Microscopic Properties Determine Macroscopic Characteristics of Proteins

A THESIS SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA BY

BENJAMIN T. HORN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

DR. ANNE HINDERLITER, DR. ALESSANDRO CEMBRAN

July, 2016

Acknowledgements

I would like to thank my advisors Dr. Anne Hinderliter, Dr. Alessandro Cembran, and Dr. Dave Thomas for training and mentoring me through my collegiate career. Without their guidance and expertise this work would not have been possible. I would like to thank former graduate students Michael E. Fealey (for his contribution to this work and for providing expertise on surviving in a research lab), Ryan Mahling (for training me how to properly purify proteins and spending time collecting CD data together at the University of Minnesota Twin Cities campus), Katie Duleavy (for her help with learning proper technique when working with lipids). I would like to thank current graduate students and past and present undergraduate members of the labs for support and help in my research, specifically: Stephanie Kobany, Ben Orpen, Anthony Meger, Justine Schramel, Michelle Botts, and Kate McMahon. I also would like to acknowledge my wife Elizabeth Horn for supporting me in my education during all of the late nights and early mornings.

I would also like to acknowledge collaborators of the Dr. Dave Thomas Laboratory for their contribution to this work: C.B. Karim, Cheng, and M.R. McCarthy

Abstract

Macroscopic characteristics of a peptide or protein are dependent on microscopic properties. One microscopic property that has an effect on the macroscopic characteristics is the oxidation of methionine. Here we show that oxidation strengthens the methionine-aromatic interaction by 0.5-1.4 kcal/mol and that oxidation plays a role in the structure and biological function of calmodulin and lymphotoxin-α/TNFR1 binding. Another microscopic property is the sequence of amino acids. We have also shown that the intrinsically disordered linker region of synaptotagmin I plays a role in modulating the binding of Ca^{2+} and the interaction with membrane. The final portion of this work determined the microscopic properties, in this case hydrophobic and electrostatic interactions within the dystrophin ABD1 to shed light on the underlying mechanism by which structural changes occur under physiological conditions. This work consisted of using thermodynamic and structural approaches and computational techniques, namely database searching and molecular dynamics simulations.

Table of Contents

List of Tables

List of Supplementary Tables

List of Figures

List of Supplementary Figures

List of Abbreviations

NMP:	1-methyl-2-pyrrolidinone
HEPES:	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
MOPS:	3-(N-morpholino) propanesulfonic acid
ABF:	Adaptive Biasing Force
AA:	Ala-Ala
AM:	Ala-Met
AM(ox):	Ala-Met (ox)
BMD:	Becker muscular dystrophy
CaM:	Calmodulin
CH:	Calponin-homology domains
CSD:	Cambridge Crystal Structural Database
CF:	Carboxyfluoroscein
ΔC_p:	Change in baseline heat capacity
CD:	Circular Dichroism
DCS:	Differential scanning calorimetry
DMS:	Dimethyl sulfide
DMSO:	Dimethyl sulfoxide
DTT:	Dithiothreitol
DEER:	Double electron-electron resonance
DMD:	Duchenne muscular dystrophy
Dys:	Dystrophin
mhc-Dys ABD1:	Dys ABD1 in which the hydrophobic patch residues and selected charged residues were mutated
mh-Dys ABD1:	Dystrophin ABD1 with the hydrophobic patch residues mutated
EPR:	Electron paramagnetic resonance
ΔH:	Enthalpy of unfolding
EGTA:	Ethylene glycol-bis(2-aminoethyl)-N,N,N',N' tetra-acetic acid
PC1:	First principal component
FRET:	Förster resonance energy transfer
ΔG:	Free energy of unfolding
HF:	Hartree-Fock
HPLC:	High-performance liquid chromatography
IDPs:	Intrinsically disordered proteins
IDRs:	Intrinsically disordered protein regions
ITC:	Isothermal titration calorimetry
LUVs:	Large unilamellar vesicles
LTα:	Lymphotoxin-α
T_m:	Melting temperature
Met^{Ox}:	Methionine sulfoxide
MD:	Molecular dynamics
NBO:	Natural Bond Orbital
ABD1:	N-terminal actin-binding domain
NMJs:	Neuromuscular junctions

NMR:	Nuclear magnetic resonance
FA:	Phen-Ala
FM:	Phen-Met
FM(ox):	Phen-Met(ox)
KCl:	Potassium chloride
PCA:	Principal component analysis
PDB:	Protein Databank
RDF:	Radial density function
RIPA:	Radio immunoprecipitation assay
REMD:	Replica exchange molecular dynamics
ABD2:	Second actin binding domain
PC2:	Second principal component
SUVs:	Small unilamellar vesicles
NaCl:	Sodium chloride
SASA:	Solvent accessible surface area
S/Ar:	Sulfur-aromatic
Syt I:	Synaptotagmin I
TFE:	Tetrafluoroethylene
TFA:	Trifluoroacetic acid
TNF:	Tumor necrosis factor
TNFR1:	Tumor necrosis factor receptor 1
Utr:	Utrophin
wt-Dys ABD1:	Wild type Dys ABD1
Utr ABD1:	Wild type Utr ABD1

Chapter 1: Studying the Microscopic Properties and Macroscopic Characteristics of a System

In both peptides and proteins, the microscopic properties shape both the secondary and tertiary structure and the physiological function of the peptide or protein. These microscopic properties can include the different amino acids present in the sequence of the peptide or protein, modifications to the amino acids such oxidation, and interactions between the amino acids. Changes in the microscopic properties effect the macroscopic characteristics and this body of work sheds light on how a variety of microscopic properties effect macroscopic characteristics. These macroscopic characteristics can include ligand binding, protein or peptide stability, secondary and tertiary structure, and biological function.

Three separate studies were done in the development of this body of work. The first observed the influence of oxidation of a methionine residue within a peptide scaffold on the strength of the methionine-aromatic interaction (Figure 1.1 A). The second studied synaptotagmin and the effect that the addition of the cytosolic linker region connecting the functional domain to the transmembrane domain has on the macroscopic characteristics of the C2A domain (Figure 1.1 B). The third study microscopic properties and resulting macroscopic characteristics of the first actin binding domain of dystrophin and it homolog utrophin (Figure 1.1 C).

To study both the microscopic properties and the macroscopic characteristics a variety of experimental and computational methods were used. Experimental approaches were used to determine thermodynamic and structural characteristics while molecular dynamics simulations, sequence analysis, and database searching provided insight into the microscopic properties of the systems being studied. Together these approaches provided a complete view of the systems being studied.

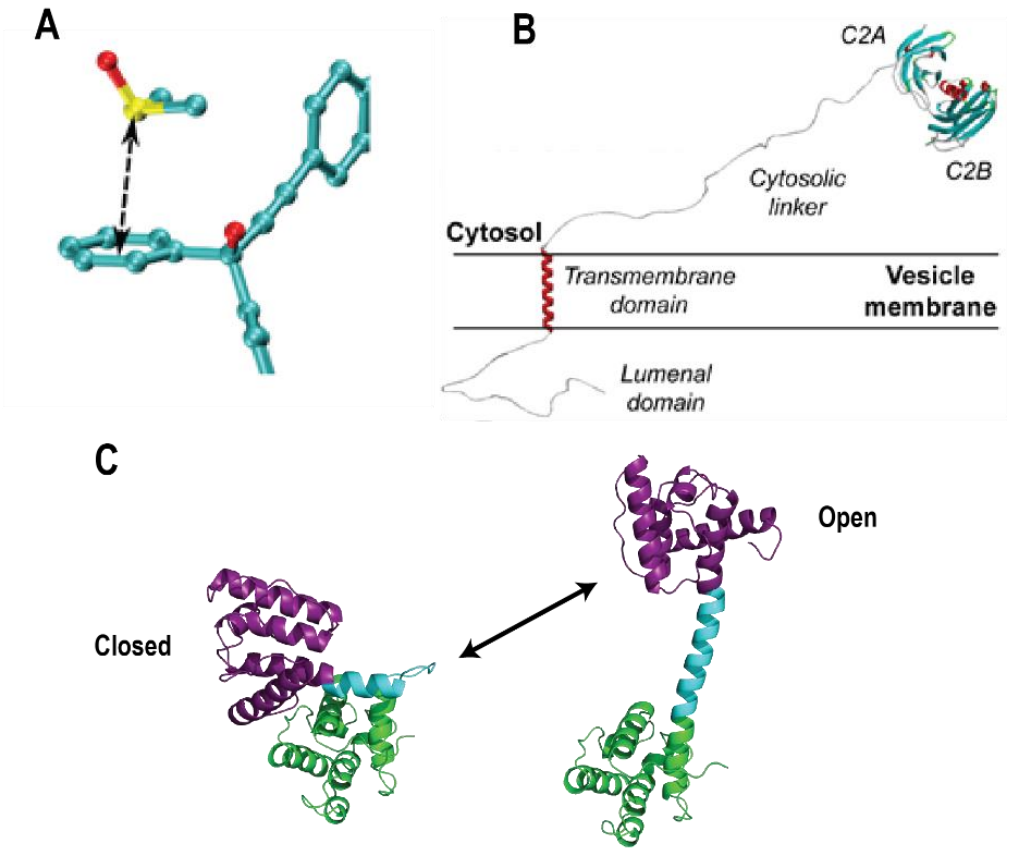


Figure 1.1. Three separate incidences in which the microscopic properties define the macroscopic characteristics. The oxidation of methionine increases the strength of the methionine aromatic interaction (A). The addition of the cytosolic linker region of synaptotagmin C2A changes the internal disorder of the protein and the interaction with its binding partners (B). The presence of hydrophobic amino acids within dystrophin's first actin binding domain create a hydrophobic patch upon the protein closing which stabilizes the conformation (C).

Chapter 2: Oxidation Increases the Strength of the Methionine-Aromatic Interaction

Note: The following chapter was reproduced in its entirety in accordance with rights given to authors by Nature Chemical Biology Andrew K. Lewis[1], Katie Dunleavy[2], Tiffany L. Senkow[1], Cheng Her[4], Benjamin T. Horn[2], Mark A. Jersett[2], Ryan Mahling[2], Megan R. McCarthy[4], Gabriella T. Perell[3], Christopher C. Valley[1], Christine B. Karim[4], Jiali Gao[3], William C. K. Pomerantz[3], David D. Thomas[4], Alessandro Cembran[2], Anne Hinderliter[2*], and Jonathan N. Sachs[1*](2016).Oxidation Increases the Strength of the Methionine-Aromatic Interaction. Nat. Chem. Bio. In Press.

[1] Department of Biomedical Engineering, University of Minnesota – Twin Cities, Minneapolis, MN 55455
[2] Department of Chemistry and Biochemistry, University of Minnesota – Duluth, Duluth, MN 55812
[3] Department of Chemistry, University of Minnesota – Twin Cities, Minneapolis, MN 55455
[4] Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota – Twin Cities, Minneapolis, MN 55455
[*]Co-corresponding authors

Project conception and production directed by J.N.S. PDB search, molecular dynamics simulations, experimental and computational LTα work, and computational CaM work were performed by A.K.L., T.L.S., C.C.V. and J.N.S. CSD search and NMR were performed by G.T.P. and W.C.K.P. Quantum calculations were performed by M.A.J., A.C, and J.G.. CD was performed by R.M., B.T.H., K.D. and A.H. Peptides were synthesized by C.B.K. CaM work was carried out by M.R.M, C.H., and D.D.T

2.1 Introduction

The methionine-aromatic motif is a promiscuous, highly stabilizing, and prevalent non-covalent interaction that forms between methionine thioether sulfurs and the aromatic rings of nearby (~5 Å) tyrosine, tryptophan, and phenylalanine residues[1-6]. Methionine-aromatic interactions have been detected in several bioinformatics studies[5,6], including our exhaustive analysis of the Protein Databank (PDB)[1]. Sulfur-aromatic (S/Ar) interactions between the C-S-C divalent sulfur motif and aromatic groups have been similarly found in the Cambridge Crystal Structural Database (CSD)[2]. The motif is substantially stronger than a purely hydrophobic contact, suggesting that it plays a unique role in stabilizing protein structure and ligand-receptor binding[1]. A thermal

unfolding experiment using a synthetic helical peptide measured the Met-Phe interaction at -0.65 ± 0.15 kcal/mol[7], and quantum calculations have revealed an interaction energy of 1-3 kcal/mol at a range of 4-6 Å between dimethyl sulfide (DMS) and benzene[4]. Our own quantum calculations expanded the analysis to include DMS interactions with phenol and indole (tyrosine and tryptophan analogs, respectively) and determined that dispersion interactions are the major contributor to stability[1].

In addition to forming highly stabilizing interactions with aromatic residues, methionine is uniquely important because its thioether can undergo modification with the addition of one or two sulfonyl oxygens to become methionine sulfoxide (Met^{Ox}) or sulfone, respectively, in oxidizing environments[8]. Elevated levels of methionine sulfoxide have been detected in diseased tissues, but methionine sulfone is rare[9]. Methionine sulfoxidation is known to regulate the activity of certain proteins, while in other cases, methionine-rich domains protect critical amino acids from oxidation by removing free radicals and regulating the local redox environment[10].

Despite this functional significance, the specific molecular mechanism through which proteins sense and respond to oxidative modification of methionine residues has not been fully characterized. It has been posited that polarization of the methionine thioether by oxidation could weaken the methionine-aromatic interaction by ablating the hydrophobic contribution[11], but this has never been tested experimentally. In the present study we will show the opposite to be true – that oxidation of methionine strengthens its interaction with aromatic residues, and that this has the potential to disrupt native conformation and alter protein function. Importantly, we show that the interaction is strengthened even in the absence of a hydrogen bond between the sulfoxide oxygen (acceptor) and donor hydrogens (on Tyr or Trp). Through quantum mechanical

calculations, we show that the interaction between the dipole moment of dimethyl sulfoxide (DMSO) and the aromatic ring quadrupole moment is the main contribution to the increased stability of the motif.

We start by analyzing both the PDB and the CSD to identify and characterize interactions between DMSO (a small molecule analog to Met^{Ox}) and aromatic groups in the crystal structures of proteins and small molecules. We then performed quantum mechanical calculations on model compounds to determine the strength of the sulfoxide-aromatic interaction, and assessed the potential contribution of hydrogen bonding on the motif in a variety of solvent environments. We then isolated and determined the strength of the specific interaction between unoxidized and oxidized methionine and the sidechain of phenylalanine by extending a previously described helix unfolding experiment and performing a double mutant cycle analysis[7]. We investigated the effects of site-specific methionine oxidation on the methionine-aromatic interaction in two biologically important proteins. First, using a combination of electron paramagnetic resonance (EPR) spectroscopy and replica-exchange molecular dynamics (REMD) simulations, we probed the effects of oxidation on the structure of calmodulin (CaM). Second, we investigated lymphotoxin-α (LTα) binding to tumor necrosis factor receptor 1 (TNFR1), which we previously showed is stabilized by a methionine-aromatic interaction between M120 of LTα and W107 of TNFR1[1].

2.2 Methods

2.2.1 Structural Bioinformatics Analysis of the Protein Data Bank

We analyzed a non-redundant subset of the PDB comprised of all structures containing at least one DMSO molecule using the Biopython toolset and a custom Python

script adapted from our previous study[1]. Structures sharing >90% sequence identity were represented by a single structure with the highest score as ranked according to RCSB PDB quality factor algorithm. Redundant structures were excluded, as they do not represent truly independent data points for the purposes of our informatics study. Distances were defined between the DMSO sulfur and the center of the aromatic ring. Distance data were collected up to a cutoff of 20 Å. Hydrogen bonds were defined as having a heavy-atom donor to sulfonyl oxygen distance of < 3.5 Å. For the radial density function (RDF), the distances from each DMSO to aromatic groups were binned and the histogram normalized by shell volume. Each RDF was normalized by the relative population of amino acids in the analyzed PDB files (divided by 9.25% for F, Y, W, 22.47% for L, I, V, and by 100% for all amino acids). The resulting plots were normalized by the mean value at the long distance, non-interacting region of the "All Amino Acids" RDF to shift the bulk density to 1 and to cast the y-axis as an enrichment factor.

2.2.2 Structural Bioinformatics Analysis of the Cambridge Structural Database

Searches were executed in the Conquest program of the Cambridge Structural Database (CSD version 5.35, 2014 release), and structural analyses were performed using the Mercury program (v 1.5). The distance was defined as in the PDB. C-CH_2-C was used as a non-interacting analog to C-S-C, following a previous CSD analysis[2] The sulfur (or carbon in control) atom to aromatic distance was constrained to 0-10 Å. For the radial density function (RDF), the distances from each DMSO (or C-CH_2-C) to all aromatic groups were binned and divided by shell volume, then normalized by the total number of contacts. The resulting RDFs were shifted to 1 as in the PDB analysis, by normalizing to the value at the flat region of the C-CH_2-C RDF. A total of 20630 DMSO-aromatic pairs

were found in 840 structures and 3093875 C-CH_2-C-aromatic pairs were found in 72745 structures.

2.2.3 ABF Simulations of DMSO and Aromatic Groups in Three Solvents

Simulations were performed using the Adaptive Biasing Force (ABF) module in NAMD 2.9 using the CHARMM 36 generalized forcefield. The adaptive biasing force (ABF) module was used so that the system could overcome free energy barriers and fully sample all conformational states. Parameters for DMS were adapted from those for ethylmethylsulfide. The distance between the DMSO thioether and the center of mass of the 6-membered ring of phenol, or indole was used as the reaction coordinate of the mean force calculation. The system potential along the reaction coordinate was negated by the adaptive biasing force to simulate a flat free energy landscape. Each system was solvated in a ~45 $Å^3$ periodic box of TIP3P water, hexane, or ethyl acetate, for a total of 18 systems. The adaptive biasing potential was applied between 2.5 and 14 Å with a binwidth of 0.1 Å and simulations were carried out for at least 7.5 x 10^6 steps. All other simulation parameters were set as in our traditional MD (see Molecular dynamics simulations of LTα and TNFR1 methods). Supplementary Table 2.4 shows the effects of solvation on the calculated interaction energy between DMS/DMSO and each aromatic compound. The calculated PMF was not used in our analysis. Supplementary Figure 2.17 shows the distance between the aromatic ring and each heavy atom of the DMSO molecule and demonstrates no strong bias for the methyl groups to interact with the aromatic ring.

2.2.4 Peptide Synthesis

To study the appropriate pair sequences in a peptide chain, the un-oxidized amino acid Met or oxidized Met was inserted into the sequence: a) Tyr-Gly-Gly-Ser-Ala-Ala-Glu-

Ala-***Aromatic***-Ala-Lys-Ala-**Met**-Ala-Arg-NH2, b) Tyr-Gly-Gly-Ser-Ala- Ala-Glu-Ala-***Aromatic***-Ala-Lys-Ala-**Met (ox)**-Ala-Arg-NH2. The peptides were assembled on Fmoc-PAL-PEG-PS resin by Fmoc chemistry using a PE Biosystems PioneerTM protein synthesis system. Single Met oxidation was accomplished by incorporation of Fmoc-Met (ox)-OH during synthesis. Standard *N*-[(dimethylamino)-1*H*-1,2,3-triazolo[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminium hexafluorophosphate N-oxide (HA TU)/ *N,N*-diisopropylethylamine (DIEA) (1/2.4 eq.) activation, in 1-methyl-2-pyrrolidinone (NMP), was applied. Fmoc deprotection was achieved with 20% piperidine in NMP. The final release of the peptides, with removal of the side chain protecting groups, were accomplished by exposure of the peptide-resin to 82.5% trifluoroacetic acid (TFA), 5% phenol, 5% thioanisol, 2.5% 1,2-ethanedithiol, 5% water (Reagent K). The peptides were precipitated with cold methyl-*t*-butyl ether, vortexed, centrifuged, decanted, and dried over argon. The dried peptides were dissolved in degassed water and purified by high-performance liquid chromatography (HPLC) using a reversed- phase C8 HPLC column. Peptide elution was achieved with a linear gradient from 0 to 34% B (95% acetonitrile / 5% water / 0.1% TFA) in 40 min at a flow rate of 2.5 mL/min with detection at 280 nm using a System Gold Beckman Coulter system. The HPLC fractions were collected and analyzed by mass spectrometry (MS).

2.2.5 Peptide Scaffold Design

The peptide scaffold utilized for defining the energetic impact of oxidation on Met-Aromatic interaction had a number of specific design elements. The helical content was optimized to reside within a range that enabled sensitive detection of its change by circular dichroism (CD). The capping box motif at the N-terminus is the sequence Ser/Thr-X-X-Glu/Gln and is a helix stop signal[33]. This stop signal has Ser or Thr as the

N-cap and a Glu or Gln residue at position i+3, specifically; the sequence Ser-Ala-Ala-Glu is the capping motif in each of the peptides of the double mutant cycle. To this capping box, the N-terminus had a tyrosine added for concentration determination. The C-terminus was blocked by changing the COO- to a CONH2 to mediate this effect of dipole destabilization and to increase helical propensity[34]

2.2.6 Electrospray Ionization (ESI)-MS

The peptides were analyzed on a LC/MSD ion trap (Agilent). ESI-MS spectra were acquired in positive ion mode. The solvent contained 40 % acetonitrile and 0.1 % formic acid. 10 to 30 μL sample were directly injected at a flow rate of 10 μL/min with a source temperature of 300 ˚C. The applied spray voltage was 3500 V, and the skimmer voltage 40 V. MS scans were acquired over an m/z range of 200 to 2000. The scan for the native mass 1499.69 Da of the sequence Tyr-Gly-Gly-Ser-Ala-Ala-Glu-Ala-**Phe**-Ala-Lys-Ala-**Met**-Ala-Arg, shows the single charged ion with *m/z* 1500.0 and the double charged ion *m/z* 750.7. The scan for the Met(O)-Phe sequence pair (1515.69 Da) shows the singly charged ion with *m/z* 1515.7, the double charged ion *m/z* 758.4, and the triple charged ion *m/z* 505.9. The observed masses for the respective molecular ions agree with the theoretical peptide scaffold design.

2.2.7 Circular Dichroism Spectroscopy

All peptides were stored in tetrafluoroethylene (TFE) after synthesis, which was removed by evaporation using Nitrogen gas prior to use. To ensure complete removal of the TFE, the sample was then put under vacuum for one hour. The dried peptide was re-constituted in 10mM KH_2PO_4, 100mM KCl at pH 7.5, and prepared for CD data collection. All CD experiments were performed on a Jasco J-815 CD Spectrometer (Annapolis, MD)

using a 0.1cm quartz cuvette at ~150μM peptide concentration. Concentrations of the peptide samples were determined using a Nanodrop Spectrometer. To display helical content and the presence of an isodichroic point ellipticity was measured between 200nm to 260 nm in 1nm increments from -2°C to 60°C for a FM, FM (ox), FA, AM, AM (ox) and AA. Once these characteristics were established for each construct, one data point was collected at 222nm at each 0.5°C or 2°C change as temperature increased from - 2°C to 60°C. Cooling melts were also collected, recording the signal at 222nm, on all peptides as temperature decreased from 60°C to -2°C. All collected data points were averaged from 3 acquisitions of the recorded ellipticity.

2.2.8 Analysis of Circular Dichroism Data

To extract the free energy of interaction from the temperature dependence of the CD, spectra were collected from the lowest experimentally accessible temperature (-2°C) to a temperature beyond which there were no observable changes in spectral shape for all peptides (60°C). The spectra are given in Supplementary Figure 2.7. Following Greenfield[18], we used the signal, Θ, at 222nm to report on changes in peptide structure. This wavelength was chosen because it reflects helical content and is also the wavelength at which the overall change in ellipticity is greatest across the temperature ramp. Following the same rationale we used previously[35], the CD signal is given by:

$$\Theta = (\Theta_{initia} + K\Theta_{final})/(1 + K) \quad (2.1)$$

where $\Theta_{initial}$ is the value at 222nm recorded at -2°C, and Θ_{final} is the value at 222nm recorded at 60°C. Because $K = e^{-\Delta G/RT}$, we rewrite this equation as:

$$\Theta(T) = (\Theta_{initial} + e^{-\frac{\Delta G}{RT}}\Theta_{final})/(1 + e^{-\frac{\Delta G}{RT}}), \quad (2.2)$$

We then fit the temperature dependent data, $\Theta(T)$, by substituting the Gibbs-Helmholtz

equation into equation 2.2,

$$\Delta G = \Delta H_{Tm}\left(1 - \frac{T}{T_m}\right) + \Delta C_p(T - T_m - T(\ln(\frac{T}{T_m}))), \quad (2.3)$$

and varying the three parameters: enthalpy (ΔH_{Tm}), heat capacity change (ΔC_p), and transition temperature (T_m).

Data for Phe-Met and Pet-Met(ox) are given for each replica in Supplementary Figure 2.5. By varying the concentrations of the peptide in each of its three replicates, the measured signal varies for each replica. The three replicates are globally fit simultaneously, an approach that improves the stringency of the parameter fits for each peptide. The free energy is then extracted from the fit at the desired temperature (T=0^{o}C in this case). The change in free energy for each leg of the double mutant cycle is reported in Supplementary Figure 2.6.

2.2.9 Double Mutant Cycle

To construct a double mutant cycle, peptides were synthesized in which the aromatic residue and Met (or Met^{Ox}) was replaced first individually, then concurrently by the neutral residue alanine (Ala). The difference in ΔG between the Met-aromatic (oxidized or un-oxidized) and its respective singly substituted construct, or between the singly substituted construct and the doubly mutated construct gives a ΔΔG. The difference of the respective ΔΔG values produces a ΔΔΔG, which indicates whether the specific interaction is contributing to the stabilization of structure. ΔΔΔG is simplified here as ΔG_{int} to match the free energy of interaction reported in the reference[7] (the two expressions are

thermodynamically equivalent, although the cited study did not use a double mutant cycle). If ΔG_{int} equals zero, the interaction observed is not contributing to the structural stability; if greater than zero, the interaction is stabilizing; and if less than zero, destabilizing[36-38].

2.2.10 Principal Component Analysis

Principal component analysis for each of the FM, FM^{Ox}, FA, AM, AM^{Ox}, and AA peptides data sets was carried out with the "prcomp" function of the program R (v. 3.2.1). For each data set only one of the three replicates was used. To increase the sensitivity of the analysis, we focused on the isodichroic point region by analyzing only the wavelengths between 200 nm and 210 nm (11 data points), while the entire temperature range was included in the analysis. The correlation matrix was calculated between the mean-centered vectors of wavelengths:

$$C_{ij}=T_i(\lambda_1,\lambda_2,\ldots,\lambda_N)\cdot T_j(\lambda_1,\lambda_2,\ldots,\lambda_N)$$

The results reported in Supplementary Figure 2.9 and Supplementary Table 2.5 show that the first two principal components describe 94% or more of the total variance, which is consistent with the assumption that the change in molar ellipticity is dominated by only two sources of signal.

2.2.11 NMR Experiments

Concentrations of peptide samples were obtained using ultraviolet absorbance of Tyrosine (ε= 1450) in 8 M urea. Peptide Y-Met^{Ox} (YGGSAAEAFAKAM(ox)AR-NH_2) was dissolved in deuterated methanol (3 mM). F-Met^{Ox} (YGGSAAEAFAKAM(ox)AR-NH_2) was dissolved in acidic (pH 3.5) D_2O, taken at 4°C. ^1H-^1H ROESY experiments (mixing time = 200 ms) were performed on a Bruker 700 MHz spectrometer with a CryoProbe: 5 mm TXI probe of Z-Gradient. ^1H experiments comparing control peptide, A-Met^{Ox} to F-Met^{Ox} in MeOD were

performed on a Bruker Advance III HD 500 MHz instrument with a 5 mm Prodigy TCI cryoprobe with z-axis gradients. Data was processed and analyzed using Bruker Topspin 3.2.

2.2.12 Dipolar EPR Spectroscopy of Spin-Labeled Calmodulin

A mammalian CaM mutant, with V136 and T146 both mutated to Cys, was prepared by site-directed mutagenesis (QuikChange II kit, Agilent, Santa Clara CA), confirmed by DNA sequencing, expressed and purified as described previously[39], and dialyzed overnight at 4°C against 10mM NaCl, 10mM Tris (pH 7.0). Spin-labeling with maleimide spin label (MSL, N-(1-oxyl-2,2,5,5-tetramethyl pyrrolidinyl) maleimide, Toronto Research Chemicals, Canada) was carried out by incubating 120 μM CaM with 480 μM MSL in CaM buffer (10 mM NaCl, 10 mM Tris, 5 mM EGTA, pH 7.0) for two hours at 22° C (followed by overnight dialysis against CaM buffer to remove unreacted label), resulting in complete labeling of the two Cys, as shown by mass spectrometry. To oxidize methionine side chains, spin-labeled CaM was incubated in 500 mM hydrogen peroxide for 30 min, followed by dialysis into CaM buffer. As shown previously under similar conditions[39], mass spectrometry verified that all of the nine methionine residues of CaM were oxidized. X-band (9.5 GHz) EPR spectra of 200 μM CaM (in CaM buffer plus 10% glycerol) were acquired at 200° K with a Bruker ER500 spectrometer, using 1 G modulation amplitude and sub-saturating microwave power (0.63 mW). EPR spectra were analyzed to determine the distribution of distances between the two spin labels, as described previously[40].

2.2.13 Ligand Preparation and Treatment

Plasmids encoding human LTα and human TNF expressing an N-terminal FLAG tag, downstream from an inducible T7 promoter were subcloned using standard cloning

techniques. Plasmids were transformed into chemically competent BL21 bacteria, plated on LB agar plates, and cultured in LB to an OD600 of ~0.5. Cultures were then cooled to 18°C, induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside, and grown for 18 hours. The bacteria suspension was centrifuged, then lysed by sonication in ice-cold PBS. The lysate was column purified using anti-FLAG M2 Affinity Gel (Sigma) according to the manufacturer protocol, dialyzed exhaustively against MilliQ H_2O, and then lyophilized. FLAG-LTα, as prepared in our laboratory, contains 4 methionine residues, of which M120 and M133 potentially play structural and functional roles in the ligand-receptor complex. M133 is buried between two folds of a β-sheet, and its precise functional role, if any, has not been established. For the LTα-M120 mutant, we mutated M133 to valine and M20 to alanine in order to match the homologous residues found in Bos Taurus LTα, which shares 73% sequence identity and 81% sequence similarity (Sequence Manipulation Suite), but lacks methionines except for its N-terminus. TNF, which also binds to and activates TNFR1 with similar potency[41] was used as a control.

Oxidative stress was applied by dissolving 8 μg of ligand in buffer containing 10 mM HEPES, 100 mM KCl, 1 mM MgCl, 1 mM CaCl, then adding 100 mM H_2O_2 to a total volume of 100 uL. The oxidation reaction was allowed to proceed for 18 hours at room temperature, then stopped by removing H_2O_2 in 2 sequential stages using Zeba 7KDa molecular weight cutoff spin columns. Removal of H_2O_2 was confirmed by Amplex red assay, and the residual concentration of H_2O_2 was determined as insufficient to induce autophagy. We measured tryptophan fluorescence to confirm that oxidation does not lead to gross unfolding of LTα, finding no shift in the emission spectrum when excited at 280 nm (Supplementary Figure 2.15).

2.2.14 Western Blot Analysis of IκBα

HEK 293 cells were maintained in DMEM supplemented with 10% FBS, penicillin/streptomycin, and 2 mM L-glutamine. Confluent cells were split at 1:12 into 6-well plates and used for experiments on day 3. Untreated and oxidized ligands were added to cells to final concentrations of 200 ng/mL and 20 ng/mL for LTα and TNF respectively, and incubated for 20 minutes at 37 °C. Cells were then washed once and gently sheared from the plate with ice-cold PBS, then centrifuged and resuspended in radioimmunoprecipitation assay (RIPA) lysis buffer. Cell lysates were normalized to equal protein concentrations, and 60 µg were loaded onto 4-12% Bis-Tris gradient gels for Western blot analysis. IκBα was detected using horseradish peroxidase (HRP) conjugated rabbit α-IκBα (Cell Signaling Technology, #9242, 1:1000) and α-Rabbit IgG (Amersham, 1:10,000). Full gels are shown in Supplementary Figure 2.16.

2.2.15 Co-Immunoprecipitation

Co-immunoprecipitation was performed using anti-FLAG M2 agarose beads (Invitrogen) according to the manufacturer instructions. Briefly, HEK 293 cells were split at 1:12 into 15 cm plates and transfected with 8 µg of plasmid encoding TNFR1 (pCMV6-XL5-TNFR1) on day 1 by calcium phosphate transfection. On day 3, cells were lifted by shearing, washed once, and resuspended in ice-cold PBS at ~10^6/mL. LTα and TNF were added at 12 and 25 ng/mL respectively, and the cells were incubated for 30 minutes at 4 °C while rocking. Cells were then washed 3 times in ice-cold PBS and resuspended in RIPA lysis buffer, then analyzed by Western blot. TNFR1 was detected using rabbit α-TNFR1 (Abcam, ab19139, 1:1000) and α-rabbit secondary, as described above. Full gels are shown in Supplementary Figure 2.16.

2.2.16 Quantum Calculations

The interaction energy between model compounds was investigated with quantum mechanical calculations performed in the gas phase. The reduced and oxidized states of methionine were modeled by dimethyl sulfide (DMS) and dimethyl sulfoxide (DMSO), respectively; the side chains of phenylalanine, tyrosine, and tryptophan were modeled by benzene, phenol, and indole. The interaction energy between compounds A and B was defined as:

$$E(AB)_{int} = E(AB)_{complex} - E(A) - E(B) \quad (4)$$

All energies were obtained from fully unrestrained structure optimizations carried out with Gaussian 09[42] using the Møller-Plesset perturbation theory at the second order (MP2/6-311+G(d,p)) and account for basis set superposition error[43]. Additional data computed at the M06-2X/6-311+G(d,p)[44] level are reported in Supplementary Table 2.1. These two levels of calculation were chosen because we previously showed that they bracket the interaction energy calculated at the CCSD(T)/6-311+G(d,p) level[1]. Notice that for each complex several minima were characterized, and the data is reported only for the lowest energy one.

The dispersion energy contribution to the complex formation was calculated as in Ref. [1] by taking the difference between the interaction energy (eq 4) computed at the MP2 and Hartree-Fock (HF) levels. In both cases the same basis set was employed, and the HF energies were calculated as single point on the MP2 optimized structures.

The effect that the variety of conformations adopted in solution has on the interaction energy was investigated by averaging the interaction energies of 128 structures randomly sampled from the minimum (sulfur – aromatic distance between 4 – 6 Å) of the ABF simulations in ethyl acetate described above. Hydrogen bonded structures were defined as having donor-acceptor distance of <2.5 Å and not-hydrogen bonded structures

had a donor-acceptor distance of >4 Å. Upon removal of the solvent, the instantaneous interaction energy was calculated at the MP2/aug-cc-pVDZ level using eq. 4 without any additional structure optimization.

A similar procedure was employed to investigate the interaction energy between the side chains of M120, Y96, and Y122 in LTα. From the MD simulations described below, 128 structures were randomly selected and the average interaction energy calculated at the MP2/aug-cc-pVDZ level as:

$$E(M_{120}Y_{96}Y_{122})_{int} = E(M_{120}Y_{96}Y_{122})_{complex} - E(M_{120}) - E(Y_{96}) - E(Y_{122}) \qquad (5)$$

The side chains were truncated at the C_β-C_γ bond and the missing valence was saturated with a hydrogen atom. Upon removal of the rest of the protein and the solvent and without any further optimization, the interaction energy was calculated at the MP2/aug-cc-pVDZ level and corrected for the basis set superposition error.

For the calculations on the ethyl acetate simulations described above, the Natural Bond Orbital[45] (NBO) charges were calculated and averaged over all 128 structures. The NBO localization scheme was tested along Mulliken[46] localization over a number of different basis sets of increasing size. Because the results showed a strong dependence of the Mulliken charges on the basis set size[47] while NBO charges did not, the NBO localization scheme was preferred.

2.2.17 Replica Exchange Molecular Dynamics of Calmodulin

The replica exchange molecular dynamics (REMD) simulation was performed in NAMD 2.9 using the CHARMM22 potential parameters[48]. Eleven residues of the last helix (136-146) of CaM (3CLN.pdb)[27] were used in the all-atom point-charge force field *in vacuo* to match conditions of the quantum calculations. Eight replicas with temperatures

exponentially spaced from 300K to 600K (300, 331.23, 365.7, 403.77, 445.8, 492.2, 543.43, 600 K) were used with the initial velocities of the atoms generated by the Maxwell-Boltzmann distribution and a timestep of 1 fs. The first 10^6 steps were performed without attempting exchanges to equilibrate, then exchanges were attempted every 1000 steps (1 ps) with an acceptance rate ~20%. Coordinates were taken every 100,000 steps (100 ps). 78,810 frames were taken for the unoxidized peptide and 74,400 frames for the oxidized peptide, corresponding to a total simulation time per replica of 985 and 930 ns, respectively. Only the replicas at 300 K were used in the analysis. Methionine-aromatic interactions were identified as having a distance of <7 Å and an angle less than 60°, consistent with previous definitions[1,2]. Hydrogen bonds were cutoff at <2.5Å from hydrogen to acceptor.

