting with a calculated $A_a$ and $A_r$ component of 0.00161 and 0.000600, respectively, for the Ig58/59 model and 0.00163 and 0.000901 for the Ig59 structure$^{58}$. For all experiments, the $^1$H chemical shifts were referenced to external DSS, the $^{13}$C shifts were referenced indirectly to DSS using the
frequency ratio $^{13}\text{C}/^{1}\text{H} = 0.251449527$ and $^{15}\text{N}$ shifts were referenced indirectly to liquid ammonia using $^{15}\text{N}/^{1}\text{H} = 0.101329118$.

Structure calculation  Interproton distance constraints were derived from 3D NOESY experiments ($^{15}\text{N}$-edited and $^{13}\text{C}$-edited 3D NOESY) as described previously. Dihedral constraints $\phi \pm 20^\circ$ and $\psi \pm 15^\circ$ for a-helix and $\phi \pm 40^\circ$ and $\psi \pm 40^\circ$ for b-sheet were included based on TALOS+ and the chemical shift index of $^{1}\text{Ha}$ and $^{13}\text{Ca}$ atoms. An ensemble of structures without dihedral restraints had a backbone rmsd of 0.85 when compared to structures with dihedral constraints. We attempted to further verify the structure by performing a H-D exchange experiment, however this Ig domain remains unfolded after lyophilization. Therefore, hydrogen bond constraints were not tested directly but instead were added into the structure only after the secondary structure was completely determined. Structures calculated without hydrogen bonds had an rmsd of 0.59 when compared to those calculated with hydrogen bonds, indicating that inclusion of these bonds did not drastically influence the overall structure. Hydrogen bond constraints of $r_{\text{HN-O}} = 1.5 \, \text{Å}$ to $2.8 \, \text{Å}$ and $r_{\text{N-O}} = 2.4 \, \text{Å}$ to $3.5 \, \text{Å}$ were included in the final stage of structure calculations, and were based off regions that were clearly in well-defined secondary structural motifs. Pseudopotentials for secondary $^{13}\text{Ca}$ and $^{13}\text{Cb}$ chemical shifts and a conformational database potential were included in the final simulated annealing structural calculations using the computer program XPLOR. Structures run with and without these pseudopotentials show an rmsd of 0.58. The internuclear dipolar coupling (in Hz) were determined from the difference in $J$ splitting between isotropic and radially compressed polyacrylamide, and were incorporated into the final structure calculation as previously described using an energy constant of 0.50. A comparison of structures run with and
without RDC measurements show an rmsd of 0.67. Q-factors were calculated by randomly removing \( \approx 10\% \) of the N-H\(^\text{\textsuperscript{N}}\) RDC data, and then comparing these values to those back-calculated from the structure. The final 20 structures were selected (from 200) based on lowest Q-values and lowest RMSD from the average, and were of high quality based on the statistical criteria listed in Table 1.

**Crystallization and X-ray diffraction** The hanging drop method with 17\% tacsimate, 0.1M HEPES pH 7.5, 4\% PEG3350 was used to obtain Ig59 crystals. Crystals were harvested and frozen in LN\(_2\) after one week using a glucose cryoprotectant and analyzed with the APS Synchotron beamline 19-ID-D. HKL2000 data processing calculated the unit cell to be P 3\(_1\)21.

**Structure refinement** Phases for these experiments were determined in PHENIX ver 1.72.2-869 via molecular replacement using PDB and reflection files from accession number 2YZ8 and (4RSV). The resulting structure was refined using PHENIX ver 1.72.2-869. Coot was used to manually rebuild the structure in iterative rounds of rebuilding and refinement in PHENIX refine, resulting in a 1.18 Å resolution structure with an \( R_{\text{free}} \) value of 0.185. The Molprobity and Coot were used to identify and correct Ramachandran plot outliers. More refinement statistics are given in Table 2.

**SAXS** Small angle X-ray scattering were performed on the F2 beamline at the Cornell High Energy Synchrotron Source (CHESS), using an X-ray source with a beam edge of 9.881 keV (1.2563 \( \text{Å} \)) and an area of 250 mm\(^2\). Protein was passed through size exclusion column immediately before SAXS measurements, and monomeric protein was loaded into a horizontal capillary tube. In the beam line, the sample was oscillated during data collection to avoid sample radiation damage. Data was collected for 3 3-minute cycles. To ensure against aggregation, 1
mg/mL, 3 mg/mL, and 5 mg/mL protein concentrations were collected. Dark field and buffer samples were subtracted from the raw protein data to obtain an $I(q)$ versus $q$ plot using the program RAW $^{63}$. Guinier plots for use in estimating the $R_g$ of obscurin Ig58/59 were then calculated. 10,000 model structures of Ig58/59 were generated using SASSIE, where residues 1 to 92 and 95 to 200 (all residues visible in the individual x-ray structures) were kept constant and residues 93-95 (the linker region) and 201-208 (the His$_6$ tail) were allowed to sample all allowed conformational space. CRY SOL was then used to back-calculate how well these structures fit with the actual SAXS data $^{64}$. Chi squared analysis and protein density map generation were calculated using SASSIE, and the protein shell was visualized using VMD.
Obscurin and titin display a directional preference to force resistance

The sarcomere, the smallest contractile unit in muscle, drives virtually all bodily motion. In order for the sarcomere to work effectively, actin and myosin filaments, along with other peripheral members of the contractile apparatus, must be properly positioned. Skeletal muscles accomplish this complex organizational task through an intricate web of scaffolding proteins that must be simultaneously pliable enough to accommodate motion yet sturdy enough for force propagation. Under the microscope, the most obvious of these sarcomeric macromolecular scaffolds are the Z-disk and the M-band. The Z-disk is perpendicular to the thin filament actin fibers, and functions to align and coordinate actin strands. Its counterpart, the M-band, organizes the myosin bundles. While the Z-disk is largely inflexible, the M-band distorts significantly upon the application of force yet regains its original structure upon muscle relaxation.

Many proteins in the M-band, including M protein, myomesin, obscurin, and titin, are organized as a series of Ig (Immunoglobulin)-like and FnIII (Fibronectin type III)-like domains, arranged in tandem and connected via semi-flexible peptide linkers. These Ig-like domains are always unique in sequence, and often bind specific molecular targets. Thus, proteins containing such structural elements likely provide elastic stability by acting as long flexible fibers that are crosslinked extensively. Implicit in this organization is that the forces holding the M-band together, at least in the aggregate, must be strong; weak protein-protein interactions would break with force, which in turn would unravel the M-band.

