Spring 2015

Biophysical characterization of naturally occurring Titin M10 mutations

Michael William Rudloff
James Madison University

Follow this and additional works at: https://commons.lib.jmu.edu/honors201019

Part of the Biochemistry Commons, Biophysics Commons, Molecular Biology Commons, and the Structural Biology Commons

Recommended Citation
https://commons.lib.jmu.edu/honors201019/111

This Thesis is brought to you for free and open access by the Honors College at JMU Scholarly Commons. It has been accepted for inclusion in Senior Honors Projects, 2010-current by an authorized administrator of JMU Scholarly Commons. For more information, please contact dc_admin@jmu.edu.
Biophysical Characterization of Naturally Occurring Titin M10 Mutations

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

by Michael William Rudloff

May 2015

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

FACULTY COMMITTEE:

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

Reader: Jonathan Monroe, Ph.D
Professor, Biology

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 Biology Symposium on April 17th.
# Table of Contents

List of Figures ...................................................... 3  
Acknowledgements .................................................. 4  
Abstract .............................................................. 5  
Introduction .......................................................... 6  
Results ................................................................. 12  
  Expression and Purification ........................................ 12  
  Initial Structural Analysis ......................................... 14  
  NMR High Resolution Structural Analysis: H56P Mutation ........ 19  
  ITC Functional Analysis: Ig1 Binding Affinity .................. 22  
  NMR High Resolution Structural and Dynamics Analysis: I57N Mutation 25  
Discussion ............................................................ 34  
Methods ................................................................. 37  
References ............................................................ 41  
Appendix A .............................................................. 45
# List of Figures

1. Muscle Contraction 6
2. Muscle and Sarcomere Structure 8
3. Model of Titin M10 Bound to Ig1 of Obscurin 11
4. SDS-PAGE of the M10 Constructs 13
5. Absorbance$_{280}$ of M10 Constructs After Size Exclusion 13
6. Circular Dichroism Analysis 15
7. Tryptophan fluorescence Analysis 16
8. 1D NMR spectra on M10 constructs 17
9. 1D and 2D NMR on H56P Mutation 18
10. Turbidity Analysis on W46K 18
11. Introduction to HSQC and Sequence Specific Assignments 20
12. HSQC Analysis and Comparison: WT and H56P Mutation 21
13. Introduction to Isothermal Titration Calorimetry 22
14. Isothermal Titration Calorimetry Analysis of Constructs 24
15. HSQC Comparison: WT and I57N Mutation 25
16. Residual Dipolar Couplings 26
17. Heteronuclear NOE Analysis 27
18. Introduction to Relaxation Dispersion Experiments 29
19. Relaxation Dispersion Analysis 30
20. HSQC Analysis and Comparison: WT-Ig1 and I57N-Ig1 32
21. Molecular Dynamics of WT and I57N Mutation 33
Acknowledgements

I would like to thank my advisor and mentor, Dr. Nathan Wright who has put in the time to not only teach me the experiments and techniques I have used during my four years of research, but also instilling a passion for science and life-long learning. I cannot express how grateful I am for the opportunities that you provided to me and the great research experiences I have had during my time at JMU. I would also like to thank Dr. Chris Berndsen, Dr. Jonathan Monroe, and Dr. Gina MacDonald for guidance and support of both my research and the writing process. This thesis would not have been possible without the love and support from my friends and family here at James Madison and at home. Thank you for always believing in me through the good and the bad of my undergraduate career.

I would also like to acknowledge the financial support from JMU start up funding, Tri-Beta Research Foundation, Jeffress Grant J-1041, NSF REU Funding, and the Research Corporation Departmental Award.
Abstract

The giant human muscle proteins titin and obscurin are important for sarcomeric organization, stretch response, and sarcomerogenesis in myofibrils. The extreme C-terminus of titin (the M10 domain) binds to the N-terminus of obscurin (the Ig1 domain) in the M-line, an interaction that is critical for sarcomere stability. The high-resolution structure of human M10 has been solved, along with M10 bound to one of its two known molecular targets, the Ig1 domain of obscurin-like protein. Multiple M10 mutations are linked to limb-girdle muscular dystrophy type 2J (LGMD2J) and tibial muscular dystrophy (TMD), however the effect of the M10 mutations on protein structure and function has not been thoroughly characterized. Here, all four naturally occurring human M10 missense mutants have been engineered and biophysically characterized in vitro with the hypothesis being the mutations ablate this critical interaction. Two of the four mutated constructs are severely misfolded and cannot bind to the obscurin Ig1 domain. One mutation, H66P, is folded at room temperature but unfolds at 37 °C, rendering it binding incompetent. The I57N mutation shows no significant structural, dynamic, or binding differences from the wild-type domain. We suggest that this mutation is not directly responsible for muscle wasting disease, but is instead merely a polymorphism found in symptomatic patients. Understanding the biophysical basis of muscle wasting disease can help streamline potential future treatments.
**Introduction**

Skeletal muscles are composed muscle cells that encompass myofibrils, which contain repeating units of the fundamental contractile apparatus, the sarcomere. Sarcomere contraction is most frequently described in the sliding filament theory, where bundles of the molecular motor myosin slide past actin in an ATP and calcium-dependent manner (Figure 1). The net effect of this concerted motion is the shortening of muscle, which allows an organism to effectively perform various movements (1; 2) (Figure 1). One facet of sarcomere assembly that has long been appreciated but never fully explained is the fact that the sarcomere must be both structurally stable to withstand the forces of contraction, yet flexible, to allow for effective range of motion (3). Multiple structural proteins, including titin and obscurin work in concert to weave a flexible yet strong cytoskeletal latticework, which in turn endows muscle tissue with these traits (3-5). Likewise, mutations in these giant muscle proteins lead to problems in sarcomere development, atypical structure, and loss of contractile ability (3-5).

Figure 1: Schematic of contraction with myosin (Red) and actin (Blue) depicted in a chain of sarcomeres or myofibril. The sliding filament theory occurs after nerve signals trigger calcium influx into the sarcomere, freeing the myosin binding site. ATP is hydrolyzed, “cocking” the myosin head group (1), myosin-ADP complex then binds to actin (2), myosin releases ADP and snaps back to resting position (3), and myosin binds ATP and releases actin (4). The result cycle is the overall contraction of the sarcomeres and thus the muscle (6).
Titin is a 3-4 MDa protein that spans half the length of the sarcomere (1.2 μm) with its N-terminus attached at the Z-disk and its C-terminus embedded in the M-line (Figure 2C) (4; 5). It is highly modular, consisting of up to 300 individual Immunoglobulin(Ig)-like and fibronectin-like domains interspersed with intrinsically disordered regions and signaling domains (reviewed in (7) and (8)). Knockdown studies show that titin is necessary for sarcomerogenesis and acts as a molecular scaffold for an array of other structural and functional proteins (9; 10). Besides these organizational and developmental roles, titin works as a direct stretch mechanosensor through its kinase domain. Its tandem Ig-like domains provide a spring-like resistance that helps protect the muscle from overstretching and prompts the sarcomere to return to optimal length after force is removed (11-14).