2.2.18 Molecular Dynamics Simulations of LTα and TNFR1

We carried out molecular dynamics simulations of oxidized or unoxidized LTα in its unbound state and unoxidized in complex with TNFR1 using the structure, 1TNR as the starting configuration[49]. Two additional systems, a Y122 interacting configuration and a Y96 interacting configuration, were constructed in which M120 began bound to Y122 or Y96, respectively, while all other residues in the 1TNR crystal structure were left essentially unchanged (minor manipulation and steepest descent minimization was required to resolve overlap in the binding pocket). The starting configurations of M120, Y96, and Y122 in these systems were taken from the oxidized ligand simulation. Each system was solvated in a box of at least 40,000 explicit water molecules modeled as TIP3P in order to fully capture solvation effects [50] and the charge was neutralized by adding K^+ and Cl^- ions. The isothermal-isobaric (NPT) ensemble was used along with the CHARMM36 force field[51-53] at a temperature of 303 K and 1 atm pressure, which were

maintained using the Langevin piston and Nosé-Hoover[54] algorithms. Parameters for Met^{Ox} were provided by Krzysztof Kuczera[55] and adapted for use in the CHARMM36 force field. Long-range electrostatics were calculated using the particle mesh Ewald method [56] with a 1.5 Å grid spacing and 4th order interpolation. Lennard-Jones interactions were switched off at a cutoff distance of 10 Å. Each system (except the Y96 and Y122 interacting configurations) was minimized using NAMD 2.8[48] for 1000 steps, then equilibrated with Cα harmonic constraints of 1, 0.1, and 0.01 kcal/mol/Å^2 for 2 ns at each stage. Dynamic trajectories were propagated with the r-RESPA algorithm [57] with a 2 fs time step and the RATTLE algorithm[58] was applied to all covalent bonds involving hydrogen. The ligand-receptor system was simulated for ~35 ns and distance and angle analyses were averaged over the entire trajectory. The ligand-only systems were simulated for a total of ~430 ns. The first 150 ns (including relaxation) were excluded from structural analyses to allow re-equilibration of the ligand after removal of the receptor chains. The Y96 and Y122 interacting systems were equilibrated with great care before running unrestrained dynamics. Each was minimized to resolve steric clashes, then equilibrated according to the scheme shown in Supplementary Table 2.6. Distances and angles were calculated from the thioether sulfur of M120 to the centers and normal vectors of the aromatic groups of Y96, Y122, and W107. Methionine-aromatic interaction cutoffs were defined as in the CaM simulations.

2.3 Results

2.3.1 DMSO-Aromatic Interactions in Crystal Structure Databases

To find examples of sulfoxide-aromatic interactions, we searched the PDB and CSD for occurrences of DMSO molecules near aromatic rings (Figure 2.1). The PDB is a repository containing crystal structures of proteins and the CSD is the analogous repository for organic and metal-organic small molecules. DMSO solubilizes tryptophan more readily than leucine, alanine, and glycine[12], so we expected to find that it specifically interacts with aromatic sidechains. Our PDB search yielded 205 unique structures containing a total of 872 DMSO molecules. Figure 2.1a shows a representative DMSO-tyrosine interaction. When compared to all amino acids, or just the aliphatic amino acids, aromatic residues are enriched 4-7 Å from DMSO sulfurs, with a maximum at 5Å, resembling the radial density function for methionine-aromatic interactions from our previous study[1] (Figure 2.1b). 451 (51.7%) DMSO molecules were found within 7 Å of the nearest aromatic residue. Only 32 of these (7.1%) are in position to accept a hydrogen bond from either tryptophan or tyrosine. Thus, hydrogen bonding does not drive DMSO-aromatic bonding in the PDB. Our CSD search similarly found an ~50% enrichment of DMSO-aromatic contacts (representative example, Figure 2.1c) compared to C-CH_2-C-aromatic pairs (Figure 2.1d), which were used previously as a non-interacting control to show the enrichment of the Met-aromatic motif[2]. Again, few DMSO/phenol or DMSO/indole pairs are in position to form a hydrogen bond (<14%).

We then parsed the frequency of methionine-aromatic interactions in the PDB <7 Å between DMSO and the center of phenylalanine, tyrosine, and tryptophan individually. Figure 2.1e shows that the interaction is more prevalent with Tyr and Trp than with Phe when normalized by the relative abundance of each of these amino acids in the analyzed structures. This trend is not explained by relative solvent accessibility, as Trp and Phe have similar solvent accessibilities, while that of Tyr is somewhat higher[13]. We calculated

a 1.3 and 1.4 fold preference for tyrosine and tryptophan over phenylalanine, respectively, translating to a ~0.3 kcal/mol increase in Met^{Ox}-Tyr and Met^{Ox}-Trp stability compared to Met^{Ox}-Phe. Therefore, these database results suggest that, although hydrogen bonding is not common in the interaction motif, DMSO interacts more favorably with tyrosine and tryptophan than with phenylalanine. This has similarly been observed in cation-π interactions in proteins, where the more electron-rich Pi systems of tryptophan and tyrosine relative to phenylalanine make them more attractive to positively charged ligands[14].

2.3.2 Oxidation Strengthens Interaction in Quantum Calculations

The interaction energies between the reduced (DMS) and oxidized (DMSO) models of methionine with the model aromatic compounds benzene, phenol, and indole are reported in Table 2.1 and Supplementary Results, Supplementary Table 2.1. The data show that upon oxidation of DMS to DMSO the complex with benzene is 0.9 kcal/mol more stable. To understand the origin of this stabilization we first tested the possible contribution of dispersion interactions. The results (see Supplementary Table 2.2) show that the dispersion energy contribution for DMS/benzene is -8.4 kcal/mol, and for DMSO/benzene it is -8.2 kcal/mol. This change in contribution is small (0.2 kcal/mol) and is more stabilizing for the DMS/benzene interaction than for the DMSO/benzene. Therefore dispersion interactions are not the source of the stronger interaction energy in the oxidized complex. Rather, the stronger interaction can be explained in terms of the interaction between the dipole moment of DMSO and the aromatic ring quadrupole moment. The calculated dipole moment for DMS is 1.8 Debye, while in DMSO the enhanced polarization due to the presence of oxygen results in a dipole moment of 5.0 Debye roughly aligned along the S=O bond. Furthermore, both DMS and DMSO maximize dispersion interactions with the

benzene ring by aligning the Me-S-Me moiety over the ring plane; as a result, in the DMS/benzene complex the orientation of the small dipole of DMS is orthogonal to the benzene quadrupole moment, while in DMSO/benzene the angle between the S=O bond and the ring is approximately 120°, which is well poised for a favorable dipole/quadrupole interaction. The Natural Bond Orbital (NBO) charge analysis for the complex and isolated species (Supplementary Table 2.3 and Supplementary Figure 2.1) supports this explanation, showing a large positive charge (+1.19 e) on the sulfur that can interact with the π electron cloud, balanced by a negative charge (-0.96 e) on the oxygen pointing away from the benzene ring (Supplementary Figure 2.2).

Table 2.1 shows that for phenol and indole the change in interaction energy upon DMS oxidation (-5.7 and -5.3 kcal/mol, respectively) is of larger magnitude than for benzene. Such a large change cannot be explained in terms of enhanced dipole/quadrupole interactions, and is a result of hydrogen bonding between the DMSO oxygen and the hydrogen of the phenol hydroxyl group or the indole amine (Supplementary Figure 2.2). To determine the relevance of hydrogen bonding in sulfoxide-aromatic interactions in a range of environments, we ran molecular dynamics simulations of DMSO and phenol or indole in hexane, ethyl acetate (EtOAc), or TIP3P water using the adaptive biasing force (ABF) module in NAMD. We found (Supplementary Figure 2.3) that DMSO hydrogen bonded with phenol and indole in hexane. Thus, in this nonpolar environment we expect the contribution of the hydrogen bond to outweigh the increased interaction energy of the S/Ar motif. Importantly, in EtOAc (a polar non-protic solvent with dielectric constant of 6, representative of the protein interior), hydrogen bonding occurred less frequently. Therefore, the S/Ar motif plays a role in the stabilization of the complex in this environment. Finally, in water hydrogen bonding between DMSO and the aromatic group was essentially absent, rather the sulfonyl and aromatic hydrogens were solvated

by water. This environment mimics interfacial residues on the protein, which are often involved in protein-protein interactions[15].

To disentangle the energy contribution due to hydrogen bonding from that of sulfur-aromatic interaction, we performed quantum mechanical calculations on configurational ensembles from our simulations in EtOAc. Benzene and DMS systems were also run to compare to the full optimization calculations. EtOAc's intermediate dielectric constant and its ability to accept hydrogen bonds allows DMSO-phenol/indole complexes to sample hydrogen-bonded and non-hydrogen-bonded conformations with similar probability. We analyzed non-hydrogen bonded and hydrogen bonded ensembles as mimics of the interactions on the protein surface and in its hydrophobic core, respectively. The results (Table 2.1, Figure 2.1f) agree with the data from the full optimization, and together indicate that (a) in the absence of hydrogen bond, oxidation of DMS to DMSO strengthens the sulfur-aromatic interaction by 0.5 - 1.5 kcal/mol, and (b) in hydrogen bonded structures (rare in proteins) the complex is stabilized by up to 8 kcal/mol.

2.3.3 Evaluation of Oxidation in a Model Peptide Scaffold

To experimentally assess the database findings and quantum results, we utilized thermal denaturation (monitored with CD) and nuclear magnetic resonance (NMR) to investigate a previously reported 15-residue peptide scaffold[7], designed such that the strength of the interaction between *i*, *i+4* amino acids determines the thermal stability of the peptide secondary structure. In that study, the Met-Phe pair was found to interact (ΔG_{int}) with 0.65 kcal/mol. We first confirmed a direct interaction between oxidized Met and Phe (Supplementary Table 2.4) using NMR. As shown in Figure 2.2a, we observed the strongest NOEs to the aryl hydrogens meta to the sidechain methylene (3,5 position) (Supplementary Figure 2.4). NOEs to the hydrogen at the 4 and 2,6 positions are present

at lower levels and overlap.

To further probe the details of the sulfoxide-aromatic (S^{Ox}/Ar) interaction in the helical state, we compared the chemical shift of the ε-S-methyl groups on the sulfoxide-containing peptide when it is in an *i* +4 relationship with phenylalanine (F-MetO) versus alanine (A-MetO) in methanol. Oxidized peptides were purified as a 1:1 mixture of diastereomers due to the new stereocenter from the sulfoxide. As such, two singlets corresponding to the methyl group of each diastereomer were observed (Figure 2.2b). DMSO methyl resonances in methanol are reported at 2.65 ppm[16]. Similarly, in our measurements for A-MetO they are found at 2.65 and 2.64 ppm. However, for F-MetO, we observe both ε-S-methyl resonances shifted upfield at 2.57 and 2.47 ppm. Interestingly, the chemical shift of the DMSO methyl resonances in benzene have also been shown to be significantly upfield at 1.68 ppm[16]. Due to ring current shielding effects from aromatic rings, we interpret these results to be consistent with the ε-S-methyl groups of oxomethionine approaching from above the aromatic ring of phenylalanine (Figures 2.1a,c). Both the NOE experiments and chemical shift information indicate an interaction between the sulfoxide group over top of the aromatic ring.

We then used CD to measure the change in the strength of the S/Ar interaction upon methionine oxidation. Our experiments reproduced the original S/Ar result for Met-Phe (ΔG_{int}=0.62 kcal/mol). The interaction strength was doubled (strengthened by 0.62 kcal/mol) upon methionine oxidation. Associated with this change in free energy, we find that oxidation increases the enthalpy of the S/Ar interaction by 1.1 kcal/mol (Supplementary Figure 2.5), consistent with the quantum result. We note that while CD spectra were used in the original S/Ar study to extract energetics of side-chain interactions, it is not possible to draw a direct correlation between peptide helicity and free

energy[7,17]. Rather, fitting of the CD spectra is necessary to extract energetics. Two aspects of our approach differed from the original study. First, we performed a double mutant cycle to isolate the methionine-aromatic interaction from possible interference due to neighboring residues (Supplementary Table 2.4, Supplementary Figure 2.6). This approach solidifies our reported interaction free energies, and is especially important in the case of oxidized methionine because of the increased likelihood of hydrogen bonds and electrostatic interactions with the other residues in the peptide.

Second, we followed Greenfield[18] by fitting the temperature-dependence of each CD spectrum (Supplementary Figure 2.7) to the Gibbs-Helmholtz equation (GHE, Supplementary Figure 2.5). We justify applying a two-state approximation and hence using the GHE by: 1) observation of complete reversibility (overlap) in the heating and cooling curves (Supplementary Figure 2.8); and 2) the presence of a molar ellipticity isodichroic point in all constructs[19] (Supplementary Figure 2.7, inset). Additionally, principal component analysis[20] of the CD spectra yielded only two components that dominated the CD spectra (Supplementary Figure 2.9 and Supplementary Table 2.5), again consistent with the two-state assumption. Together with the addition of the double mutant cycle, the fact that we obtained nearly identical values for the Met-Phe energy as in the original study[7] (compare 0.62±0.09 to 0.65 kcal/mol) reinforces the robustness of our approach. Thus, our peptide study lends experimental support to the conclusions drawn from the quantum calculations: oxidation increases the strength of the methionine-aromatic interaction.

2.3.4 Oxidation Rearranges Met-Aromatic Contacts in Calmodulin

In order to test the impact of the increased strength of the S^{Ox}/Ar motif in proteins, we first investigated the effects of methionine oxidation in an already well-

characterized protein, CaM. The C-terminal helix of CaM has a high methionine content and is sensitive to oxidative stress *in vivo,* which alters the protein's structure and function[21-26]. In the crystal structures of CaM, both in its calcium-bound (halo, 3CLN)[27] and unbound (apo, 1CFD)[28] forms, a C-terminal methionine forms an S/Ar interaction with F141 (M144 in apo (Figure 2.3a) and M145 in halo). It has been speculated previously that oxidation at M144 and M145 might destabilize the protein by altering an important, nearby hydrogen bond between Y138 and E82 that connects the short C-terminal helix to the central linker[24,29]. No specific molecular mechanism for this effect has been proposed or tested.

Based on our results above, one possibility is that oxidation of M144 and/or M145 increases the likelihood of an S^{Ox}/Ar interaction with Y138, which would be manifest in a spatial rearrangement within the C-terminal domain that should be observable by EPR. As M144/145 and Y138 are on opposite ends of a short helix, such an interaction would require helical unwinding, an effect that has been observed using NMR[30]. Replica exchange molecular dynamics simulations (REMD) of the isolated, Ca^{2+}-free (apo) C-terminal peptide fragment (residues 136-146) of CaM were used to guide the design and interpretation of experimental EPR measurements on full-length protein. We simulated both the unoxidized and doubly oxidized (at M144 and M145) forms. Because of the expectation of helical unwinding, a computational investigation of full-length CaM is hampered by limitations in sampling efficiency. Thus, while the peptide fragment is an abstraction of the real system, it is nevertheless a useful tool because it accelerates the sampling of relevant conformational space. Furthermore, the simulations were useful in 1) establishing V136 and T146 as appropriate sites for spin-labels to test the impact of

oxidation on the full-length protein; and 2) providing molecular scale details unavailable from EPR that correlate distance changes to the potential presence of an S^{Ox}/Ar motif.

In the simulation of the unoxidized peptide, an S/Ar interaction formed between Y138/M144 or Y138/M145 only 3.3% of the time. On the other hand, an S^{Ox}/Ar interaction formed 22.7% of the time when both M144 and M145 were oxidized (Figure 2.3b, Supplementary Figure 210). In only 1.2% of these interactions did a hydrogen-bond form (Supplementary Figure 2.11). As can be seen in Figure 2.3b, the S^{Ox}/Ar interaction with Y138 leads to a spatial rearrangement that brings together residues V136 and T146. The distribution of the distance between the Cα atoms of those two residues is shown in Figure 2.3c. Likewise, the frequency of the native interaction between M144 and F141 is also increased upon oxidation (49% *vs.* 12% of frames, Supplementary Figure 2.10). Thus, oxidation increased the total percentage of simulated frames in which an S/Ar interaction occurs by 3.7-fold (72% *vs.* 15%). Based on Boltzmann's law, this change reflects a roughly 1.6 kcal/mol increase in the free energy of the interaction upon oxidation. This value compares well to the results of our peptide experiment (0.62 kcal/mol) and supports the underlying conclusion of increased interaction strength upon oxidation of the S/Ar motif, even absent hydrogen bonds. It is important to note that the simulations of the short, unoxidized peptide fragment sample unfolded conformational states that are not represented by the crystal structure (where residues 138-146 are helical) and may not be accessible in a simulation of the full-length protein. Nonetheless, the emergence of a second population of states containing the S^{Ox}/Ar motif (with Y138)—only in the oxidized simulation—motivated experiments to explore whether this subpopulation exists in the full-length protein when oxidized.

To test the effects of oxidation of CaM's C-terminal helix in the full-length protein, we used dipolar EPR spectroscopy to detect intramolecular structural perturbations at submicromolar [Ca2+]. The intramolecular distance distributions detected by EPR are shown in Figure 2.3d. As shown in Supplementary Figure 2.12, the EPR spectrum of the unoxidized protein was best fit to a single broad Gaussian distance distribution, centered at 13.7 Å. After methionine oxidation, the data were best fit to a two-population model in which one population is centered at a similar distance to that observed in the absence of oxidation (14.0 Å), and there is a new population at shorter distance (8.9 Å). These experimental EPR results confirm key predictions from MD simulations (Figure 2.3c): the long-distance population becomes narrower, and a new well-ordered short distance emerges. It is not surprising that the EPR distance distributions do not match precisely those predicted by MD simulation, since the MD simulations report distances between Cα atoms, while EPR distances are between nitroxide groups at the ends of flexible side chains. EPR spectra do not directly detect Cα-Cα distances, but changes in such distances have been shown to correlate well with EPR data[31]. The convergence of the simulated and experimental results strongly suggests that upon oxidation of CaM a new S^{Ox}/Ar interaction forms between either M144/Y138 or M145/Y138. Work still remains to definitively show that this minor component in the overall accessible conformational space (Figures 2.3c and 2.3d) is responsible for disruption of the Y138/E82 hydrogen bond and alters the protein's function.

2.3.5 Oxidation of LTα Prevents Ligand/Receptor Binding

We investigated the effects of oxidation on the bioactivity and binding of LTα and TNF in live cells. We have previously shown that the interaction between LTα and TNFR1 is stabilized by a methionine-aromatic interaction via M120 of LTα and W107 of TNFR1.

Mutation of M120 to alanine interrupts ligand binding, resulting in a >10 fold loss in ligand function[1]. Thus, we hypothesized that the increased strength of the S/Ar interaction upon oxidation should stabilize binding via the M120/W107 interaction. Tumor necrosis factor (TNF) is structurally homologous to LTα and similarly activates TNFR1, but notably lacks methionine residues. Therefore, we expected its action to be unaffected by oxidative stress. Western blot analysis (Figure 2.4a) showed that untreated LTα and TNF efficiently induced downstream signaling, as has been definitively established. Surprisingly, LTα pretreated with H_2O_2 failed to induce IκBα degradation.

In order to isolate the impact of oxidation of M120 on LTα function and binding, we mutated the methionines at residues 20 and 133 to remove their susceptibility to oxidization (see Methods for details). This mutant ligand—denoted here as M120 or $M^{Ox}120$ when oxidized—triggered IκBα degradation in the unoxidized form to the same extent as wild-type (indicating that two other mutated methionine residues, M20 and M133, are not important for activity). Site-specific oxidation ($M^{Ox}120$) rendered LTα inactive (Figure 2.4a). Co-immunoprecipitation experiments explain this loss in function as a result of lost binding of the $M^{Ox}120$ ligand to TNFR1 (Figure 2.4b). Thus, we conclude that oxidation of LTα inhibits its activity by disrupting the critical M120-W107 methionine-aromatic interaction, thereby preventing binding to TNFR1.

To explore the molecular basis for this loss in binding, and to explain this apparent contradiction (decreased binding despite increased S/Ar interaction strength), we used molecular dynamics simulations and quantum calculations. Analysis of the ligand/receptor crystal structure (1TNR) revealed aromatic residues in the ligand (Y96 and Y122) that are near to and on the same chain as M120 (Figure 2.5a). Given their spatial proximity, it is likely that M120 forms S/Ar contacts with either or both of these tyrosines. MD simulations of the unbound ligand showed this to be the case in approximately 70% of the simulated

frames (the M120-Y122 interaction was approximately twice as frequent as the M120-Y96 interaction). Oxidation of M120 increased the interaction strength of M120 with Y96—again using the Boltzmann equation we calculate an increased affinity of 0.35 kcal/mol. This value is lower than that described above for CaM, but once again is consistent with our general conclusion that oxidation strengthens S/Ar interactions. Interestingly, there was no oxidation-induced change in the frequency of M120/Y122 contacts. These data raise the possibility that Y122 and Y96 interact differently with M120 under oxidative conditions. We confirmed that hydrogen bonding was not a major factor in the interaction (Supplementary Figure 2.13). We again used the simulations to generate an ensemble of configurations for quantum energy calculations. A snapshot from the oxidized ensemble is shown in Figure 2.5b. As was the case with our calculations on the non-hydrogen-bonded ensemble of DMS/DMSO-phenol configurations, oxidation increased the strength of the S/Ar interaction by 1-2 kcal/mol (Figure 2.5c).

We hypothesized that the strengthened $M^{Ox}120$/Y96 interaction could lock the ligand in a configuration that prevents $M^{Ox}120$/W107 interaction, thereby interrupting binding. To test this, we first needed to show that the addition of a sulfonyl oxygen to M120 in the crystal structure binding configuration does not itself destabilize the binding pocket by steric overlap or electrostatic repulsion. We first simulated the unoxidized crystal structure of the ligand-receptor complex. Figure 2.5d highlights the stability of the M120/W107 interaction, and shows the infrequent instances of M120 interaction with either Y122 or Y96. We then started a simulation of the ligand-receptor complex in a configuration in which the $M^{Ox}120$/Y122 pair was pre-formed (using a configuration generated from the unbound ligand simulation). Y96 was allowed to remain in its receptor backbone binding position. The binding cavity organization reverted to that of the crystal

structure configuration after 50ns (Figure 2.5d) and was stable throughout the remainder of the simulation (Supplementary Figure 2.14). Therefore, we conclude that loss of ligand-receptor binding is not caused by intrinsic steric clashes with $M^{Ox}120$ that disrupt the binding pocket. We also conclude that $M^{Ox}120$/Y122 interactions do not prevent $M^{Ox}120$/W107 interactions and ligand binding.

We then started a second oxidized ligand-receptor simulation, this time with a pre-formed $M^{Ox}120$-Y96 pair. We observed two distinct effects of oxidation, both of which are consistent with loss of receptor binding. In two of the three chains (of the symmetric ligand trimer), $M^{Ox}120$ remained stably bound to Y96, preventing it from forming the critical S/Ar motif with W107 (Figure 2.5d) and also preventing Y96 from binding the receptor backbone. In these cases, the $M^{Ox}120$/Y96 pairing (Figure 2.5e), along with the surrounding binding pocket, remained stable throughout the simulation. In the remaining chain, Y96 released from $M^{Ox}120$, and in this process disrupted the binding pocket while ejecting the receptor. Collectively, the results suggest that increased stability of the $M^{Ox}120$/Y96 pair prevents the critical M120/W107 interaction and destabilizes the bound state.

2.4 Conclusion

Oxidative stress plays a prominent role in a number of normal and pathological biological functions. However, the precise chemical and physical mechanism through which oxidative modification of proteins influences their structures and functions is largely unknown. We have shown that oxidation of methionine strengthens the methionine-aromatic interaction motif, independent of hydrogen bonding, by at least 0.5 kcal/mol and up to 1.5 kcal/mol. In the vast majority of instances in the PDB and CSD as well as the two proteins studied here in detail, the aromatic group does not form a stable hydrogen

bond with the sulfoxide due to solvation of both the donor and acceptor by water. Hydrogen bonding may further enhance the interaction in hydrophobic and aprotic environments, however we do not observe that in our two test cases where the motif is solvent exposed. In LTα-TNFR1 and CaM, we have shown that competing interactions between methionine sulfoxide and nearby aromatic residues contributes to the modulation of protein structure and function.

To conclude, non-covalent interactions are the cornerstone of biological molecular recognition events. For bioactive small-molecule development, which exploits these interactions, sulfur is the third most commonly incorporated heteroatom in pharmaceuticals next to nitrogen and oxygen.[32] The sulfoxide-aromatic interaction provides a new handle for tuning affinities of small molecules for chemical probe and therapeutic development as well as altering protein function. Moreover, the dynamic nature of methionine oxidation provides a reversible switch that can be employed for introducing responsive molecules to changes in redox environment. The energy range of these altered interactions corresponds to twofold to tenfold change in the equilibrium constant, which together with the ambivalent character of sulfoxide functional groups, the dynamic nature of their formation, and their prevalence in biology and pharmaceutics, offers medicinal chemists and chemical biology a useful tool to probe biological systems.

Table 2.1: Interaction energies for DMS or DMSO with benzene, phenol and indole. Results in kcal/mol are computed for fully optimized structures and for conformational ensembles from MD simulations in EtOAc at the MP2/6-311+G(d,p) level and corrected for basis set superposition error. The corresponding structures are reported in Supplementary Figure 2.2 and uncorrected interaction energies and M06-2X data are shown in Supplementary Table 2.1.

		Benzene	Phenol	Indole
Full optimization	DMS	-2.4	-4.9	-4.7
	DMSO	-3.3	-10.6	-10.0
Molecular dynamics simulations	DMS	-0.9	-1.0	-1.3
	DMSO Non H-bonded	-1.4	-2.2	-2.7
	DMSO H-bonded		-9.0	-7.5

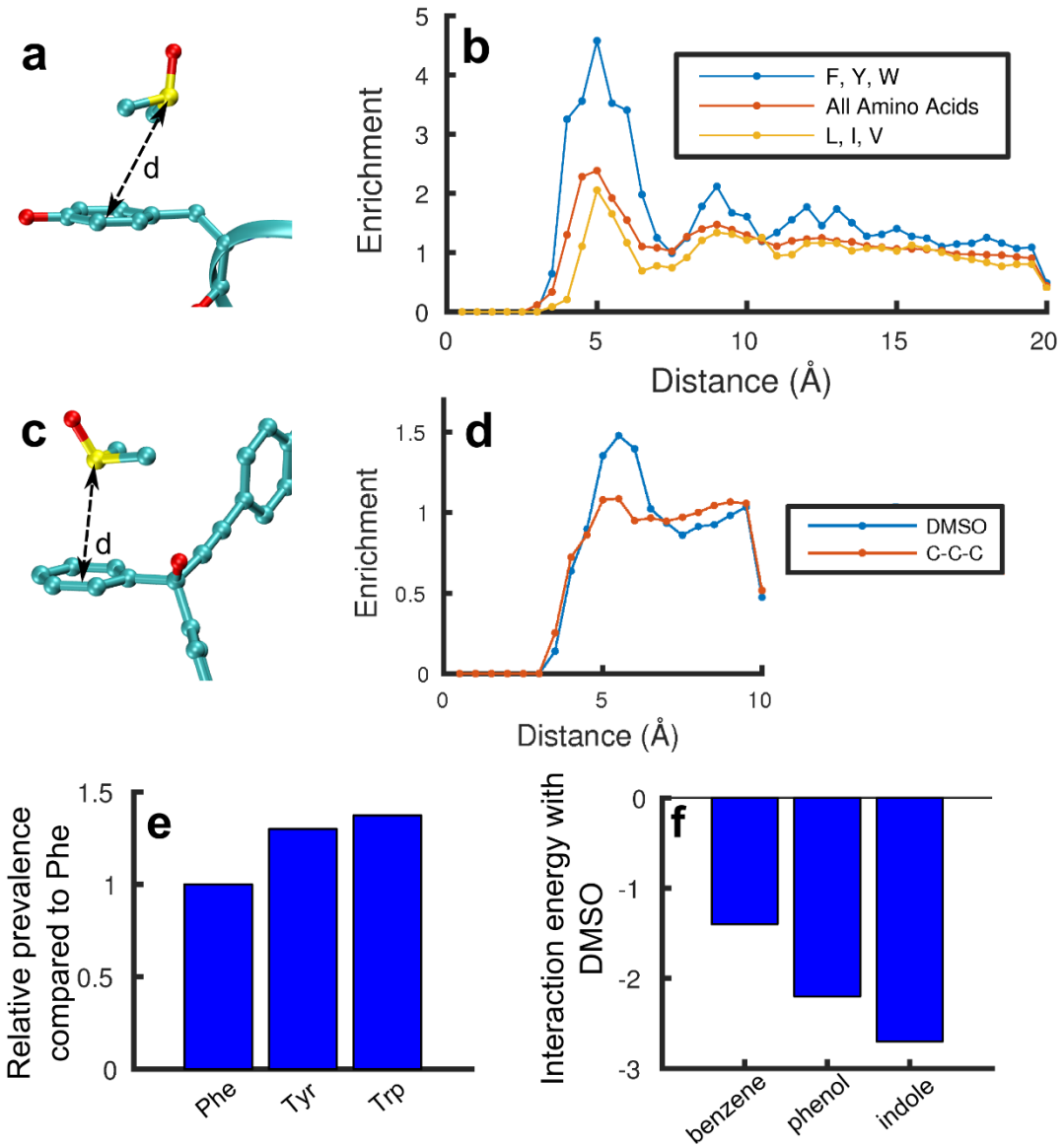


Figure 2.1. Structural informatics search of the CSD and PDB and interaction energy. A representative snapshot of a DMSO-aromatic interaction taken from the PDB (PDB: 4KAD) (a). The radial distributions of all amino acids and aromatic amino acids is plotted relative to DMSO, showing a clear enrichment of aromatic residues 4-7 Å from DMSO (b). We compare to the radial density of all amino acids and the aliphatic amino acids, Leu, Ile, and Val, relative to DMSO. A representative snapshot of a DMSO-aromatic interaction taken from the CSD (CSD: AWUHEF) (c). The radial density of aromatic groups relative to DMSO sulfur shows an enrichment around 4-7 Å (d). We compare to the radial density of aromatic groups relative to C-CH_2-C motifs, which are non-interacting. Frequency of DMSO interactions with each aromatic amino acid are plotted. These values are normalized against the abundance of each aromatic group in the PDB subset analyzed and shows enrichment of interactions with tyrosine and tryptophan compared to DMSO-phenylalanine interactions (a). Interaction energies calculated at the MP2/6-311+G(d,p) with CP-corrections for the complexes between benzene, phenol, and indole and DMSO. The structures exclude hydrogen bonding and are calculated using ensembles of configurations from the simulations as described (b).

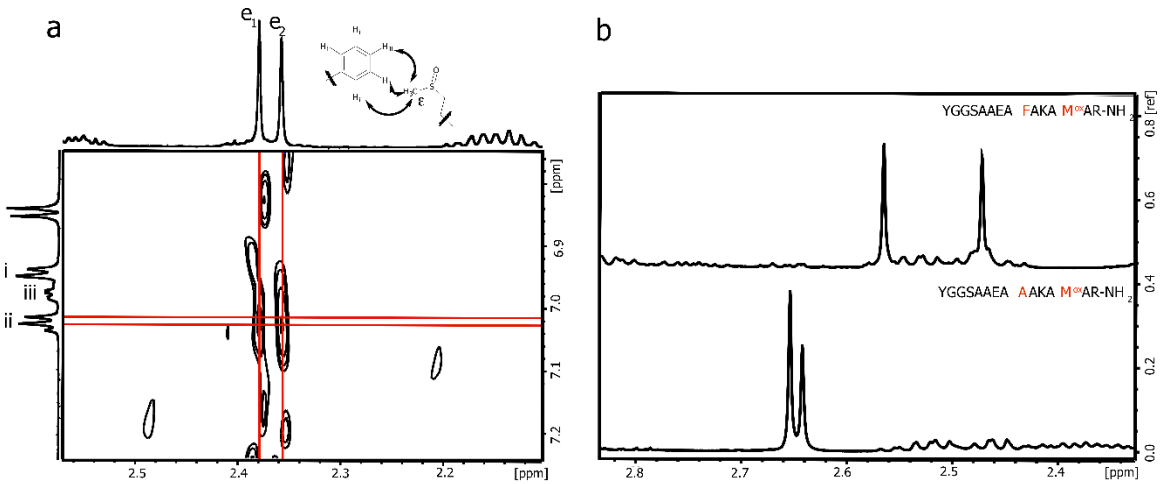


Figure 2.2. ^{1}H-^{1}H ROESY NMR and chemical shift analysis of α-helical peptide mixtures of two diasteriomers containing the *R* and *S* sulfoxide of methionine. ^{1}H-^{1}H ROESY NMR correlating NOEs from the two diastereomeric ε-methyls of oxomethionine to the aromatic resonances of phenylalanine. The strongest NOEs correlating with the aryl protons of phenylalanine Hb, are indicated by thicker arrows (a). Comparative analysis of F-Met^{Ox} and A-Met^{Ox} ^{1}H NMR experiments indicating an upfield shift of the ε-methyl resonances of oxomethionine when an aromatic ring is present relative to a methyl group (b).

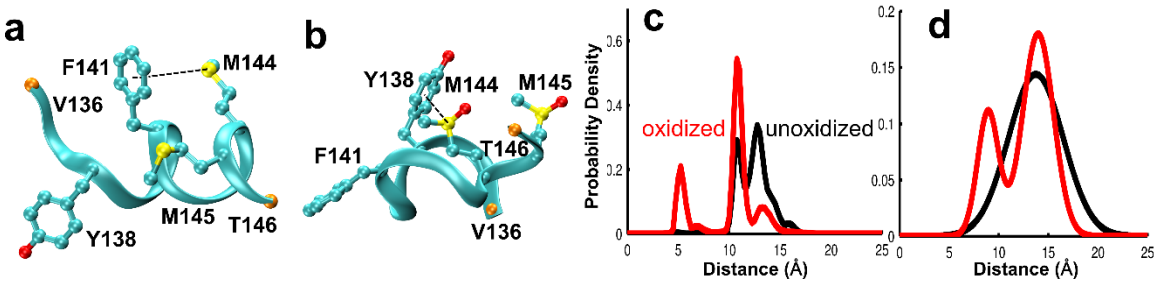


Figure 2.3. REMD and EPR measurements of CaM. Structure of unoxidized apo-CaM fragment (residue 136-146, PDB:1CFD) where F141 interacts with M144 (a) compared to a representative structure of the Y138 aromatic interacting with oxidized M144 (b) from REMD. Distance distribution calculated between Cα atoms of T136/V146 from REMD (c) and best-fit models of the distance distribution from dipolar EPR spectroscopy at submicromolar [Ca^{2+}] with maleimide spin labels at Cys residues 136 and 146 (d).

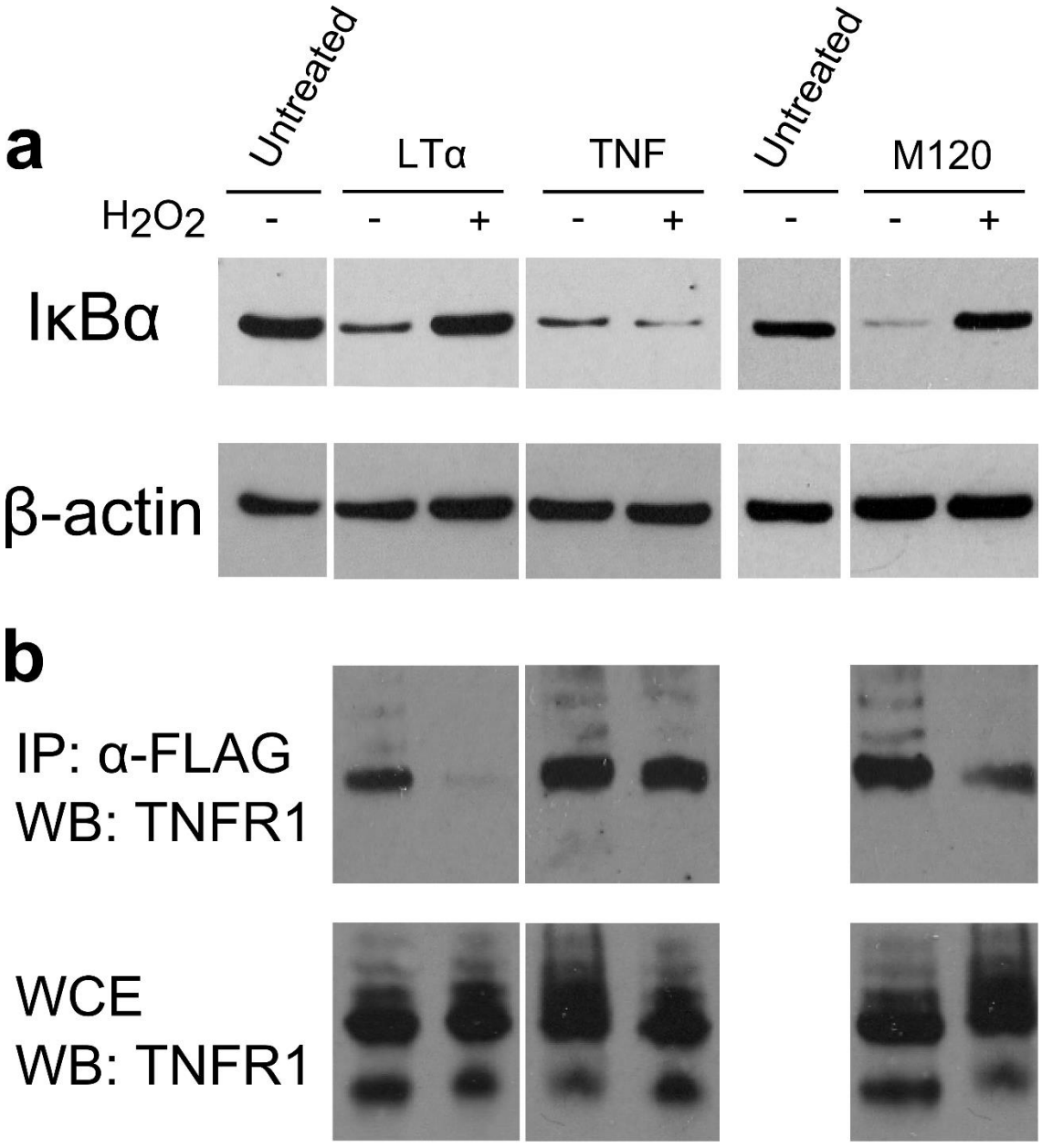


Figure 2.4. Oxidation of LTα, but not TNF ablates its interaction with TNFR1. Western blot analysis of IκBα degradation in response to treatment with oxidized (+) and unoxidized (-) LTα and TNF (a). LTα bioactivity is ablated after exposure to oxidative stress. This effect is not observed in TNF, where oxidized and unoxidized ligand trigger IκBα degradation to the same extent. Exposure to oxidative stress similarly prevented M120 LTα from triggering IκBα degradation. Co-immunoprecipitation of TNFR1 with oxidized (+) and unoxidized (-) ligands (b). LTα efficiently pulls down TNFR1 when untreated. Treatment of LTα under oxidative stress causes dramatic reduction in the amount of receptor bound. TNF is again unaffected by oxidative stress. The LTα M120 mutant again behaves the same as wild-type, indicating that oxidation of M120 is responsible for the loss of binding. The Western Blots shown are representative of at least 2 independent experiments. Full gels are shown in Supplementary Figure 2.16.