Two of the proteins anchored in the M-band, titin and obscurin, are critical for global sarcomeric organization. Titin (3-4 MDa) is the largest polypeptide in humans, and
performs multiple roles in the sarcomere including setting the overall sarcomere length and acting as a stretch sensor\textsuperscript{6,38,75-77}. The 10 C-terminal Ig-like domains of titin are imbedded in the M-band, where they interact with proteins including myomesin, M protein, and obscurin\textsuperscript{32,43,73}. Obscurin partially localizes to the M-band through interactions with titin, some variants of myosin binding protein C (MyBP-C), and myomesin in its N-terminal 1\textsuperscript{st}, 2\textsuperscript{nd}, and 3\textsuperscript{rd} Ig domain, respectively\textsuperscript{48,51-52,73,78}. The C-terminus of obscurin binds to small ankyrin, and through these interactions obscurin is the only known link between this cytoskeletal component of the contractile apparatus and the surrounding sarcomplasmic reticulum membrane and transverse tubule structures\textsuperscript{4,50}.

The extreme N-terminus of obscurin (Ig1) binds directly to the extreme C-terminus of titin (M10)\textsuperscript{51-52,78}. The high resolution structures of one M-band domain of titin (M10) bound to a close cousin of obscurin, obscurin-like Ig1 (OL1), reveal the M10/OL1 complex exists in an antiparallel Ig-Ig formation, with the protein-protein binding surface consisting of extensive inter-protein backbone hydrogen bonds within a large hydrophobic binding interface\textsuperscript{48,52,78}. NMR and more recent X-ray studies show that Ig1 also binds to M10 in this same manner (Rudloff in press,\textsuperscript{52}). With this head-to-tail molecular arrangement, there are two ways in which a force initiated outside of these domains can influence this complex. In one instance, the domains can be peeled apart with a lateral force. In another, they can be sheared apart with a longitudinal force. The directionality of the force depends on the obscurin/titin complex orientation within the M-band, which in turn is dictated by how the full obscurin and titin proteins interact with a myriad of other M-band binding partners (Figure 1A). Given the need for components of the M-band to remain bound under high mechanical forces, a detailed
understanding of how titin and obscurin behave when pulled from different directions may give insight into the absolute titin/obscurin orientation within the M-band. Here we use simulation to probe how the direction of applied force affects the stability of the titin/obscurin complex. We use SMD to study both the strength and the directional dependence of the interactions between both the Ig1 and OL1 domain of obscurin and the M10 domain of titin.

**Results**

The Ig domains in the M10-OL1 structure are in an antiparallel orientation (Figure 1A). Given a) this head-to-tail structure, b) the long filamentous overall architecture of both obscurin and titin, and c) that mechanical force exerted on this complex must be initiated distally, we reasoned that there are two ways in which the domains can be separated. In the first scenario, other molecules do not significantly influence the orientation of the complex. In this situation one would expect a pulling force to peel the two domains apart from each other (Figure 1A, top). This has been experimentally tested on M10/OL1 via AFM. Alternatively, one or both domains may be held rigidly in place. The titin M10 domain is separated from the next Ig-like domain by a presumably flexible linker approximately 100 residues in length. However, the obscurin Ig1 domain is separated from the neighboring Ig2 domain by only a couple of residues, and thus if Ig2 were immobile, Ig1 would have much more limited mobility. In this scenario, shear force is required to separate the titin and obscurin domains (Figure 1A, bottom). Using SMD, we investigated how much force would be required to separate M10 from OL1 in both models.
Figure 1. M10/OL1 peeling vs. shearing models. A) Schematic of the two models by which the titin M10 domain could be separated from the OL1 structure. Top is a peeling model, involving a gradual dissociation of the domains from each other. Bottom is a shearing model, which requires all inter-protein interactions to break almost simultaneously. B) Force-distance trace of the OL1-M10 complex showing all points (gray) and a running average (black) in which 40 points (20 fs of data) were averaged. The top panel is data from the peeling model and the bottom is from the shearing model. C) Force-distance trace of the shearing model (black) and the peeling model (red). D) Work-distance trace of the shearing model (black) and the peeling model (red). In B), C) and D) the ‘0’ position is the point of domain separation.
Initial force-distance profiles show significant noise, despite using a small spring force constant ($K = 1.0 \text{ kcal mol}^{-1} \text{ Å}^{-2}$) (Figure 1B). To partially mitigate this noise and reveals differences in the direction of pulling, a running average of 40 data points (20 fs) was plotted (Figure 1C) \cite{33-34}. In this analysis, the maximum force required to shear this complex apart approaches 350 pN while the force required to peel the domains is roughly 75 pN less. Subsequent simulations always showed this same trend. To examine whether the molecular mechanism of domain separation could explain the differences in the maximum force, we next observed the change in work with separation distance (Figure 1D). While this comparison is normally used to calculate free energy, it can also provide insight into how many independent energy-requiring events are necessary to break OL1 away from M10. This analysis clearly shows that shearing is accomplished in one prolonged step while peeling is a two step process with a long lived intermediate (arrow, Figure 1D). Additionally, the total amount of work to separate the domains is less in the peeling model than in the shearing model. An examination of the force v. distance graph (Figure 1C) provides an explanation for these discrepancies. In the shearing model, a sustained force of greater than 100 pN is exerted on this system between distances of roughly 20 and 10 Å prior to domain separation. This region does not rival the maximum force, however it nonetheless indicates a prolonged period where work is applied to the system in order to separate the domains. The peeling model does not have this feature. Instead, the energy steps required to peel the domains apart take place in two distinct, shorter steps, resulting in less overall work being put into the system over a longer distance. This trend of the peeling model requiring less overall work was seen in all subsequent simulations.
Next, we examined the relationship of the energy steps in both models to molecular events. Backbone hydrogen bonds between Glu92, Tyr94, and Ala96 of OL1 and Val21, Thr23, and Ala25 of M10 initially hold OL1-M10 together (in red, Figure 2A). These bonds form an interprotein antiparallel beta sheet, and are surrounded by extensive hydrophobic interactions consisting of residues Pro11, Pro12, Phe14, Phe17, Ala93, Tyr94, Ala95, and Ala96 of OL1 and