Titin’s function is dependent on secure adherence of both ends of the molecule to the Z-disk and M-line of the sarcomere. While the N-terminus is embedded in the Z-disk, the C-terminus is secured at the M-line, in part through an interaction between its own M10 domain and the N-terminal domain of either obscurin or the closely related obscurin-like protein (termed Ig1 and O11, respectively) (Figure 2) (10). Obscurin has a similar modular architecture to titin, and ranges in size from 720-900 kDa (15). Obscurin binds to both titin and small ankyrin, and is the only known link between the contractile apparatus and the surrounding sarcoplasmic reticulum/T-tubules (Figure 2) (15-17). While obscurin’s function is still not fully understood, it likely provides structural support and longitudinal organization for mature sarcomeres in addition to helping organize the sarcomere during development (18; 7; 19).
Figure 2: A) Schematic of a skeletal muscle with major components identified (20). B) Electron micrograph of a myofibril consisting of interconnected sarcomeres with important landmarks identified (21). C) Simplified schematic of giant muscle proteins involved in the sarcomere. Titin is colored magenta, obscurin is colored blue, actin is colored green, and myosin is depicted as gray hook-shaped proteins extending from M-line (22).
Various missense and deletion mutations affecting the C-terminus lead to the truncation of titin and are associated with muscle degeneration and malformation (23-25). However, four missense mutations in the extreme C-terminal domain of titin (the M10 domain) that do not truncate the protein are implicated in limb girdle muscular dystrophy 2J (LGMD2J) and tibial muscular dystrophy (TMD), likely due to the ablation of the critical titin-obscurin binding event at the M-line of the sarcomere (26-30).

LGMD2J and TMD are gradual muscle wasting diseases that can progress to eventual wheelchair confinement of the patient (31). LGMD2J involves the breakdown of skeletal muscles in the arms, legs, and to some extent, the torso. The onset of this condition occurs within the first two decades of life and can lead to loss of function of arms and legs by age 20. TMD is characterized by weakness and atrophy of the lower legs, specifically the anterior tibialis muscle and other pedal dorsiflexors (32; 31). This condition occurs between ages 30-50 and can lead to either the impediment or prevention of walking and running (26; 30). While the exact molecular process for either condition is not known, disease progression always involves the molecular breakdown of specific skeletal muscles (31). Muscle biopsies taken from patients afflicted with LGMD2J and TMD show atypical sarcomeric organization and the appearance of fatty deposits in the muscle tissue. It is speculated that this sarcomere remodeling process is triggered, at least in part, as a consequence of their inability to withstand forces inherent from repeated contractions and stretchings.

To date, previous studies have linked four naturally occurring missense mutations of the M10 domain found in various European populations or families, 37EVTW→VKEK, H56→P, L66→P, and I57→N, to the propagation of LGMD2J and TMD (26-28; 30). 37EVTW→VKEK was found in a Finnish population and was the first M10 mutation associated with these diseases.
Recent studies have shown that patients heterozygous for this mutation exhibit TMD while patients homozygous for this mutation exhibit the more severe LGMD2J (27). The mutants H56→P, L66→P, and I57→N were identified later in Italian, French, and Belgian families respectively, all of which appear to propagate TMD with relatively similar severity (26-28; 30).

Recent crystal structures of the titin M10 domain in complex with Ol1 show a unique Ig-Ig binding interface, consisting of five hydrogen bonds making up a interdomain antiparallel β-sheet surrounded by a large hydrophobic network (Figure 3) (33; 34). When mapped to the M10 structure, mutations are located distally in both sequence and three-dimensional space to the titin-obscurin binding site (Figure 3) (33). Thus, it is not clear how such mutations would affect the M10-obscurin binding event, or if such mutations could be the molecular basis for LGMD2J/TMD. In order to better understand the relationship between titin mutations and muscle wasting disease, we characterized all known human M10 missense mutations using biochemical, biophysical, and computational methods. The modular design of titin, consisting of independently stable domains linked together, very similar to boxcars of a train, allowed us to work with M10 and Ig1 domains in vitro separate from the full-length protein. This allowed for accurate structural and functional analysis on the just the target M10 and Ig1 domains. While most mutations result in a misfolded version of M10 at physiologic temperatures, one mutation, I57N, is structurally, functionally, and dynamically nearly indistinguishable from the WT M10 domain.
Figure 3: Model of M10 domain of titin (green) bound to OI1 of obscurin-like-1 (grey), adapted from (32). Both structures represent examples of the Ig domains that compose obscurin and titin. Amino acid sequence with numbering system used in this report shown at the bottom. Finnish mutation (37EVTW→VKEK) is colored blue, Italian mutation (H56→P) is colored yellow, Belgian mutation (I57→N) is colored red, and French mutation (L66→P) is colored orange.
Results

Expression and Purification

We expected all four of the disease-associated M10 mutations to at least partially disrupt the hydrophobic core of the M10 domain; the Finnish cohort mutation replaces two hydrophobic residues involved in the hydrophobic core fold with two charged residues (37EVTW→VKEK; here referred to as E37V), the Italian and French mutations substitute a proline residue into well-formed beta sheets (H56P and L66P, respectively), and the Belgian mutation replaces a buried isoleucine with an asparagine (I57N). We thus expected each of these mutant domains to be at least partially misfolded. Constructs were expressed at 37 °C in *E. coli* via a codon-optimized recombinant expression system and initial purification with Ni-NTA His-bind resin showed significant differences in protein expression among the M10 mutant constructs (Figure 4). The WT and I57N variant expressed robustly, the L66P and H56P variants expressed at lower levels, and the E37V mutation exhibited the lowest expression. Each construct was then further purified to homogeneity via size exclusion chromatography, where all the M10 constructs except the E37V mutation eluted as a monomer (Table 1 and Figure 5). Even when diluted to a concentration less of than 1 µM, the E37V protein eluted at an apparent weight of a dimer and, once pure, gradually precipitated out of solution over several days at 4 °C, suggesting inherent instability relative to WT.
Figure 4: SDS-PAGE of whole cell fractions and post-nickel column fractions of WT M10 and all M10 mutants (~10kDa), showing expression levels at 37 °C. Target domain can be visualized below the 12 kDa mark in all lanes.