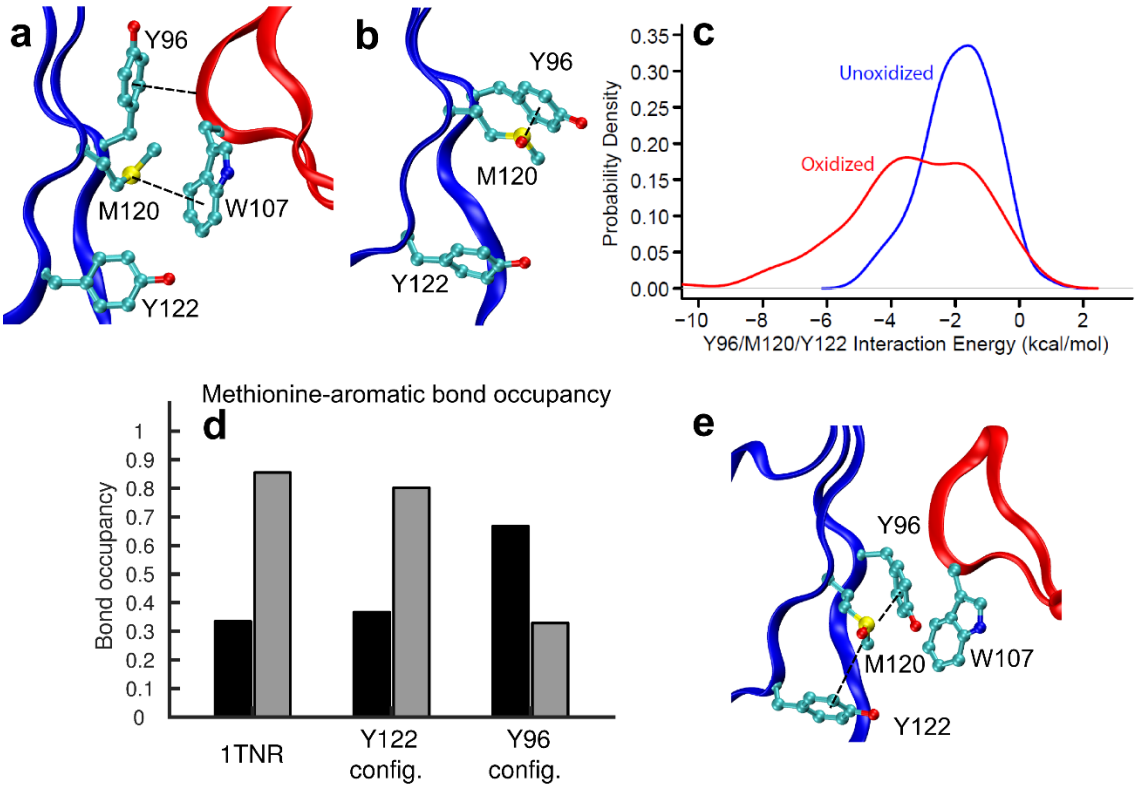


Figure 2.5. Molecular dynamics simulations of LTα with M120 oxidized show that competitive interaction of M120 with Y96/Y122 prevents its interaction with W107. In the receptor bound state, Y96 is folded upward so that it interacts with the backbone of the receptor, allowing M120 to form a stable methionine-aromatic interaction with W107 (a). When Met120 is oxidized, it draws Y96 downward so that it protrudes into the binding pocket (b). The histogram of the interaction energy between M120 and Y96/Y122 calculated from 128 structures taken from the MD simulations shows enhanced interaction due to oxidation. The histogram was smoothed with the Gaussian kernel-smoothing function “density” of the program “R” (c). The bond occupancy of the M120-Y96/Y122 (black) and M120-W107 (gray) interactions is shown for 3 systems: the unmodified crystal structure (1TNR), the M120-Y122 interacting configuration, and the M120-Y96 interacting configuration (d). M120 interacts stably with W107 in 1TNR. In the M120-Y122 interacting system, M120 interacts somewhat more with Y96/Y122, but still primarily with W107. In the M120-Y96 interacting system, the bond occupancy is reversed - that is, M120 interacts primarily with Y96/Y122 and the W107 interaction is blocked. A snapshot of the Y96 interacting system with Y96 blocking the M120-W107 interaction is shown (e).

2.5 Supplementary Information

Supplementary Table 2.1: Interaction energies for DMS and DMSO with benzene, phenol, and indole. Computed interaction energies in kcal/mol for the complexes of benzene, phenol, and indole with DMS and DMSO (structures are shown in Supplementary Figure 2.1). M06-2X and MP2 data are reported for structures fully optimized at the corresponding levels of theory with the 6-311+G(d,p) basis set. Hartree-Fock data are reported for benzene only as single point energy calculations on the MP2 optimized structures. Interaction energies corrected for basis set superposition error are reported in parenthesis.

	Benzene		Phenol		Indole	
Complex	M06-2X	MP2	M06-2X	MP2	M06-2X	MP2
DMS	-4.6 (-4.0)	-6.0 (-2.4)	-8.3 (-7.3)	-9.3 (-4.9)	-7.3 (-6.3)	-9.5 (-4.7)
DMSO	-6.2 (-4.9)	-7.3 (-3.3)	-16.0 (-14.4)	-16.0 (-10.6)	-13.6 (-12.1)	-15.5 (-10.0)

Supplementary Table 2.2: Dispersion energy interactions. Dispersion energy interactions are calculated in kcal/mol for the complexes of benzene with DMS and DMSO as the difference between the MP2/6-311G(g,p) interaction energies and the HF/6-311+G(d,p) calculated as single point on the MP2 minima.

	MP2	HF	Dispersion Energy
DMS	-6.0	2.4	-8.4
DSMO	-7.3	0.9	-8.2

Supplementary Table 2.3: NBO Charges analysis of the benzene complexes with DMS and DMSO, and of the isolated species based on the MP2/6-311+G(d,p) electron density.

	DMSO/Benzene	DMS/Benzene	DMSO	DMS	Benzene
O	-0.96		-0.96		
S	1.19	0.19	1.18	0.20	
C	-0.77	-0.70	-0.76	-0.70	
H	0.22	0.21	0.22	0.21	
H	0.22	0.21	0.21	0.19	
H	0.21	0.19	0.21	0.19	
C	-0.76	-0.70	-0.76	-0.70	
H	0.22	0.21	0.22	0.21	
H	0.21	0.19	0.21	0.19	
H	0.21	0.19	0.21	0.19	
C	-0.20	-0.20			-0.20
C	-0.21	-0.21			-0.20
C	-0.20	-0.19			-0.20
C	-0.20	-0.21			-0.20
C	-0.21	-0.20			-0.20
C	-0.20	-0.20			-0.20
H	0.21	0.20			0.20
H	0.20	0.20			0.20
H	0.21	0.20			0.20
H	0.21	0.20			0.20
H	0.20	0.20			0.20
H	0.21	0.20			0.20

Supplementary Table 2.4. All peptides used in CD thermal studies, their appropriate abbreviation and their sequences. Only residues 9 and 13 were varied, and these positions are shown in bold.

Peptide	***Abbreviation***	***Construct Sequence***
Phe-Met	FM	YGGSAAEA-**F**-AKA-**M**-AR-NH_2
Phe-Met (ox)	$FM_{(ox)}$	YGGSAAEA-**F**-AKA-**M (ox)**-AR-NH_2
Phe-Ala	FA	YGGSAAEA-**F**-AKA-**A**-AR-NH_2
Ala-Met	AM	YGGSAAEA-**A**-AKA-**M**-AR-NH_2
Ala-Met (ox)	$AM_{(ox)}$	YGGSAAEA-**A**-AKA-**M (ox)**-AR-NH_2
Ala-Ala	AA	YGGSAAEA-**A**-AKA-**A**-AR-NH_2

Supplementary Table 2.5. Principal Component Analysis. The details of PCA for each peptide are shown as standard deviations, proportion of variance, and cumulative proportion of variance.

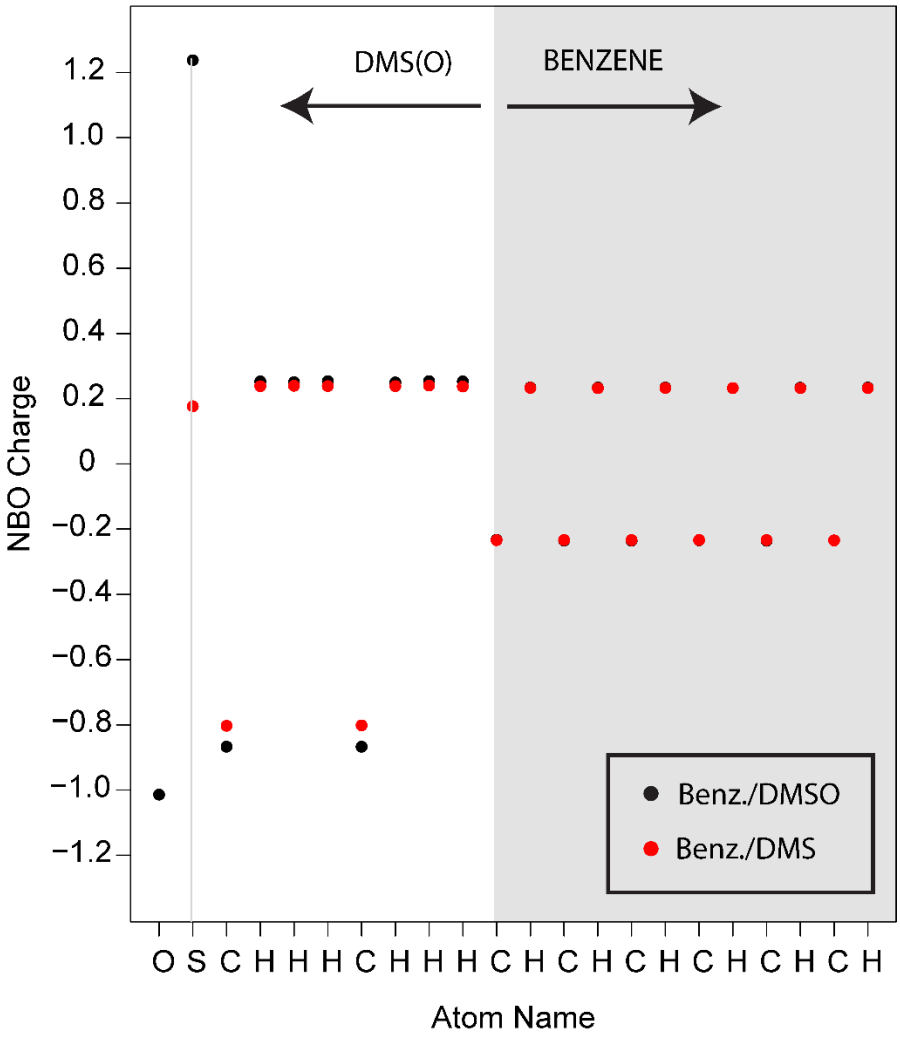
FM	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
σ	29.55	19.16	6.30	1.99	1.40	0.85	0.62	0.59	0.37	0.27	0.00
Var Prop.	0.68	0.29	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cum. Prop.	0.68	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
FMx	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
σ	17.12	15.89	5.03	1.62	1.46	0.93	0.67	0.55	0.36	0.32	0.00
Var Prop.	0.51	0.44	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cum. Prop.	0.51	0.94	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
FA	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
σ	10.00	5.69	0.66	0.46	0.36	0.29	0.23	0.21	0.18	0.14	0.00
Var Prop.	0.75	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cum. Prop.	0.75	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AM	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
σ	24.56	23.44	1.30	0.87	0.59	0.52	0.42	0.31	0.23	0.16	0.00
Var Prop.	0.52	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cum. Prop.	0.52	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AMx	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
σ	22.72	15.39	0.65	0.39	0.36	0.25	0.21	0.20	0.12	0.12	0.00
Var Prop.	0.68	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cum. Prop.	0.68	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AA	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
σ	38.38	28.45	2.60	1.19	1.07	0.80	0.47	0.42	0.27	0.19	0.00
Var Prop.	0.64	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cum. Prop.	0.64	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Supplementary Table 2.6. Relaxation schedule for the $M^{Ox}120$-receptor complex simulations. Restraint spring constants are given in kcal/mol/Å^2.

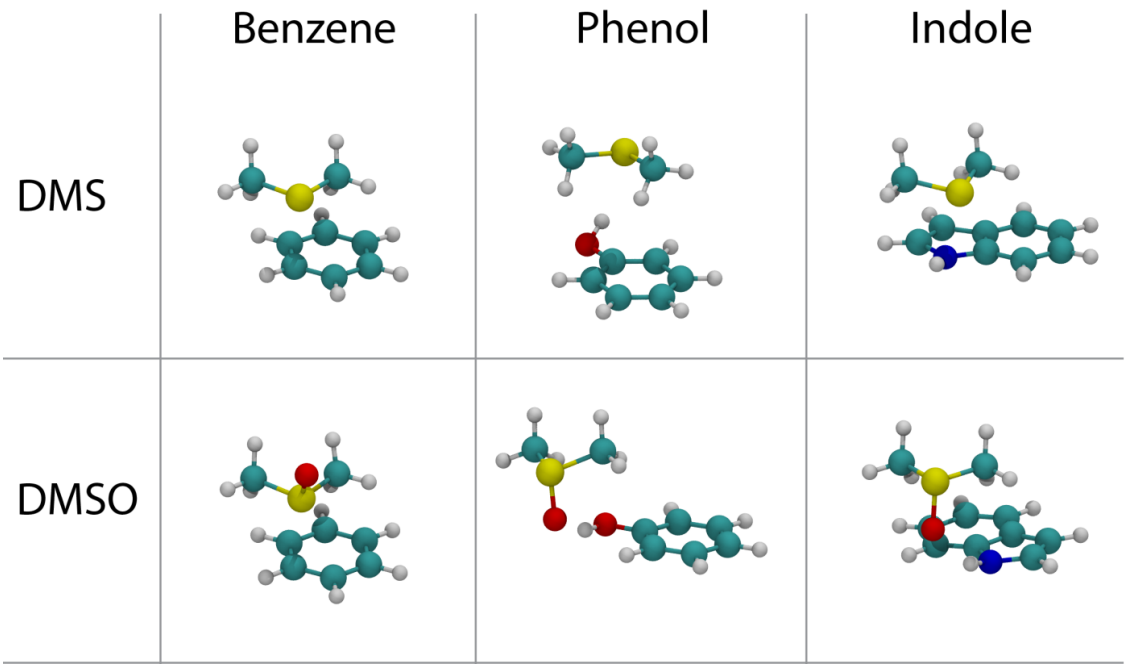
Stage	Receptor		Ligand		Steps
	Backbone	Sidechain	Backbone	Sidechain	
1	1	1	1	1	10^6
2	0.1	1	1	1	10^6
3	0.01	1	1	1	10^6
4	0	1	1	1	10^6
5	0	0.1	1	1	10^6
6	0	0.01	1	1	10^6
7	0	0.005	1	1	10^6
8	0	0	1	1	10^6
9	0	0	0.1	1	10^6
10	0	0	0.01	1	10^6
11	0	0	0	1	10^6
12	0	0	0	0.1	10^6
13	0	0	0	0.01	10^6
14	0	0	0	0.005	10^6
15	0	0	0	0	10^6

Supplementary Table 2.7: Effects of solvation on quantum mechanical interaction energy. The effect of solvation on quantum mechanical interaction energies has been tested for the complex of benzene with DMS and DMSO. The interaction energy in kcal/mol for the complexes optimized in gas phase at the M06-2X level was calculated in the gas phase and with single point PCM calculations to describe the effect of solvation in EtOAc and water. The basis set employed was 6-311+G(d,p). The ΔE_{int} entry shows that the effect of solvation is less than 0.4 kcal/mol.

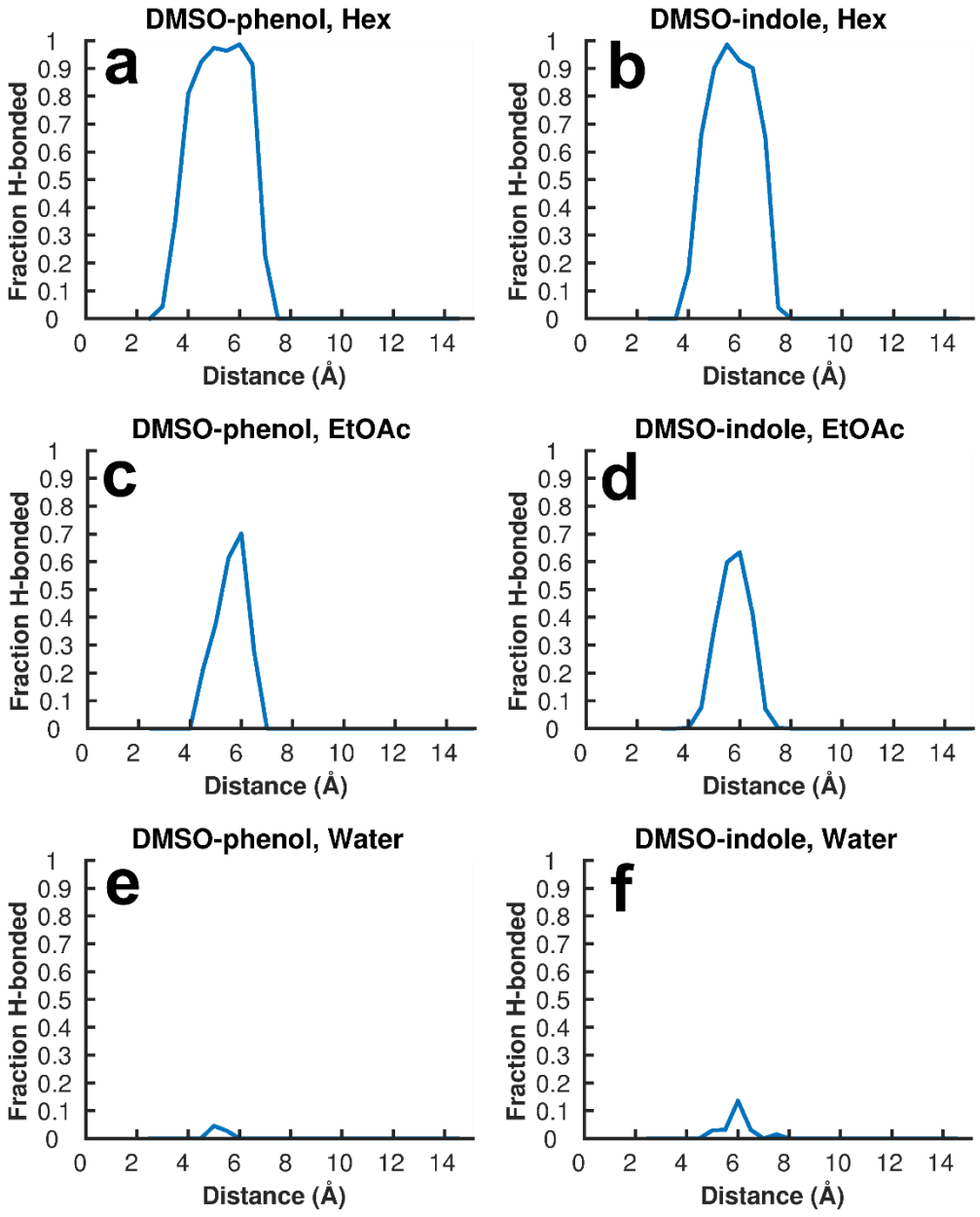
	Gas	EtOAc	Water
DMS	-4.7	-3.9	-3.6
DMSO	-6.2	-5.1	-4.7
ΔE_{int}	-1.5	-1.2	-1.1



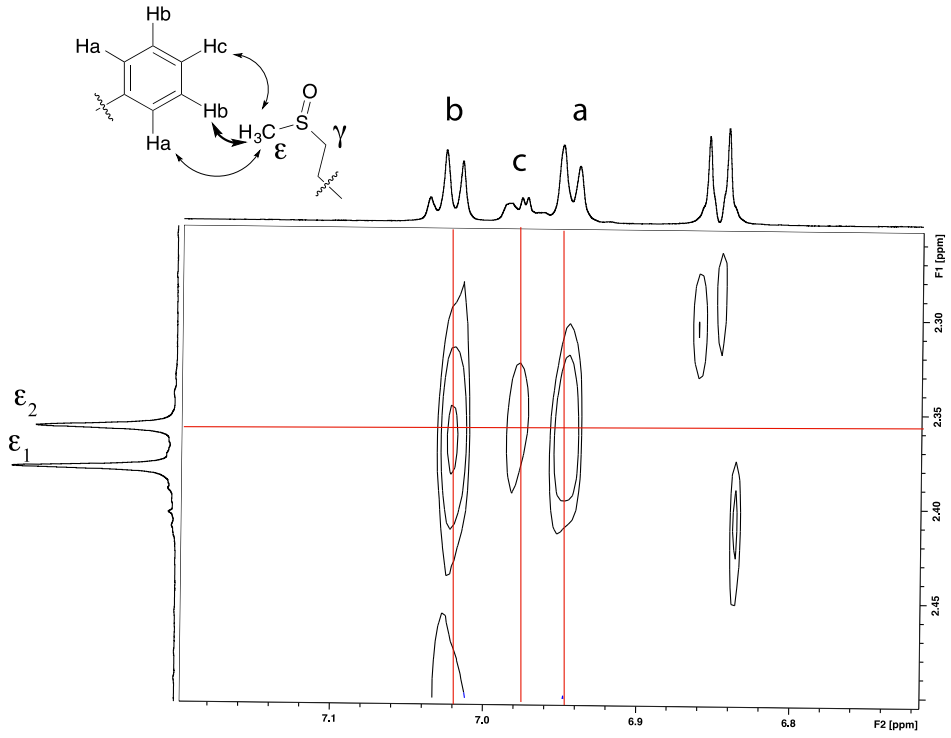
Supplementary Figure 2.1: NBO average charges for the ABF simulations. NBO charges were computed as the average for the 128 structures for the simulations of DMS/DMSO and benzene in EtOAc using the MP2/6-311+G(d,p) electron density. In red the charges for benzene/DMS are reported, and in black for benzene/DMSO.



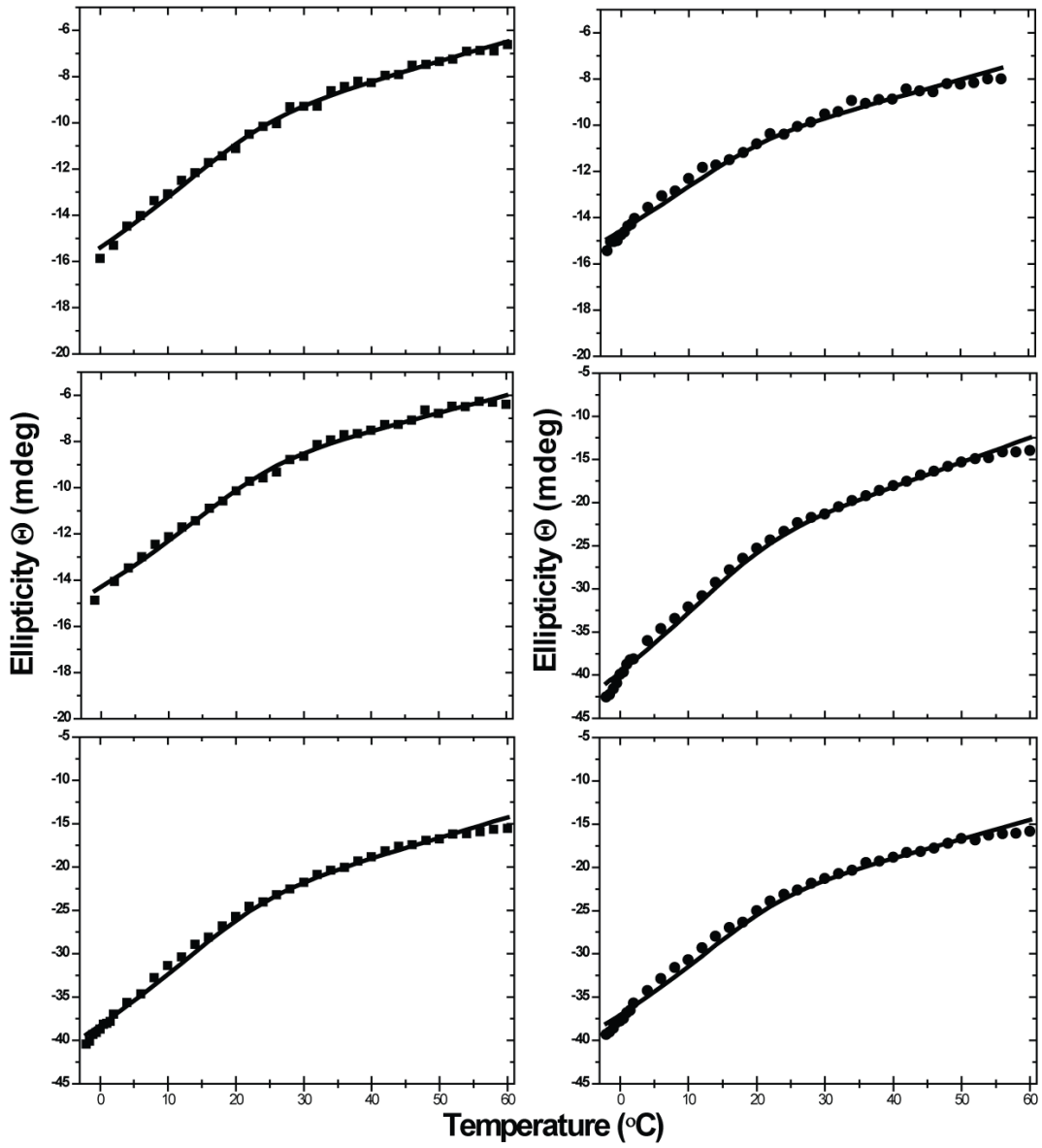
Supplementary Figure 2.2. Minimum energy structures. The minimum energy structures for the complex between DMS or DMSO and benzene, phenol, and indole, calculated from full optimizations at the MP2/6-311+G(d,p) level.



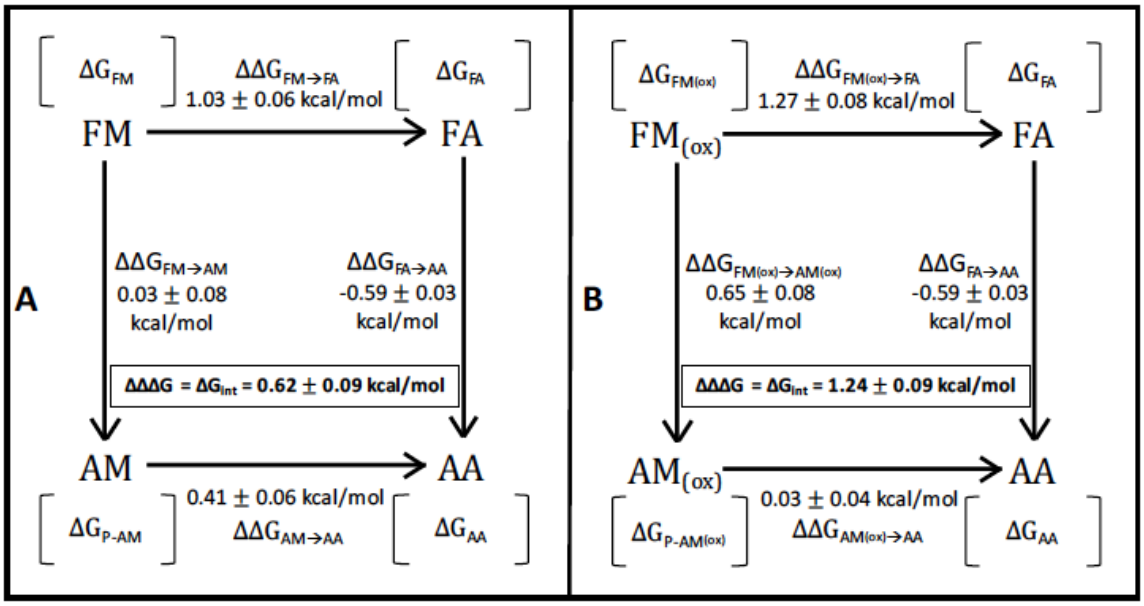
Supplementary Figure 2.3. Fraction of structures in which the DMSO sulfonyl accepts a hydrogen bond from the aromatic. Structures were extracted from the ABF simulations of DMSO and aromatic amino acid analogs in three solvents. The fraction of structures that are hydrogen bonded (H-acceptor distance < 2.5Å) along all sulfur-aromatic distances is plotted. Hydrogen bonding occurred in nearly all configurations that were in range in hexane (a, b), fewer configurations in EtOAc (the frequency of hydrogen bonding was reduced by ~60%) (c, d), and almost none of the configurations in TIP3P water (e, f). EtOAc is able to accept hydrogen bonds from the aromatic group and TIP3P is able to donate and accept hydrogen bonds, which results in the lower prevalence of bonding between the DMSO and aromatic group.



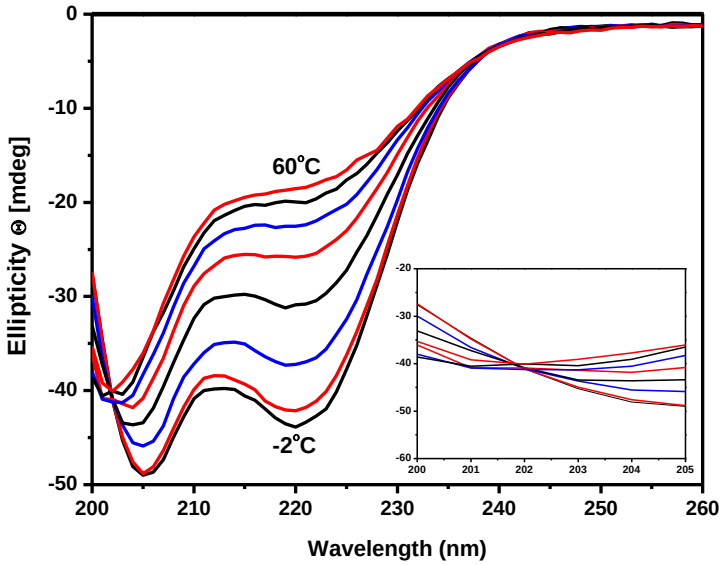
Supplementary Figure 2.4. A portion of the F-Met^{Ox} ^{1}H-^{1}H ROESY NMR spectrum taken in acidic (pH 4.0) D_2O at 4ºC with reversed axis. NOEs shown between the ε-S-methyl sulfoxide diasteriomeric resonances and the aromatic protons of phenylalanine. The strongest NOEs are observed with Hb at the 3,5 position of phenylalanine. The upfield doublet is from one of the N-terminal tyrosine resonances.



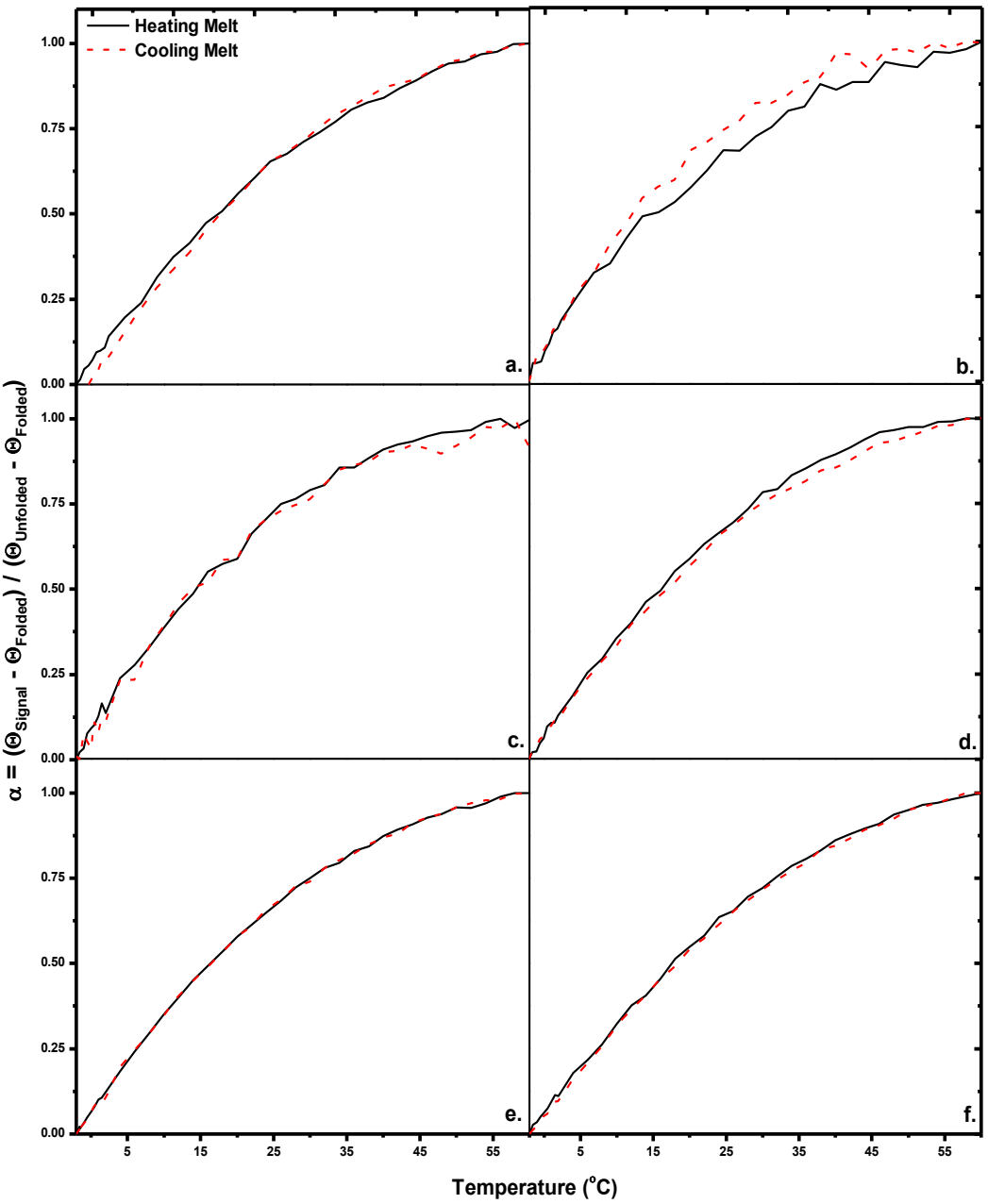
Supplementary Figure 2.5: Representative data sets and global fits for three replicas of Phe-Met (Left, solid squares) and Phe-Met^{Ox} (Right, solid circles). Signal was acquired at 222nm. For each peptide, a black solid line represents the global fit to all three datasets using the Gibbs Helmholtz equation, as described in Methods. The fitting parameters are, for Phe-Met: ΔH_{Tm} = 26.0 ± 0.7 kcal/mol, ΔC_P = 0.2 ± 0.2 kcal/mol°C and T_m = 18.3 ± 0.3 °C; and for Phe-Met^{Ox} are: ΔH_{Tm} = 27.1 ± 0.1 kcal/mol, ΔC_P = 0.21 ± 0.02 kcal/mol°C and T_m = 21 ± 1°C were determined for Phe-Met^{Ox} peptide. Raw ellipticity signal varies for each peptide sample set due to variations in each individual peptide experiment (~150μM concentration). Free energy values were extracted from the global fit at T=0°C and are presented in Supplementary Figure 2.6.



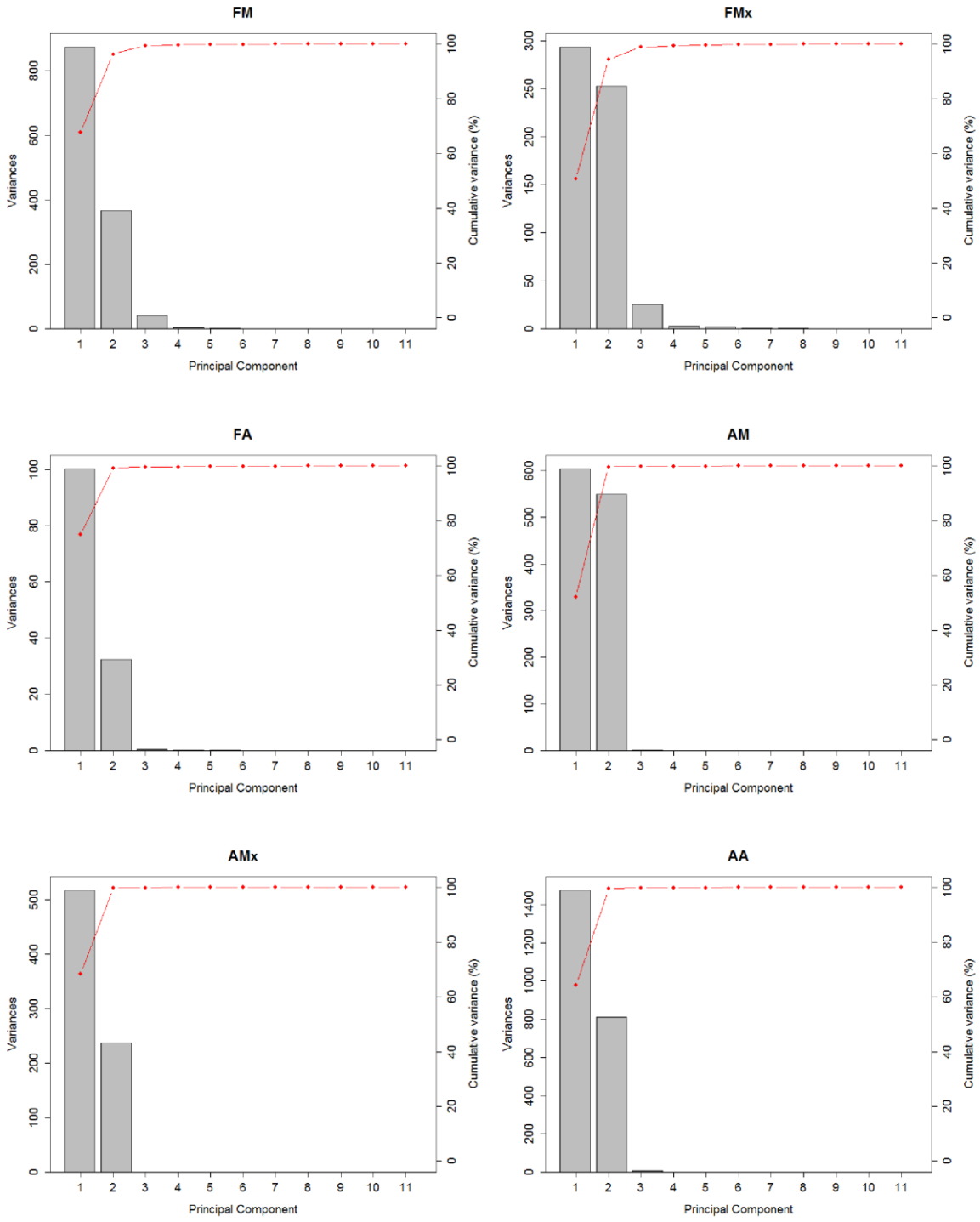
Supplementary Figure 2.6. Double-mutant cycle. Thermodynamic double mutant cycle of FM (a) and $FM_{(ox)}$ (b) peptides. Nomenclature is: Wild-type peptide (FM and $FM_{(ox)}$), singly mutated peptide (FA, AM and $AM_{(ox)}$) and doubly mutated peptide (AA). F, M and A represent phenylalanine, methionine (un-oxidized or oxidized) and alanine, respectively. Free energy of each peptide denaturation is represented as ΔG at each step in the cycle. ΔΔG values are determined by the difference in ΔG denaturation values of peptides at 0°C, which were used to determine the ΔΔΔG. For simplicity, and as described in the Methods, we assign the nomenclature ΔG_{int} to reflect the change in interaction free energy between Met/Phe and $Met_{(ox)}$/Phe (S^{Ox}/Ar- S/Ar = 1.24 - 0.62 = 0.62 kcal/mol). Thus there is a doubling of the strength of the interaction upon oxidation. Error was calculated through a total differential approach, where the derivatives of the GHE with respect to ΔH, ΔCp and Tm were calculated. The error for each thermal denaturation ΔG is the sum of the products of each derivative and their respective individual error values. Error was then calculated with ΔG values for each side of the thermodynamic cycle for ΔΔG and ΔG_{Int}.