**Figure 2.** A) The initial equilibrated M10-OL1 model, showing the termini and the five interprotein H-bonds in red. Residues involved in initial hydrophobic interactions are drawn as spheres. B) 20 Å pre-break in the shearing simulation. Atoms involved in native hydrogen bonds are colored in red (oxygen) and blue (nitrogen) C) 20 Å before breakage in the peeling simulation. D) % H-bonds (black) and hydrophobic interactions (red) in the shearing model. E) % H-bonds (black) and hydrophobic interactions (red) in the peeling model. Arrows correspond to maximum force peaks in Figure 1B.
Pro11, Val21, Leu22, Thr23, Val24, Ala25, and Ala27 of M10 (spheres, Figure 2A-C). In both the shearing and peeling simulations, these native hydrogen bonds are broken early in the simulation (Figure 2D-E, first arrow). In the shearing model, new transient backbone hydrogen bonds then re-form with residues further down the opposite beta strand. This rupture/reformation pattern repeats in a predictable pattern, and coincides with the high force peaks in the shearing force/distance graph (Figure 1C). Having to break several series of hydrogen bonds explains the large amount of work required to shear the OL1-M10 domains apart (Figure 2B and 2D). The second round of hydrogen bond breaking coincides with a rapid loss of hydrophobic contacts between the two subunits (second arrow, Figure 2D). Since this event does not require a significantly higher force, it seems that hydrophobic interactions make only a minor contribution to mechanical stability than might have been expected.

The peeling model initially follows the same pattern as the shearing model. However after an initial decline in the number of hydrophobic contacts and backbone hydrogen bonds, these values stabilize during a period in which no work is being done on the system (Figure 1D, 2C, 2E). Here, this intermediate complex is metastable and resembles a molten globule, with extensive hydrophobic contacts. At this point the OL1 and M10 domains have pivoted around the interdomain hydrophobic region and the two Ig structures are perpendicular (compare Figures 2A and 2C). This twisting motion precludes reformation of backbone interdomain hydrogen bonds and continues until the domains are parallel before they fully separate. Several sidechain-sidechain and sidechain-backbone hydrogen bonds form and then break during this time. Unlike in the shearing model, these hydrogen bonds do not form in a predictable repeating pattern. These transient hydrogen bond breakages correspond to a broad force increase around -
10 Å and another around -5 Å (Figure 1C).

In both models, the force required to break hydrogen bonds dominates the energy landscape. Analysis of both trajectories provides an initial study into the limitations of hydrophobic interactions to resist mechanical stress. While such interactions resist force, they clearly play an ancillary role here. Without specific bonds, hydrophobic interactions can glide over other hydrophobic surfaces, creating a more malleable interaction surface. In the peeling model, this causes rotation of the two domains. During this twisting motion sidechain hydrogen bonds can form, but these are sequence dependent and less numerous than the transient backbone hydrogen bonds in the shearing model. As a result, this twisting action overall requires less work to separate the two domains.

M10 is the only known titin domain that has multiple binding partners; both obscurin and obscurin-like bind to this region, depending on the cell type79. To test whether the obscurin Ig1 domain behaves in a similar fashion to the OL1 domain, we modeled the obscurin sequence on to the OL1 structure (46% identity) within the M10-OL1 complex, equilibrated this structure, and ran a 6.4 ns molecular dynamics simulation on this complex. The model changed little over the final 2.5 ns, and the resulting RMSD, when compared to the original structure, was low (0.324 Å) (Figure S1). Recently, a crystal structure of the M10-Ig1 complex (pdb 4UOW) was released, and this new structure aligns to our modeled complex with a backbone RMSD of 0.37852. There is virtually no difference in the interdomain interface between the model complex and the crystal structure, lending validity to the practice of using such model structures for SMD. SMD shearing experiments show that our model behaves similarly to the M10-OL1 crystal structure (Figure 3A). The higher force in these traces is likely the result of using a larger spring constant
(K=20.0 kcal mol\(^{-1}\) Å\(^{-2}\) in this figure compared to K=1.0 kcal mol\(^{-1}\) Å\(^{-2}\) in figures 1 and 2). SMD experiments using an energy-minimized I57N M10 mutant produced data similar to WT. This mutation is linked to muscular dystrophy, but has a very similar structure and thermodynamic binding profile to the wild-type protein (Rudloff, in press). The work-distance relationship of these mutated complexes also show similar patterns (Figure 3B), and closely resembles the WT shearing work-distance profile in both scale and shape (Figure 1D). The fact that none of these changes drastically influences the overall force or work profile of the M10-OL1 interaction suggests that this binding event is particularly robust and further supports the hypothesis offered in Rudloff et al. that this mutation does not directly cause muscular dystrophy.

**Figure 3.** A) Force-distance traces of titin-obscurin models. The ‘0’ distance is the point of domain separation. Red is the original OL1-M10 complex, green is an energy minimized model of the OL1-I57N complex, black is a model of the obscurin Ig1 model-M10 complex, and gray is a model of the obscurin Ig1 model-I57N complex. B) Work-distance traces of the titin-obscurin models, using the same coloring as in part (A).
Discussion

The M-band stretches yet remains intact when subjected to strong physical forces. To accommodate this unique characteristic, M-band proteins must be both flexible and strongly anchored to their binding partners. Flexibility is likely a consequence of the modular architecture of M-band proteins, while strength is likely derived from protein-protein interactions. Being firmly anchored in the M-band would seem to be a prerequisite for titin and obscurin. These two proteins extend outside the M-band. Loose association with M-band binding partners could result in significant protein mislocalization upon force application. Yet previous AFM studies show that modest force (30 pN) is sufficient to break the titin/obscurin interaction. Here, we present molecular dynamics data that provides insight into this apparent discrepancy. As observed previously in several other theoretical and experimental works, the strength of the obscurin-titin interaction is directionally dependent. The interaction more strongly withstands force when ‘sheared’ apart than when ‘peeled’ apart. Thus the strength of this interaction depends, in part, on the protein complex orientation relative to the sarcomere.