Figure 5: Size exclusion absorbance profiles of M10 constructs. WT, H56P, L66P, and I57N elute from a G75 column at the same molecular weight. E37V, which does not contain tryptophan, elutes at a higher molecular weight (denoted by asterisks). All protein was judged pure via SDS-PAGE gel.

Table 1: WT M10, mutant constructs, and Ig1 constructs with various molecular weights and the fraction numbers of eluted protein from size exclusion chromatography. Ig58-59 and Ig1-3 were used as molecular weight controls.

<table>
<thead>
<tr>
<th>Construct</th>
<th>WT M10 (10.7kDa)</th>
<th>I57N (10.7kDa)</th>
<th>L66P (10.7kDa)</th>
<th>H56P (10.7kDa)</th>
<th>W37K (10.7kDa)</th>
<th>Ig1-3 (39kDa)</th>
<th>Ig58-59 (21.5kDa)</th>
</tr>
</thead>
</table>
Initial Structural Analysis

Circular Dichroism (CD) was used to initially study the secondary structure of all constructs. Proteins are composed of chiral amino acids that can warp plane-polarized light if oriented in ordered structures such as those found in β-pleated sheets and α-helices. CD emits circular plane-polarized light through a solution of protein sample and measures the differences in absorption between left and right circular polarized light (mdeg) with respect to wavelength. Secondary structure produces known unique CD profiles that increases and decreases in strength depending on the quantity of the respective structure (Figure 6A).

Low expressing mutations (E37V, L66P, and H56P) were induced at 30 °C for initial structural analysis as previous studies had shown apparently normal folding with lower temperatures (34). While E37V and L66P exhibited aberrant circular dichroism (CD) spectra, H56P and I57N appeared surprisingly similar to the WT spectra (Figure 6B). Specifically, E37V and L66P spectra displayed a less defined minimum at 217 nm, a shoulder at 208 nm, and negative ellipticity near 200 nm. These CD spectral features are all typical of unfolded Ig-like domains (35). In contrast, the WT, I57N, and H56P constructs produced the characteristic β-pleated sheet minimum at 217 nm, which were in agreement with the β-pleated sheet rich high-resolution structure (32, 33) (Figure 6B). These folding trends were verified through fluorescence of the single tryptophan moiety present in all M10 constructs except E37V, as this mutation replaced the target tryptophan residue with lysine (Figure 7A). H56P, I57N, and WT exhibited blue-shifted tryptophan emission near 322 nm, indicative of this residue being in a well-formed hydrophobic pocket (Figure 7A) (36). L66P emitted most strongly near 345 nm (red-shifted), conversely suggesting increased solvent accessibility to the hydrophobic pocket and thus perturbed structure (Figure 7A). WT, I57N, and H56P also produced fully dispersed 1D
NMR spectra with well defined downfield peaks indicative of folded structure while L66P and E37V mutations exhibited more centered and less defined 1D NMR spectra indicative of misfolded structure (Figure 8). Together, these data show that the L66P and E37V mutants are misfolded while I57N and H56P appear to fold in a native-like conformation.

Figure 6: A) Known CD profiles of protein secondary structure (37). B) CD spectra of WT M10, E37V, L66P, I57N, and H56P at 15 µM concentration. C) CD signal at 205 nm of H56P (Yellow Diamond), WT (Black Square), and I57N (Red Circle) with respect to temperature, showing the thermal unfolding of these Ig-like domains.

As seen in Figure 6B, folded and unfolded Ig domains exhibited the largest difference in CD signal between 200 and 210 nm. To study the thermal stability of H56P, I57N, and WT, we
monitored the CD$_{205}$ signal with respect to increasing temperature. I57N and WT started to unfold significantly above physiological temperature (approximately 53 and 58 °C, respectively), however H56P unfolded near 37 °C (Figure 6C). Upon unfolding, all constructs displayed CD spectra similar to that of L66P and did not refold with lower temperatures. These unfolding temperatures were similar to those measured using tryptophan fluorescence and NMR spectroscopy (Figure 7B-C and 9). Tryptophan emission profiles exhibited red-shifting due to solvent accessibility to the hydrophobic pocket at 40 °C compared to WT, which exhibited red-shifting around 60 °C (Figure 7B and C). 1D and 2D NMR experiments confirmed that H56P misfolded over time (Figure 9A and B) and after heating to 37 °C (Figure 9C).

Figure 7: Tryptophan fluorescence emission profiles after excitation at 285 nm. A) Comparison of M10 constructs. WT, I57N, and H56P exhibit maximum emission at 322 nm. L66P exhibits maximum emission at 345 nm. B) Raw data showing the change in tryptophan fluorescence of H56P and WT with respect to increasing temperature. C) Maximum wavelength of tryptophan emission for H56P and WT with respect to increasing temperature.
Figure 8: 1D NMR of all M10 mutations at 25 °C. WT, I57N, and H56P (A-C) show well-distributed spectra with defined peaks, while L66P and E37V (D-E) give poorly defined, spectrally condensed data. Red boxes indicate ppm ranges used in evaluation of folding.
To further assess aggregation, the high-tension voltage signal of each construct was plotted against sample temperature. This measurement is a well-known indicator of turbidity, and is used for monitoring protein aggregation (38; 39). E37V showed increased turbidity beginning at 65°C suggesting significant aggregation at higher temperatures (Figure 10). None of the other variants exhibited this behavior (data not shown). When taken together, CD, fluorescence, NMR, size exclusion, and expression data indicated that WT and I57N behave as well-folded, monomeric Ig-like domains.

Figure 9: H56P unfolding. H56P exhibits spectra indicative of a folded protein at 25 °C (A), however the spectra becomes significantly less well dispersed upon incubation at 20 °C for 3 days (B). The H56P construct is initially folded at 25 °C as judged by an HSQC spectra (C, black). Red boxes indicate ppm ranges used in evaluation of folding. The HSQC spectrum at 37 °C of the same sample indicates the mutated protein is in an unfolded or aggregated state at 37 °C (C, yellow).

Figure 10: CD spectra of HT voltage of E37V mutation taken at 280 nm comparing the turbidity to temperature.
H56P maintained a native-Ig like fold at lower temperatures but unfolded near 37 °C. The L66P variant was monomeric but misfolded at all temperatures, and E37V was misfolded at all temperatures with a strong tendency to aggregate.