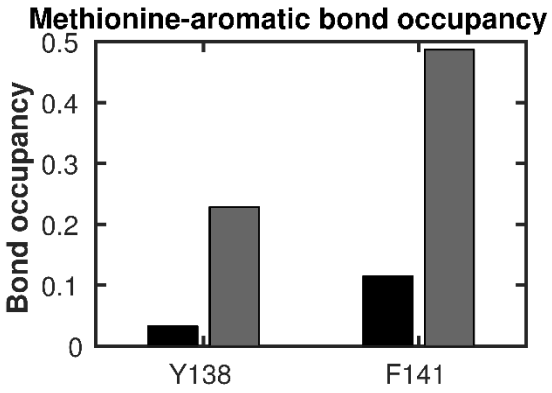
Supplementary Figure 2.7. Combined Ala-Met spectra observed from 200 to 260nm at increasing temperatures from -2°C to 60°C. Crossing of all spectra at ~202nm indicates isodichroic point seen in inset plot. This plot is representative of the all peptide constructs used in the double mutant cycle, where an isodichroic point is also observed.



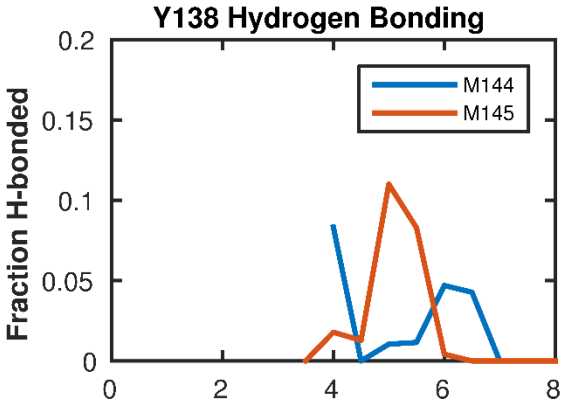
Supplementary Figure 2.8. Overlaid heating and cooling melt curves for FM, FM^{Ox}, FA, AM, AM^{Ox}, and AA peptides. Raw Elliptical signal (mdeg) at 222nm was normalized from 0-1 for each peptide construct. Heating curves are represented by a black solid line, and cooling curves are represented by a red dashed line. Ellipticity displayed as a fraction of unfolded peptide. FM heating and cooling melt (a), FM^{Ox} heating and cooling melt (b), FA heating and cooling melt (c), AM heating and cooling melt (d), AM^{Ox} heating and cooling melt (e), and AA heating and cooling melt (f).



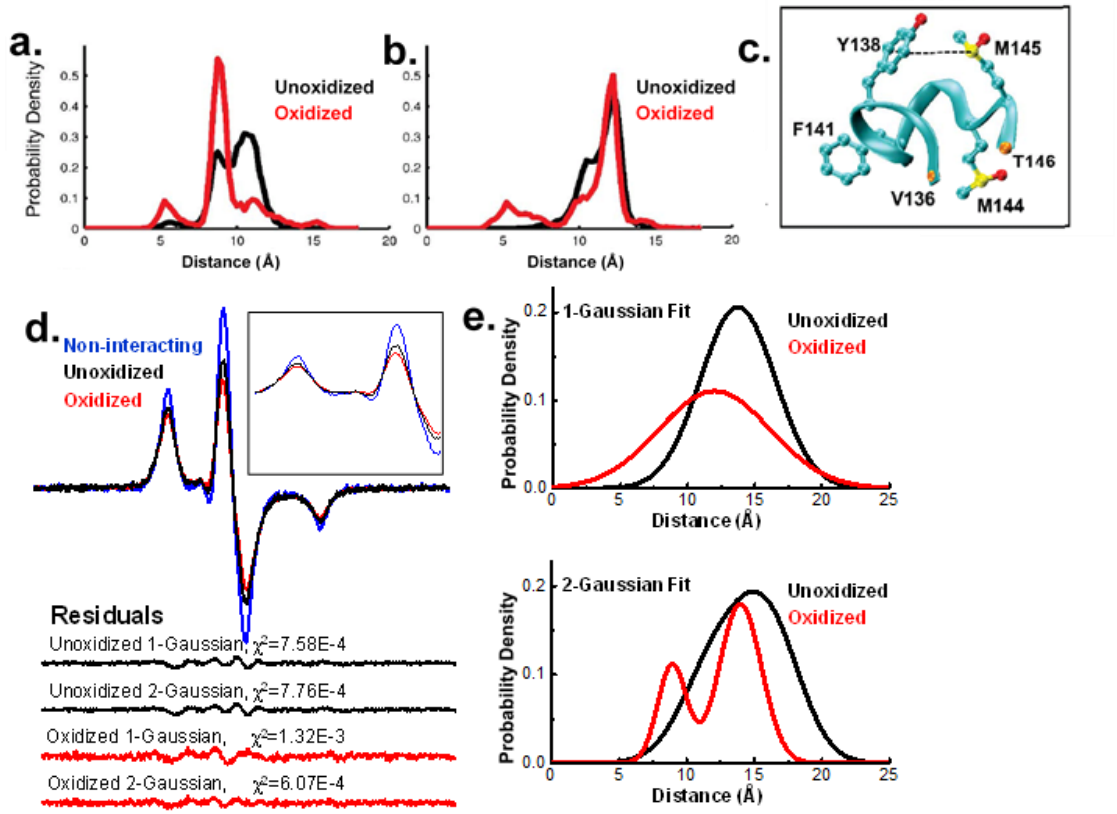
Supplementary Figure 2.9. Principal Component Analysis. The variance for all principal components is shown as gray bars, and the cumulative percent variance is shown as red lines for all peptides.



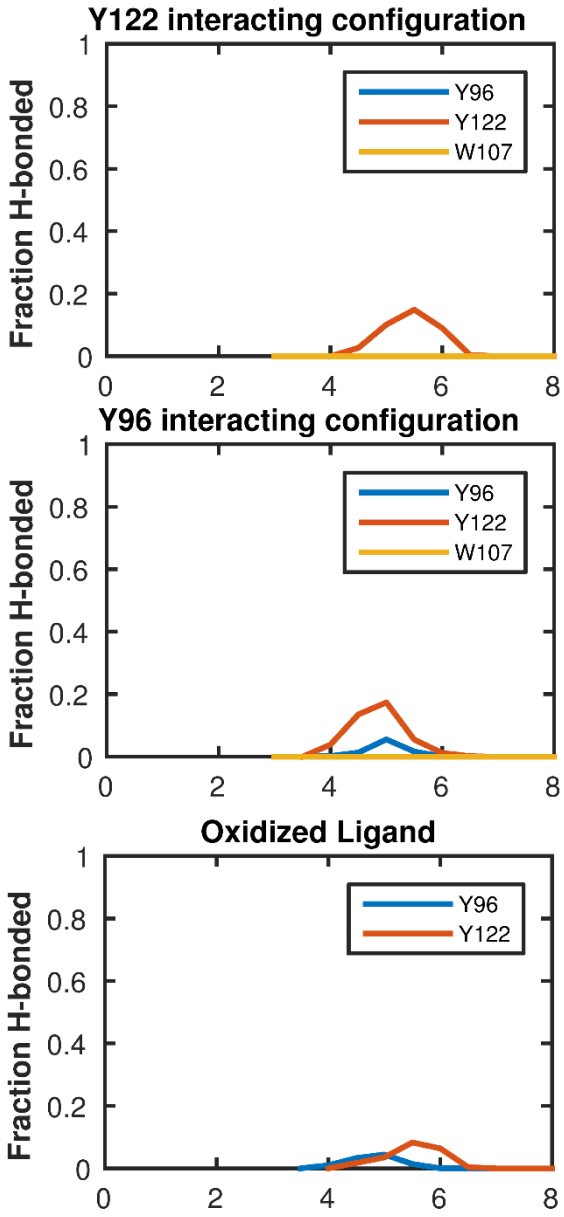
Supplementary Figure 2.10. S/Ar interactions in CaM simulations. The fractional occupancy of the unoxidized (black) and oxidized (gray) methionine aromatic interaction involving either M144 or M145 from the REMD simulation



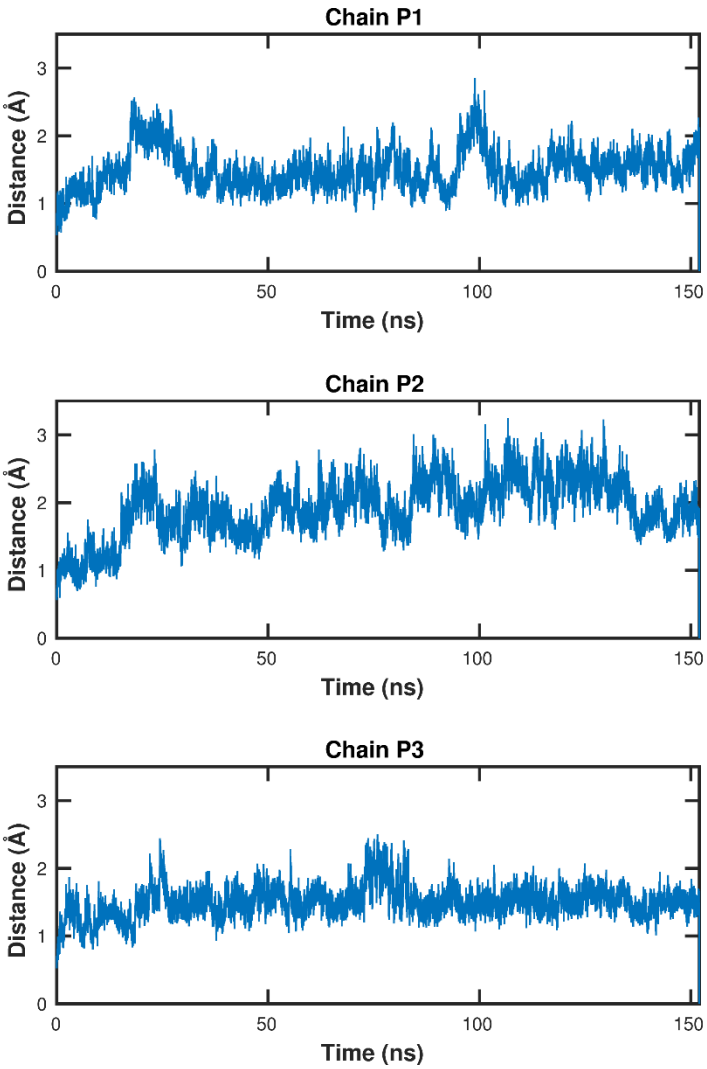
Supplementary Figure 2.11. Hydrogen bonding of $M^{Ox}144$ and $M^{Ox}145$ with Y138 in the calmodulin simulations. The histogram shows that only a small fraction of Met-aromatic interactions involve a hydrogen bond. The methionine-aromatic distances were binned, and then the fraction of structures that were hydrogen bonded in each bin was plotted.



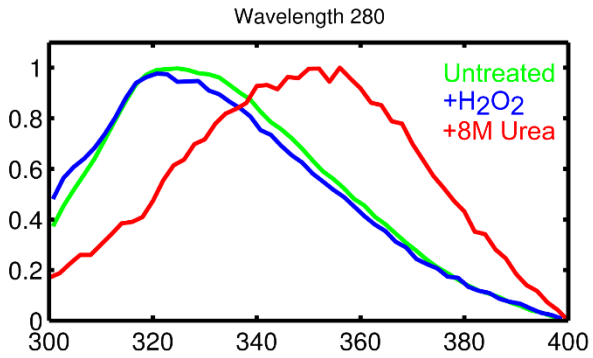
Supplementary Figure 2.12. Distribution of distances in CaM, from MD simulation, between sulfoxide and center of aromatic ring of Y138-M144 (a) and Y138-M145 (b). (c). A snapshot from MD simulation showing the interaction between Y138 and M145. (d) Representative EPR spectra (200 G scan width. Each spectrum has been normalized to the same number of spins (by dividing by the double integral). Inset shows expanded view of a 50-G portion of the low-field spectra) of CaM spin-labeled at V136 and Y138, showing residuals to fits. The "non-interacting" spectrum was obtained by summing spectra of the two singly-labeled samples, prepared from single-Cys mutants. (e) Distribution of distances from fits to EPR spectra. For unoxidized CaM, residuals show that the 2-Gaussian fit is no better than the 1-Gaussian fit, which is thus shown in Figure 2.3. For oxidized CaM, the 2-Gaussian fit is superior and is thus shown in Figure 2.3. Statistics for EPR fits (mean ± SEM, n = 4) are as follows. Unoxidized 1-component fit, central distance = 13.7 ± 1.4 Å, χ-squared = 7.65E-4 ± 1.25E-4. Unoxidized 2-component fit, χ-squared = 7.49E-4 ± 0.82E-4. Oxidized, 1-component fit, χ^2 = 14.5E-4 ± 1.67E-4. Oxidized, 2-component fit, χ^2 = 6.22E-4 ± 1.88E-4, long distance (center of distribution) = 14.0 ± 1.2 Å, short distance = 8.9 ± 0.9 Å, mole fraction of short distance population was 0.26 ± 0.06.



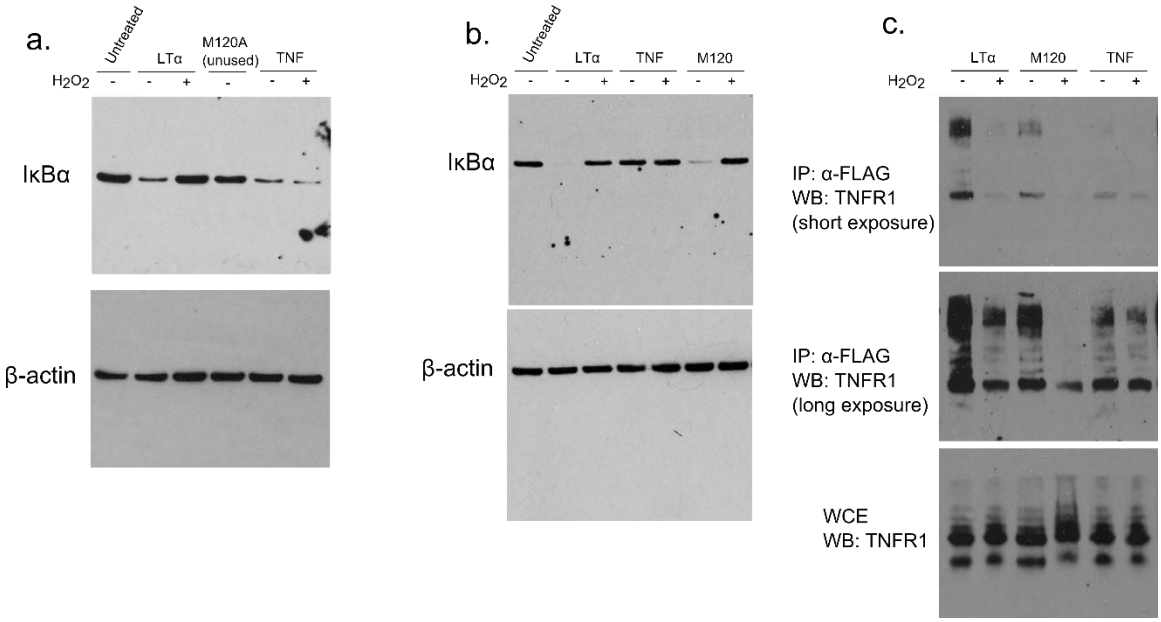
Supplementary Figure 2.13. Hydrogen bonding in the molecular dynamics simulations of LTα. The fraction of frames in which hydrogen bonding occurs for the range of methionine-aromatic distances. The methionine-aromatic distances were binned, then the fraction of structures that were hydrogen bonded in each bin were plotted. In all three simulations, the fraction of structures that are hydrogen bonded is small.



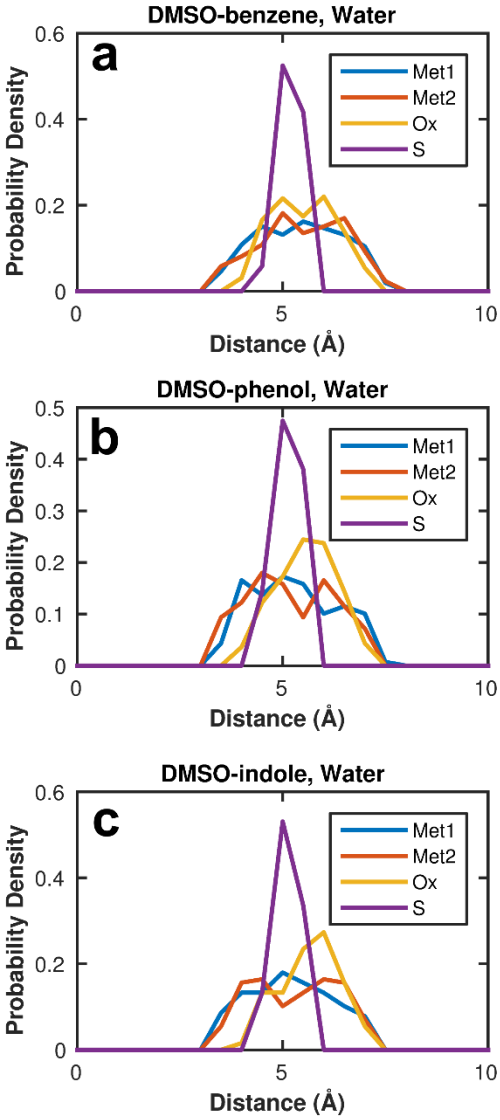
Supplementary Figure 2.14. RMSD of the binding pocket for the M^{Ox}120-Y122 interacting configuration. The RMSD is shown for the binding pocket (all ligand and receptor residues within 15 Å of M^{Ox}120) for the 3 ligand chains. All three binding pockets are stable.



Supplementary Figure 2.15. Tryptophan fluorescence of LTα after exposure to oxidative stress. LTα was left untreated or oxidized as in cell experiments, and its fluorescence emission spectrum was measured after ~18 hours. The unfolded control sample was dissolved in 8M urea for several hours before measuring its emission spectrum. Fluorescence emission wavelength scans from 300 – 400 nm were taken on a Varian Cary Eclipse Fluorescence spectrometer at an excitation wavelength of 280 nm and a scan rate of 30 nm/min. Excitation and emission slits were set to 5 nm. The emission spectra were normalized between 0 and 1. Total unfolding in 8M urea causes tryptophan residues to become exposed to polar solvent, resulting in a red-shifted emission spectrum (red). Exposure of LTα to oxidative stress does not result in red-shifting of the emission spectrum (blue) compared with the untreated LTα spectrum (green), indicating that the ligand is stably folded.



Supplementary Figure 2.16. Full gels from Western Blots and immunoprecipitation. IκBα degradation is induced by unoxidized LTα and unoxidized or oxidized TNF. Oxidation of LTα prevents it from inducing IκBα degradation. M120A is a reduced-binding mutant negative control and was not the subject of this investigation (a). A second iteration of the same experiment is shown with the M120 control (b). The full immunoprecipitation blots are shown with short and long exposure times (c).



Supplementary Figure 2.17. Histogram of distances from each DMSO atom to the center of the aromatic ring. Non-hydrogen-bonded DMSO-aromatic pairs from the ABF simulations were analyzed and show no strong bias for the methyl groups to interact with the aromatic ring. Benzene (a), phenol (b), and indole (c).

<u>Chapter 3:</u> Synaptotagmin I's Intrinsically Disordered Region Interacts with Synaptic Vesicle Lipids and Exerts Allosteric Control Over C2A

[1]Department of Chemistry and Biochemistry, University of Minnesota Duluth, Duluth, MN

[2]Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Minneapolis, MN

Funding sources: The material is based upon work supported by the National Science Foundation CAREER Award Grant Number (MCB-0845676) and Chancellor's Small Grants (UMD) (both to AH). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Funding for NMR instrumentation was provided by the Office of the Vice President for Research, the Medical School, the College of Biological Science, NIH, NSF, and the Minnesota Medical Foundation.

MEF, RWM, AMR, KD, KJL, and BH purified proteins. MEF, RWM, KJL, and BH collected and analyzed DSC data. MEF collected and analyzed NMR and CD data. AMR, KD, and SEGK collected and analyzed ITC data. MEF, RWM, AMR, KD, KJL, and AH designed experiments. MEF and AH wrote the paper.

3.1 Introduction

Synaptotagmin I (Syt I) is a vesicle-localized integral membrane protein composed of a short lumenal domain, single transmembrane helix, a cytosolic linker region followed by tandem C2 domains (termed C2A and C2B) that bind acidic phospholipid and calcium ion (Ca^{2+}) (Figure 3.1A). Syt I has been identified as the Ca^{2+} sensor for neuronal exocytosis, providing a biochemical link

between the influx of Ca^{2+} induced by propagating action potentials and the fusion of synaptic vesicle and plasma membranes that underlies neurotransmission.[59] How exactly Ca^{2+} ligation of Syt I's C2 domains leads to synchronization of the exocytotic machinery has been studied extensively, but the mechanism for coupled ligation and fusion still remains incompletely understood. We believe that insight into the underlying mechanism requires investigation of allostery in the propagation and modulation of binding signals; an understanding of how binding in one region of the Ca^{2+} sensor influences other distal locations that participate in regulatory protein-protein and/or protein-lipid interactions (and vice versa) is crucial to unraveling coupled Ca^{2+} binding and membrane fusion. Towards this goal, we have shown previously that C2A and C2B are negatively allosterically coupled to one another indicating that binding events in each C2 domain reciprocally regulate one another.[35] It is less clear, however, if the stretch of ~60 amino acids between Syt I's transmembrane helix and C2A plays an important allosteric role in function (Figure 3.1A, amino acid sequence).

Until recently, little attention was given to this juxta-membrane linker's role in neurotransmission as the adjacent C2 domains were assumed to be the primary effectors. However, in recent single vesicle docking and content mixing assays, wherein full-length Syt I was a reconstituted component, missense mutations that perturbed the charge distribution of the linker were shown to attenuate pore opening, a critical step for membrane fusion. Additionally, large deletions of the linker were found to disrupt vesicle docking.[60] Consistent with these observed *in vitro* disruptions of function, subsequent *ex vivo* studies performed on murine phrenic nerve-hemidiaphragm preparations showed that treatment of these neuromuscular junctions (NMJs) with small peptide derivatives encompassing residues 80-98 of the juxta-membrane linker caused significant inhibition of neurotransmitter release; treated nerves were unable to elicit muscle contraction when stimulated.[61] In *Drosophila melanogaster*, deletion of this juxta-membrane linker abolishes evoked release of neurotransmitter.[62] All of the above findings indicate that the linker region plays an important role in synaptotagmin biology and is worthy of additional inquiry. In this study, we probe juxta-membrane linker function in the context of allosteric regulation.

One approach to assessing allosteric regulation and function is through examination of protein folding.[63,64] When looking at the amino acid sequence of this juxta-membrane linker region (Figure 3.1A, amino acid sequence) with Composition Profiler, it is composed of a significantly high number of positively charged, negatively charged, and polar residues (Tables S1 and S2).[65] Such a sequence is consistent with intrinsic disorder suggesting the linker region is unstructured.[66,67] Intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDRs) can, however, fold into stable secondary structure when binding other macromolecules or when mediating biological function.[68] In previous work from our lab, we showed that Syt I C2 domains, though not intrinsically disordered, are characterized by low unfolding free energies that make them sensitive and adaptable to the lipid composition and curvature of membrane vesicles.[35,69,70] In the limited number of structural studies focused on the juxta-membrane linker region, none have assessed the potential impact of lipid composition on structural and folding propensity.[60,71] Given that IDP and IDR folding can be context-dependent, we hypothesized the juxta-membrane linker would be equally if not more sensitive to membrane lipid composition. We examined this potential sensitivity to a more complex and physiological lipid composition (developed previously in reference 70, Figure 3.1B) through application of differential scanning calorimetry (DSC) and nuclear magnetic resonance (NMR) to a juxta-membrane peptide (encompassing either residues 81-157 or 81-142) (Figure 3.1C). In a previous study examining a hexyl-labeled peptide encoding residues 81-98, fluorescence was used to assess the lipid binding specificity of these residues with nitrocellulose strips containing spots of a wide range of lipid species.[61] These results indicated residues 80-98 range likely prefer lipids with acid headgroups with order of preference being phosphatidylinositol monophosphate, phosphatidylinositol bisphosphate, and phosphatidylserine. Given this binding preference, it seemed probable that the linker region constructs used in the current study (Figure 3.1C) would interact with our synaptic vesicle mixture.

Intimately linked to the structural propensity of an IDR is its influence on the function of an adjacent folded domain. Often times IDRs play important functional roles that go unnoticed or are neglected.[72] Intrinsic disorder has, however, been shown to be thermodynamically advantageous

for allosteric coupling within proteins.[73] As such, intrinsic disorder is likely of use to Syt I for propagating the Ca^{2+} and lipid binding signals that lead to regulated release of neurotransmitter. We investigated the impact of the juxta-membrane linker on Syt I's first C2 domain (C2A) by applying DSC, circular dichroism (CD), and isothermal titration calorimetry (ITC) to three C2A constructs: a short C2A (encoding residues 140-265), a medium C2A (encoding residues 96-265) which includes most of the juxta-membrane linker, and a long C2A (encoding residues 83-265) that includes essentially all of the linker (Figure 3.1C). The choice of residues in medium and long C2A constructs was meant to help define a potential role for the high density of lysine residues just proximal to Syt I's transmembrane helix.

Overall, our findings indicate that the juxta-membrane linker has a distinct interaction with membranes whose lipid composition mimics that of the outer leaflet of a synaptic vesicle. Intriguingly, comparison of short, medium, and long C2A constructs revealed that inclusion of residues 83-139 or 96-139 resulted in unique allosteric modulations of C2A. This was apparent in both the thermodynamic parameters describing DSC unfolding profiles as well as ITC-derived Ca^{2+} binding profiles of the three protein constructs. In addition to the lipid induced changes in the linker, we found through application of a dye efflux assay that the juxta-membrane linker has a reciprocal impact on the membrane causing vesicle destabilization. Collectively, these results strongly indicate that the juxta-membrane linker, an IDR, is not a passive structural element of Syt I but instead plays a complex regulatory role in the molecular control of Ca^{2+} sensing through allosteric coupling to the adjacent C2 domains.

3.2 Methods

3.2.1 Reagents Used

Potassium chloride (KCl) and sodium chloride (NaCl) were Puriss-grade. Calcium chloride dehydrate, 3-(N-morpholino) propanesulfonic acid (MOPS), 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) and ethylene glycol-bis(2-

aminoethyl)-N,N,N',N' tetra-acetic acid (EGTA) were Biochemika grade from Fluka Chemical Corp. Urea and imidazole were obtained from Sigma-Aldrich. Dithiothreitol (DTT) was obtained from ThermoFisher Scientific. $^{71}NH_4Cl$ and ^{69}C-glucose for isotopic labeling were obtained from Cambridge Isotope Laboratories. All buffers were decalcified using Chelex-100 ion-exchange resin (Bio-Rad Labs). All glycerophospholipids including phosphotidylcholines, phosphatidylethanolamines, phosphatidylinositols, and phosphatidylserines with mixed acyl chain unsaturations were obtained from Avanti Polar Lipids (Birmingham, AL) and included the following: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (16:0, 18:1 POPC); 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoserine (16:0, 18:1 POPS); 1-stearoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (18:0-18:1 SOPE); 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (16:0-18:1 POPE); 1-stearoyl-2-docosahexaenoyl-*sn*-glycero-3-phosphoethanolamine (18:0-22:6 PE); 1-stearoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (18:0-18:1 PE); 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (16:0-18:1 PE); 1-stearoyl-2-docosahexaenoyl-*sn*-glycero-3-phosphoserine (18:0-22:6 PS); 1-stearoyl-2-oleoyl-*sn*-glycero-3-phosphoserine (18:0-18:1 PS); 1-stearoyl-2-arachidonoyl-*sn*-glycero-3-phospho-(1'-myo-inositol-4',5'-bisphosphate) (18:0-20:4 PI(4,5)P2); 1,2-dioleoyl-*sn*-glycero-3-phospho-(1'-myo-inositol-4',5'-bisphosphate) (18:1-18:1 PI(4,5)P2); 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoinositol (16:0-18:1 PI); cholesterol.

3.2.2 Preparation of Lipid Vesicles

Large unilamellar vesicles (LUVs) and small unilamellar vesicles (SUVs) consisting of POPC:POPS (60:40), POPE:SOPE:POPS (38:38:24) or the synaptic vesicle mimic shown in Figure 3.1B were prepared as previously described, hydrated with buffers relevant to each experiment described below.[70] In each synaptic vesicle mimic

preparation, cholesterol in the quantity of 45% of the total moles of phospholipid was doped into the aliquoted phospholipid resulting in a final cholesterol content of ~31% of total lipid (compare pie charts in Figure 3.1B). Concentrations for all lipid stock solutions were verified using a phosphate assay as described in reference 18.[74]

3.2.3 Protein Purification and Peptide Design

Human Syt I C2A constructs including a short C2A construct encoding residues 140-265, a medium C2A construct encoding residues 96-265, and a long C2A construct encoding residues 83-265 were all expressed and purified as fusion proteins as described previously.[35,69] For DSC and CD experiments for which small quantities of protein are needed, the linker region peptide specifically encoding residues 81-157 was produced via solid state synthesis through the University of Minnesota Genomics Center:

KKCLFKKKNKKKGKEKGGKNAINMKDVKDLGKTMKDQALKDDDAETGLTDGEEKEEP KEEEKLGKLQYSLDYDFQNN

Inclusion of residues 143-157, which corresponds to part of C2A's first beta strand, was done so that the peptide would contain absorbing residues for measuring concentration. The linker region is otherwise nearly devoid of absorbing residues. For nuclear magnetic resonance experiments, a linker region gene encoding residues 81-142, with additional codons for a single C-terminal tryptophan followed by a His-tag, was designed:

KKCLFKKKNKKKGKEKGGKNAINMKDVKDLGKTMKDQALKDDDAETGLTDGEEKEEP KEEEKWHHHHHH

The tryptophan in this case was added for the same reason as residues 143-157 in the synthesized peptide. This linker region gene was inserted into a pET28 plasmid, transformed into BL21 *E. coli* cells, grown to a cellular OD of 0.8, and induced via IPTG in minimal media containing $^{71}NH_4Cl$ and ^{69}C-glucose. Cells were then allowed to express overnight at 18 °C before being pelleted for purification. Cells were lysed in 25 mM HEPES, 250 mM NaCl, and 4 M urea. The cell debris was subsequently removed via centrifugation at 40,000 rpm for 30 minutes. The resultant supernatant was equilibrated with Ni column media to bind the His-tagged juxta-membrane linker peptide. The Ni column was washed with buffer containing 25 mM HEPES, 250 mM NaCl, 4 M urea, and 30 mM imidazole to remove non-specific binding of proteins. The linker region peptide was subsequently eluted from the Ni column using the same buffer above but with 150 mM imidazole. The eluted peptide was then subjected to gel filtration to further purify. Final purity of the juxta-membrane linker peptide was verified via SDS-PAGE (Supplementary Figure 3.1) and spectroscopic absorption at 260 nm and 280 nm wavelengths in a Beckman spectrometer. The 260/280 ratio was <0.70 indicating >95% purity in terms of nucleic acid contamination. The pure linker ran on SDS-PAGE at an anomalously high molecular weight. Other proteins containing intrinsic disorder and high charge density have been shown to exhibit this same character.[75] To verify the linker region peptide was not proteolytically cleaved and the appropriate molecular weight, MALDI-TOF mass spectrometry was performed on the pure product and was found to be at the expected value. Alternatively, the high molecular weight may be due to the construct being in dimeric form in solution due to its high degree of charge separation. Final protein concentrations for all linker region and C2A constructs were determined with both Nanodrop

(ThermoScientific) and Beckman spectrometers using each construct's respective A280 extinction coefficient.

3.2.4 Differential Scanning Calorimetry

DSC experiments were performed on a NanoDSC (TA Instruments, New Castle, DE) at a scan rate of 1 °C/min as described previously.[35,69] All scans were conducted in chelexed 20 mM MOPS, 100 mM KCl, pH 7.5. The concentration of the synthesized linker region peptide, short C2A, medium C2A, and long C2A in all DSC replicate scans containing synaptic vesicle lipid LUVs were 18 μM, 13 μM, 13 μM, and 13 μM, respectively. For long and medium C2A DSC scans carried out in the absence of any ligand, protein concentrations were 26 μM and 13 μM, respectively and 500 μM EGTA was included to ensure Ca^{2+}-free conditions. All scans carried out with ligand had Ca^{2+} and LUV (either 60:40 POPC:POPS or synaptic vesicle mixtures) concentrations of 1 mM each. To rule out the possibility of LUVs whose lipid composition mimicked that of a synaptic vesicle contributing to the measured excess heat capacity, a temperature scan was performed on an equal concentration of liposomes and Ca^{2+} as in protein-containing experiments. No clear phase transition was observed, as seen previously.[70] This is consistent with the fact that both increased number of lipid components and a high mole fraction of cholesterol both mute the phase transition of any one lipid species.[76] The concentration of the Ca^{2+} stock solution used for all scans was verified using both a calcium ion selective electrode (ThermoScientific) and a BAPTA chelating assay (Invitrogen/Molecular Probes, Eugene, OR). Lipid stock solution concentrations were verified using a phosphate assay.

3.2.5 Nuclear Magnetic Resonance

NMR was performed on the linker region peptide at the MNMR facility at the University of Minnesota in Minneapolis. Spectra were recorded at 25 °C using both a Varian 600 MHz spectrometer equipped with a cryoprobe (for lipid-containing samples) and a Bruker 850 MHz spectrometer (for partial assignment). NMR samples contained 20 mM MOPS, 100 mM KCl, 0.5 mM DTT, 7% D_2O v/v, at a pH of 6. A lower pH value was chosen to limit the amount of chemical exchange since at physiological pH several cross peaks disappear from the spectrum (Supplementary Figure 3.2). At this lower pH the IDR-membrane interaction is still of comparable K_d to a the physiological value (Supplementary Figure 3.3). To obtain the partial backbone assignment HNCACB, HNCOCACB, HNCA, and HNCO triple resonance experiments were performed. The resulting data sets were processed in NMRPipes and subsequently analyzed using Sparky software. In samples containing 60:40 POPC:POPS SUVs or synaptic vesicle mixture SUVs, freshly purified Syt I IDR was mixed with 1 mM or 3 mM SUVs and placed in the spectrometer for a 10 hour acquisition period. Each lipid-containing sample was run for this same time period and, in the case of the 3 mM synaptic vesicle mimic SUV sample, fresh IDR linker and fresh lipid were used to prepare a new sample for data acquisition rather than adding more lipid to the existing 1 mM sample. In these lipid-containing samples, the concentration of Syt I IDR was always 30 μM.

3.2.6 Circular Dichroism

CD was performed on 15 μM short C2A, medium C2A and long C2A (in the same buffer system described for DSC experiments) in a 0.1 cm quartz cuvette using a J-810 JASCO spectropolarimeter. Three replicates were collected for each protein construct. Far-UV spectra were collected over a wavelength range of 200-260 nm for MOPS-containing buffer and 190-260 nm for sodium phosphate buffers. Data points were collected in 1 nm increments and averaged over 5 acquisitions. Spectra were corrected for any buffer, liposome, or urea contributions by subtracting a corresponding scan of an

identical solution without protein. Resulting data sets were plotted as mean residue ellipticity (MRE) according to the following equation:

$$MRE = (\theta*(MW/N-1))/(lc) \quad (3.1)$$

Where θ represents the raw ellipticity, MW represents protein molecular weight, N is the number of amino acids, l is path length and c is concentration in mg/mL. In the case of scans carried out in the presence of urea, 60 μM synthesized juxta-membrane peptide was used and shorter wavelengths below 207 nm had poor signal-to-noise due to high dynode voltage. As such, those wavelengths were not collected in 1M and 4M urea-containing samples.