The forces calculated in these SMD simulations are significantly higher than those measured in AFM. This is a common critique of SMD simulations and can be at least partially attributed the higher pulling speeds required in SMD due to computational limitations; this work used a speed of 10 Å/ns whereas previous AFM pulling speeds for this same complex were 10^{-5} Å/ns. To better mirror AFM work, both the ‘shearing’ and the ‘peeling simulation were conducted at ten-fold slower speed (1 Å/ns). While the maximum forces were smaller when the complex was pulled at lower a slower velocity (250 pN for shearing and 175 pN for peeling; total work of 50 kcal/mol for shearing and 30 kcal/mol for peeling), the values were still higher.
than for AFM (Figure S2). Of note, while this slower-velocity simulation provides greater resolution to the molecular events involved in domain-domain separation, the overall mode of separation remains the same. The shearing trajectory occurs as a concerted, one-step mechanism over a long duration while the peeling mechanism has two distinct energy-requiring steps, each of a shorter duration. Likewise, the mechanisms of hydrogen bond and hydrophobic interaction breakage are similar for both trajectories at this slower velocity.

Many Ig domains unravel when subjected to forces between 150-400 pN\textsuperscript{86-89}. In the shearing simulations presented here, the obscurin Ig1 domain but not the M10 domain partially unravels before the obscurin-titin interaction is broken (Figure S3). While this is perhaps not surprising, it none-the-less provides an interesting example where the protein-protein interaction is mechanically stronger the forces holding together an individual domain \textsuperscript{86-87, 90-91}. It should be noted, however, that the forces required to either unfold an Ig domain or to shear apart the titin/obscurin complex are likely much stronger than the myofibril would ever experience except in situations of extreme muscle overstretching \textsuperscript{47}.

The Ig-like domains in the I-band region of titin have been proposed to sequentially ‘burst’ upon severe overstretching, and this bursting in turn is an intrinsic mechanism to help protect myofibrils from overextension \textsuperscript{3, 47, 75, 86}. The work presented here supports this hypothesis. If properly situated, the M10-Ig1 interaction is strong enough that it could remain intact while other parts of titin unfold. Upon subsequent muscle relaxation and protein refolding, titin would then be correctly oriented to aid in reorganization of the recovering myofibril \textsuperscript{41, 86, 92}.

While the next predicted Ig domain in titin is roughly 100 residues away from M10, the obscurin Ig-like domains are each separated by very short (1-2 residue) linkers. Similar
organizations in other proteins show that this kind of orientation affords each domain only limited flexibility \(^{37,49}\). Also, the numerous protein binding sites that are located near the obscurin N-terminus suggest that obscurin may have restricted motion within the M-band \(^{6}\). When combined with the SMD simulations presented here, it is reasonable to suspect that the titin-obscurin complex could be held in a fixed orientation, and that this orientation could maximize the force required to separate titin from obscurin. Given how extensively crosslinked individual members of the M-band cytoskeleton are, another possibility is that the M-band withstands high force load due to high collective protein-protein avidity, and not strong mechanical affinity. In this scenario the obscurin-titin orientation would not need to be fixed relative to the sarcomere. There is currently no data about the interaction strength between most components of the M-band. Therefore, this affinity versus avidity argument will be the subject of future research.

**Materials and Methods**

*MD simulations* All MD simulations were performed with the PMEMD module of the Amber 12 MD software package, the AMBER ff12SB force field, and a generalized Born implicit solvent \(^{93-95}\). Mutations were incorporated into the OL1 and M10 structure in coot using the ‘mutate’ function. The resulting structures were energy minimized, and equilibrated, until the RMSD of the structure was roughly constant for 3 ns. The non-bonded interaction cutoff distances were set at 100 Å for MD and 150 Å for steered molecular dynamics (SMD). For equilibrium simulations, constant temperature (T = 300K) was enforced using a Langevin
A thermostat with a collision frequency of 1 ps$^{-1}$. The integration time step was 2 fs and all covalent bonds to hydrogen were held fixed with the SHAKE algorithm$^{96}$.

Simulations were run with a small force constant ($K$ in Eq. 1). This force constant was chosen because the size of $K$ is directly proportional to the expected thermal fluctuations in the external force by the one-dimensional Boltzmann distribution of a harmonically bound particle:

$$\sigma_F^2 = k_B T K$$

Using this force constant the thermal fluctuations in the force are roughly equal to those expected in AFM experiments ($\sigma_F \approx 50 \text{pN}$). Further experiments were completed with $\sigma_F \approx 240 \text{pN}$ to see if a larger force constant would have any effect on the force required to separate the domains. These experiments yielded the same results in both the shearing and the peeling models.

Force was applied along a reaction coordinate defined as the distance between the $\alpha$-carbon at the N-terminus of M10 and the $\alpha$-carbon at the C-terminus of OL1 for SMD shearing simulations. In peeling simulations force was applied along a reaction coordinate defined as the distance between the $\alpha$-carbon at the N-terminus of M10 and the $\alpha$-carbon at the C-terminus of OL1$^{11}$. Unless otherwise stated, the protein was pulled at a constant velocity of 10 Å/ns. The SMD spring constant was set to either 1.0 or 20 kcal mol$^{-1}$ Å$^{-2}$.

**Data Analysis** All analysis was done via AmberTools12 and gnuplot$^{93}$. Hydrogen bonds were calculated via the ptraj module and used a distance and angle cutoff of 3.2 Å and 120°. Hydrophobic contacts were calculated with the ‘distance’ tool in cpptraj and were defined with a cutoff distance of 8 Å between the centers of mass of given hydrophobic residues. Protein were deemed to be separated via manual inspection, defined as when the two domains no longer had any atoms within 3 Å of each other. This point was normalized as 0 for all experiments.
Protein images for this manuscript were rendered in Pymol, and data traces were analyzed in Microsoft excel.