**NMR High Resolution Structural Analysis: H56P Mutation**

To determine the structural cause of why H56P exhibited decreased thermal stability, we sequence-specifically assigned the WT M10 construct and compared these shifts to the H56P chemical shifts using multidimensional heteronuclear NMR. With the sequence of each construct known, a heteronuclear single quantum coherence (HSQC) experiment was performed on $^{15}$N-labeled protein to produce a two-dimensional spectrum exhibiting nitrogen and hydrogen shifts for amino acid residues (Figure 11A). Using various three-dimensional NMR experiments (HNCO, HNCACO, HNCACB, and CBCACONH) the $\alpha$-carbon, $\beta$-carbon, and carbonyl-carbon shifts for an amino acid (own) and the amino acid that precedes it (previous), can be determined (Figure 11B). These experiments are named based on the path of magnetization transfer through the protein. For example, the HNCO experiment induces a magnetization transfer from the H-N bond to the C=O bond, providing chemical shifts for the carbonyl carbon in the process. The HSQC can be subsequently assigned in a sequence specific-fashion by finding the amino acid residue with “own” shifts identical to the target residue’s “previous” shifts. Figure 11 provides an example of this concept with three peptides of the M10 sequence, Lys 46, Ile 47, and His 48. The “previous” chemical shifts from Ile 47 match the “own” shifts of Lys 46. Alternatively the “own” peaks of Ile 47 match the “previous” peaks of His 48 (Figure 11C). Thus it can be
concluded that the peaks highlighted in HSQC are correctly assigned with their respective amino acids (Figure 11A).

Figure 11: A) Example of a labeled HSQC with three residues highlighted. B) Segment of three amino acids from the M10 primary sequence consisting of Lys 46 (Green), Ile 47 (Magenta), and His 48 (Red) from HSQC. The nitrogen (N), amino hydrogen (NH), α-carbon (Cα), β-carbon (Cβ), and carbonyl carbon (C') for each amino acid is labeled with its respective color. Chemical shift values for the “own” atoms and atoms from previous amino acid (-1) from various multidimensional experiments showing how the primary sequence can be matched to HSQC peaks.
The sequence-specifically assigned HSQCs of WT M10 and H56P constructs were completed at 25 °C (Figure 12A and B). Due to its inherent instability, H56P HSQC peaks were verified using a 3D $^{15}$N-edited nuclear overhauser effect spectroscopy (NOESY) at 16 °C. HSQC peaks corresponding to the mutated site (H56P), Ile 67, Gln 72, and Gln 74 were present in the wild-type spectra but not in that of the mutant. Of the rest remaining amide peaks, the largest chemical shift perturbations were localized near the mutation site (Figure 12C and D). Thus, the H56P mutation disrupted the outer beta sheet and the surrounding residues, but not the rest of the protein. This perturbation in structure accompanied the much lower denaturation temperature shown above (Figure 6, 7B-C, and 9).

Figure 12: A) Fully assigned HSQC of WT M10. B) Fully assigned HSQC of H56P. Conditions for NMR are 20 mM Tris pH 7.5, 50 mM NaCl, 0.3 mM NaN$_3$, 10% D$_2$O, 25 °C. C) Chemical shift differences between WT and H56P. Residues with chemical shifts greater than 2x the average chemical shift differences are colored in yellow and residues with greater than 3x the average chemical shift difference are colored in red. The mutation site is colored purple. D) Chemical shift differences of H56P compared to WT, mapped onto the M10 X-ray structure (PDB 2Y9R).
ITC Functional Analysis: Ig1 Binding Affinity

ITC was used to evaluate the ability of WT and mutant constructs to bind a physiologically relevant M10 target, the Ig1 domain of obscurin, was then evaluated using isothermal titration calorimetry (ITC). This instrument functions by injecting small increments of highly concentrated protein into a sample cell containing a relatively dilute binding partner. The heat of the two proteins interacting per each injection is measured with respect to the reference cell containing water (Figure 13A). As the protein in the sample cell becomes saturated, the heat per injection decreases until full saturation is achieved and no binding is observed. This produces the titration curve or binding curve shown in Figure 14C and D. This binding curve can be presented with enthalpy (ΔH) on the y-axis and molar ratio (titrant/titrate) on the x-axis. The ΔH of binding is the maximum heat change, the ΔG is the slope of the titration curve, and the stoichiometry of binding (n) is the inflection point of the titration curve (Figure 13B). With temperature and protein concentration known, the entropy (ΔS) and binding affinity (Kd) can be calculated by the two following Gibbs free energy relationships.

1) ΔG = ΔH − (TΔS)
2) ΔG = RTln(Kd)

Figure 13: A) The basic instrumental set up of isothermal titration calorimeter. B) Example of ITC binding curve depicting where various variables are derived by ITC analytic software (40).
All constructs used for ITC were expressed at 37 °C as unfolding events were nonreversible thus the structures used should be similar to those found at physiological relevant temperatures. No binding curve was observed for L66P and H56P indicating these constructs did not bind obscurin Ig1 (Figure 10A and B). Instead, these mutations exhibited only slight amount of heat per injection, which was determined to be heat due to friction after comparison with control buffer into buffer experiments. Sauer et al. recently published results showing that H56P can bind to Ig1 however this difference can almost certainly be attributed to the differing expression temperatures (37 °C compared to 20 °C) and the fact that H56P irreversibly denatures at near-physiological temperature (Figure 6C) (34). Both WT and I57N were capable of binding to obscurin Ig1 with a 1:1 stoichiometry in the low µM range, with similar enthalpy and entropy values (Figure 14C and D and Table 2). Based on these findings, the I57N mutation does not significantly alter the obscurin binding interface as this construct successfully bound Ig1 in a manner very similar to WT. We conclude that the H56P and L66P mutations are unfolded at physiological temperature, and render the M10 domain incapable of binding normal cellular targets. Due to weak expression and instability, this experiment was not performed on E37V although the CD and NMR structural results suggested this construct would also be incapable of binding obscurin.
Figure 14: Raw data showing the isothermal titration calorimetry curves for M10 constructs. L66P (A) and H56P (B) do not exhibit binding but instead show a slight amount of heat per injection. Control experiments showed that this is due to the friction of injection and not due to any protein-protein interaction. I57N (C) and WT (D) produce the expected endothermic binding curves.

Table 2: Summary of ITC results of M10 plus M10 mutants bound to obscurin Ig1. All data were collected at 16°C, 20 mM Tris pH 7.5, 50 mM NaCl.