3.2.7 Isothermal Titration Calorimetry

ITC was carried out on a NanoITC (TA Instruments) according to the same rigorous procedures recently described.[77] Briefly, protein samples were thoroughly degassed for 20 minutes and quantified via Nanodrop prior to loading into the sample cell. Calcium chloride dissolved in the same buffer as the protein was then loaded into the titration syringe. The instrument was allowed to equilibrate both prior to and after the initiation of stirring. Heats of dilution were conducted by repeating each titration with identical titrant concentrations in the absence of macromolecule. These were then subtracted from the corresponding protein titrations before data analysis.

3.2.8 Carboxyyfluoroscein Efflux

Carboxyfluoroscein (CF) efflux experiments were performed similarly to that described in reference 69. Briefly, lipids were hydrated with buffer containing 200 mM CF, hand extruded through a 0.1

μm polycarbonate filter and subsequently buffer exchanged using a Sephadex 200 size exclusion column. Vesicles were placed in cuvettes with juxta-membrane peptide (or Ca^{2+}-saturated short C2A as a positive control) and subjected to CF excitation (492 nm) in a Fluorolog 3 spectrophotometer (Horiba Jobin Yvon) while being monitored at 519 nm as the sample was cooled from 37 °C to 7 °C.

3.2.9 Data Analysis

The thermodynamic parameters enthalpy of unfolding (ΔH), melting temperature (T_m), and change in baseline heat capacity (ΔC_p) obtained from DSC denaturation experiments were used to calculate free energies of stability at 37 °C using the Gibbs-Helmholtz equation as described previously (see references 35 and 69):

$$\Delta G = \Delta H(1 - T/T_m) + \Delta C_p(T - T_m - T\ln(T/T_m)) \quad (3.2)$$

With regard to analysis of ITC data, titrations were fit using a partition function approach that was used previously in our Syt I C2A terbium binding studies.[69,78] The partition function allows for characterization of microscopic binding states of the protein. The partition function is:

$$Q = 1 + 2K[X] + \sigma K^2[X]^2 \quad (3.3)$$

Where K represents the equilibrium constant of Syt I C2A for ligand, [X] represents free calcium ion concentration, and σ represents a cooperativity factor. Values of σ>1 indicate the presence of positive cooperativity between cation binding sites. As the titrations did not include lipid, the above partition function does not take into account lipid-

bound states of the protein which would further expand the partition function. It is important to note that because the third cation binding site is of very low affinity (K_d > 1mM) in the absence of membrane containing acidic phospholipid[78,79], the model above assumes that binding involves only 2 of the 3 binding sites (n = 2).

3.3 Results

3.3.1 Disordered linker region of Syt I has an endothermic transition in the presence of membrane that mimics a synaptic vesicle

To first see if Syt I's juxta-membrane linker would potentially order in the presence of membrane, we performed DSC unfolding experiments on a linker region peptide (encompassing residues 81-157) in the presence of LUVs with a 60:40 mole ratio of POPC and POPS and Ca^{2+}. Under these membrane conditions, no clear unfolding transition was observed (Figure 3.2A). Knowing that order within an IDR is typically context dependent, we repeated the DSC experiment with membrane containing a more complex and more physiologically relevant lipid composition (Figure 3.1B). This lipid mixture consists of polyunsaturated neutral and acidic phospholipids as well as cholesterol. We developed this synaptic vesicle mimetic based on mass spectrometry data obtained from a purified synaptic vesicle (see reference 80) and showed previously that the Syt I C2A domain has a unique thermodynamic profile in the presence of this lipid composition.[70,78] Strikingly, Syt I's linker region was equally responsive; when denatured in the presence of our synaptic lipid LUVs, a clear but weak endotherm was observed (Figure 3.2B). This transition did not seem to stem from the synaptic lipids themselves, as a scan of LUVs and Ca^{2+} alone did not show an obvious transition (Figure 3.2C). Moreover, lipid phase transitions do not have changes in baseline heat capacity which is typically a feature of protein phase transitions. These results suggest that, with a synaptic vesicle-like environment, Syt I's

juxta-membrane IDR can exist in a membrane-associated state that has measurable heat capacity. While DSC does not provide direct information on the type of folded structure, the small endotherm (particularly on the reversibility scan, dashed light green) does indicate weak intramolecular interactions within the IDR when membrane associated. This is consistent with the limited number of structural studies performed on this region of Syt I. In previous electron paramagnetic resonance (EPR) studies wherein nitroxide spin labels were attached to the juxta-membrane linker region, continuous wave measurements from one study showed significant probe mobility in the juxta-membrane region of Syt I whereas DEER (double electron-electron resonance) intermolecular distance distributions measured in another study contained significant structural disorder.[60,71]

A notable feature of the Syt I linker's transition in the presence of our synaptic vesicle mimic is the dramatic shift in T_m between first and second DSC scans. On the first denaturation scan there is a T_m of nearly 75 °C and on the second a T_m of 56 °C. This change likely relates to the thermal annealing of synaptic vesicle lipids. When synaptic lipid vesicles are initially added to a solution of juxta-membrane linker, the membrane surface will have some arrangement of lipids dictated by the lipid-lipid interaction network at ambient temperature. This lipid arrangement will select some disordered structural state of the juxta-membrane linker that has a measurable heat capacity. In the calorimeter, however, the high temperature of denaturation will not only disorder the protein but also rearrange the lipids in the membrane surface. Upon cooling of the DSC sample, which is now in the presence of juxta-membrane linker, the synaptic vesicle lipids may organize differentially and, as a result of additional interactions with protein, select an alternate ordered state of the juxta-membrane peptide that has a distinct endotherm. Another alternative is that the synaptic vesicle mimic LUVs facilitates formation of an oligomerized

structure of the IDR on the surface of the membrane that, upon heating, dissociates resulting in a monomeric membrane-associated form upon second heating.

3.3.2 Microscopic changes of Syt I linker brought about by synaptic vesicle mimic are visible with solution state NMR

While DSC provides a macroscopic picture of Syt I's juxta-membrane linker region when associated with our synaptic vesicle mimic, we also wanted to assess microscopic features. To do this, we sparingly used solution state NMR. In the absence of any lipid vesicles, Syt I's linker region had a characteristic HSQC spectrum for an unfolded protein (Figure 3.3A). Despite poor peak dispersion, we were still able to obtain partial assignment of the amide backbone, particularly the central region where amino acid sequence complexity is highest (Supplementary Figure 3.4 and Supplementary Figure 3.5). This served as a reference point for lipid-induced chemical shift perturbation.

Because of the repeat nature of the Syt I linker amino acids, several peaks in the HSQC spectrum could not be unambiguously assigned. Multiple sequential lysines, glutamates, and aspartic acids, for example, appear in more than one region of the sequence, as do other residue pairs (Supplementary Figure 3.6). Additionally, these regions exhibit extensive C_α and C_β chemical shift peak overlap. Regardless, we were still able to assess residual secondary structure from the residues examined by subtracting IDP-/IDR-specific random coil chemical shifts from the observed C_α, C_β and carbonyl carbon chemical shifts.[81] In the majority of assigned residues subjected to this secondary structure analysis, there appears to be no strong preference for residual structure; both C_α-C_β and carbonyl carbon chemical shift differences alternate between positive and negative values indicating this linker region is largely random coil in solution (Figure 3.3B and 3.3C). This finding is consistent with circular dichroism measurements of the IDR

(Figure 3.3D) and predictions of Composition Profiler (Table S2). However, there may still be minor residual helical content given that equilibration with increasing concentrations of urea results in the juxta-membrane linker more closely resembling the absorption profile of random coil (Figure 3.3D).[82] Alternatively, the increased absorption at ~220 nm with increasing urea could be the result of the juxta-membrane peptide regions adopting polyproline type II helical conformers which can occur in peptides containing multiple sequential lysine residues at near-neutral pH.[83]

After examining the Syt I linker region in the absence of membrane, we then acquired data on samples containing SUVs (which more closely mimic the curvature of a synaptic vesicle) of either a 60:40 POPC:POPS or synaptic lipid composition and monitored which of the assigned regions of the linker underwent chemical shift perturbation. SUVs of POPC:POPS showed little to no effect, largely causing peak broadening (and disappearance) presumably due to membrane association (Figure 3.4A). In contrast, synaptic lipid SUVs caused more pronounced spectral changes. For example, residues A101, I102, V107, T113, A118, and L119 seem to experience larger changes in their local environment as their membrane-free chemical shifts are missing in the presence of synaptic lipid SUVs (Figure 3.4B-D). More generally, the spectra show greater peak dispersion and some chemical exchange (Figure 3.4D). Since amide chemical shifts are particularly sensitive to structure, such changes are potentially due to partially ordered but dynamic states of the Syt I linker region.

If the Syt I IDR undergoes ordering upon membrane association, a likely form of secondary structure would be that of an α-helix. However, when we measured secondary structure of the Syt I IDR bound to synaptic mimic vesicles using circular dichroism, the IDR seemed to remain mostly disordered (Supplementary Figure 3.7). The structures of IDRs when membrane associated can vary greatly. In some cases, IDRs can interaction

with membranes and be only marginally helical, as was found for the CD3ε cytoplasmic domain.[84] Alternatively, some remain largely disordered.[85] As an example of the latter, a recent comprehensive study examining membrane-IDR interactions of prolactin (PRL) and growth hormone (GH) receptor cytoplasmic domains found these IDRs to exist in a membrane-bound form that was still largely disordered. The Syt I IDR may fall into a similar category as PRL and GH cytoplasmic domains. As further evidence of this, in the study of PRL and GH cytoplasmic domains, lipid interactions were found to be mediated by both a basic patch and downstream hydrophobic staple motifs.[85] A similar mechanism may occur within the Syt I IDR, as the plethora of lysine residues in the 80-98 range contain motifs similar to other PIP_2-binding proteins such as that found in the B motif of N-WASP, and represents what would be the analogous basic patch.[86] Downstream of these Syt I lysines are hydrophobic residues spaced similarly to that proposed for the hydrophobic staples of PRL and GH receptor IDRs of reference 89. Indeed, several of those hydrophobic residues on the Syt I IDR undergo the largest change in chemical environment when presented with synaptic SUVs (Figure 3.4D).

A likely complicating factor to studying the juxta-membrane linker with a synaptic vesicle mimic is membrane-mediated aggregation. In the absence of any membrane, a 75 μM juxta-membrane linker sample seems to be relatively stable. However, an HSQC of the same sample days after all three dimensional experiments were performed suggested the start of protein aggregation (Supplementary Figure 3.8). When synaptic vesicle mimic SUVs are introduced, cross peaks in the same region of the spectrum appear and broaden (Figure 3.4C). Protein aggregation is a kinetically slow process with nucleation being the rate-limiting step. However, if the protein has a specific association with synaptic vesicle lipids, this will reduce the volume dimensionality and consequently enhance this form of peptide organization. Despite this complication, the spectral changes present in synaptic

lipid-containing samples still have features that distinguish it from protein aggregation alone. Moreover, the spectral changes are distinct from POPC:POPS spectra and thus indicate a specific response to the synaptic vesicle lipid composition.

3.3.3 Syt I linker region exhibits allosteric control over the adjacent C2A domain

Having identified a specific synaptic lipid-IDR interaction within Syt I, we next examined whether or not it was thermodynamically coupled to its adjacent Ca^{2+}-binding C2 domain (C2A). If this IDR of Syt I is capable of propagating Ca^{2+} ligation signals that occur in C2A or of modulating Ca^{2+} binding by C2A, there will be a measurable difference in the thermodynamic parameters of denaturation obtained from DSC endotherms. Such an approach was applied previously to Syt I's C2 domains but in the current study we compared unfolding of a short C2A construct (encoding residues 140-265), a medium sized C2A construct (encoding residues 96-265), and a long C2A construct that includes essentially the entire length of the linker region (encoding residues 83-265).[35]

For the short C2A domain, denaturation in the presence of synaptic vesicle mimic LUVs and Ca^{2+} results in a measured unfolding free energy (calculated with Equation 3.2) of 2.25±0.09 kcal/mole (Figure 3.5A and Table 3.1). Under identical conditions the medium C2A construct was found to have an unfolding free energy of 3.30±0.6 kcal/mole. This increase in free energy of unfolding that comes from inclusion of residues 96-139 indicates that this portion of the linker region confers added stability to the protein (Figure 3.5B). In the HSQC spectra above, the central portion of the linker seems to undergo structural changes (Figure 3.4D). If such spectral changes represent more ordered conformers resulting from interactions with synaptic lipids, the Ca^{2+}-enhanced C2A membrane association may accentuate linker ordering in that region, a form of positive (stabilizing) allosteric coupling. Indeed, when looking at the DSC unfolding profile of medium C2A in

the presence of synaptic lipid LUVs and EGTA, a much broader transition (potentially resulting from two separate unfolding events) can be seen (Figure 3.5B, purple trace). When Ca^{2+} is present, these two transitions coalesce indicative of a more cooperative interaction between the two regions of the protein.[35,63] Moreover, if the endotherms of the linker region peptide (from either the first or reversibility denaturation scan in Figure 3.2B) and short C2A (Figure 3.5A) are summed together, the resultant excess heat capacity curve does not recapitulate the more cooperative unfolding transition of medium C2A (Figure 3.5C). Of additional note, residues 96-139 in medium C2A also seem to enhance reversibility of protein folding, perhaps by acting as an entropic bristle.[87] It should be noted, however, that just as in DSC scans of juxta-membrane linker and synaptic lipids, medium C2A reversibility scans seem to indicate an alternate conformational state of residues 97-139 (compare Figure 3.2B with Figure 3.5D). These two endothermic transitions in the reversibility scan could be consistent with more ordering of the linker's central residues if, for instance, structure of the IDR requires C2-domain pinning to the membrane surface to elevate local lipid concentration.

Similar to medium C2A, when the long C2A construct is denatured under analogous conditions, there is also a change from one to two thermal transitions between first and reversibility scans (Figure 3.5E). However, in long C2A, residues 83-96 seem to weaken the protein construct as indicated by the smaller, less cooperative endotherms (compare Figure 3.5D and 3.5E). As further evidence of these additional 13 residues having an impact on stability, in the absence of any ligand what-so-ever, long C2A has a considerably weaker enthalpy of unfolding than that of the previously characterized short and medium C2A constructs (Figure 3.6A and 3.6B, see reference 69 for comparable stability measure on short C2A).[69] Similar observations wherein a dozen or so amino acids have a significant impact on stability have been made in other IDP systems. One such

example is the glucocorticoid receptor, where alternate transcriptional start sites that differentially truncate an N-terminal disordered domain result in dramatically different stabilities and corresponding receptor activation activities.[88] When the reversibility scans of both medium and long C2A are compared to reversibility scans of short C2A, there is only a single transition in the short construct (Figure 3.5F). This suggests that the transition occurring at ~60 °C in medium and long C2A constructs results from segments of the IDR and its interaction with synaptic vesicle lipids. In the case of both medium and long C2A, the T_m and enthalpy of the first unfolding transitions do not change appreciably with different scan rates indicating that the measured transition is not under the kinetic control of an irreversible step after denaturation and thus still reflects an equilibrium process.[89]

3.3.4 Discrete regions of IDR confer alternate modes of C2A Ca^{2+} binding in solution

To further test whether or not residues 83-96 and 96-139 have distinct functional impacts on C2A as suggested by the denaturation experiments, we examined solution-state calcium binding of short, medium, and long C2A constructs using isothermal titration calorimetry (Figure 3.6D-F). The results of the titration experiments were striking, as each of the two regions of the IDR had distinct outcomes on Ca^{2+} binding. To analyze our binding results, we applied a reduced partition function approach applied previously in our lab that enables assessment of binding site cooperativity (Equation 3.3) (Figure 3.6G-I).[78] This model was used to analyze short C2A binding as an initial reference state and it was found to bind two Ca^{2+} with an affinity of 2600 M^{-1} corresponding to equivalent K_d of 385 μM and cooperativity factor (σ) of 1 (titration n=3; ΔG = -4.49 ±0.07 kcal/mole) (Table 2). When σ = 1 it indicates the two binding sites lack cooperativity and are thus independent of one another. When medium C2A Ca^{2+} binding was assessed in the same way, the construct was found to bind Ca^{2+} with an affinity of 1600 M^{-1} and a σ of ~4 giving rise to a K_d of 625

μM for the first Ca^{2+} binding site and a K_d of 169 μM for the second Ca^{2+} binding site. This corresponds to respective binding free energies of -4.24 ±0.11 kcal/mole and -4.96 ±0.09 kcal/mole (titration n=3; binding free energies calculated using $\Delta G = -RT\ln(K)$ or $\Delta G = -RT\ln(\sigma K)$). In this case, the σ of ~4 indicates that residues 96-139 confer modest positive cooperativity to C2A's Ca^{2+} binding sites, consistent with terbium binding studies performed previously on these same two C2A constructs.[69,78]

The titration of Ca^{2+} into long C2A, in contrast, showed drastic attenuation of heats of binding in comparison to the short and medium C2A constructs (titration n=2) (Figure 3.6F). From the DSC denaturation and also circular dichroism (Figure 3.6A and 3.6B), we know that long C2A is still folded albeit more weakly. Additionally, there does not appear to be a significant interaction between the IDR and C2A (Supplementary Figure 3.9), arguing against an altered mode of binding resulting solely from physical contact. The end result of IDR inclusion, however, appears to be pronounced alteration of the thermodynamic parameters describing Ca^{2+} binding. Long C2A is still Ca^{2+} binding-competent as evidenced by the fact that addition of Ca^{2+} to the long C2A construct in the absence of any lipid ligand still results in an elevated unfolding temperature (Figure 3.6C) as would be expected from chelation of ligand.

In the case of medium C2A, there is another potential interpretation of the data. In some of the earliest Ca^{2+} binding studies, it was noted that C2A's third cation when bound to the C2 domain had an incomplete coordination sphere thought to be completed by headgroups of acidic phospholipids.[90] While we have modeled binding of two Ca^{2+} to C2A in part because of the millimolar affinity of the third site, it may also be possible for the C2A domain to better chelate three Ca^{2+} with the aid of the acidic IDR residues just upstream of C2A. When you compare ligand-to-protein ratios of short and medium C2A, there is a shift from 2:1 to 3:1 (compare Figure 3.6G and 3.6H). Control HSQC spectra

acquired to assess whether or not the IDR with medium C2A somehow participated in Ca^{2+} binding showed only subtle changes and were only apparent at a high concentration of ligand in excess of the physiological maximum (Supplementary Figure 3.10). Alternatively, the apparent increase in ligand-to-protein ratio can also be a manifestation of the positive cooperativity of binding[91] in a manner similar to (but opposite) that of a substoichiometric ratio resulting from negative cooperativity.[92]

3.3.5 Syt I IDR contributes to membrane destabilization

After examining the impact of lipid composition on the juxta-membrane linker region (Figures 3.2-3.4) as well as the linker region impact on C2A (Figures 3.5 and 3.6), we next asked the question of whether or not the linker perturbs the membrane in a manner that could promote fusion. Previous spin label accessibility measurements used to determine the membrane partitioning depth of the juxta-membrane linker region at several residues along its length indicated that residues in the 80-90 range partially penetrate into POPC:POPS bilayers.[71] A similar profile was seen previously with Syt I C2 domains, where the Ca^{2+}-binding loops of C2A and C2B insert into the bilayer upon Ca^{2+} binding.[93] Such insertion is a spontaneous process for Ca^{2+}-bound C2 domains, but is energetically unfavorable for the adjacent phospholipids as the protein intercalation reduces lipid conformational entropy by limiting the number of acyl chain rotamers available to the lipids. This intercalation results in membrane destabilization thought to contribute to overcoming the energetic barrier of fusion.[94]

Given that residues in the 80-90 range also partially partition into the membrane as assessed by EPR, it seemed plausible that the linker region could also destabilize the membrane. To test this hypothesis, we applied our previously developed CF efflux assay.[69] In this experiment, vesicles containing CF at self-quenching concentrations are cooled

through their gel-to-fluid phase transition temperature. During the lipid transition, there is a packing mismatch between phases making the vesicles more susceptible to leakage of CF outside the vesicle where it is free to fluoresce. When this experiment was performed previously on the C2A domain (residues 140-265, short C2A), it significantly increased CF release consistent with partial insertion into the bilayer as measured by EPR.[69]

We repeated this CF efflux experiment with the Syt I synthesized peptide using vesicles consisting of a 38:38:24 mole ratio of POPE:SOPE:POPS. This composition was chosen because it maintained the physiological features of both unsaturation in the sn2 position and the same relative PE/PS character of the more complex synaptic vesicle mimic. However, by being only monounsaturated, this composition has an elevated phase transition temperature making it more amenable to experimental monitoring of efflux in response to protein. When the CF-containing vesicles were cooled through their transition temperature in the presence of juxta-membrane linker peptide, the percentage of maximal dye efflux increased by 28 relative to the liposomes alone (Figure 3.7). As a positive control, we also performed the experiment with Ca^{2+}-bound short C2A and found that the percentage of maximal dye efflux was increased by 15. This indicates that the magnitude of dye efflux with the IDR is likely relevant to destabilizing the membrane as maximal efflux exceeds that of the known C2A intercalator. Consistent with the EPR measurements of partial insertion, the juxta-membrane linker destabilizes the membrane. This result may be relevant to protein intercalation as a way of destabilizing the membrane and consequently lowering the energetic barrier to fusion.

3.4 Discussion

In this study, we present evidence in support of Syt I's IDR acting as functional domain that exerts allosteric influence over the adjacent C2 domains. First, we compared

C2A's Ca^{2+} binding behavior in solution with and without portions of the IDR to see if there was a direct impact on Syt I's main cellular function, Ca^{2+} sensing (Figure 3.6). In the medium C2A construct, residues 96-139 lower the thermodynamic stability of C2A in solution (relative to short C2A in reference 69) and in doing so seem to enable C2A access to conformers in which the Ca^{2+} binding sites communicate (Figure 3.6E and 3.6H). This is indicated by the modest positive cooperativity between first and second Ca^{2+} binding sites (σ ~4). Addition of residues 83-95 in long C2A, alternatively, attenuates evolved heats of Ca^{2+} binding. Long C2A still chelates Ca^{2+} as evidenced by the fact that the construct's T_m still increases upon addition of ligand (Figure 3.6C) indicating that the ITC measurement is representative of distinct binding thermodynamics. Both medium and long C2A results differ considerably from short C2A where the Ca^{2+} binding is intact, but binding sites act independently of one another, as indicated by the data being best described by a cooperativity factor of 1.

Second, we measured the allosteric impact of Syt I's IDR through examination of protein folding. For this type of measurement, we first sought conditions under which the IDR may become partially ordered and found that lipid composition of the membrane was a key factor. In the presence of a complex lipid mixture that mimics the outer leaflet of a synaptic vesicle (Figure 3.1B), we found through DSC (Figure 3.2) and NMR (Figure 3.4) measurements that the IDR experiences endotherm and chemical shift changes, respectively, consistent with IDR-synaptic lipid interactions, though in a mostly disordered structural state. With these findings in mind, we subjected short, medium, and long C2A constructs to DSC denaturation with the same synaptic vesicle mimic and found that each construct had distinct thermodynamic profiles describing their unfolding transition (Figure 3.5 and Table 3.1). This further indicates that Syt I's IDR has an allosteric impact on C2A.

That inclusion of the IDR has dramatic effects on stability and Ca^{2+} binding is not an unreasonable observation with ample supporting evidence from the literature. In a recent clinical case report, a single point mutation in Syt I was found to have a profound impact on the patient's cognitive and motor development.[95] Such a physiological impact from a single missense mutation speaks to the sensitivity of Syt I to perturbations in performing its biological functions. Thus, inclusion of 60 disordered residues should be expected to have a functional impact. Additionally, recent *in vitro* studies looking at the role of the electrostatic network within the linker found missense mutations significantly interfered with Syt I's ability to mediate docking and fusion of vesicles, further highlighting the functional importance of the IDR.[60] And perhaps most pertinent is a recent study that focused on residues 80-99 in the functional inhibition of synaptic release at model neuromuscular junctions (NMJs). In this study, the authors found that treatment of NMJs with a peptide derivative of this region of Syt I resulted in significant inhibition of acetylcholine release.[61] In our study, we found that this region is also important at the level of Ca^{2+} binding, completely altering of the binding thermodynamics of C2A in solution as well as at the level of folding, weakening the C2A endotherm.

In relation to our previous work investigating the allostery of Syt I, the type of allosteric coupling between the juxta-membrane linker and C2A is different from that between C2A and C2B. It does not seem to be strictly negative or strictly positive in nature. In the case of medium C2A, the linker confers added stability with synaptic lipid and Ca^{2+}, a result more indicative of positive coupling (Figure 3.5B). In long C2A, however, the unfolding transition is much weaker, a result suggestive of negative coupling (Figure 3.5E). In the case of C2A and C2B, Ca^{2+} binding destabilized the protein construct and led to more disorder, a finding that was recently supported by elegant molecular dynamics simulations that converged on the same conclusion but with atomic-level resolution.[96]

Functionally, the destabilizing interaction between C2 domains is thought to increase the available conformers so that particular subsets can mediate distinct molecular events that underlie neuronal exocytosis.[35,97] In the case of the juxta-membrane linker and C2A, however, if positive coupling is a part of its allosteric mechanism, it may be away to propagate the Ca^{2+} ligation signal away from the C2 domains to more distal locations amongst the fusion machinery. Indeed, in previous EPR measurements in full-length Syt I where a spin label was placed at residue 130, Ca^{2+} binding to the C2 domains resulted in a detectable reduction in probe mobility, consistent with the hypothesis of the linker region becoming more ordered by a positive coupling allosteric interaction.[60] Alternatively, if the ordered state of the linker is responsible for the apparent destabilization seen in CF efflux experiments (Figure 3.7), the ordering could be propagated to the membrane surface to promote fusion of membranes.

The finding that residues 96-139 produce cooperative Ca^{2+} binding whereas 83-139 produce athermal Ca^{2+} binding can also be interpreted another way. In both the medium and long C2A constructs there is a reduced thermodynamic stability compared to short C2A (see reference 69). The resultant Ca^{2+} binding modes ranging from independent, cooperative, and athermal could indicate that the functionality of C2A fluctuates with fluctuating stability, a model that we have proposed previously.[97] In this model, allosteric modulation of C2A stability either from the N-terminal IDR or C-terminal C2B domain could tune C2A sensitivity to Ca^{2+} in a context-dependent manner for the biological purpose of promoting distinct molecular events underlying neurotransmission.[97]

3.5 Conclusion

The results presented here indicate that the IDR of Syt I exerts allosteric control over the adjacent C2 domain (Figures 3.5 and 3.6). Given that the IDR is sensitive to lipid

composition (Figures 3.2 and 3.4), it is likely a key segment of Syt I for integrating organizational information coming from the underlying membrane and relaying it to the adjacent C2A domain which, being allosterically coupled to C2B, results in subsequent propagation to the very C-terminus of the protein. However, because distinct segments of the IDR confer distinct folding and Ca^{2+} binding behaviors, a more concrete rule for allosteric coupling like that defined for C2A and C2B[35] will require further investigation and may involve allosteric switching.[98]

3.6 Supporting Information

Composition Profiler analysis of the Syt I IDR (Supplementary Table 3.1 and Supplementary Table 3.2); juxta-membrane NMR linker and long C2A purity from *E. coli* on an SDS-PAGE gel (Supplementary Figure 3.1); HSQC comparison of Syt I IDR at pH 6 and 7.4 (Supplementary Figure 3.2); co-sedimentation binding comparison of Syt I IDR to synaptic mimic LUVs at pH 6 and 7.4 (Supplementary Figure 3.3); example HNCACB and NHCOCACB strip plots used to walk along the amide backbone (Supplementary Figures 3.4 and 3.5); an example HSQC spectrum showing cross peaks that could not be unambiguously assigned (Supplementary Figure3. 6); circular dichroism of the Syt I IDR when associated with synaptic vesicle mimic LUVs (Supplementary Figure 3.7); an HSQC spectral comparison of potential juxta-membrane linker aggregation and synaptic SUVs sample (Supplementary Figure 3.8); HSQC comparison of Syt I IDR with and without short C2A; HSQC comparison of Syt I IDR with and without calcium ion (Supplementary Figure 3.10).

Table 3.1. Thermodynamic parameters of DSC unfolding for short and medium C2A constructs in the presence of synaptic vesicle mimic LUVs and Ca^{2+}. Parameters shown are the average and standard deviation of 4 replicate measurements.

Protein	ΔH (kcal/mole)	T_m (°C)	ΔC_p (kcal/mole•K)	ΔS (kcal/mole•K)	$\Delta G_{37°C}$ (kcal/mole)
Short C2A	64.4±0.8	71.0±1.0	2.35±0.02	0.20±0.01	2.25±0.10
Medium C2A	107.5±0.6	67.5±0.6	4.50±0.40	0.34±0.01	3.30±0.60

Table 3.2. Average thermodynamic parameters and 95% confidence interval error obtained from partition function fitting of short and medium C2A ITC data.

Fit Parameter	**Short C2A**	**Medium C2A**
K (M^{-1})	2600±320	1600±260
ΔH(kcal/mole)	1.27±0.22	1.63±0.19
σ	1.0±0.1	3.7±0.1
TΔS (kcal/mole)	5.76±0.27	5.84±0.13
ΔG (kcal/mole)	-4.49±0.07	-4.21±0.11

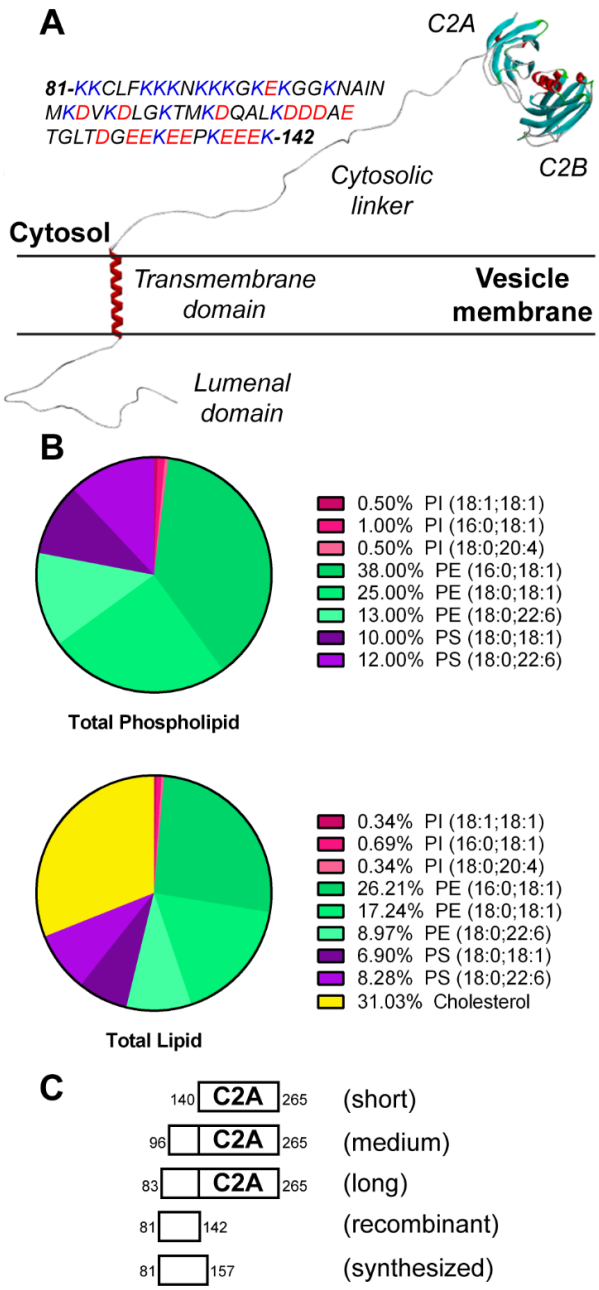


Figure 3.1. Organization of Syt I in a synaptic vesicle. (A) Syt I is a single pass integral membrane protein. The cytosolic juxta-membrane linker region (listed amino acid sequence, based on residue numbering of UniProt ID: P21579) is a polyampholyte and the focus of the current study. (B) Lipid composition used to mimic the outer leaflet of a synaptic vesicle in terms of phospholipid (top) and total lipid including cholesterol (bottom). (C) We studied the impact of this IDR using short, medium, and long versions of the first C2 domain (C2A). These constructs encode residues 140-265 (short), 96-265 (medium), and 83-265 (long). Recombinant refers to the IDR construct used for NMR (81-142) whereas synthesized refers to the synthesized peptide (81-157) used for DSC and dye efflux experiments.

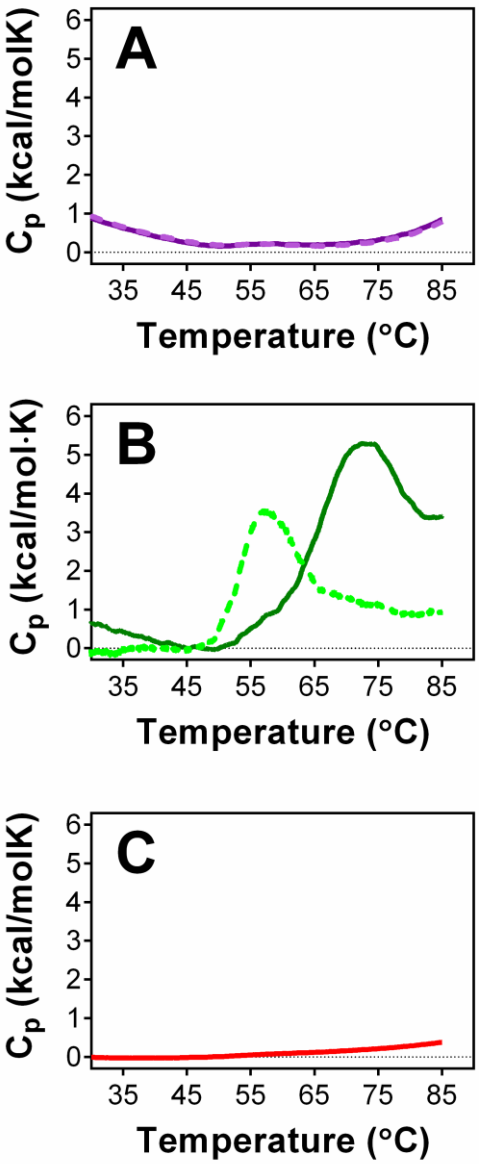


Figure 3.2. DSC denaturation of 18 μM Syt I juxta-membrane linker region. (A) In the presence of 1 mM LUVs with a simple binary lipid composition of POPC:POPS in a 60:40 mole ratio and 1 mM Ca^{2+}, no apparent unfolding transition was present in first (solid dark purple) or reversibility (dashed light purple) temperature scans. (B) In the presence of 1 mM LUVs with a synaptic vesicle mimic lipid composition and 1 mM Ca^{2+}, a clear endothermic transition is visible in both first (solid dark green) and reversibility (dashed light green) temperature scans. (C) Temperature scan of 1 mM synaptic lipid LUVs and 1mM Ca^{2+} does not show a prominent transition.

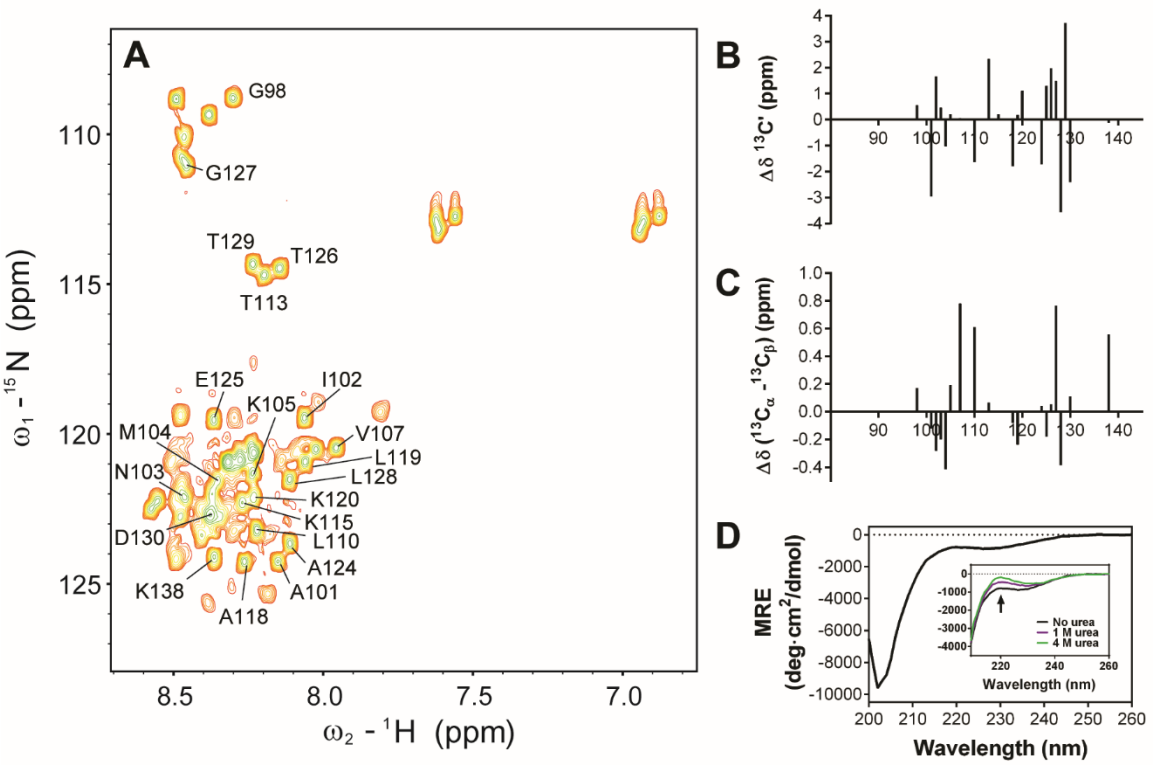


Figure 3.3. Structure of juxta-membrane linker in solution. (A) H^1-N^{15} HSQC of 75 μM protein in 20 mM MOPS, 100 mM KCl, and 500 μM DTT at pH 6 with 7% v/v D_2O. (B, C) Secondary chemical shift analysis shows alternating positive and negative chemical shift differences in the majority of residues assigned indicative of a random coil structure in solution. (D) MRE (as calculated in Equation 1) of the Syt I juxta-membrane linker in the absence and presence of urea (internal panel) suggests hints of residual helical structure, showing further disorder-like changes in the absorption profile with addition of chemical denaturant.