**Figure S1.** RMSD traces of molecular dynamic simulations of all models used in this manuscript. All models were allowed to equilibrate until the RMSD plot stabilized for at least 2.5 ns.
Figure S2. Force-distance (A) and work-distance (B) plot of OL1-M10 pulled at 1 Å/ns. In both plots red represents the peeling model and black represents the shearing model. While the maximum force and work is lower when the complex is pulled more slowly, both trajectories nonetheless retain the same overall trace shape as when the system is pulled at greater velocities. C) % original H-bonds (black) and hydrophobic interactions (red) in the peeling model. D) % original H-bonds (black) and hydrophobic interactions (red) in the shearing model, show a single concerted unraveling step, even at lower pulling velocities.
Figure S3. RMSD profiles of individual domains from the sheared trajectories in figure 3. Note that most domains remain close to their original structure when sheared apart. OL1 exhibits more perturbation, which is the result of the more peripheral beta sheets being significantly changed at the start of the trajectory.
Re-examination of the Structure and Elasticity of the Ig65-Ig70 Segment of Titin

Titin, the longest protein in the human genome and the largest known polypeptide, sets the sarcomere length within striated muscle cells. Titin’s ability to both stretch and recoil is fundamental in preventing muscle overstretching, and helps myocytes return to their original length. In this capacity, titin acts as a mechanosensor. Recent studies have examined the mechanosensing capabilities of tandem Ig and Fn-III-like domains that make up over 90% of titin. While many of these domains are closely associated with other parts of the cytoskeleton, there are ~100 tandem domains in the I-band that are relatively free from extensive target protein interactions (reviewed in). Several recent papers have shown that this I-band region acts as a modified entropic spring whereby each domain behaves as a link in a chain connected by 3-5 residue interdomain linker ‘hinges’. The further a polypeptide is stretched, the more energy additional stretching requires. Through extensive molecular dynamics (MD) simulations, Lee et al. took this idea further and developed a comprehensive mathematical model where tandem Ig domains behave as a kind of entropic-enthalpic spring, also known as the modified worm-like model (mWLM). In this model, the Ig domains can move stochastically relative to one another, though there is also some breaking and forming of transient domain-domain and domain-linker non-covalent bonds. The mWLM agrees extremely well with the experimental data and explains titin’s stretch resistance at very low forces (1-5 pN). At slightly higher forces (~5 pN), the mainly disordered PEVK, N2A, and N2B regions of titin straighten out in an entropic-enthalpic spring mechanism. At extremely high forces (>50 pN), discrete titin Ig and Fn-III domains completely unfold, although the physiological relevance of such an event is uncertain. Thus, titin displays a
continuum of soft elasticity with the application of force 34.

This nuanced view of titin’s flexibility at low forces runs somewhat counter to the crystal structure of 6 tandem titin domains, I65-70 (I6), which depict well-formed interactions between the linker regions and Ig domains 49. Such fixed interdomain interactions lead to the ‘carpenter’s ruler’ model. This model reconciles several incongruences that arise from thinking of titin as a worm-like model: data collected by cryo-EM show large segments of the I-band region to be in straight lines with only occasional bend points, and SAXS data indicate a mostly elongated structure 49, 106-108. Additionally, the persistence length of the titin I-band region is longer than a worm-like model would suggest 49, 109.

In an attempt to resolve the discrepancy between the worm-like and ‘carpenter’s ruler’ models, we re-analyzed the published crystal structure data 37, 49. Many of the linker/domain interactions that are present in the original model and contribute to an elongated overall structure are either absent or ambiguous in our structure. Examination of symmetry mates suggest that the extended conformation of titin is not based on stabilizing domain-linker interactions but based instead on crystal packing interactions.

Molecular dynamics simulations, performed by ourselves and others, support the theory that this region of titin can be elongated, however the energy difference between a slightly bent and a straight molecule are below the threshold of thermal noise 37. Together, these X-ray and MD analyses structurally validate the mWLM. Titin likely forms transient interdomain interactions, which explains the protein’s resistance to stretch and its persistence length. These interactions, combined with limited flexibility due to steric hindrance between Ig domains, serve to keep titin in a mostly elongated orientation. While there is computational and circumstantial
evidence for stabilizing interdomain interactions (26), their presence in this crystal structure is not definite.

**Results**

The original structure of I65-I70 is hook-shaped; the first two domains are oriented 114 degrees relative to the final four domains, which are nearly linear \(^{49}\). The high B factors of the published structure prompted our re-analysis of the electron density in 3B43 to determine if the sections with high B-factors had well-defined electron density. A 1.5 \(\sigma\) cutoff was used to fit the model in Coot. Continuous, unambiguous density for the backbone and side chains was required; all other residues and side chains were removed (Table 1 for fitting statistics and Figure 1A). This newly refined structure has an RMSD of 0.6 Å compared to the original with several important differences. The overall B-factor is significantly lower than the original structure (88.4 Å\(^3\) vs 118.5 Å\(^3\); Figure 1A inset). Furthermore, domain Ig68 is almost completely absent. In our model, interactions between the Ig domains and linker residues, which were previously postulated to stabilize the structure in an elongated form, are either poorly-defined or absent.

**Table 1.** I6 crystallography statistics. Data collection methods and statistics are found in \(^{49}\).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spacegroup</td>
<td>P6(_{5}22)</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>(a=b=141.43\text{Å}, c=166.01\text{Å}, \alpha=\beta=90^\circ, \gamma=120^\circ)</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>16.99-3.30</td>
</tr>
<tr>
<td>Solvent content, %</td>
<td>67.1</td>
</tr>
<tr>
<td>Matthews coefficient</td>
<td>3.74</td>
</tr>
<tr>
<td>Completeness, %</td>
<td>98.6</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>14953</td>
</tr>
<tr>
<td>No. of protein atoms in asymmetric unit</td>
<td>3371</td>
</tr>
<tr>
<td>R factor/R free, %</td>
<td>28.4/31.0</td>
</tr>
<tr>
<td>RMSD bond length, Å</td>
<td>.0084</td>
</tr>
<tr>
<td>RMSD bond angle, °</td>
<td>1.992</td>
</tr>
<tr>
<td>Avg B factor (B factor range), Å(^3)</td>
<td>88.5 (19.9-200.2)</td>
</tr>
<tr>
<td>Wilson B factor, Å(^3)</td>
<td>82.6</td>
</tr>
<tr>
<td>Ramachandran favored, %</td>
<td>90.6</td>
</tr>
<tr>
<td>Ramachandran outliers, %</td>
<td>2.2</td>
</tr>
</tbody>
</table>
**Figure 1.** A) Cartoon representation of Ig65-70, colored by B-factor. Inset is the distribution of B-factors within the model.  B) Linker region between Ig67 and Ig68, showing a lack of interactions that keep the molecule rigid.  C) Example of the extensive packing of Ig66 against its symmetry mate.  D) Global view of the crystal packing of Ig65-Ig70 relative to symmetry mates.  E) Global view of the crystal packing of the 2RIK structure with central molecules colored via B-factors (range from 20-65 Å³).
(Figure 1B). These residues were removed due to poor electron density (see Materials and Methods for details). Analysis of the entire structure reveals a positive correlation between extensive crystal packing interactions and low B-factors (Figure 1A and C; Figure S2).