<table>
<thead>
<tr>
<th></th>
<th>WT M10</th>
<th>Belgian (I57N)</th>
<th>French (L66P)</th>
<th>Italian (H56P)</th>
<th>Finnish (W37K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH</td>
<td>44.61 ± 7.3 KJ/mol</td>
<td>51.80 ± 5.6 KJ/mol</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>1.223 ± 0.128</td>
<td>0.933 ± 0.0759</td>
<td>No Binding</td>
<td>No Binding</td>
<td>N/A</td>
</tr>
<tr>
<td>K_d</td>
<td>1.15x10^{-5} ± 1.0x10^{-6} M</td>
<td>2.06x10^{-6} ± 5.0x10^{-7} M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔS</td>
<td>248.9 ± 18.95 J/mol*K</td>
<td>216.6 ± 20.79 J/mol*K</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
NMR High Resolution Structural and Dynamics Analysis: I57N Mutation

Expression data, CD, and functional analysis provide no compelling evidence as to why the Belgian mutation would result in muscular dystrophy. The Belgian mutation is located distally to the obscurin/obscurin-like binding site thus it is possible that this mutation locally distorts the M10 domain, and this does not affect obscurin binding but instead interferes with another, yet-unknown protein interaction. In order to more thoroughly study this possibility, the chemical shifts of I57N were compared to those of WT M10 via sequence-specific assignments described above (Figure 11). Surprisingly, these spectra displayed no significant amide chemical shift differences suggesting highly similar if not identical high-resolution structures. Careful inspection revealed novel weak side-chain Asn peaks in the I57N construct. These peaks are present due to the new amine group brought in by the Ile→Asn mutation (Figure 11).

Figure 15: Overlay of fully assigned I57N HSQC, (red), onto the fully assigned HSQC of WT M10 (black). The new nitrogen-containing side chain of the Ile→Asn mutation is highlighted with a blue rectangle.
In parallel, over 60 N-H residual dipolar couplings (RDCs) were collected on both I57N and the WT M10 domain in order to evaluate and compare peptide backbone angles between constructs and the previously published crystal structure. These RDC values were obtained by measuring the chemical shift differences between isotropic protein in solution and anisotropic protein in a polyacrylamide gel (Figure 16A). The backbone angles of the M10 structure were then produced by PALES (prediction of alignment from structure) analytic software and compared to the known 2Y9R X-ray structure. Virtually all RDC splittings, including those near the I57N mutation, were nearly identical with the exception of the more dynamic N-terminus (Figure 16A and B). Both RDC data sets agree with existing crystal structure as indicated by the Q-value (Q-value=0.30 for WT and Q-value=0.30 for I57N mutation). This value represented the agreement between the model, 2Y9R X-ray structure in this case, and the NMR data where lower
numbers indicate a higher agreement. These data independently verified that I57N has no significant backbone structural deviations from WT and that the presented NMR data are in agreement with published crystal structure. The RDC data also suggested that the I57N mutation has similar backbone dynamics compared to WT.

As further verification of the backbone dynamics, fast-motion heteronuclear NOE experiments providing information on backbone energies and movement were performed. This experiment identified the amino acid residues that had high backbone energies and motions in the nano-pico second timescale. The data are reported as a ratio of signal strength between NOE1, providing baseline signal, and NOE2, which decreases with respect to increased energy/dynamics of the peptide backbone. This ratio has a range of 1.0 to -1.0 indicating a low- and high-energy backbone respectively. Virtually all residues exhibited NOE2/NOE1 values between 0.8 and 1.0 suggesting a highly stable and immobile structure in the nano-pico second timescale (Figure 17). Furthermore, both I57N and WT M10 constructs exhibit residues with similar or identical NOE values suggesting highly similar fast

Figure 17: Heteronuclear NOE values for WT (black squares) and Belgian construct (white circles). Both Ig domains exhibit very little fast-timescale motions, and have correspondingly large hetNOE values. Red asterisk indicates the location of I57N mutation.
timescale backbone dynamics.

Slower movements of tertiary structure in the micro-milli second timescale were evaluated for the I57N construct with relaxation dispersion experiments in order to determine if the mutation causes slow timescale instabilities. Relaxation dispersion utilizes the spin echo principle of NMR to determine if a protein is changing between two different states (41). The spin echo is a method of refocusing the signal strength that was lost due to transverse dephasing. This occurs after an atom’s magnetic dipole is forced from a parallel to perpendicular alignment with respect to the external magnetic field of the NMR by a 90° pulse. The dephasing is then refocused with a 180° pulse allowing a strong signal to be collected (Figure 18A). If a protein is switching between two states, A and B for example, and changes state between the 90 and 180° pulses, the signal will not refocus adequately. This is due to different transverse relaxation rates between the two states and results in a weaker signal. Relaxation dispersion experiments gradually increases the pulse frequency thus decreasing the time between the 90° and 180° pulses, and decreasing the likelihood of a conformational change happening before refocusing occurs (Figure 18B)(41). This signal strength is reported as an R2 effective value, which is inversely proportional to the strength of the signal and is plotted with relation to increasing pulse frequency. Therefore, a protein switching between state A and state B in the micro-milli second timescale would exhibit a plot similar to that in Figure 18C.
Figure 18: A) Direction and strength of atomic dipoles in an NMR during a spin echo experiment where B° represents the magnetic field applied by the NMR. B) Relaxation dispersion experiment with a protein switching between state A (Red) and state B (Blue). A 90° pulse of spin echo is represented by solid black line and an 180° pulse is represented by dashed line. C) R₂ effective value with respect to increasing relaxation dispersion frequency for a protein undergoing slow exchange (42).
Residues of the binding interface, those surrounding the mutation site, and around the mutation site itself were sampled and matched to a model of “no exchange” by NESSY analytic software indicating the protein was not experiencing slow timescale motions/conformational changes (Figure 19) (Appendix A). The model of “no exchange” makes sense as the R\textsubscript{2} effective values did not appear to change with respect to increased pulse frequency for sampled residues (Figure 19B) (43). As the sampled residues and all other tested residues were not found to be undergoing conformational changes, it appeared that the mutation does not cause slow-timescale instabilities that could interfere with the binding of known and unknown targets. Thus high-resolution structure, fast timescale, and slow timescale dynamics could not explain how this protein would propagate disease.