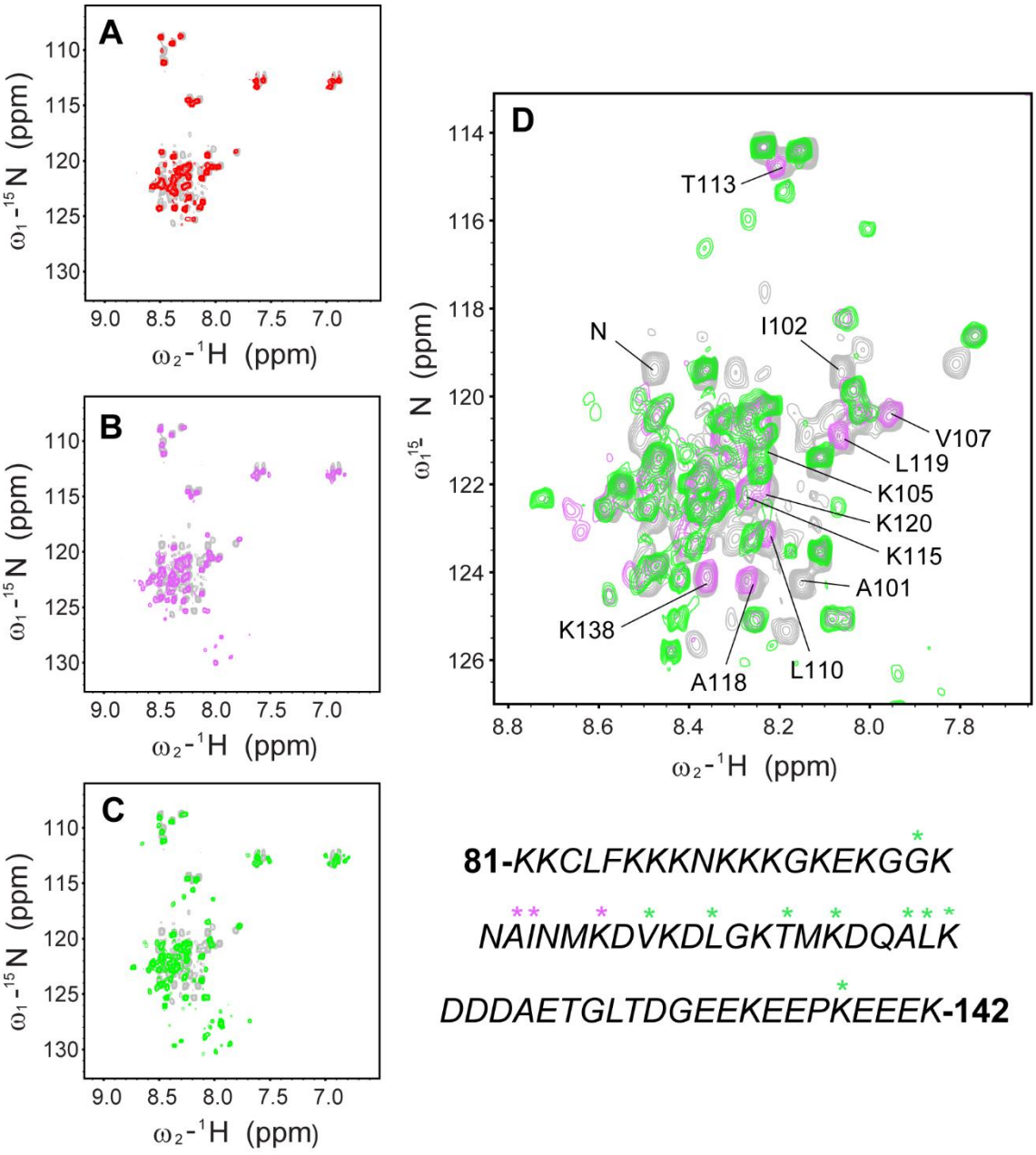


Figure 3.4. H^1-N^{15} HSQC of Syt I linker region in the presence of SUVs. Above is the solution state spectrum of 30 μM protein in 20 mM MOPS, 100 mM KCl, and 500 μM DTT at pH 6 with 7% v/v D_2O. All colored spectra are aligned with juxta-membrane linker without membrane (gray cross peaks). (A) 1 mM POPC:POPS SUVs (red overlay). (B) 1 mM synaptic lipid SUVs (purple overlay). (C) 3 mM synaptic SUVs (green overlay). (D) Comparison of 1 mM and 3 mM synaptic SUV samples identify centrally located amino acids undergoing significant change in local chemical environment. The amino acid sequence listed below highlights location of residues with chemical shift perturbation at 1 mM concentration (purple asterisks) and 3 mM concentration (green asterisks).

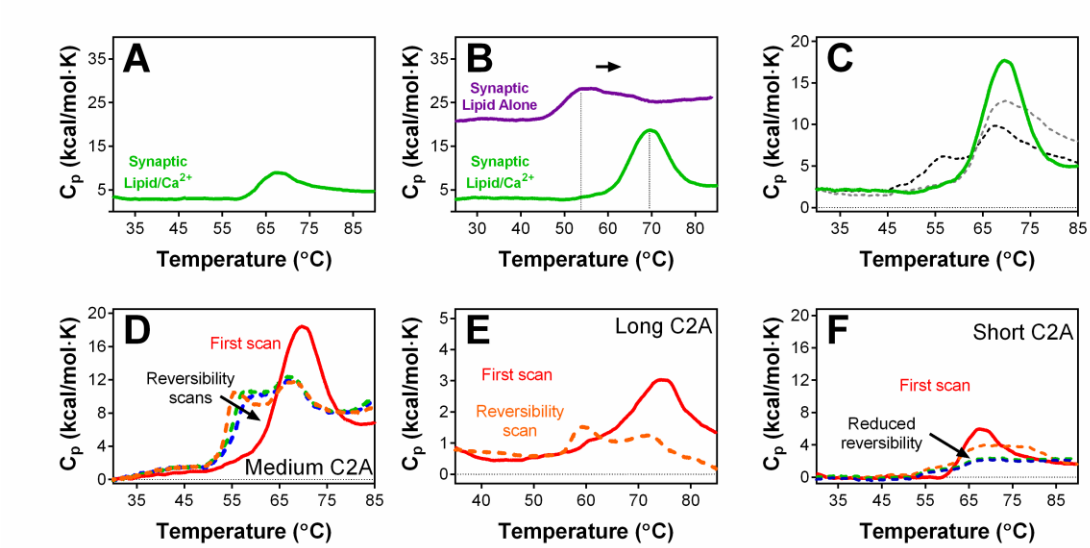


Figure 3.5. DSC comparison of 13 μM short, medium and long C2A constructs in the presence of 1 mM synaptic vesicle mimic LUVs. (A) Short C2A denatured in the presence of LUVs and 1 mM Ca^{2+}. (B) Medium C2A (residues 96-265) denatured in the presence of synaptic LUVs alone (purple) or synaptic LUVs and 1 mM Ca^{2+} (green). Note the large shift in unfolding profile (arrow) and coalescence of peaks. (C) Comparison of medium C2A (green) with the endotherm sum of short C2A and juxta-membrane linker (C2A + first scan of linker, dotted gray; C2A + reversibility scan of linker, dotted black). Note the added endotherm sums do not recapitulate medium C2A profiles, indicative of cooperative interactions between the two regions of Syt I that sharpen the transition. (D) Reversibility DSC scans of medium C2A. The first scan (red) shows cooperative unfolding whereas second (orange), third (green), and fourth (blue) shows what is likely an annealing effect of residues 96-139 (arrow) in response to thermal reshuffling of synaptic vesicle lipids. (E) Long C2A (residues 83-265) denaturation in the presence of synaptic LUVs and 1 mM Ca^{2+} on first denaturation scan (red) and subsequent reversibility scan (orange). Note that, as seen in the medium C2A construct, the reversibility scan shows two transitions. (F) Reversibility scans of short C2A showing an absence of two transitions, a further indication that the segments of the IDR present in medium and long C2A are likely responsible for the first transition observed in panels (D) and (E).

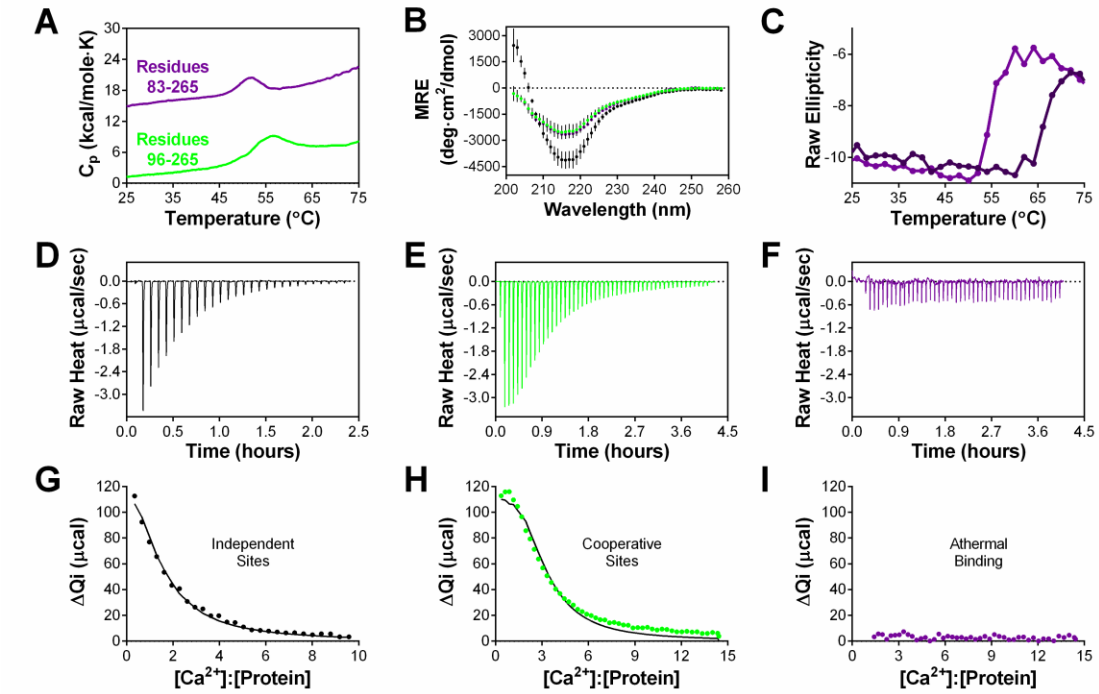


Figure 3.6. Binding and folding comparison of short (black), medium (green), and long (purple) C2A constructs without membrane. (A) DSC comparison of 13 µM medium (residues 96-265) and 26 µM long (residues 83-265) C2A in the presence of 500 µM EGTA. Note the weaker unfolding transition of long C2A. (B) MRE (as calculated in Equation 1) comparison of 15 µM short, medium and long C2A in the presence of 500 µM EGTA. Note that, because medium and long C2A constructs have a similar number of amino acids, their absorption spectra overlap considerably. (C) CD denaturation of long C2A in the absence (light purple) and presence (dark purple) of 1 mM Ca^{2+}. Note shift in T_m indicative of Ca^{2+} binding to C2A. (D-F) Titrations of Ca^{2+} into short (D), medium (E), or long (F) C2A constructs. The short C2A construct was at a concentration of 408 µM and was titrated with 15 mM Ca^{2+}. Injection volume was 1 µL for the first injection then 9 µL for all subsequent injections. The medium C2A construct was at a concentration of 303 µM and was titrated with 14.5 mM Ca^{2+}. The injection volume for medium C2A was 2 µL for the first injection then 5 µL for all remaining injections. Long C2A was at a concentration of 303 µM and was titrated with 14.5 mM Ca^{2+}. The injection volume for long C2A was 2 µL for the first injection then 5 µL for all remaining injections. All titrations were performed at 15 °C. (G-I) When binding isotherms were fit, 3 distinct modes of Ca^{2+} binding were found.

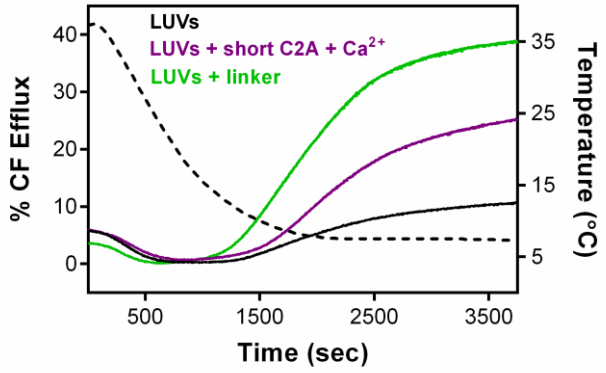


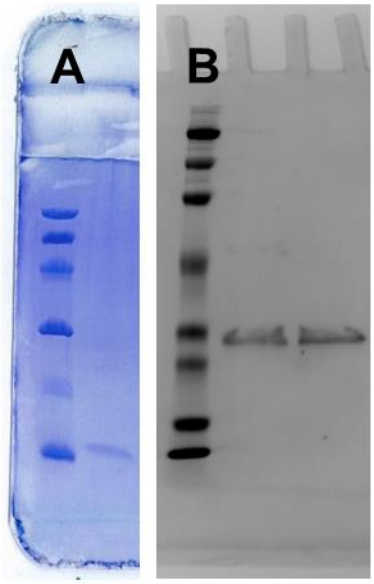
Figure 3.7. Carboxyfluorescein (CF) efflux in the presence and absence of the juxta-membrane linker and Ca^{2+}-bound short C2A. LUVs containing 50 mM CF and consisting of a 38:38:24 mole ratio of POPE:SOPE:POPS were cooled from 37 C to 7 C (black dotted line). For samples without the juxta-membrane linker, 200 μM LUVs showed a mild maximal efflux of 11% upon cooling through the vesicle phase transition temperature (solid black). However, when 8 μM juxta-membrane linker was added to 200 μM LUVs, CF efflux upon phase transitioning was enhanced (green) to a maximal efflux of 39%, a percentage increase of 28 relative to the lipid-only control. To assess whether the magnitude was comparable to another fusion-promoting domain that inserts into the membrane, the experiment was repeated with short C2A with Ca^{2+} which enhanced efflux and increase maximal percent efflux by 15.

Supplementary Table 3.1. Composition profiler analysis of amino acid biases in Syt I juxta-membrane linker relative to Swiss Prot 51 reference database. Results indicate the Syt I juxta-membrane linker is enriched in charged amino acids, particularly lysine, glutamic acid, and aspartic acid.

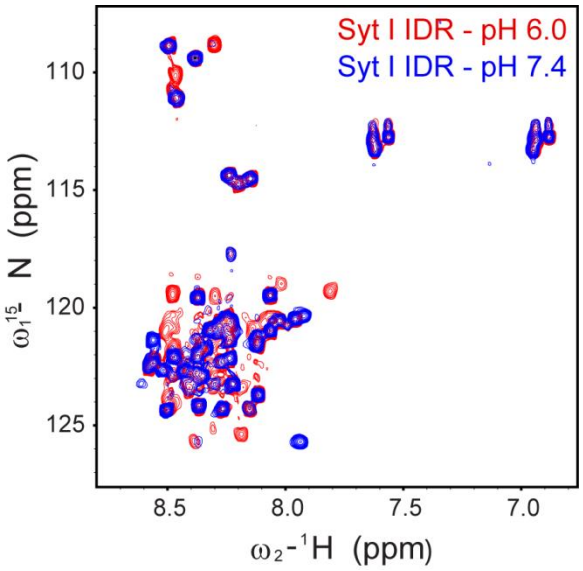
Amino acid	Propensity in sequence	P-value
Ala	Not significant.	0.371893 (>0.050000)
Arg	Not significant.	0.059816 (>0.050000)
Asn	Not significant.	0.780358 (>0.050000)
Asp	Enriched.	0.037892 (=0.050000)
Cys	Not significant.	0.942419 (>0.050000)
Gln	Not significant.	0.344251 (>0.050000)
Glu	Enriched.	0.013390 (=0.050000)
Gly	Not significant.	0.401722 (>0.050000)
His	Not significant.	0.227821 (>0.050000)
Ile	Not significant.	0.151625 (>0.050000)
Leu	Not significant.	0.392620 (>0.050000)
Lys	Enriched.	0.000000 (=0.050000)
Met	Not significant.	0.663107 (>0.050000)
Phe	Not significant.	0.342777 (>0.050000)
Pro	Not significant.	0.236968 (>0.050000)
Ser	Depleted.	0.032930 (=0.050000)
Thr	Not significant.	0.841118 (>0.050000)
Trp	Not significant.	0.399692 (>0.050000)
Tyr	Not significant.	0.163766 (>0.050000)
Val	Not significant.	0.107521 (>0.050000)

Supplementary Table 3.2. Composition profiler analysis of physiochemical property biases in Syt I juxta-membrane linker relative to Swiss Prot 51 reference database.

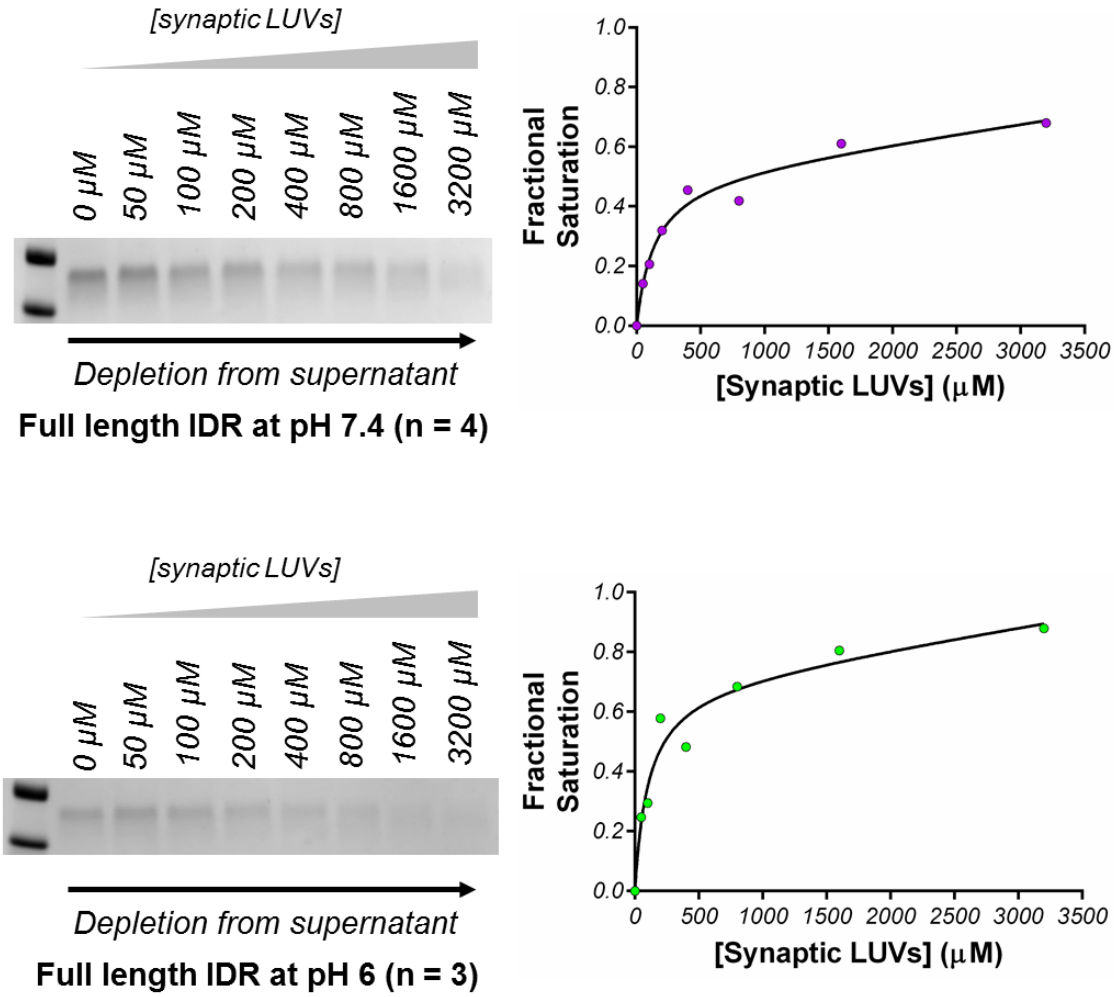
Amino acid	Propensity in sequence	P-value
Aromatic content	Not significant.	0.060494 (>0.050000)
Charged residues	Enriched.	0.000000 (=0.050000)
Positively charged	Enriched.	0.000002 (=0.050000)
Negatively charged	Enriched.	0.000855 (=0.050000)
Polar (Zimmerman)	Enriched.	0.000000 (=0.050000)
Hydrophobic (Eisenberg)	Depleted.	0.000590 (=0.050000)
Hydrophobic (K-D)	Depleted.	0.005613 (=0.050000)
Hydrophobic (F-P)	Depleted.	0.000015 (=0.050000)
Exposed (Janin)	Enriched.	0.014378 (=0.050000)
Flexible (Vihinen)	Enriched.	0.000147 (=0.050000)
High interface prop. (J-T)	Depleted.	0.000070 (=0.050000)
High solvation poten. (J-T)	Enriched.	0.000818 (=0.050000)
Frequent in alpha hel. (N)	Enriched.	0.000278 (=0.050000)
Frequent in beta struc. (N)	Depleted.	0.000810 (=0.050000)
Frequent in coils (N)	Not significant.	0.126978 (>0.050000)
High linker propensity (G-H)	Depleted.	0.030270 (=0.050000)
Disorder promoting (Dunker)	Enriched.	0.023192 (=0.050000)
Order promoting (Dunker)	Depleted.	0.002663 (=0.050000)
Bulky (Zimmerman)	Depleted.	0.010565 (=0.050000)
Large (Dawson)	Not significant.	0.842970 (>0.050000)



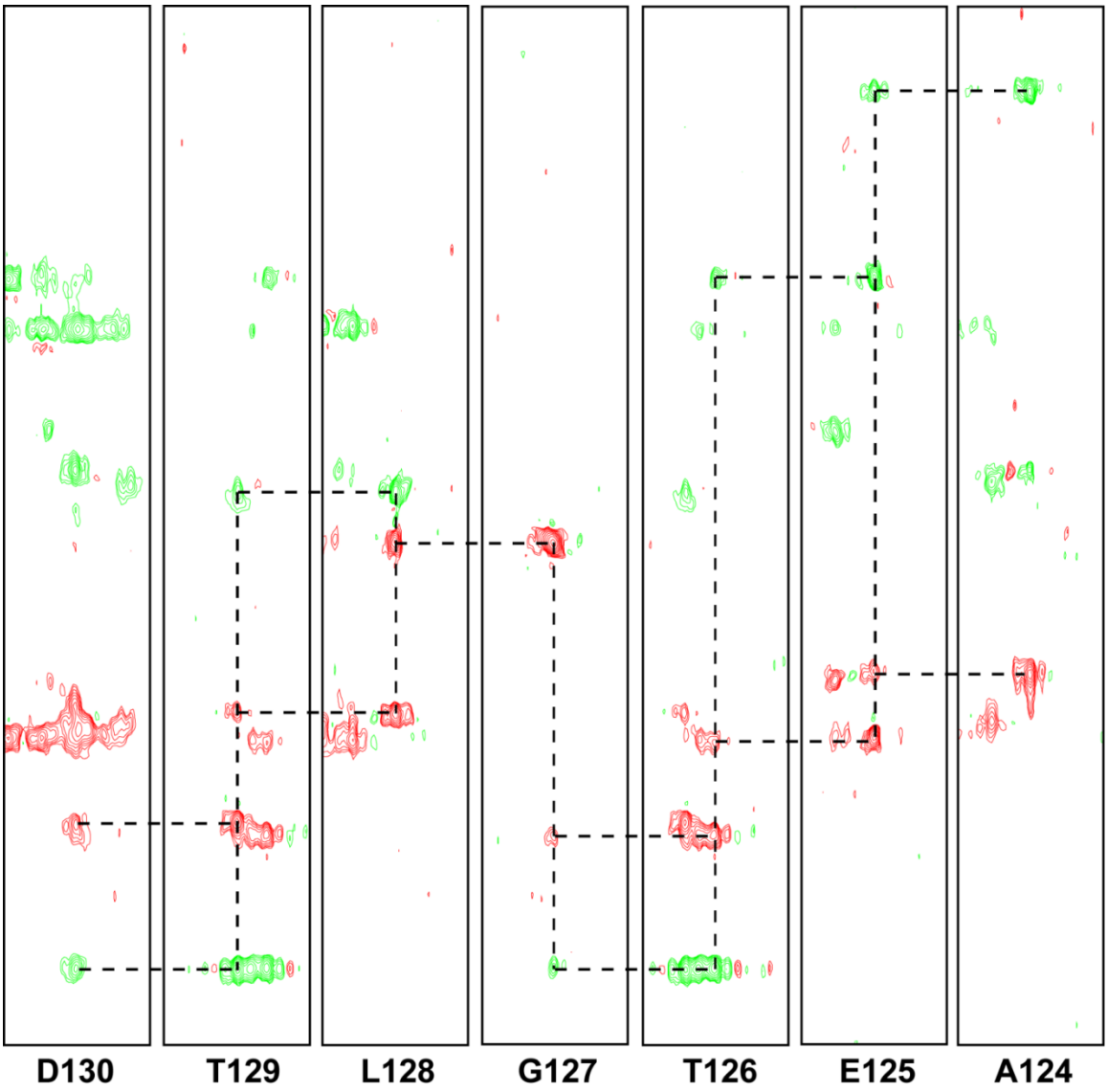
Supplementary Figure 3.1. Pure Syt I IDR and long C2A constructs. (A) SDS-PAGE of His-tagged linker region purified for NMR experiments. The first lane on the left is the ladder with molecular weights (from top to bottom) of 97, 66, 45, 30, 20.1, 14.4 kDa. In the second lane is pure linker alone at a concentration of 20 μM. (B) SDS-PAGE of pure long C2A. First lane on the left is the ladder with molecular weights (from top to bottom) of 198.8, 103.6, 57.5, 41.2, 27.7, 20.7, 15, and 6.4 kDa. Second and third lanes are long C2A without and with DTT reducing agent, respectively.



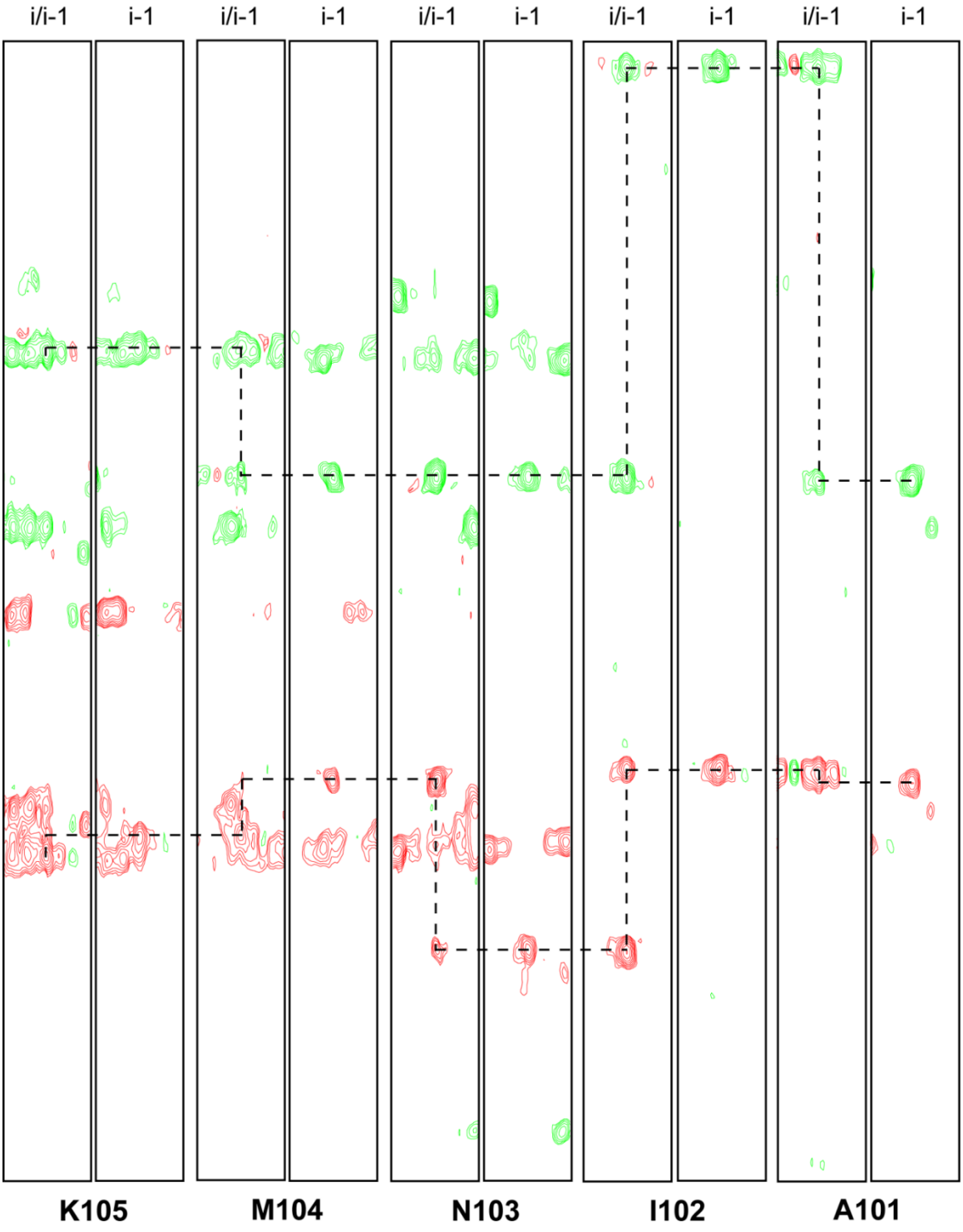
Supplementary Figure 3.2. Comparison of Syt I IDR at different pH values at 25 °C. Both NMR samples contained 60 μM IDR, 1 mM DTT, 1 mM EGTA and 7% v/v D_2O in a buffer consisting of 20 mM MOPS and 100 mM KCl. Note loss of cross peaks in sample whose pH is physiological.



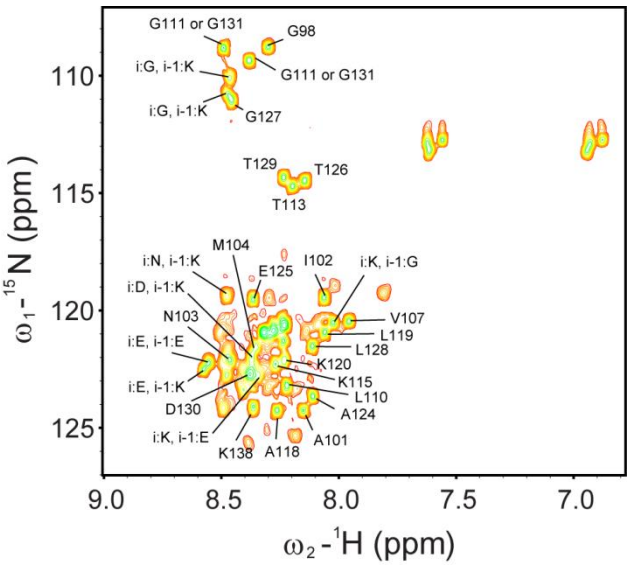
Supplementary Figure 3.3. Comparison of Syt I IDR peptide (encoding residues 81-157) binding to synaptic vesicle mimic LUVs. Co-sedimentation of Syt I IDR with LUVs was performed to assess impact of pH. After a 30 minute incubation of 15 µM Syt I IDR with increasing concentrations of LUVs, samples were spun down at 72,000 rpm in a TLA 100 rotor for 1 hour at 22 °C. Depletion of IDR from the supernatant was then used to assess binding. Shown on the left are representative gel images from four or three replicate co-sedimentation assays. On the right are the corresponding binding curves derived from each (plotted as the average of all replicates). The K_d for IDR binding to synaptic LUVs at pH 6 and 7.4 were 138±32 and 117±84, respectively.



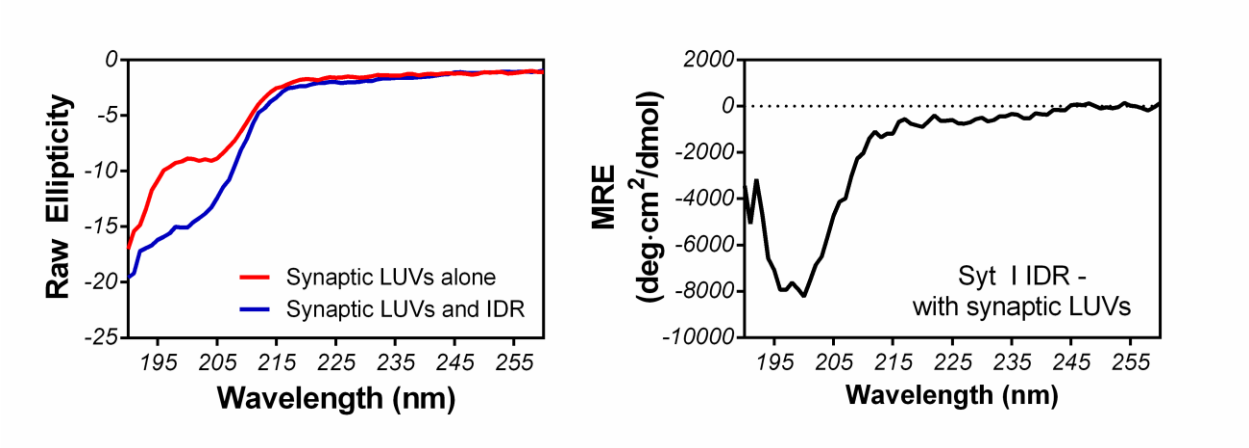
Supplementary Figure 3.4. Example HNCACB strip plots showing a walk along the peptide backbone.



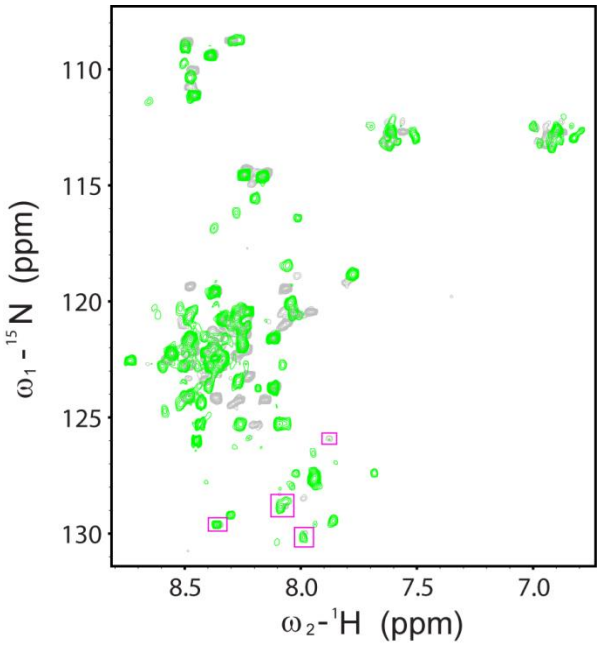
Supplementary Figure 3.5. Example HNCACB and HNCOCACB stip plots showing a walk along the peptide backbone.



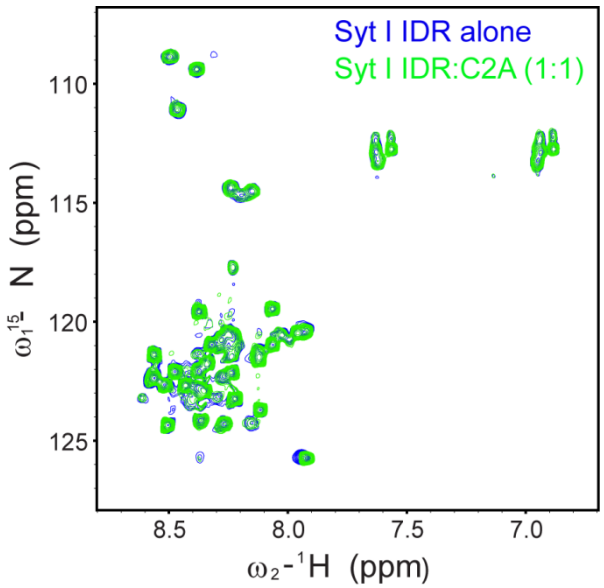
Supplementary Figure 3.6. Peaks that could and could not be unambiguously assigned.



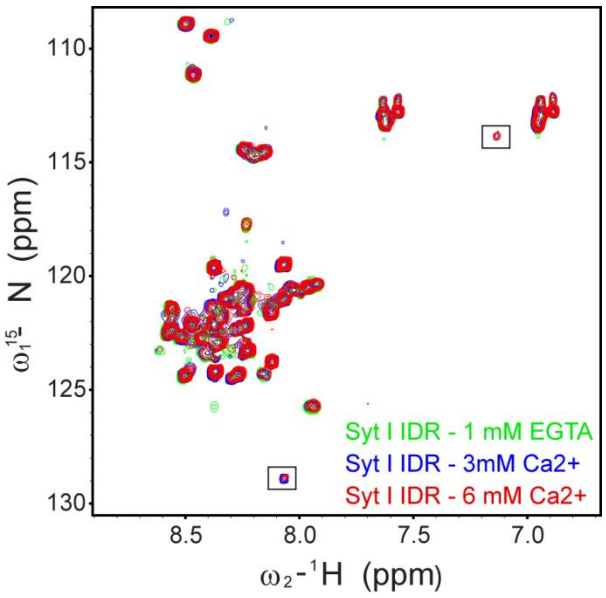
Supplementary Figure 3.7. Circular dichroism of Syt I IDR in the presence of synaptic vesicle mimic LUVs. On the left is the absorption spectrum for 1 mM synaptic mimic LUVs alone (red) and the raw, uncorrected absorption of 1 mM synaptic mimic LUVs and 15 μM Syt I IDR. The lipids absorb strongly at shorter wavelengths, likely due to high prevalence of polyunsaturated acyl chains. When corrected for the synaptic mimic LUVs (spectrum on the right), the Syt I IDR still retains significant disorder when membrane-associated. The buffer for these experiments consisted of 10 mM sodium phosphate at a pH of 7.4.