From this analysis, we reasoned that the overall shape of this region of titin could be explained by crystal packing interactions and not through inter-protein interactions (Figure 1D). Since titin is predicted to be moderately flexible, this hypothesis could explain the dichotomy between MD simulations and the crystal structures. We therefore investigated the structure of Ig68-69-70 (2RIK), which was previously solved at 1.6 Å resolution to determine if crystal packing may have also contributed to its extended conformation. As can be seen in Figure 1E, this construct has extensive packing interactions against its neighbors, and has correspondingly lower overall B-factors. Overlays of the same domain in 2RIK and 3B43 show that, while the tertiary structure of the Ig domains are virtually identical, the orientation of the Ig domains relative to each other vary (Figure S1). We conclude that the overall shape of I6 observed in the 3B43 crystal structure is most likely dependent on crystal packing, not inter-protein domain-domain or domain-linker interactions.

To examine the flexibility of titin Ig65-Ig70, we performed a series of MD and SMD simulations on the original titin structure. Model equilibration using implicit solvent conditions and the AMBER ff12SB force field showed that each domain displayed significant motions relative to its neighbor over 10 ns (Figure S2). This is despite the Ig domains themselves being essentially static (average RMSD per domain over 10 ns is about 1.9 Å). This result agrees with similar experiments performed by Lee et al. To further probe the stiffness of this section, we brought together the two ends of this structure together at a rate of 15 Å/ns, from 232 Å to 141
Å. Analyses show that the straight part of the original structure (Ig67-Ig70) bent readily, and that it required less than 0.6 kcal/mol of work (which is a rough cut-off for noise attributable to random thermal motion) to decrease the end-to-end distance to 190 Å (Figure 2A-D). Smaller-diameter models required progressively more work. These data suggest that while each domain is mobile relative to its neighbor, mobility only extends to roughly 140 degrees, in agreement with previous data \(^{37}\). At smaller angles, the Ig domains begin to run in to each other; thus this molecule is intrinsically partially elongated. These data agree well with previously published MD bending data, as well as experimental SAXS data, showing that the average length of the I6 region is 220 Å \(^{37,49}\).

**Discussion**

Here we re-analyze the Ig65-Ig70 crystal structure and suggest that the X-ray data do not support the ‘carpenter’s ruler’ model of titin in solution. Instead, these data agree with subsequent MD simulations showing that titin domains move mostly independently of each other, with only transient inter-domain or domain-linker interactions \(^{34,37}\). This is particularly evident in Ig68, whose dynamic movement makes it virtually invisible in the structure. The presence of motion in this region is likely caused by a lack of symmetry mates to pack against, resulting in higher-than-average B-factors.
Figure 2. A) Force (in kcal/mol/Å) v. distance graph of Ig65-70 (PDB 3B43). B-D) Images of Ig65-70 at B) 0 C) 23 and D) 45 Å of compression.

Our new model of Ig65-70 does not explain EM data that indicate that this region of titin is mostly elongated. Several possibilities could explain this discrepancy. First, it could be that...
transient interdomain and domain-linker interactions, while weak, short-lived, and not present in the crystal structure, are present in sufficient quantities as to exert a substantial stabilizing force on the molecule. These interactions could be overwhelmed by thermal motion and crystal packing interactions and might not be seen in crystal structures. Such interactions are built into the mWLM and seem to fit the wet lab stretching data\textsuperscript{34, 37}. Previous SMD experiments had problems in explicit solvent with hydrodynamic drag\textsuperscript{37}. Thus, another possible explanation is that titin, once straight, tends to stay straight because of insurmountable solvent drag. Lastly, it may be that in the conditions under which EM experiments were conducted, titin is particularly amenable to forming straight lines.

MD simulations argue for a mWLM, yet previous a crystallography structure suggest a more rigid molecule. In this work, we identify crystal packing artifacts as contributing to the apparent rigidity of the published crystal structure. Thus, the crystallography data in fact also suggest a dynamic system. This more nuanced view of the X-ray data reconciles this structure with the subsequent MD simulations.

**Materials and Methods**

*Structure refinement* PDB and reflection files from accession numbers 3B43 and 2RIK were used for these experiments. For 3B43, the structure was refined using PHENIX ver 1.72.2-869. Coot was used to manually rebuild the structure in iterative rounds of rebuilding and refinement in PHENIX refine\textsuperscript{110}. Criterion for inclusion of an amino acid was continuous, clearly defined electron density at a sigma value greater than 1.5 for at least two residues. The
Molprobity server and Coot were used to identify and correct Ramachandran plot outliers\textsuperscript{111-112}. Refinement statistics are given in Table 1.

**MD simulations** All MD simulations were performed with the PMEMD module of the Amber 12 MD software package, the AMBER 12SB force field, and a generalized Born implicit solvent\textsuperscript{113-115}. The non-bonded interaction cutoff distances were set at 100 Å for MD and 150 Å for steered molecular dynamics (SMD). For equilibrium simulations, constant temperature (T = 300K) was enforced using a Langevin thermostat with a collision frequency of 1 ps\textsuperscript{-1}. The integration time step was 2 fs and all covalent bonds to hydrogen were held fixed with the SHAKE algorithm.

SMD simulations fixed the α-carbon at the N-terminus of Ig65 (A1) and applied force to the α-carbon at the C-terminus of Ig70 (Q567), moving it closer to the N-terminus\textsuperscript{11}. In this way the protein was compressed at a constant velocity of 15 Å/ns. The SMD spring constant was set at 10 (kcal mol\textsuperscript{-1} Å\textsuperscript{2})

**Figure S1.** Alignment of 2RIK (cyan) and 3B43 (green), showing that the placement of the Ig domains are different between the two structures
Figure S2. Crystal packing between A) Ig69 and B) Ig70 and their respective symmetry mates.