![Figure 19: A) Model of mutant M10 domain with sequence specifically assigned amino acids in green and unassigned amino acids in grey. Thr 25 and Cys 28 (Purple) are located in the Ig1 binding site, Asn 57 (Red) represents the I57 mutation, and Val 38, Lys 46, and His 56 (Blue) represent residues surrounding the mutation site. B) R\textsubscript{2} effective values for each residue with respect to increasing pulse frequency. A “no exchange” model best fit all data as determined via the NESSY program.](image-url)
While the I57N and WT M10 constructs are virtually identical in solution, they may exhibit slightly different target binding. To test this we titrated unlabeled Ig1 into $^{15}$N-labeled M10 and monitored the N-H chemical shift changes (Figure 20A). Mapping these chemical shifts onto the M10 crystal structure show that a majority of the largest chemical shifts occur at the target binding site, as defined in the crystal structure (Figure 20C and E). Specifically, the WT residues with the largest chemical shifts include Ala 11, Leu 12, Ser 17, Ile 18, Thr 25, Ala 29, His 56, and Asp 70 and the I57N mutation residues with the largest chemical shifts include Asp 15, Ser 17, Thr 25, Ala 29, and Asp 70. Ala 11, Leu 12, and His 56 are not present in the I57N HSQC, likely due to either the size of the complex (24 kDa) or exchange broadening effects. The resulting spectra of WT-Ig1 and I57N-Ig1 are virtually indistinguishable (Figure 20B, D, F). Likewise, when either of the unlabeled M10 constructs are titrated into $^{15}$N-labeled Ig1, the resulting HSQCs are identical (data not shown). Together, these data strongly suggest that the I57N mutation does not influence normal titin/obscurin structure or target binding.
Figure 20: Chemical shifts changes of WT M10 and I57N mutation. (A) An HSQC overlap of labeled M10 (black) and M10 bound to Ig1 (Cyan). (B) The extensive overlap between the sequence-specifically assigned Ig1-bound M10 (black) and Ig1-bound I57N (Red). The HSQC chemical shift changes of M10 (C) and I57N (D) mapped onto the M10-Oll X-ray structure (PDB 2WP3). The dotted oval represents where Ig1/Oll binds to M10. Red coloring represents chemical shifts above 2x standard deviations from the average unbound ⇝ Ig1 bound chemical shift change. (E) and (F) show the unbound ⇝ bound chemical shifts on the M10 sequence. Red coloring represents chemical shifts 2x standard deviations from the average unbound ⇝ Ig1 bound chemical shift change.
Molecular Dynamics Simulations

Residue I57 is part of the hydrophobic core of the M10 domain. We expected the substitution with a hydrophilic moiety to perturb the protein structure and/or lead to increased protein dynamics around this site. To better visualize why this mutation has little effect on structure, we performed a molecular dynamics (MD) computer simulation on both the M10 crystal structure and an energy-minimized I57N model. The WT I57 is situated at the edge of the hydrophobic pocket and is partially solvent accessible (Figure 21A). When this residue is changed to asparagine, the methylene group of the Asn occupies the same space as the Ile $\gamma$2-methyl group, and the NH$_2$ is oriented towards the solvent (Figure 21B). None of the surrounding residues show any significant reorientation, nor does the overall structure change appreciably (RMSD=0.36 Å). These simulations agreed with the NMR data, and supported the idea that this Ig-like fold can accommodate the hydrophobic to polar substitution.

Figure 21: Comparison between WT M10 (A) and I57N (B) sidechain packing. The I57N substitution was made using YASARA in the crystal structure of M10 (PDB 2Y9R) and MD simulation was then run in explicit solvent until the RMSD of the overall protein structure was stable for 2 ns.
Discussion

Here, we present a structural and functional explanation of why various M10 mutations cause skeletal muscle disease. Experiments such as protein expression tests, CD, temperature denaturation, and size exclusion chromatography were used to gain insight into why mutations in independently folded titin domains might cause disease. We expect that similar methods can be applied to other diseased-linked titin domains, and that techniques such as these will be useful for understanding why such mutations lead to disease.

The H56P (Italian) and L66P (French) mutations are associated with TMD, a disease characterized by weakness and atrophy of the anterior tibialis muscle and other pedal dorsiflexors (26; 27; 30). Both mutations involve the insertion of a proline into a β-sheet. These changes lead to a loss of target binding, probably due to domain unfolding but not to aggregation. Sauer et al. reported that the H56P mutation is correctly folded and has similar target protein binding affinities to WT when expressed at 20°C (34). Our data show that the sequence-verified Italian mutation remains folded at room temperature but not at physiological temperatures. This finding reconciles why this mutation can, under some conditions, have near-normal target binding and CD spectra, yet contributes to TMD.

The heterozygous E37V is also linked to TMD and the homozygous Finnish mutation has been associated with a particularly severe condition, LGMD2J, where disease onset begins in childhood and gradual muscle weakness and atrophy of arms and legs eventually leads to wheelchair confinement within the first two decades of life (32; 27). Correspondingly, the E37V construct readily degrades, is misfolded, and is prone to aggregation. This may explain why previous studies on affected muscle biopsies show a reduced amount of the titin C-terminus in patients with the homozygous Finnish mutation (29).
M10 domains containing the E37V, H56P, and L66P mutations are misfolded at physiological temperatures. While this is attributed to disruption of the Ig-like fold, a major caveat is that the proteins in this study were recombinantly expressed in *E. coli*, and thus may fold differently in human myofibrils. However, while this remains a theoretical possibility, we consider it unlikely. All constructs used in this study express small single domains that are known to independently fold, are codon optimized, and are not predicted to undergo post-translational modifications, thus making them excellent candidates for bacterial expression (44). Additionally, E37V and L66P remain incorrectly folded after urea unfolding/refolding protocols, yet WT, I57N, and H56P can be unfolded and then correctly refolded at 4 °C, as judged by CD (data not shown) (45). Thus while chaperones may induce these mutants to fold correctly in human cells, it seems unlikely that they could remain folded.

The I57N mutation associates with TMD, with the onset of disease similar to that of the H56P and L66P mutations. This mutation does not change the surface of the protein, and so has been thought to disrupt the hydrophobic pocket of the Ig fold (28). However, while this change leads to a slight decrease in melting temperature, no other biochemical or structural differences between this variant and the wild-type protein were observed. Analysis of M10 domain crystal structure revealed that the mutation site is located at the edge of a β-sheet with the Ile 57 side chain projecting inward toward a hydrophobic region (Figure 21A) (33). Our MD analysis shows that Asn 57 can orient its side chain to face the solvent without perturbing the overall secondary and tertiary structure (Figure 21B). Our structural simulation, NMR, and ITC data show no significant differences in either structure or binding between the two protein constructs (Figures 14-19). These results strongly suggest that the I57N mutation is not directly responsible for TMD. There are two possibilities as to why this variant is disease-linked. First, this mutation
may be a polymorphism, and simply co-segregates with the disease. This co-segregation could be due to either other mutations within titin, or possibly mutations in other M-line proteins. Alternatively, it may be that this mutation influences other factors such as transcriptional regulation or cellular localization. While all other known M10 disease-causing mutations can be explained through aberrant protein structure/function, a broader characterization of titin and obscurin, involving domains other than M10 and Ig1, are likely needed to explain the root causes of why the I57N mutation is associated with TMD.
Methods