Supplementary Figure 3.8. Comparison of juxta-membrane linker with (green) and without (gray) 3 mM brain lipids at 25 °C. Note that magenta boxes indicate the cross peaks that are the same between linker alone after potential onset of aggregation and linker with synaptic vesicle SUVs.



Supplementary Figure 3.9. HSQC comparison of Syt I IDR with and without short C2A at 25 °C. The Syt I IDR alone was at a concentration of 60 μM and the sample containing short C2A contained 60 μM of both IDR and C2A for a 1:1 molar ratio. Both samples contained 1 mM DTT, 1 mM EGTA and 7% v/v D2O in a buffer consisting of 20 mM MOPS and 100 mM KCl adjusted to a pH of 7.4.



Supplementary Figure 3.10. HSQC comparison of Syt I IDR with and without Ca^{2+} at 25 °C. The Syt I IDR alone (green spectrum) was at a concentration of 60 µM and contained 1 mM DTT, 1 mM EGTA and 7% v/v D2O in a buffer consisting of 20 mM MOPS and 100 mM KCl adjusted to a pH of 7.4. Ca^{2+}-containing samples were of identical concentration, but in place of 1 mM EGTA had either 3 mM Ca^{2+} (blue spectrum) or 6 mM Ca^{2+} (red spectrum).

Chapter 4: Structural and Thermodynamic Changes in the First Actin-Binding Domain of Dystrophin

Experimental work was done by the author and Michael E. Fealey. The author conducted and analyzed the DSC experiments and the molecular dynamics simulations. Michael E. Fealey conducted and analyzed the time-resolved FRET and DEER experiments.

4.1 Introduction

Dystrophin (Dys) is a large muscle protein of 427 kDa. Structurally, Dys consists of an N-terminal actin-binding domain (termed ABD1), 24 spectrin-like repeats that form the rod domain and house the second actin binding domain (ABD2), and a C-terminal region containing a cysteine-rich domain that binds the dystroglycan complex. The N- and C-terminal ends of the protein provide a biochemical picture of how Dys can link the costameric F-actin network to both the sarcolemmal membrane and extracellular matrix via the dystroglycan complex. This location in the cytoskeletal network of muscle is crucial for maintaining membrane integrity, a fact that is proved of vital importance in the context of Duchenne and Becker muscular dystrophies (DMD and BMD, respectively). In DMD and BMD, a multitude of mutations (missense and nonsense alike) in Dys lead to the development of muscle tissue degeneration. Patients with these forms of muscular dystrophy have elevated cytosolic calcium levels, indicative of an extracellular leakage of the ion through the sarcolemma due to focal membrane tears. Similarly, Dys-deficient model organisms, particularly *mdx* mice, show the same elevated cytosolic calcium phenotype.[99] The biophysical mechanism by which Dys protects the sarcolemma from mechanical stress is still an active research question. Dys contains several spectrin-like repeat domains, triple helix bundles, between the N-terminal ABD1 and C-terminal ZZ domain that anchors it to the dystroglycan complex. Atomic force microscopy has been

used to unfold these types of protein domains from spectrin and found they have a low energy barrier to unfolding which likely serves as a mechanical means to dissipate transduced force propagating through the costamere.[100] However, independent of such a mechanism is the impact Dys has on the actin cytoskeleton alone. Previously it was shown that the large scale bending and twisting motions of F-actin are greatly restricted when both Dys and its homologous partner utrophin (Utr) bind.[101-102] Intriguingly, amplitude of these large scale motions is reduced, but the rate of motion is increased indicating that both Dys and Utr binding impart resilience to the actin cytoskeleton. Such an effect could be another means by which the two proteins dampen the laterally transduced force of contraction within the costamere.

While both Dys and Utr were found to restrict the amplitude of motion, the proteins were found to do so differentially with Utr being the more restrictive of the two proteins. How such functional differences arise are unclear, but they likely arise from differences in the structural binding modes of each protein's actin-binding domains. Previously, the Utr ABD1 was shown to undergo a conformational opening upon association with F-actin; Utr ABD1's adjacent calponin-homology domains (CH) transitioned from a more closed compact state to an open extended state forming a well ordered and stable complex with F-actin. We believe that the Dys ABD1 undergoes a similar structural transition as Utr ABD1, but with distinct regulatory nuances that could contribute to a less stable Dys ABD1-F-actin complex and, in turn, a smaller restrictive contribution to overall filament motions.

To test this hypothesis (Figure 4.1), we used a combination of biophysical techniques to elucidate the Dys ABD1 mode of binding including the pulsed EPR method DEER, time-resolved FRET, molecular dynamics simulations, and differential scanning calorimetry. We also conducted molecular dynamics simulations on Utr ABD1 to provide microscopic

insight into the current experimental findings. By applying these methods in parallel, we were able to define a mechanism of Dys ABD1 binding that is distinct from Utr and a possible source of reduced actin filament restriction.

4.2 Methods

4.2.1 Protein Expression, Purification, and Labeling

A plasmid encoding the mouse Dys ABD1 protein fragment (residues 8-246) was transformed into BL21 DE3 *E. coli* and subsequently growth to an OD of 0.7 at which point protein expression was induced with 2 mM IPTG per liter of growth. Expression was allowed to occur overnight at room temperature. After expression, cells were pelleted at 10,000 rpm. Cells were subsequently lysed in 50 mM Tris pH 8, 20% w/v sucrose, 1 mM EDTA and 1 mM DTT by incubation with lysozyme for 1 hour at 4 °C followed by freeze-thaw cycling in a dry ice isopropanol bath. The resulting lysate was then incubated with DNase I for 1 hour at 4 °C. All cell debris was removed from solubilized protein by centrifugation at 18,500 rpm in a SS-34 Sorvall rotor and was subsequently sterile filtered through a 0.2 µm filter to ensure complete removal of particulate. The lysate was then loaded onto a GE 5 mL SP ion-exchange column and eluted over a linear gradient (0-500 mM) of NaCl buffered by 10 mM Tris pH 8, 1 mM EGTA, and 1 mM DTT. Elution fractions containing the ABD1 fragment were then concentrated and further purified using an S100 sepharose gel-filtration column. Protein fractions pooled after size exclusion were then dialyzed in buffer containing 10 mM Tris, 100 mM NaCl, 2 mM $MgCl_2$ and 1 mM DTT at pH 8, buffer conditions under which actin-bundling does not occur.[103] Final protein purity was verified by SDS-PAGE. A Dys ABD1 construct containing identical residues as

described above but with an additional TEV-cleavable His tag was also generated for differential scanning calorimetry (DSC) experiments. This construct was expressed and purified in a similar manner to the Dys ABD1 construct used for spectroscopy except lysis was carried out using sonication and a Ni NTA column was used to purify the protein from the initial lysate. For the Dys ABD1 construct used in DSC experiments, protein samples were dialyzed into a solution of 20 mM MOPS and 100 mM KCl adjusted to a pH of 7.5 because MOPS-buffered solutions do not undergo as large of a pH change (compared to Tris) with the change in temperature that occurs during a DSC experiment.

For Dys ABD1 protein used in DEER experiments, a 5-fold excess (to cysteine residue concentration) of spin label was added to the protein sample in which DTT was removed and allowed to incubate for 3 hours at room temperature. Free label was removed by 4 rounds of dialysis in 4 liters of solution. The resultant labeled protein was spin counted to determine the label concentration which, when compared to the protein concentration, resulted in ~80% labeling efficiency. Spin labeled samples were then loaded into quartz capillary tubes (1.1 mm ID, 1.6 mM OD, 15 µL sample volume) and flash frozen in liquid nitrogen and stored at -80 °C until use. In the case of Dys ABD1 samples for FRET, donor dye was first added to Dys ABD1 in a 1.8-fold excess of cysteine and allowed to incubate at room temperature for 2 hours. Excess donor was removed by dialysis after which a 4-fold access of acceptor dye was added to the sample.

4.2.2 Double Electron-Electron Paramagnetic Resonance (DEER)

We performed DEER experiments to measure inter-probe distances in the range of 2-6 nm. Measurements were made on a Bruker E580 spectrometer (Billerica, MA). The resulting DEER waveform was analyzed using the model-independent Tikhonov fit. The

Tikhonov distribution was then fit to multiple Gaussian distributions as recently described[104].

$$\rho(r) = 1/(\sigma\sqrt{2\pi}) \exp((r-R)^2/(2\sigma^2)) \quad (4.1)$$

$$\sigma = (FWHM)/(2\sqrt{\ln(2)}) \quad (4.2)$$

4.2.3 Differential Scanning Calorimetry (DSC)

DSC experiments were performed on a NanoDSC (TA Instruments, New Castle, DE) at a scan rate of 1 °C/minute using micromolar protein concentrations. Buffers consisted of 20 mM MOPS, 100 mM KCl at a pH of 7.5. Solutions were extensively degassed under vacuum with gentle stirring prior to loading into the calorimeter. This helps prevent release of air bubbles from solution during the experiment. From the thermodynamics parameters enthalpy of unfolding (ΔH), melting temperature (T_m), and change in baseline heat capacity (ΔC_p), we calculated a free energy of unfolding (ΔG) using the Gibbs-Helmholtz equation:

$$\Delta G = \Delta H(1-T/T_m) + \Delta C_p(T-T_m-T\ln(T/T_m) \quad (4.3)$$

4.2.4 Time-resolved FRET

Time-resolved FRET was performed on a custom-built spectrometer that measures direct waveforms. Waveforms are described as a sum of exponential functions:

$$FD(t) = \sum_i^n Aie^{-\tau it} \quad (4.4)$$

Collected waveforms were fit to multi-exponential functions using an in house program called FargoFit as described previously. The donor-only sample was first fit to exponential decay models of increasing complexity until there was no significant improvement in the χ^2 value of the fit. The resultant amplitude and lifetime decay parameters of the multi-exponential fit were then fixed and used as starting parameters (in addition to the weighted R_0) to fit donor-acceptor samples. In donor-acceptor samples, only the inter-probe distance (R) and full width half-max (FWHM) were allowed to vary in the fit.

4.2.5 Molecular Dynamics Simulations

Simulations were conducted on four different constructs: wild type Dys ABD1 (wt-Dys ABD1), dys ABD1 with the hydrophobic patch residues mutated (mh-Dys ABD1), dys ABD1 in which the hydrophobic patch residues and selected charged residues were mutated (mhc-Dys ABD1), and wild type Utr ABD1 (Utr ABD1). From the crystal structure of Dys ABD1 (PDB 1DXX[105], residues 9-246) a single chain, with the positions of crystal waters preserved, was selected for the CHARMM36[106] and CHARMM22STAR[107] wt-Dys ABD1 simulation starting structure. The mh-Dys ABD1 structure was created from the wt-Dys structure with the following residues being mutated to glutamines: V120, V123, M124, I127, M128, A129, I139, L140, F236, and V243. The mhc-Dys ABD1 structure was created from the mh-Dys ABD1 structure with the following residues being mutated to lysine: D182, E197, and E216. All mutations were made using PyMOL and simulations for the mutated structures were only conducted using CHARMM36. For all Dys ABD1 structures the protonation state of histidine residues 36, 72, 104, 117, 174, and 198 were set to neutral with a proton on the Nε. From a single chain in the crystal structure of Utr ABD1

(PDB 1QAG[108]) the structure used in the Utr-ABD1 simulations was developed. Using PyMOL the Utr-ABD1 chain was extended to comprise residues 25-261. Prior to simulation using CHARMM36 and CHARMM22STAR the protonation state of histidine residues 29, 101, 120, 133, 192, and 213 were set to neutral with a proton on the Nε and the protonation state of histidine residues 88 and 190 were set to neutral with a proton on the Nδ.

Periodic boundary conditions were used in all simulations as the structures were placed in 11x11x80 nm boxs, solvated with TIP3P[109] water, brought to electroneutrality and provided KCl salt at an ionic strength of 150 mM. GROMACS 5.0.6[110] was used for both setup and productions with H-bonds constrained with the LINCS[111] algorithm. Equations of motion were propagated with a time step of 2.0 fs and neighbor searching was performed every 40 steps. Electrostatic interactions were treated with the particle-mesh Ewald[112] algorithm with a real cut-off of 1.0 nm and a grid of approximately 0.1 nm. Van der Waals interactions were switched off between 1.2 nm and 1.0 nm. A constant temperature of 300.0 K was maintained with the V-rescale algorithm[113] applied to the protein and solvent groups independently. A constant pressure of 1 atm was maintained with the Parrinello-Rahman barostat[114].

The equilibration of the system for each force field consisted of a minimization simulation that concluded when the maximum force was less than 1000.0 kl mol^{-1} nm^{-1} or after 50000 steps. Following the minimization, a constant volume and temperature simulation was conducted lasting 1ns. Prior to starting the replicate simulations, a constant pressure and temperature simulation was conducted where the force constant was stepped down to zero over the course of 3ns. Five separate systems were generated using each force field, totaling 10 simulations, from the endpoint structure of the constant pressure and

temperature simulation for both wt-DYS ABD1 and Utr ABD1. For the mh-Dys ABD1 and mhc-Dys ABD1 five separate systems were generated using the CHARMM36 force field. In each system atom velocities were randomly assigned using the Maxwell-Boltzmann distribution. Each system was simulated for 500ns while saving the trajectories every 5ps. For analysis, a total of 5μs of simulation time was analyzed for wt-Dys ABD1 and Utr ABD1 and a total of 2.5μs of simulation times was analyzed for mh-Dys ABD1 and mhc-Dys ABD1 .

To determine the main motions in the conformational dynamics during the trajectory, principal component analysis (PCA)[115] was carried out for the combined trajectories. Using GROMACS the mass-weighted covariance matrix was calculated, diagonalized and the Eigen vectors were analyzed to produce the first and second principal component. Additionally, the distance between the alpha-carbon atoms of residues 120 and 239 was measured using the molecular visualization program Visual Molecular Dynamics[116]. This distance was measured for the entire 5 microsecond combined trajectory. Solvent accessible surface area (SASA)[117] was measured for the total protein, the hydrophobic residues, and the hydrophilic residues, and each residue individually using GROMACS. The differences in SASA between the open conformation of ABD1 and the closed conformations were plotted on the closed conformation's representative structure using PyMOL.

4.3 Results

4.3.1 Dys ABD1's structure transitions from a closed to open state, but with significant structural disorder

To test the hypothesis that the CH domains of Dys ABD1 undergo a conformational opening upon binding of F-actin, we performed DEER experiments on a spin labeled

ABD1 fragment. Application of DEER allows us to make a direct comparison to existing structural data on the homologous Utr ABD1 fragment.[118] Using site-directed mutagenesis, we removed all native cysteine residues in the Dys ABD1 construct and engineered in new cysteines at residue positions 120 and 239. Based on the domain-swapped dimer crystal structure previously determined for Dys ABD1[105], these residue locations would be amenable to detecting a change in inter-domain distance upon F-actin association. After labeling the protein with spin labels, we measured inter-probe distance with and without F-actin. In the absence of F-actin, DEER measurements identified a short inter-probe distance, indicative of the CH1 and CH2 domains of ABD1 being in a more compact closed state (Figure 4.2, top). This is apparent in the DEER echo amplitude that shows rapid decay and mild oscillation in the time domain, both of which are features resulting from short inter-probe distance. In samples containing F-actin, the DEER echo amplitude decays more slowing. The oscillations initially present in the absence of F-actin are dampened out in F-actin containing samples indicative of significant structural disorder (Figure 4.2, middle and lower panels).

When comparing these structural results to the previously examined Utr ABD1, several features stand out. First, both ABDs seem to share a common mechanism of opening upon binding to F-actin. Second, the Dys ABD1 contains much more structural disorder than seen previously in Utr ABD1. This is evident in the mild or complete lack of oscillations present in the echo amplitude decay which, in Utr ABD1, were very well resolved. Lastly, unlike the Utr ABD1, Dys ABD1 does not undergo a complete shift to the open structural state. This is evident by the fact that a shorter inter-probe distance persists even in the presence of a four-fold molar excess of actin. Collectively the DEER data identifies a common binding mode for Dys and Utr ABD1, but also highlights differences that may represent distinct structural nuances in the ABD-actin interaction.

4.3.2 The Structural Disorder of Dys ABD1 Stems from an Intrinsic Instability

To further investigate the binding mode of Dys ABD1, particularly the structural disorder indicated by DEER, we subjected the Dys ABD1 fragment to thermal denaturation using DSC. One explanation for the conformational heterogeneity seen spectroscopically is protein instability. Given that DEER measurements indicate a significant amount of structural disorder, we would anticipate a low thermodynamic stability of the protein.

To test the hypothesis that Dys ABD1 is intrinsically unstable, we performed denaturation experiments using differential scanning calorimetry (Figure 4.3). In these experiments, heat capacity of the protein colloidal suspension is measured as a function of increasing temperature. As the protein unfolds, the measured heat capacity rises and falls giving rise to an excess heat capacity peak that can be integrated to determine the enthalpy of unfolding (ΔH).[64-119] The temperature at which half this integrated area is reached is the melting temperature (T_m). The difference in baseline heat capacity (ΔC_p) is a measure of the change in hydrophobic residue exposure. In the native state, hydrophobic residues are buried in the core of the protein and are thereby protected from solvent. In the unfolded state, however, the hydrophobic residues are exposed to solvent. To compensate for the energetically unfavorable exposure of hydrophobic residues to polar solvent, water molecules form highly ordered water cages around exposed residues to limit their interaction with bulk solvent. These ordered water cages have a slightly higher heat capacity relative to water molecules surrounding a folded protein and, as a result, elevate the baseline heat capacity of the unfolded state of the protein. Empirically, the ΔCp parameter has been shown to be linearly related to the molecular weight of a protein.[22]

When these thermodynamic parameters are measured for Dys ABD1, we can make two important conclusions. First, Dys ABD1 is a weakly held-together protein indicated by the fact that its enthalpy of unfolding (which is a direct measure of the intramolecular interactions holding a folded domain together) and the free energy of unfolding are low for a protein of similar molecular weight. The 2 kcal/mole unfolding free energy is ~1/5 that of other similar molecular weight proteins whose stability has been determined with DSC.[120-121] This marginal stability is not uncommon and has been seen in other systems.[35, 122] In relation to the spectroscopic data measured above, this marginal stability is consistent with the above DEER measurements showing that each distance distribution had a large degree of structural disorder (Figure 4.2). Such a stability-structure correlation also exists for the protein synaptotagmin 1. Like Dys ABD1, synaptotagmin 1 contains two adjacent domains (C2 domains) that are weakly held together.[123] Just as in Dys ABD1, inter-C2 domain distances measured by DEER indicate significant structural disorder with essentially featureless decays.[124] Second, the ABD1 protein has a notably larger change in baseline heat capacity (Figure 3). This observation suggests that in the folded-to-unfolded transition there is a larger change in exposure of hydrophobic residues than that predicted based on molecular weight alone.[120]

4.3.3 TR-FRET Converges on a Closed-to-Open Transition of Dys ABD1

While the DEER data indicates an open-to-closed structural transition, there is another possible explanation for the longer inter-probe distance, namely the actin-repeat; two adjacent ABD1 molecules could be associated with the filament in a more compact structural state but be in close enough proximity that the spin labels experience inter-molecular dipolar interactions. To rule this possibility out, we performed a complementary

TR-FRET experiment which requires significantly lower protein concentrations that would prevent such an inter-molecular probe interaction from occurring.

Using the same cysteine labeling sites that were utilized in the DEER experiments above, we attached the dyes IAEDNAS and fluorescein to Dys ABD1 and made the analogous structural measurements. Compared to donor-only labeled Dys ABD1, the fluorescence lifetime of AEDANS is shorted in donor-acceptor labeled Dys ABD1. This shortened lifetime is indicative of energy transfer and a more closed compact structural state of the CH domains (Figure 4.4). When these fluorescence decays are fit to multiexpoential functions, the inter-probe distance parameters that best describe the data converge on an average distance of 2.0 nm and σ = 1.8. When a sample containing Dys ABD1 and F-actin at a 1:100 molar ratio is subjected to the same measurement, an increase in lifetime is observed indicating the CH domains of Dys ABD1 transition to a more open structural state (Figure 4.4) where the inter-probe distance becomes 3.0 nm and a σ = 2.5. In both methods, there appears to be a longer distance that likely corresponds to a more open structural state of the ABD1 fragment. However, just as in DEER, the distance distributions are broad. This likely stems from structural disorder but may also arise in part from the fluorescent probes themselves which have longer linkers between site of thiol attachment and the probe itself.

4.3.4 Molecular Dynamics Simulations Indicate Hydrophobic Contacts Stabilize the Closed Structural States of Dys ABD1

While both DEER and TR-FRET support the structural model of a closed-to-open transition upon binding F-actin, neither provide atomic-level molecular mechanisms that explain Dys ABD1's conformational preferences. In an attempt to gain atomic-level insight,

we performed molecular dynamics simulations on the Dys ABD1 crystal structure and examined the free energy landscape of the protein. In this experiment, a monomer from the domain-swapped dimer crystal structure[105] was used as the starting point of the simulation.

During the course of the simulations, the wt-Dys ABD1 structures sampled multiple closed conformations with the CH domains interacting. Principal component analysis determined the two main motions of the structures during the course of the simulations to be the collapsing of the CH domains into a closed compact structural state (PC1) and the rotation of the CH domains about each other (PC2). Plotting PC1 vs. PC2 and using the Boltzmann weighting function shows a relatively flat free energy surface when in the closed conformation and there is an energy barrier that needs to be crossed to reach the open conformation (Figure 4.5). Measurement of the distance between the alpha-carbon of V120 and L239 was determined over the course of all wt-Dys as a comparison to the DEER and TR-FRET measurements (Figure 4.6). In all but one simulation the structures converged on a closed compact conformation. Calculation of the probability density for the distance over the course of all wt-Dys ABD1 simulations show the presence of a partially open conformation. However, measurement of the distance between the center of mass of the CH domains displays that indeed the conformations sampled during the simulations can be summarized as a closed compact state and an open state (Figure 4.7). The fact that numerous closed conformers were possible is consistent with both DEER and TR-FRET measurements of a structurally heterogeneous closed state, validating the MD simulations and allowing us to examine the atomic details of these simulated conformers.

Upon closer inspection, all closed structural states of the wt-Dys ABD1 undergo a decrease in solvent accessible area as compared to the open conformation (Figure 4.8,

left). In each of these structures, it seems that there is a gain of hydrophobic-hydrophobic contacts between CH domains or CH1 and the ABD1 linker region occurs. Mapping the difference between the solvent accessible surface area (SASA) per residue for the open and the closed conformation shows a series of hydrophobic residues that undergo a decrease in SASA in the closed conformation (Figure 4.9). These hydrophobic residues (V120, V123, M124, I127, M128, I139, F236, and V243) form a hydrophobic patch that is present at least in partiality in the array of closed compact conformations. Such a pattern could indicate that the closed structural states are stabilized in part by interactions involving hydrophobic side chains. A similar mechanism was recently proposed for the ABD of βIII-spectrin.125

To test the presumption that hydrophobic residues stabilize the closed structural states we conducted a series of simulations in which the hydrophobic residues found to be in the hydrophobic patch along with other selected hydrophobic residues around the patch were mutated to glutamines (mh-Dys ABD1). Though we hypothesized that the structures in the simulations would move towards and stay in an open conformation, this was not observed. To test to see if the closing of the structures during the simulations was due to electrostatic interactions we mutated residues within the second calponin homology domain (CH2) to create a net positive charge. Residues D182, E197, and E216 were mutated to lysine. This structure, mhc-Dys ABD1, then had a net positive charge on the first calponin homology domain (CH1), the linker region and CH2. Analysis of the simulations using radius of gyration shows that when the hydrophobic patch is mutated the simulations still sample closed conformations similar to those sampled in the wt-Dys ABD1 simulations, however, the conformations sampled were not as compact. When the charged residues are also mutated the simulations sample a range of partially open conformations (Figure 4.10).

4.3.5 Molecular Dynamics Simulations Display a Difference in the Conformational Ensemble of Utrophin ABD1 as Compared to Dystrophin ABD1

Previous studies have been done observing the structural states of Utr ABD1. To gain atomistic insight into these states and make comparisons to Dys ABD1 simulations were conducted on a modified crystal structure of Utr ABD1. The modification entailed adding amino acids to the crystal structure to better compare to the construct in which the experimental work had been done.

As with Dys ABD1, PCA was conducted on the simulations of Utr ABD1 to determine the main structural motions of the system during the simulation. These results were consistent with Dys ABD1 in the manner that the first principal component described the closing of the structure into a closed conformation and the second principal component described the rotation of the CH domains about each other. Through plotting PC1 vs. PC2 and using the Boltzmann weighting function (Figure 4.11) the free energy landscape of the simulations was observed. Similar to Dys ABD1 there are both open and closed conformations. However, the total space that the energy landscape takes up is less for Utr ABD1 than what was observed for Dys ABD1. In addition, there is no distinct energy barrier to cross to reach the open conformation and there is a partially open conformation that was highly sampled (Figure 4.11, group 3). Calculation of the total solvent accessible surface area for the highly sample conformations determined through the PCA analysis shows that in Utr ABD1 does not reach the level of compactness that Dys ABD1 reached in its simulations (Figure 4.8, Right). In further support of the Utr ABD1 closed conformation not being as compact is the comparison of the CH center of mass measurement for the simulations (Figure 4.12). Although both Dys and Utr ABD1 sampled

open and closed states, in general the CH domain center of mass distance was greater for Utr ABD1 than for Dys ABD1. Another observed difference between the simulations that can be seen in Figure 4.12 is the presence of a shoulder in the curve for the closed conformation. This is indicative of the presence of a highly sampled partially closed conformation.

4.4 Discussion

The diverse biophysical methods employed in this study were geared towards identifying and understanding underlying molecular mechanisms for structural transitions in Dys's ABD1. We found that, much like Dys's homologue utrophin, the ABD1 can exist in two main structural states: a compact state where the CH domains are in close apposition, and a more extended open state that is stabilized by actin. Unlike our previous study of utrophin ABD1, Dys ABD1 did not shift completely to its extended open state; a significant mole fraction of ABD1 remained in a more compact state even at actin concentration 20 times K_d (Figure 4.2 and Figure 4.4). This indicates that the equilibrium between closed and open states more strongly favors the closed state when compared to the previously studied utrophin ABD1. Why does Dys ABD1 prefer a compact state, even in the presence of excess actin? DSC data and MD simulations point to a potential mechanism.

As discussed above, the change in baseline heat capacity of a protein is largely a measure of the change in solvent accessibility of hydrophobic residues. The magnitude of this change was shown empirically to be linearly related to protein molecular weight. When a globular protein's change in baseline heat capacity deviates from its calculated value, it can provide additional macroscopic information about the protein. In our case, the larger

than expected change in baseline heat capacity indicates that there are more hydrophobic residues protected in the native folded state which creates a larger relative change in the amount of ordered water which subsequently elevates the unfolded baseline heat capacity value (Figure 4.3). This is supported by the fact that both spin and FRET label inter-probe distances indicate a compact state of the protein. This is also consistent with previous pyrene excimer experiments using different labeling sites.

The other important feature of the data presented comes from the MD-detected closure of ABD1 as an energy-minimized conformational state. If Dys ABD1's preferred conformational state is a compact closed state so that it can protect hydrophobic residues, such a mechanism could be relevant to Dys's function overall. If we consider Dys in the full context of a contracting and relaxing myocyte, we can imagine that the forces exerted on the protein during relaxation could change the conformation to a more extend state, not unlike the open state of ABD1 as measured here. If Dys needs to return to its compact structural state as part of its structural function, a restoration mechanism would likely be required. In this case the energetically unfavorable exposure of hydrophobic residues could provide an energetic driving force that allows it to re-adopt its closed conformation. Between the two types of structural change, namely F-actin induced opening and hydrophobic residue exposure-induce condensation, a potential regulatory mechanism emerges for governing structural change that underlies function.

In the case of the previously studied utrophin ABD1 wherein there was a large and complete opening of tandem CH domains, it was proposed that structural opening would enable better restriction of rotational motion of F-actin. This, in turn, could be a mechanism related to diffusing lateral force experienced at the sarcolemmal during muscle contraction and relaxation. Dys did not restrict-actin rotational dynamics to the same extent as

utrophin. The reduced opening of Dys's tandem CH domains measured here correlates with this functional hypothesis and could be a contributing factor to the mechanism of reduced restriction of F-actin dynamics. Additionally, when considering the molecular dynamics simulations, the high degeneracy of closed structural states suggest that even when Dys ABD1 is associated with F-actin, conformational heterogeneity may allow the CH domains to occupy more compact states afforded by the large number of hydrophobic-hydrophobic contacts in a way that doesn't result in steric clashes with the actin filament. Recent work suggesting that the CH1 domain of Dys ABD1 is primarily responsible for the F-actin interaction supports such a claim (Mallela, Biochemistry 2015).

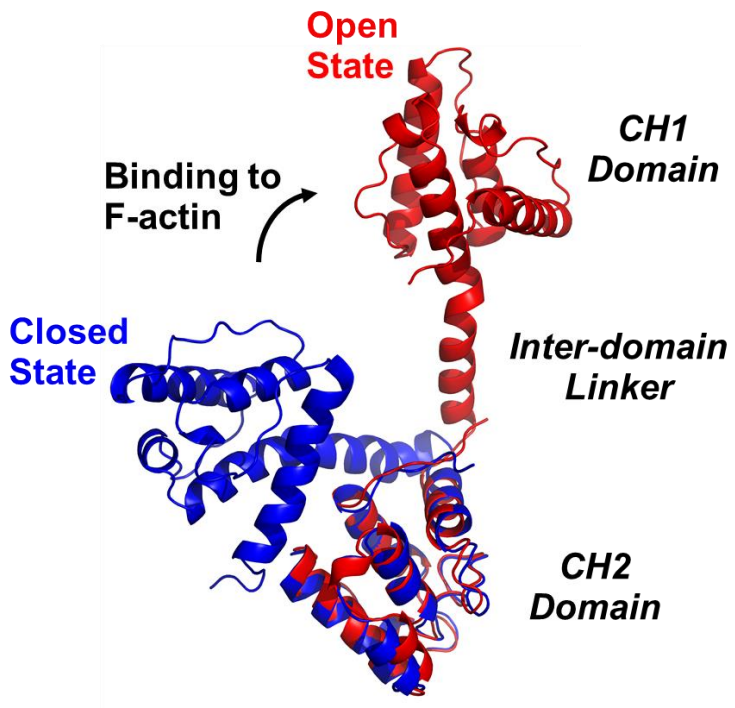


Figure 4.1. Proposed structural model for Dys ABD1 upon binding F-actin. In the absence of actin, the two adjacent CH domains are closely packed (closed state). In the presence of actin, CH1 and CH2 become more separated (open state).

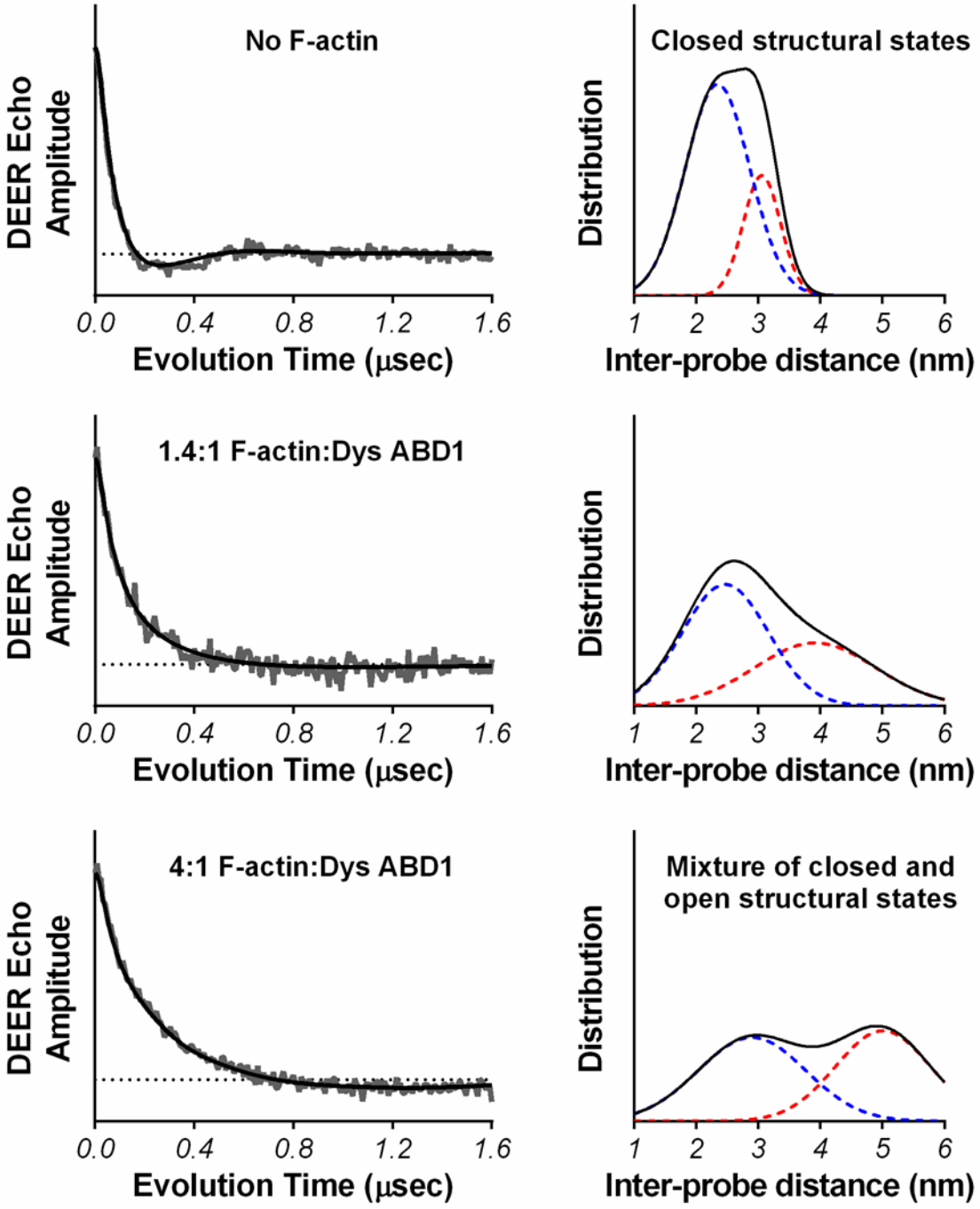


Figure 4.2. DEER data on 80μM MSL-labeled Dys ABD1 in the presence of increasing F-actin. Left: time-domain decays. Right: derived distance distributions. Tikhonov distributions indicate F-actin shifts inter-probe distance within Dys ABD1 toward a more open structural state.

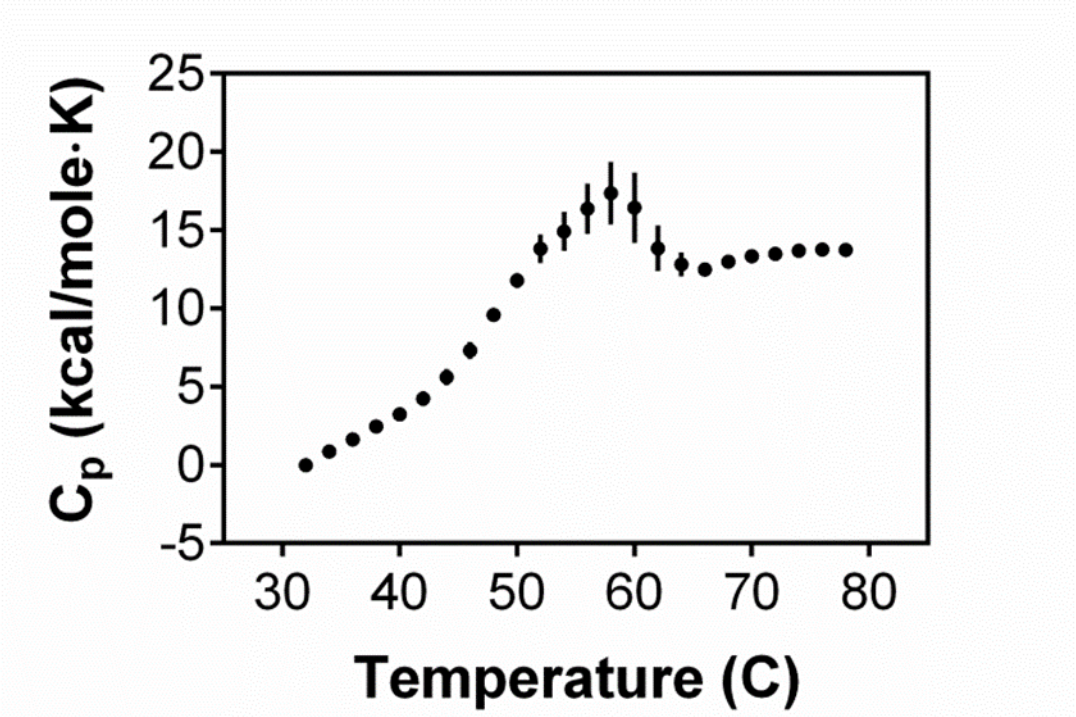


Figure 4.3. DSC denaturation of Dys ABD1. A 10 μM sample of Dys ABD1 was denatured in 20 mM MOPS, 100 mM KCl at a pH of 7.5. Plotted is the average and standard deviation of 3 replicate endotherms.