Figure S3. Short linker regions of I6 between A) Ig66-67, B) Ig67-68, and C) Ig69-70.
**Figure S4.** Snapshots of MD simulations at A) 2, B) 5, and C) 9 ns of equilibration showing the straightening and bending of the final four domains of I6. The models were aligned from residue 98 to 288.
References


94. Hornak, V.; Robert, A.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. L.
    Comparison of multiple Amber force fields and development of improved protein backbone

95. Nguyen, H.; Roe, D. R.; Simmerling, C. L. Improved Generalized Born Solvent Model
    Parameters for Protein Simulations. *Journal of Chemical Theory and Computation* **2013**, *9*,
    2020-2034.

96. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian

97. Matsubara, S.; Maruyama, K. Role of connectin in the length-tension relation of skeletal and


99. Granzier, H. L.; Labeit, S. The giant protein titin: a major player in myocardial mechanics,

100. Ma, K.; Wang, K. Malleable conformation of the elastic PEVK segment of titin: non-co-

    C.; Labeit, D.; Gregorio, C. C.; Granzier, H.; Labeit, S. The complete gene sequence of titin,
    expression of an unusual approximately 700-kDa titin isoform, and its interaction with


108. Marchetti, S.; Sbrana, F.; Toscano, A.; Fratini, E.; Carla, M.; Vassalli, M.; Tiribilli, B.; Pacini, A.; Gambi, C. M. beta-Connectin studies by small-angle x-ray scattering and single-


Appendix I: Transformation

- Add 1.0µL DNA in microcentrifuge tube on ice
- Thaw BL21 cells on ice
- Add 40 µL BL21 to DNA tube
- Wait for 20 min on ice
- Heat shock in 42°C water bath for 30 sec and put back on ice for 1 min
- Add 400 µL LB
- Put in 37°C incubator for 30 min-1hr
- Plate 200µL onto kanamycin agar
Appendix II: Protein Growth

Make starter culture the day before:
• Add Kanamycin to the a small flask of LB that has been autoclaved (depends on the size of the growth)
• With the tip of a pipet scrape a colony from gel dish and add to the flask
• Let shake in the incubator overnight at 37°C

Start growth the next morning:
• Add kanamycin to each flask of LB which have be autoclaved (1mL to 1L)
• Pour equal amounts of starter culture to each flask
• Let shake in the incubator at 37°C

Induce 3-5 hours later:
• When the OD of the solution is 0.6 at 600nm add about 0.24g of IPTG

Spin down about 4 hours later:
• Pour bacteria into centrifuge tubes making sure to balance them
• Spin for 10 min at 4°C with the specifications of whatever rotor you are using
• Pour off liquid
• Scrape pellet and put in 50mL conical tubes
• If not purifying right away freeze in -80°C freezer
Appendix III: Protein Purification

- Add 100 µL PMSF to pellet
- Resuspend pellet in binding buffer using the vortex
- Put the resuspended pellet into a metal centrifuge tube and place in ice bath
- Sonicate with the needle as close to the bottom without it actually touching the bottom for 30 sec per mL with 15 sec on and 15 sec off to make sure it doesn’t overheat
- Centrifuge at 14,000 rpm for 45 min at 4°C
- Wash nickel column with about 5 volumes of binding buffer (*never let the column run dry!!)
- Run the supernatant from the centrifugation over the column and collect fractions of about 75-100 drops/tube
- Run through 25 mL binding buffer (record what fraction you change buffer)
- Run through 80 mL wash buffer
- Run through 40 mL elution buffer
- Run a sample of each buffer fraction plus all of the elution fractions on a gel to see if you have pure protein!
- Concentrate
- Then you may have to run size exclusion if not completely pure
  - Take off liquid to the top of the resin
  - Slowly pipet in concentrated protein around edges
  - Rinse the microcentrifuge tube and add that to the column
  - Let sink into resin
  - Slowly add size exclusion buffer
  - Let run for about 8 hours collecting a fraction about every 7 minutes
  - Take OD of tubes at 280nm to see where the protein is
  - Run gel of tubes with high OD to see if you have pure protein!
Appendix IV: Molecular Dynamics

Create a username (Dr. Sumner did this)

To log in type in X-11 terminal:
```bash
ssh -X username@faust.chemistry.jmu.edu
```
NOTE: faust is the supercomputer in Burress, flamel is Dr. Sumner’s computer

Make a folder in your home directory:
```bash
mkdir nameoffolder
```

Copy a pdb from the computer to the supercomputer:
```bash
scp model.pdb username@faust.chemistry.jmu.edu:~/foldername
```

LEAP: NOTE: you are using AMBER 12SB forcefield
```bash
xleap -x -f $AMBERHOME/dat/leap/cmd/leaprc.ff12SB
model =loadpdb “modelname.pdb”
set default PBRadii mbondi3
saveamberparm model modelname.prmtop modelname.inpcrd
quit
```

Minimization:
```bash
emacs modelname.in
obscurin
&cntrl
imin =1,
maxcyc =5000,
cyc =500,
ntb =0,
igb =8,
cut =99,
/
Save: control x, control s
Exit: control x, control c
```

NOTES:
- make sure you hit enter after the backslash
- change obscurin to be what you want
- ntb=0 means you are using implicit solvent
- igb=8 tells it what forcefield you are using (ff12SB)
- cut=99 you can change- tells it how many angstroms away each atom can see
sander –O –i modelname.in –o modelname.out –c modelname.inpcrd –p
modelname.prmtop –r modelname-min.rst

To visualize in VMD:
   vmd modelname.prmtop
   in VMD: file new molecule
   load prmtop file
   filename: .rst file
   file type: Amber7 restart

**Equilibration:**
   emacs modelname-equil.in
   modelname-equil
   &cntrl
   irest=0, ntx=1, ig=-1,
   imin = 0, ntb = 0,
   igb = 8, ntp = 1000, ntwx = 1000,
   ntt = 3, gamma_ln = 1.0,
   tempi =0.0, temp0 = 300.0,
   nstlim = 100000000, dt = 0.002,
   cut = 100,
   ntwr=1000,
   NTC=2, NtF=2,
   ioutfm=1, ntxo=2
/
Save and quit
   NOTE: If starting equilibration for the first time (for that model) irest=0 and
   ntx=1. If restarting equilibration irest=1 and ntx=5

   emacs submit-script.sh (which has been copied into folder)
   at bottom:
   mpirun -np 24 pmemd.MPI -O -i modelname-equil.in -o modelname-equil.out -c
   modelname-min.rst -p modelname.prmtop -r modelname-equil.rst -x modelname-
   equil.mdcrd