Protein Isolation

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)) or HMS174(DE3) using a pET24a vector system (Novagen, San Diego CA) in a manner similar to (46). The E37V, H56P, and L66P constructs were induced at 30 °C with IPTG at an OD$_{600}$=0.6-0.8 while all other constructs were induced at 37 °C. Due to similar molecular weights of all the constructs, plasmids from the induced cells were sequenced before each experiment. Cells were sonicated and centrifuged in a buffer containing 50 mM phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 1 mM PMSF. The resulting cleared supernatant was passed over Ni-NTA His-bind resin (Novagen). The column was washed extensively with 50 mM imidazole, 300 mM NaCl, and 50 mM phosphate buffer and protein was then eluted with this same buffer plus 250 mM imidazole. Fractions containing the M10 or Ig1 protein domains were then concentrated in 5000 MWCO concentrators (Corning SpinX, Tewksburg MA) and applied to a Sephadex G75 (Sigma; St. Louis MO) size exclusion chromatography column in 50 mM Tris pH 7.5, 20 mM NaCl, 0.35 mM NaN$_3$ (G75 buffer). Pure protein, as determined by SDS-PAGE, was then once again concentrated in a 5000 MWCO concentrator. Concentrations were determined by both absorbance at 280 nm and BCA assays (Thermo Scientific, Waltham, MA).

Protein Refolding

Protein and urea were mixed until the final urea concentration was 7.5 M. A rapid refolding procedure was used, where 1mL of the protein-urea solution was added dropwise into a 40 mL solution of 20 mM Tris pH 7.5, 100 mM NaCl, 1mM DTT, and 10 mM Arginine at 4 °C.
Arginine is a frequently used refolding additive, and can act as a chaperone-like molecule (47).

**Circular Dichroism**

All CD and fluorescence experiments were conducted in a 3 mm pathlength cuvette at 15 µM protein in 20 mM phosphate buffer, pH 7.5 and 50 mM NaF. Samples were measured in triplicate at temperatures from 10 °C to 85 °C in either 2.5 or 5 °C intervals on a Jasco J-810 spectrophotometer. In fluorescence experiments samples were excited at 285 nm and emission was recorded from 300-400 nm using a 4 nm band pass.

**Isothermal Titration Calorimetry**

Protein was expressed at 37 °C for ITC experiments. A Nano ITC (TA instruments, New Castle, DE) was equilibrated with G75 buffer until stable baseline was obtained. Experiments were performed at 16 °C with 1 mL of 50 µM Ig1 in the sample cell and 1mM M10 WT or mutant in the injection syringe. Typically, 5 µL were injected over 20 injections. Reversing the injection and sample proteins resulted in severe Ig1 aggregation, and was thus not used in the analysis. Titration curves were analyzed via NANO Analyze software to determine ΔH, ΔS, ΔG, kD, and stoichiometry.

**NMR**

All NMR experiments were conducted 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 experiments were collected at 25°C in 20 mM Tris pH 7.5, 20 mM NaCl, 0.35 mM NaN3, and 0.5-2.5 mM protein with 10% D2O. We conducted a 2D HSQC and standard triple resonance experiments including HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, C(CO)NH, HCCCONH, 15N-edited TOCSY, and 15N-edited NOESY. 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 (48), extended in the indirect dimension via linear prediction, and the resulting spectra were analyzed via Sparky (49). Chemical shifts for the titin M10 domain, M10 bound to obscurin Ig1, obscurin Ig1, and Ig1 bound to M10, and the H56P M10 mutation were deposited in BMRB under accession numbers 25305, 25303, 25301, 25304, and 25406, respectively.

Standard Bruker IPAP experiments using 256 pts for each $T_i$ dimension were used to collect RDC data in isotropic and axially-compressed 5.5% acrylamide gel samples, as previously described (50). We used the program PALES for RDC alignment tensor fitting with a calculated $A_a$ and $A_r$ component of 0.00188 and 0.000386, respectively (51). $^1H$-$^15N$ NOE ratios were acquired in an interleaved experiment, with 16 scans per direct point and 200 $T_i$ points. The NOE ratios were collected with a 3 second presaturation period followed by a 2 s saturation delay, while the control experiment had an equivalent 5 s delay. For all experiments, the $^1H$ chemical shifts were referenced to external DSS, the $^{13}C$ shifts were referenced indirectly to DSS using the frequency ratio $^{13}C/^1H = 0.251449527$ and $^{15}N$ shifts were referenced indirectly to liquid ammonia using $^{15}N/^1H = 0.101329118$.

Relaxation dispersion data was analyzed by first extending in the indirect dimension using linear prediction using a standard Bruker pulse program. Correlation peaks with $S/N > 10$ were selected and peak heights were measured by fitting Gaussian curves to the transformed data and taking the maximum peak height. A reference spectrum was acquired for all relaxation dispersion measurements (without a CPMG period), followed by eleven spectra containing a constant CPMG period, $T_{cp}$, where the time between each CPMG pulses varied (52) and the effective field, $n_{CPMG}$, is defined by $1/4T_{cp}$. At each effective field, an $R_{2eff}$ was calculated from
the ratio of two signals $I_{CP}$ and $I_0$, where $I_0$ is the intensity of the peak in the reference spectrum and $I_{CP}$ is measured at the end of the $T_{CP}$ period. Spectra were recorded with $T_{CP} = 60$ ms or 100 ms and $n_{CPMG}$ equal to 25, 50 (x2), 75, 100, 200, 400, 800 (x2), and 1000 Hz. The $^{15}$N dispersion experiments were recorded with 256 and 2048 complex points in $F1$ and $F2$ dimensions, respectively and with 8-24 scans. All dispersion data were processed with NMRPipe and Sparky and then analyzed using the program NESSY, where the data were fit to a Bloch-McConnell equation assuming a 2-site exchange model (53; 43). A standard NESSY protocol was followed as described in the manual, and uncertainties in simulations were determined by Monte-Carlo simulation (53; 43). Such an analysis showed no exchange in any residues of either WT or the I57N mutant M10 domain at 60 MHz at 25°C.

All NMR titrations were conducted using ~1 mM $^{15}$N-labeled protein, with the addition of buffer-matched unlabeled protein. A standard HSQC was then performed on this mixture. All changes in chemical shift proved to be in slow exchange, and unlabeled protein was added until none of the original (unbound) peaks were present, typically at roughly 1.5X concentration compared to the labeled protein. Due to the large chemical shift perturbations, the target-bound HSQC needed to be reassigned using standard triple resonance experiments (HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, C(CO)NH, HCCCONH, $^{15}$N-edited TOCSY, and $^{15}$N-edited NOESY).