   NOTE: If restarting equilibration change appropriately to equil2, etc.

qsub submit-script.sh (submits job)
qstat (to see if it is running -r)
ls –ltr (see what files it is writing and the time)

If need to quit job: first find job number from qstat (the first number on the left) and then
type qdel #
If need to delete core files: rm –f core*
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
```

To view better in VMD:

```
Display: orthographic, depth cueing off, settings-near clip as small as possible
Graphics: drawing method-new cartoon
```

**Steered Molecular Dynamics (pulling or compression):**

```
emacs modelname-equil.in
modelname-equil
&cntrl
irest=1, ntx=5,ig=-1,
imin = 0, nt b = 0,
igb = 8, npr = 1000, ntwx = 1000,
ntt = 3, gamma_ln = 1.0,
tempi =300.0, temp0 = 300.0,
stlim = 100000000, dt = 0.001,
cut = 999,
jar=1,
ntwr=1000,
NTC=2, NtF=2,
ioutfm=1, ntxo=2
/
&wt type='DUMPfreq', istep1=1000, /
&wt type='END', /
DISANG=dist.RST
DUMPAVE=dist_vs_t
LISTIN=POUT
LISTOUT=POUT
```

NOTE: change model name, temperature, nstlim, dt, and cut as see fit
nstlim is number of steps- change to create your velocity
dt is time for each step (0.001=1 ps)

NOTE: for explicit solvent:

```
ntb=1
igb=0
cut=8
```
dt=0.002

write dis.RST file

For constant velocity:
\&rst iat=5,2754, r2=59.3, r2a=659.3, rk2=10.1/

NOTE: 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 distance
rk2 is the spring constant

For constant force:
\&rst iat=5,2754, r1=1., r2=1., r3=1., r4=5., rk2=10., rk3=-0.54/

NOTE: iat is still the two atoms that are being pulled
r1-4 stay the same
rk3=\-F/8

force needs to be in kcal/Åmol. Example: 500pN[(1 kcal/Å)/4.18x10^25 pN](6.02x10^23)=7.2 kcal/Åmol

emacs submit-script.sh
change appropriately - example:
mpirun -np 24 pmemd.MPI -O -i 5859model-equil.in -o 5859model-pull.out -c 5859model-equil3.rst -p 5859model.prmtop -r 5859model-pull.rst -x 5859model-pull.mcrd

qsub submit-script.sh
cat mdinfo (tells you how many ns a day it is going)

To get work plots:

plot ‘dist_vs_t’ u($1-startdistance):4 w l
set xri[0:25] (whatever part you want to plot)
replot
f(x)=m*x+b
fit f(x) ‘dist_vs_t’ u($1-startdist):4 via m,b
plot ‘dist_vs_t’ u ($1-startdist):4 w l,f(x)

Hydrogen-bond analysis:

need analyze-hbond.ptraj file:

trajin 5859model-pull.mcrd 1 17769 1

#-- Donors from standard amino acids
donor mask :GLN@OE1
donor mask :GLN@NE2
donor mask :ASN@OD1
donor mask :ASN@ND2
donor mask :TYR@OH
donor mask :ASP@OD1
donor mask :ASP@OD2
donor mask :GLU@OE1
donor mask :GLU@OE2
donor mask :SER@OG
donor mask :THR@OG1
donor mask :HIS@ND1
donor mask :HIE@ND1
donor mask :HID@NE2

#-- Acceptors from standard amino acids
acceptor mask :ASN@ND2 :ASN@HD21
acceptor mask :ASN@ND2 :ASN@HD22
acceptor mask :TYR@OH :TYR@HH
acceptor mask :GLN@NE2 :GLN@HE21
acceptor mask :GLN@NE2 :GLN@HE22
acceptor mask :TRP@NE1 :TRP@HE1
acceptor mask :LYS@NZ :LYS@HZ1
acceptor mask :LYS@NZ :LYS@HZ2
acceptor mask :LYS@NZ :LYS@HZ3
acceptor mask :SER@OG :SER@HG
acceptor mask :THR@OG1 :THR@HG1
acceptor mask :ARG@NH2 :ARG@HH21
acceptor mask :ARG@NH2 :ARG@HH22
acceptor mask :ARG@NH1 :ARG@HH11
acceptor mask :ARG@NH1 :ARG@HH12
acceptor mask :ARG@NE :ARG@HE
acceptor mask :HIS@NE2 :HIS@HE2
acceptor mask :HIE@NE2 :HIE@HE2
acceptor mask :HID@ND1 :HID@HD1
acceptor mask :HIP@ND1,NE2 :HIP@HE2,HD1

#-- Backbone donors and acceptors for this particular molecule
#  N-H for prolines do not exist so are not in the mask
#
donor mask @O
acceptor mask :95-121,123-136,138-159,161-182@N :95-182@H
#Terminal residues have different atom names
donor mask @OXT
acceptor mask :1@N :1@H1
acceptor mask :1@N :1@H2
acceptor mask :1@N :1@H3
#-- series hbt is just a placeholder to ensure we get the full analysis. If you don't
#have the word series you don't get a full analysis.
hbond print .05 series hbt

NOTES: the two numbers after .mdcrd in the first line are the range of steps that you
want to analyze. You can put # before any in the acceptor or donor mask depending on
what you’re looking for. acceptor mask @N is all residues minus prolines. You can also
change these numbers to only look at certain amino acids.

ptraj _____prmtop <analyze-hbond.ptraj> analyze-hbond.out

Saving a pdb from VMD:
open .mdcrd (or whatever you want) in VMD
go to frame you want to save
click on the molecule name
File-save coordinates
type “all” into selected atoms
choose pdb for filetype
change first and last to be frame # you want to save (stride=1)
save all at once
click save
put in file name and change directory if needed

To move from supercomputer to ours:
open new terminal on our computer
scp –p username@faust.chemistry.jmu.edu:~/folderitsin Documents/folder
(folder on our computer it will be put in)