**MD Simulations**

All MD simulations were performed using the YASARA 12.4.1 software package, the AMBER03 force field, and explicit solvent (with 150 mM NaCl) in a 10x10 nm box size at 25°C (54). The I57N substitution was done with the swap function and the resulting model was then energy minimized.
References


6. The Muscular System.


20. Structure of Skeletal Muscle


40. Isothermal Titration Calorimetry to Study Biomolecular Interactions


Appendix A: NESSY Manual (Relaxation Dispersion Analysis)

1. Relaxation Dispersion Parameters for Topspin.
   a. Delay constant: 0.06-0.1ms
   b. Pulse frequencies: 0Hz, 25Hz, 50Hz, 75Hz, 100Hz, 150Hz, 200Hz, 400Hz, 800Hz, 1000Hz, 50Hz, 800Hz. (you need to have 12 pts)
2. Process each pulse frequency as an individual file. Processing Instructions are in the name of Relaxation Dispersion Parameter titles. Name plane files from 999 down. So 999 will contain 0Hz plane, 998 will contain the 25Hz plane, etc. (tf3… tf2.. r23.. save as a new PROCNO)
3. Move experiment file onto the s-drive and then onto lab computer under its designated protein domain. It is useful to create a folder named something like RelaxDispDynamics to organize multiple trials of the experiment.
4. Convert each individual plane file into a sparky file using X11 terminal
   a. Open X11 and type csh. Press enter.
   b. Access experiment folder on lab computer.
      i. Ex) cd Documents/nmr/M10/RelaxDispDynamics/211. Press enter. (211 is the experiment folder moved over from Topspin files, you can rename this if you like)
   c. Access the pdata folder and open the folder for the first plane.
      i. Ex) cd pdata/999. Press Enter.
      ii. The 999-988 folders should contain a 2rr script that can be used to create .ucsf files.
d. Create .ucsf file in folders 999-998, each file will be named after the pulse plane contained in the folder.
   
   i. Ex) bruk2ucsf 2rr M10-0Hz.ucsf. Press enter.

   ii. Use the command cd ../ to move up a directory out of the 999 folder then enter the 998 folder and use the same command shown above to create another.ucsf file.

   iii. NOTE: Remember to differentiate the titles between the two repeated pulses (50 and 800). Ex) M10-50Hz.ucsf and M10-50Hz-2.ucsf

5. Open the .ucsf files created in sparky and label the respective peaks with their corresponding amino acid. Once one spectra is labeled the “select” option on the sparky menu can be selected and used to select all peaks. Type “oc” for ornament copy, select an unlabeled spectra and type “op” for ornament paste. Always use the peak center function (“pc”) to center ornaments.

   a. NOTE: Check spectra after using peak center to make sure ornament peaks did not center on top of one another. Manual manipulation might be necessary for problem peaks.

6. List table (“lt”) to view assigned peaks and click on options. Select the Data Height box and click apply. The peak height should now be listed. Save all tables as a .LIST files by clicking save.

   a. For organization it would be helpful to create a folder named after your experiment in the sparky list folder. Move your list files into this folder.
b. Make a folder on the S-drive under your name to contain all of your Relaxation Dispersion experiments. Create a folder for the experiment you are currently making. Create two more folders inside one for data and one for analysis.
   i. Ex) S-drive -> Michael -> RelaxationDispersion -> 211 -> 211data and 211analysis.
   ii. Copy your .LIST files created in sparky into the “_____data” folder.
7. On the lab Dell (optiplex 9010) access the .LIST files from your “_____data” folder and open them in Microsoft Excel. NOTE: you will likely have to right click and select “open with” select excel. All of the data will be in one column; do not move them to individual columns.
   a. Evaluate data in excel and make edits if necessary.
      i. Delete data not labeled with an amino acid.
      ii. Check data to ensure there are no repeats of amino acids that would throw off numbering in NESSY (such as two R55s or an S55 and an R55)
   b. Save final excel spreadsheet as a “text delimited” file and add a number in front of the title representing its position in the sequence of pulses. This will allow NESSY to move data sets into the program in the correct order.
      i. Ex) 1-M10-0Hz.txt, 2-M10-25Hz.txt, 3-M10-50Hz.txt, 4-M10-75Hz.txt, etc.
8. On the lab Dell (optiplex 9010) go to the C-drive and open the NESSY application. C-Drive -> nessy app -> nessy -> nessy (application).

9. Set the Select Project Folder line to your “______analysis” folder. This is where data analysis from NESSY will be placed. Next add PDB file of your structure in the Select PDB file line. NOTE: the sequence of your PDB MUST match the sequence you will load into NESSY otherwise the pymol models will be inaccurate.

   a. Load Protein Sequence: Select “Import” and click on load FASTA sequence. You can either upload the code or copy and paste. By copying and pasting you have the option to edit the sequence it there are some differences. Here, you can also designate which amino acid is the start in your sequence if there is a numbering difference.
10. Click on the Data Experiment #1 tab to set up the experiment. The sequence you imported should be in the “sequence” column.

   a. Click the “Set Up” button to enter the delay constant and pulse frequencies for the experiment. The column titles should update to represent their respective frequencies.

   b. Click on “Import” and select the “multiple peak list” option.

   i. Set Data separator to space. The start of the Data Sets, Residue #, and Data Height columns should be correct. (The number is the column number where residue, height, data sets etc. is located)
ii. Select the load button and select all twelve of your tab-delimited files to import at one time. Press “Import”.

iii. The data is now located in the columns.

NOTE: NESSY does not like negative values so try and find them here and delete them. If you miss one you will be directed to the column and residue of the missed negative value when you start the experiment. Also check to make sure the data in the column matches the data from the text file you imported (Ex. Check that the 50Hz column is in fact data from the 50Hz text file).

11. Click the “Analysis” tab and click the “analyze data” button. If you forgot a negative value you will be notified here. If everything is ok the “Start Calculation” button will be available. Click it to start the experiment.

a. NOTE: Sometimes errors will occur or the program will freeze on a particular residue. If this occurs restart the calculation. If the problem persists, quit NESSY.
and reset the experiment. Delete the problem residue before starting next calculation.

12. When calculation is finished results will be available under the “results” tab and info on model selection will be under the “Summary” tab. Everything found under these tabs will have automatically been added to your “___ analysis” folder.

a. NOTE: NESSY will match each residue with a model of best fit based on various statistical and model evaluation tests however manual inspection of residue data and statistical output is necessary to ensure proper data interpretation.

b. Statistical results found under Analysis tab.