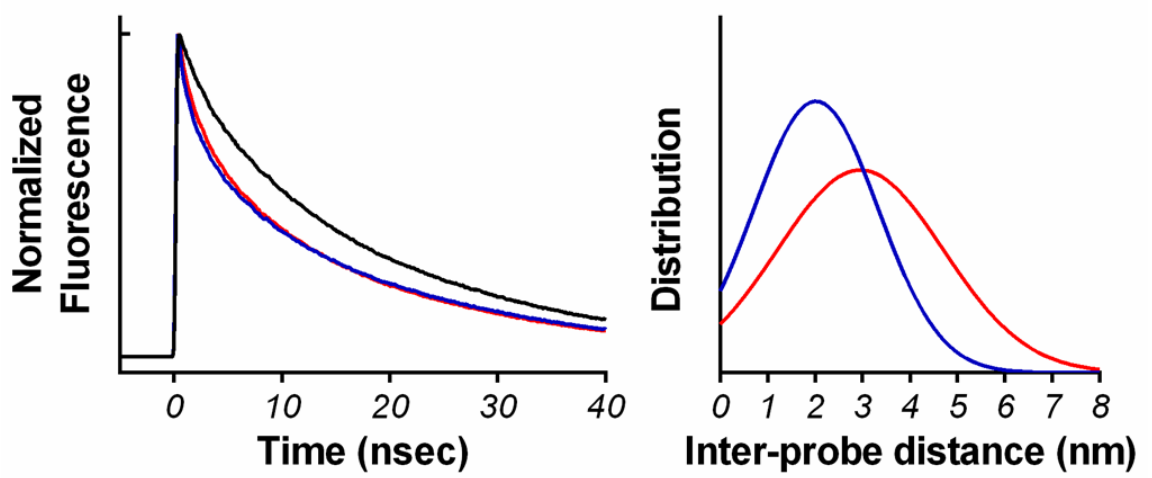


Figure 4.4. TR-FRET measurement of closed-to-open structural transition of Dys ABD1. (Top left) AEDANS-fluoroscein dye pair detection of closed state of Dys ABD1. (Top right) AEDANS-Fluoroscein detection of open state of Dys ABD1.

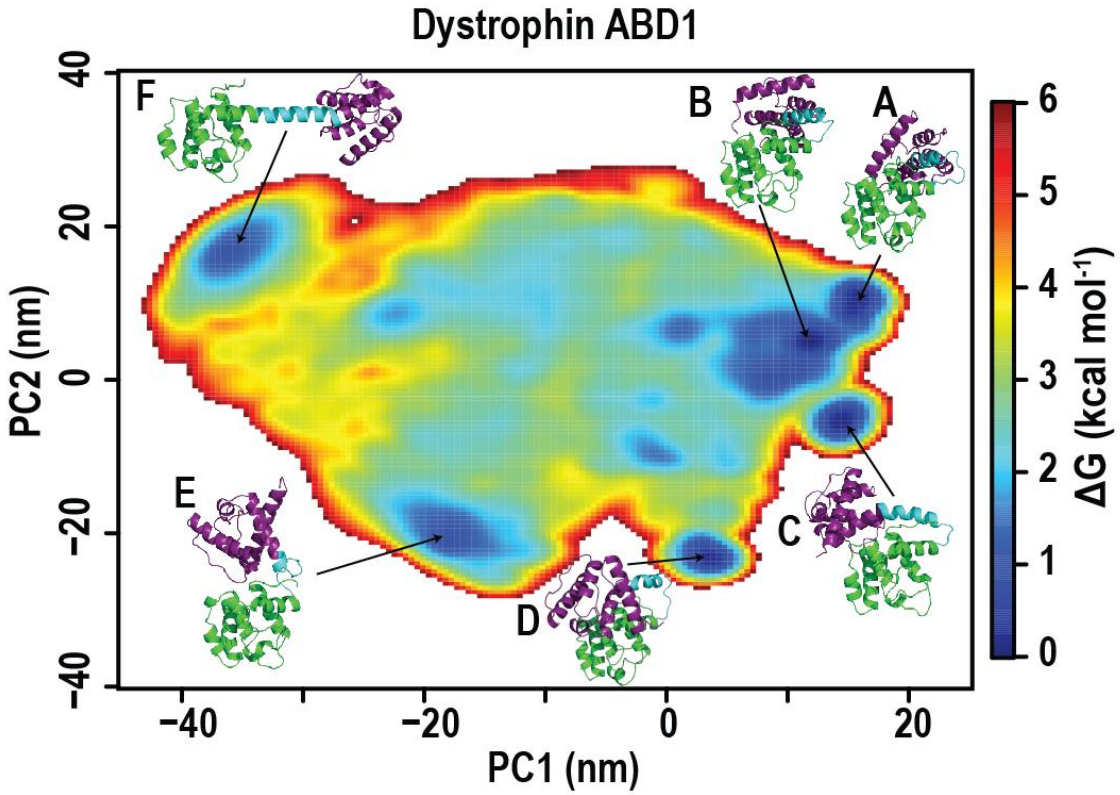


Figure 4.5. Two-dimensional plot of principal component 1 and principal component 2 derived from the Principal component analysis of wt-Dys ABD1 simulations. The free energy landscape was calculated using a Boltzmann weighting function where the lowest energy levels correspond to conformations that were highly sampled during the simulations. Structures A-E are representative structures for the highly sample closed conformations and structure F is a representative structure of the open conformation. In each structure the first CH domain is shown in purple, the linker region in blue and the second CH domain in green.

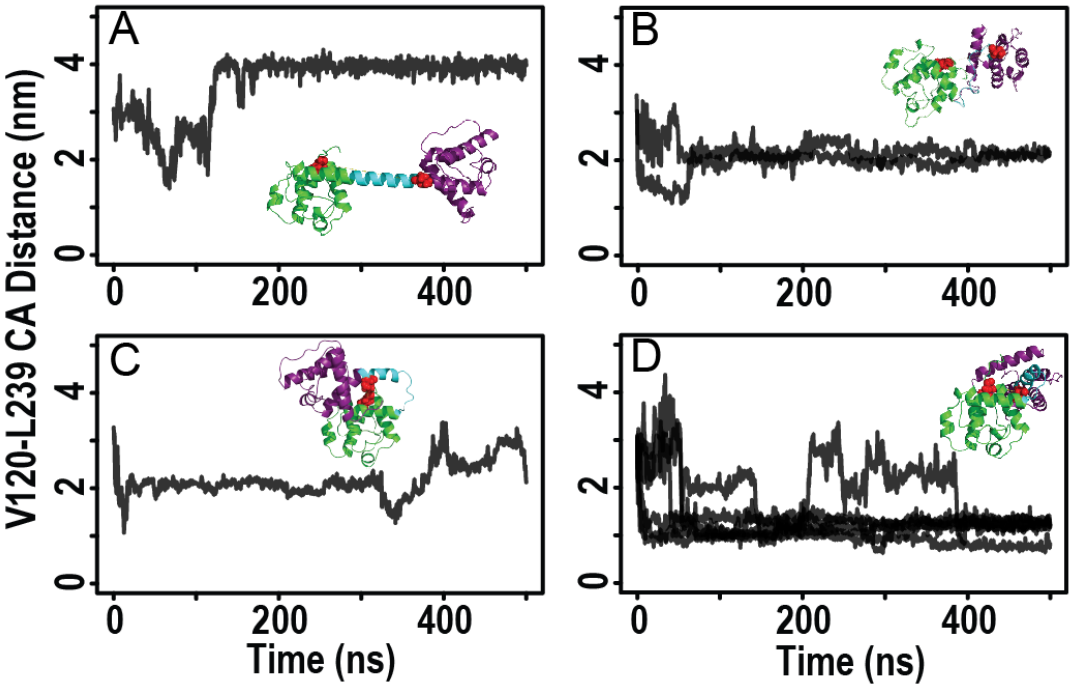


Figure 4.6. Measurement of the distance between the alpha-carbons of V120 and L239 (shown in red) of wt-Dys ABD1 during the course of the simulations. A single simulation (A) converged at an open conformation while the other nine simulations (B-D) converged within a closed compact conformation.

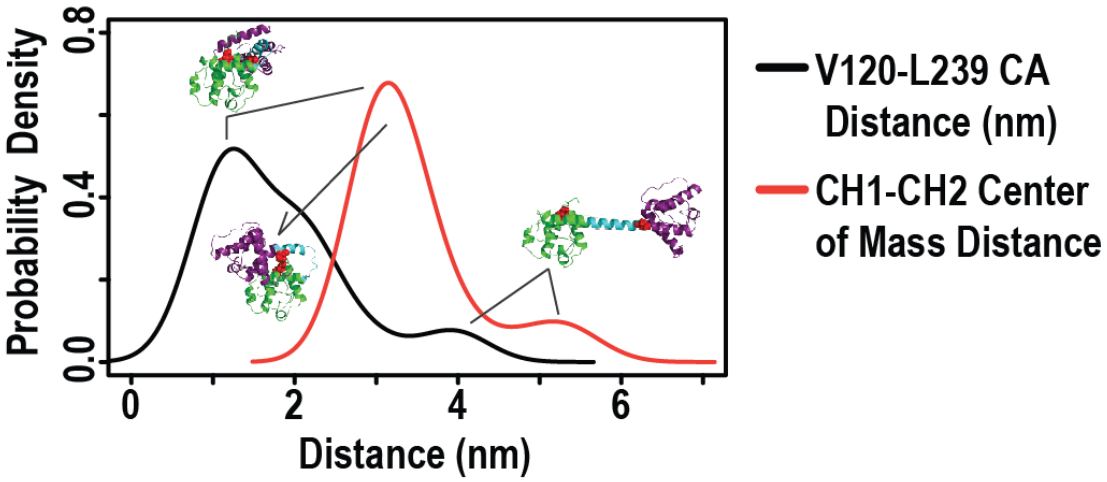


Figure 4.7. Probability density of the V120-L239 alpha-carbon distance (black) and the CH domain center of mass distance (red). Measurement of the V120-L239 alpha carbon distance shows a small shoulder around 2nm indicating a partially open conformation. However, this is an artifact of the location of V120 and L239 within the structure. When the distance between the center of mass for the CH domains is measured, the probability density plot shows that the conformations sampled during the simulations can be summarized as closed compact and open.

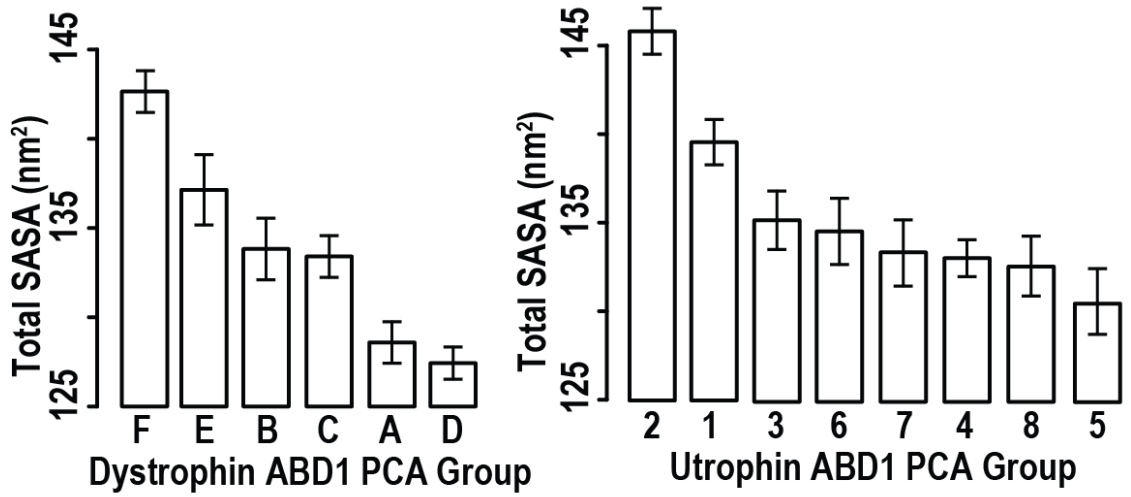


Figure 4.8. Total solvent accessible surface area (SASA) for the highly sampled conformations determined from the principal component analysis for wt-Dys ABD1 (Left) and Utr ABD1 (Right). Total SASA was calculated for every structure in the group and then averaged. Error bars represent one standard deviation.

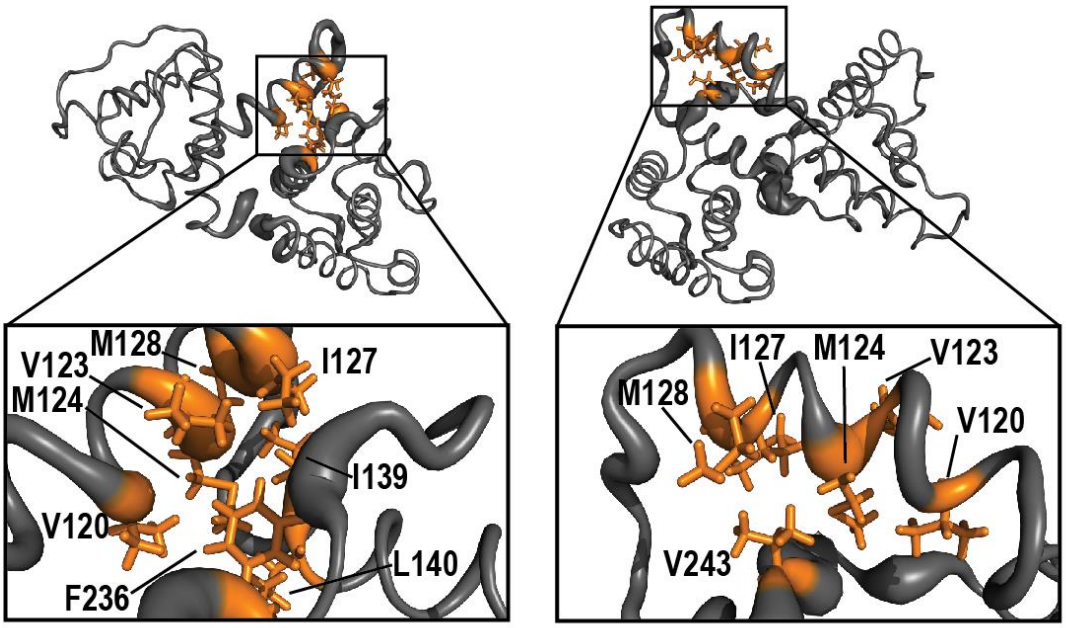


Figure 4.9. Dystrophin ABD1 hydrophobic patch. A hydrophobic patch develops in the closed conformation of wt-Dys ABD1. Mapping the difference in the SASA per residue over the structure shows a decrease in SASA for the hydrophobic residues highlighted in orange as compared to the open conformation. The two structures are representative of the differing closed conformations and the orange residues are representative of those found to have a decrease in SASA for all highly sampled closed conformations.

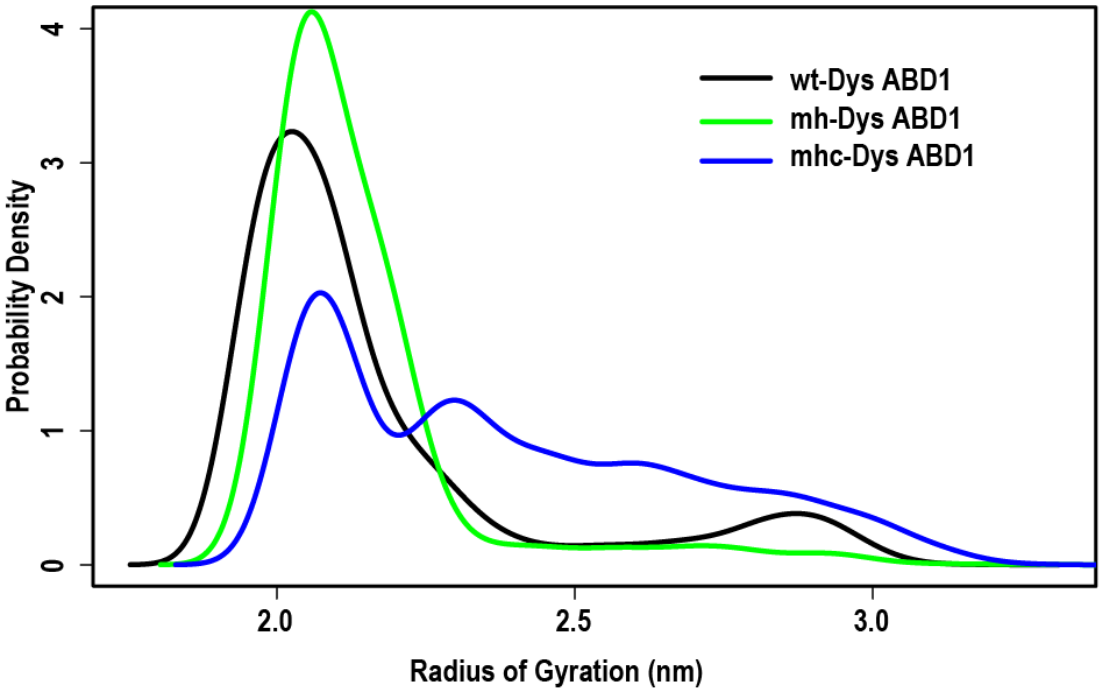


Figure 4.10. Analysis of the radius of gyration for the wild type Dys ABD1 (black), the Dys ABD1 construct in which the hydrophobic patch was mutated to glutamine residues, (green), and the Dys ABD1 construct which in addition to the hydrophobic residues being mutated, the net charge of the CH2 domain was changed by mutating three negatively charged residues to lysine residues. The probability density was calculated for the entirety of the simulation.

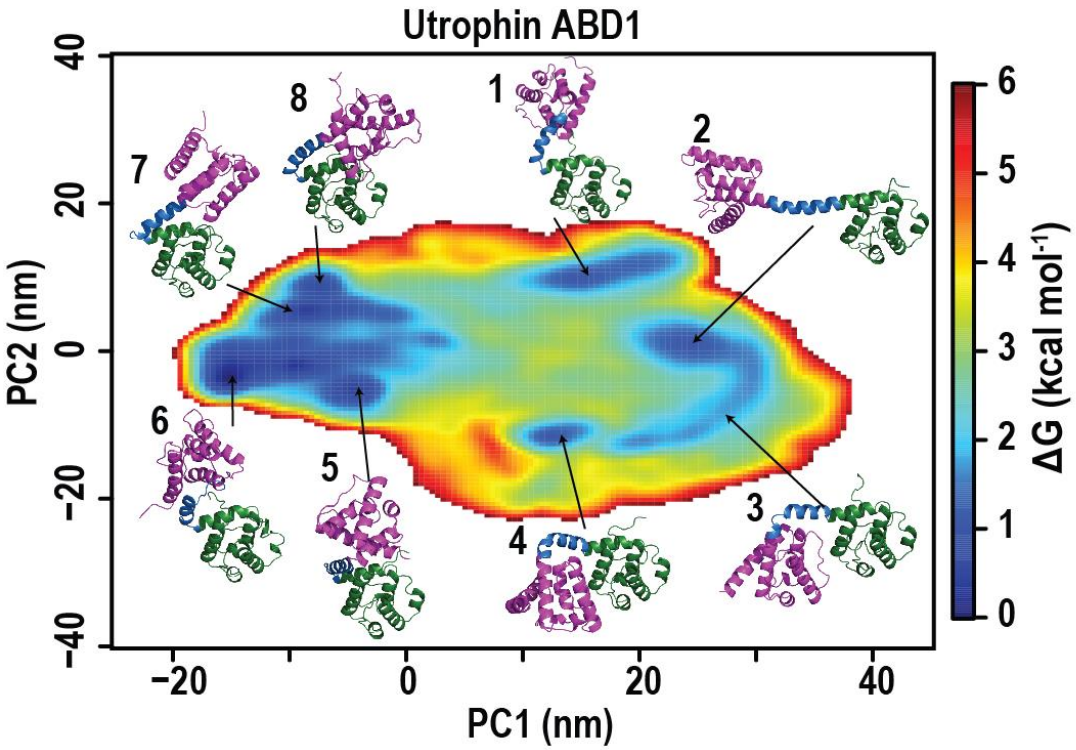


Figure 4.11. Two-dimensional plot of principal component 1 and principal component 2 derived from the Principal component analysis of Utr ABD1 simulations. The free energy landscape was calculated using a Boltzmann weighting function where the lowest energy levels correspond to conformations that were highly sampled during the simulations. Structures 1 and 4-8 are representative structures for the highly sample closed conformations, structure 3 is a representative structure for the partially open conformation and structure 2 is a representative structure of the open conformation. In each structure the first CH domain is shown in purple, the linker region in blue and the second CH domain in green.

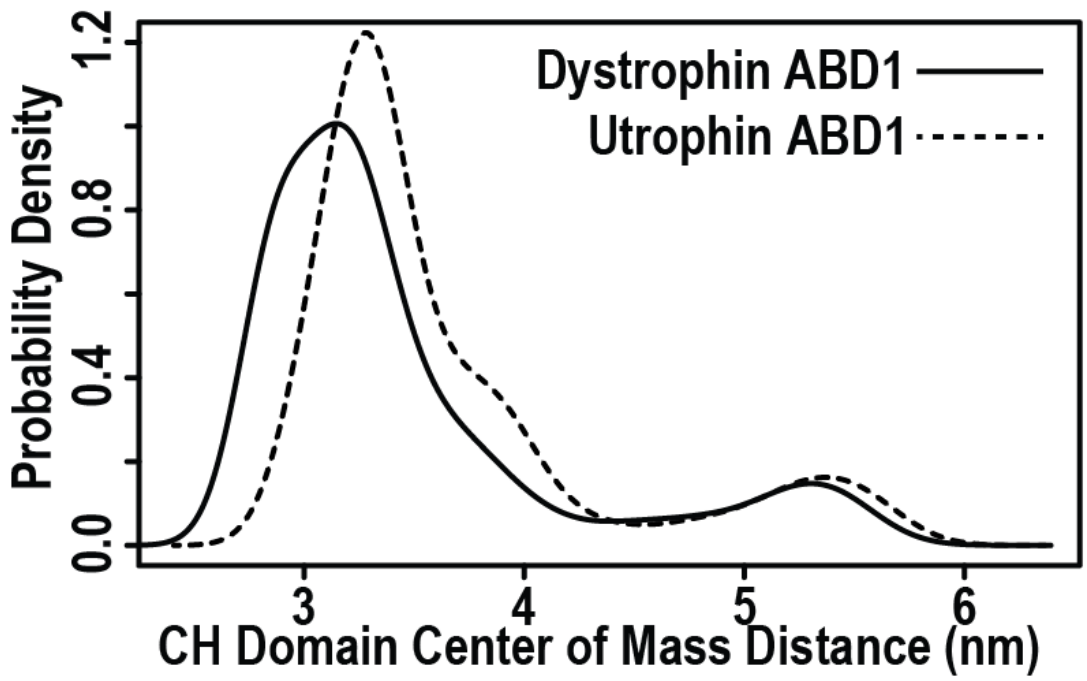


Figure 4.12. Calponin homology domain center of mass distance comparison for Dys ABD1 (solid line) and Utr ABD1 (dashed line). Note the shift in the density distribution of Utr ABD1 indicating a less compact structure and the presence of a shoulder for Utr ABD1 near 3.75 nm indicating a partially open conformation that was not observed for Dys ABD1.

References

1. Valley, C.C. et al. The Methionine-aromatic Motif Plays a Unique Role in Stabilizing Protein Structure. J Biol Chem 287, 34979-91 (2012).
2. Zauhar, R.J., Colbert, C.L., Morgan, R.S. & Welsh, W.J. Evidence for a strong sulfur-aromatic interaction derived from crystallographic data. Biopolymers 53, 233-248 (2000).
3. Reid, K.S.C., Lindley, P.F. & Thornton, J.M. Sulphur-aromatic interactions in proteins. FEBS Letters 190, 209-213 (1985).
4. Pranata, J. Sulfur–Aromatic Interactions: A Computational Study of the Dimethyl Sulfide–Benzene Complex. Bioorganic Chemistry 25, 213-219 (1997).
5. Morgan, R.S. & McAdon, J.M. Predictor for sulfur-aromatic interactions in globular proteins. International journal of peptide and protein research 15, 177-180 (1980).
6. Morgan, R.S., Tatsch, C.E., Gushard, R.H., McAdon, J. & Warme, P.K. Chains of alternating sulfur and pi-bonded atoms in eight small proteins. International journal of peptide and protein research 11, 209-217 (1978).
7. Viguera, A.R. & Serrano, L. Side-chain interactions between sulfur-containing amino acids and phenylalanine in α-helices. Biochemistry 34, 8771-8779 (1995).
8. Shechter, Y., Burstein, Y. & Patchornik, A. Selective oxidation of methionine residues in proteins. Biochemistry 14, 4497-503 (1975).
9. Hoshi, T. & Heinemann, S.H. Regulation of cell function by methionine oxidation and reduction. Journal of Physiology-London 531, 1-11 (2001).
10. Levine, R.L., Berlett, B.S., Moskovitz, J., Mosoni, L. & Stadtman, E.R. Methionine residues may protect proteins from critical oxidative damage. Mech Ageing Dev 107, 323-32 (1999).
11. Kim, G., Weiss, S.J. & Levine, R.L. Methionine oxidation and reduction in proteins. Biochimica et biophysica acta 1840, 901-5 (2014).
12. Arakawa, T., Kita, Y. & Timasheff, S.N. Protein precipitation and denaturation by dimethyl sulfoxide. Biophysical Chemistry 131, 62-70 (2007).
13. Rose, G.D., Geselowitz, A.R., Lesser, G.J., Lee, R.H. & Zehfus, M.H. HYDROPHOBICITY OF AMINO-ACID RESIDUES IN GLOBULAR-PROTEINS. Science 229, 834-838 (1985).
14. Dougherty, D.A. Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science 271, 163-8 (1996).
15. Bogan, A.A. & Thorn, K.S. Anatomy of hot spots in protein interfaces. Journal of Molecular Biology 280, 1-9 (1998).
16. Gottlieb, H.E., Kotlyar, V. & Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. J Org Chem 62, 7512-7515 (1997).
17. Butterfield, S.M., Patel, P.R. & Waters, M.L. Contribution of aromatic interactions to alpha-helix stability. J Am Chem Soc 124, 9751-5 (2002).
18. Greenfield, N.J. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat Protoc 1, 2527-35 (2006).
19. Marqusee, S. & Baldwin, R.L. Helix stabilization by Glu-...Lys+ salt bridges in short peptides of de novo design. Proc Natl Acad Sci U S A 84, 8898-902 (1987).

20. Pearson, K. LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2, 559-572 (1901).
21. Yin, D., Kuczera, K. & Squier, T.C. The sensitivity of carboxyl-terminal methionines in calmodulin isoforms to oxidation by H(2)O(2) modulates the ability to activate the plasma membrane Ca-ATPase. Chemical research in toxicology 13, 103-10 (2000).
22. Sacksteder, C.A. et al. Tertiary structural rearrangements upon oxidation of Methionine145 in calmodulin promotes targeted proteasomal degradation. Biophys J 91, 1480-93 (2006).
23. Gao, J., Yao, Y. & Squier, T.C. Oxidatively modified calmodulin binds to the plasma membrane Ca-ATPase in a nonproductive and conformationally disordered complex. Biophys J 80, 1791-801 (2001).
24. Gao, J. et al. Loss of conformational stability in calmodulin upon methionine oxidation. Biophysical journal 74, 1115-34 (1998).
25. Boschek, C.B., Jones, T.E., Smallwood, H.S., Squier, T.C. & Bigelow, D.J. Loss of the calmodulin-dependent inhibition of the RyR1 calcium release channel upon oxidation of methionines in calmodulin. Biochemistry 47, 131-42 (2008).
26. Bartlett, R.K. et al. Oxidation of Met144 and Met145 in calmodulin blocks calmodulin dependent activation of the plasma membrane Ca-ATPase. Biochemistry 42, 3231-8 (2003).
27. Babu, Y.S., Bugg, C.E. & Cook, W.J. Structure of calmodulin refined at 2.2 A resolution. J Mol Biol 204, 191-204 (1988).
28. Kuboniwa, H. et al. Solution structure of calcium-free calmodulin. Nat Struct Biol 2, 768-76 (1995).
29. Mukherjea, P., Maune, J.F. & Beckingham, K. Interlobe communication in multiple calcium-binding site mutants of Drosophila calmodulin. Protein Sci 5, 468-77 (1996).
30. Anbanandam, A. et al. Mediating molecular recognition by methionine oxidation: conformational switching by oxidation of methionine in the carboxyl-terminal domain of calmodulin. Biochemistry 44, 9486-96 (2005).
31. Jeschke, G. DEER Distance Measurements on Proteins. Annual Review of Physical Chemistry, Vol 63 63, 419-446 (2012).
32. Ilardi, E.A., Vitaku, E. & Njardarson, J.T. Data-Mining for Sulfur and Fluorine: An Evaluation of Pharmaceuticals To Reveal Opportunities for Drug Design and Discovery. Journal of Medicinal Chemistry 57, 2832-2842 (2014).
33. Jimenez, M.A., Munoz, V., Rico, M. & Serrano, L. HELIX STOP AND START SIGNALS IN PEPTIDES AND PROTEINS - THE CAPPING BOX DOES NOT NECESSARILY PREVENT HELIX ELONGATION. Journal of Molecular Biology 242, 487-496 (1994).
34. Vasquez, M., Pincus, M.R. & Scheraga, H.A. HELIX COIL TRANSITION THEORY INCLUDING LONG-RANGE ELECTROSTATIC INTERACTIONS - APPLICATION TO GLOBULAR-PROTEINS. Biopolymers 26, 351-371 (1987).
35. Fealey, M.E. et al. Negative Coupling as a Mechanism for Signal Propagation between C2 Domains of Synaptotagmin I. Plos One 7, 11 (2012).
36. Horovitz, A. Double-mutant cycles: A powerful tool for analyzing protein structure and function. Folding & Design 1, R121-R126 (1996).

37. Horovitz, A., Serrano, L., Avron, B., Bycroft, M. & Fersht, A.R. STRENGTH AND COOPERATIVITY OF CONTRIBUTIONS OF SURFACE SALT BRIDGES TO PROTEIN STABILITY. Journal of Molecular Biology 216, 1031-1044 (1990).
38. Serrano, L., Bycroft, M. & Fersht, A.R. AROMATIC AROMATIC INTERACTIONS AND PROTEIN STABILITY - INVESTIGATION BY DOUBLE-MUTANT CYCLES. Journal of Molecular Biology 218, 465-475 (1991).
39. Balog, E.M., Lockamy, E.L., Thomas, D.D. & Ferrington, D.A. Site-specific methionine oxidation initiates calmodulin degradation by the 20S proteasome. Biochemistry 48, 3005-16 (2009).
40. Klein, J.C. et al. Actin-binding cleft closure in myosin II probed by site-directed spin labeling and pulsed EPR. Proceedings of the National Academy of Sciences of the United States of America 105, 12867-72 (2008).
41. Etemadi, N. et al. Lymphotoxin α induces apoptosis, necroptosis and inflammatory signals with the same potency as tumor necrosis factor. FEBS J 280, 5283-97 (2013).
42. Frisch, M.J. et al. Gaussian 09, Revision D.01. (Wallingford CT, 2009).
43. Boys, S.F. & Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. Molecular Physics 19, 553-566 (1970).
44. Zhao, Y. & Truhlar, D.G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. Journal of Chemical Physics 125(2006).
45. Reed, A.E., Curtiss, L.A. & Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. Chemical Reviews 88, 899-926 (1988).
46. Mulliken, R.S. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I. The Journal of Chemical Physics 23, 1833-1840 (1955).
47. Cramer, C.J. Essentials of computational chemistry: theories and models, (John Wiley & Sons, 2013).
48. Phillips, J.C. et al. Scalable molecular dynamics with NAMD. J Comput Chem 26, 1781-802 (2005).
49. Banner, D.W. et al. Crystal structure of the soluble human 55 kd TNF receptor-human TNF beta complex: implications for TNF receptor activation. Cell 73, 431-45 (1993).
50. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. Journal of Chemical Physics 79, 926-935 (1983).
51. MacKerell, A.D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. Journal of Physical Chemistry B 102, 3586-3616 (1998).
52. MacKerell, A.D., Feig, M. & Brooks, C.L. Improved treatment of the protein backbone in empirical force fields. Journal of the American Chemical Society 126, 698-699 (2004).
53. Best, R.B. et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. J Chem Theory Comput 8, 3257-3273 (2012).

54. Martyna, G.J., Tobias, D.J. & Klein, M.L. Constant-Pressure Molecular-Dynamics Algorithms. Journal of Chemical Physics 101, 4177-4189 (1994).
55. Jas, G.S. & Kuczera, K. Free-energy simulations of the oxidation of C-terminal methionines in calmodulin. Proteins: Structure, Function and Genetics 48, 257-268 (2002).
56. Darden, T., Perera, L., Li, L.P. & Pedersen, L. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. Structure with Folding & Design 7, R55-R60 (1999).
57. Tuckerman, M., Berne, B.J. & Martyna, G.J. Reversible Multiple Time Scale Molecular-Dynamics. Journal of Chemical Physics 97, 1990-2001 (1992).
58. Andersen, H.C. Rattle - a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. Journal of Computational Physics 52, 24-34 (1983).
59. Sudhof, T. C. (2004) The synaptic vesicle cycle, Annu Rev Neurosci 27, 509-547.
60. Lai, Y., Lou, X., Jho, Y., Yoon, T. Y., and Shin, Y. K. (2013) The synaptotagmin 1 linker may function as an electrostatic zipper that opens for docking but closes for fusion pore opening, Biochem J 456, 25-33.
61. Caccin, P., Scorzeto, M., Damiano, N., Marin, O., Megighian, A., and Montecucco, C. (2015) The synaptotagmin juxtamembrane domain is involved in neuroexocytosis, FEBS Open Bio 5, 388-396.
62. Lee, J., and Littleton, J. T. (2015) Transmembrane tethering of synaptotagmin to synaptic vesicles controls multiple modes of neurotransmitter release, Proc Natl Acad Sci U S A 112, 3793-3798.
63. Luque, I., Leavitt, S. A., and Freire, E. (2002) The linkage between protein folding and functional cooperativity: two sides of the same coin?, Annu Rev Biophys Biomol Struct 31, 235-256.
64. Freire, E. (1995) Differential scanning calorimetry, Methods Mol Biol 40, 191-218.
65. Vacic, V., Uversky, V. N., Dunker, A. K., and Lonardi, S. (2007) Composition Profiler: a tool for discovery and visualization of amino acid composition differences, BMC Bioinformatics 8, 211.
66. Chen, J. W., Romero, P., Uversky, V. N., and Dunker, A. K. (2006) Conservation of intrinsic disorder in protein domains and families: II. functions of conserved disorder, J Proteome Res 5, 888-898.
67. Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., and Dunker, A. K. (2001) Sequence complexity of disordered protein, Proteins 42, 38-48.
68. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, Nat Rev Mol Cell Biol 6, 197-208.
69. Gauer, J. W., Sisk, R., Murphy, J. R., Jacobson, H., Sutton, R. B., Gillispie, G. D., and Hinderliter, A. (2012) Mechanism for calcium ion sensing by the C2A domain of synaptotagmin I, Biophys J 103, 238-246.
70. Rice, A. M., Mahling, R., Fealey, M. E., Rannikko, A., Dunleavy, K., Hendrickson, T., Lohese, K. J., Kruggel, S., Heiling, H., Harren, D., Sutton, R. B., Pastor, J., and Hinderliter, A. (2014) Randomly organized lipids and marginally stable proteins: a coupling of weak interactions to optimize membrane signaling, Biochim Biophys Acta 1838, 2331-2340.

71. Lu, B., Kiessling, V., Tamm, L. K., and Cafiso, D. S. (2014) The juxtamembrane linker of full-length synaptotagmin 1 controls oligomerization and calcium-dependent membrane binding, J Biol Chem 289, 22161-22171.
72. Uversky, V. N. (2015) Unreported intrinsic disorder in proteins: disorder emergency room, Intrinsically Disordered Proteins 3, 1-15.
73. Hilser, V. J., and Thompson, E. B. (2007) Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins, Proc Natl Acad Sci U S A 104, 8311-8315.
74. Kingsley, P. B., and Feigenson, G. W. (1979) The synthesis of a perdeuterated phospholipid: 1,2 dimyristoyl-sn-glycero-3-phosphocholine-d72, Chem. Phys. Lipids. 24.
75. Iakoucheva, L. M., Kimzey, A. L., Masselon, C. D., Smith, R. D., Dunker, A. K., and Ackerman, E. J. (2001) Aberrant mobility phenomena of the DNA repair protein XPA, Protein Sci 10, 1353-1362.
76. McMullen, T. P., Lewis, R. N., and McElhaney, R. N. (1993) Differential scanning calorimetric study of the effect of cholesterol on the thermotropic phase behavior of a homologous series of linear saturated phosphatidylcholines, Biochemistry 32, 516-522.
77. Gauer, J. W., Knutson, K. J., Jaworski, S. R., Rice, A. M., Rannikko, A. M., Lentz, B. R., and Hinderliter, A. (2013) Membrane modulates affinity for calcium ion to create an apparent cooperative binding response by annexin a5, Biophys J 104, 2437-2447.
78. Kertz, J. A., Almeida, P. F., Frazier, A. A., Berg, A. K., and Hinderliter, A. (2007) The cooperative response of synaptotagmin I C2A. A hypothesis for a Ca2+-driven molecular hammer, Biophys J 92, 1409-1418.
79. Ubach, J., Zhang, X., Shao, X., Sudhof, T. C., and Rizo, J. (1998) Ca2+ binding to synaptotagmin: how many Ca2+ ions bind to the tip of a C2-domain?, Embo J 17, 3921-3930.
80. Takamori, S., Holt, M., Stenius, K., Lemke, E. A., Gronborg, M., Riedel, D., Urlaub, H., Schenck, S., Brugger, B., Ringler, P., Muller, S. A., Rammner, B., Grater, F., Hub, J. S., De Groot, B. L., Mieskes, G., Moriyama, Y., Klingauf, J., Grubmuller, H., Heuser, J., Wieland, F., and Jahn, R. (2006) Molecular anatomy of a trafficking organelle, Cell 127, 831-846.
81. Tamiola, K., Acar, B., and Mulder, F. A. (2010) Sequence-specific random coil chemical shifts of intrinsically disordered proteins, J Am Chem Soc 132, 18000-18003.
82. Lee, S. H., Cha, E. J., Lim, J. E., Kwon, S. H., Kim, D. H., Cho, H., and Han, K. H. (2012) Structural characterization of an intrinsically unfolded mini-HBX protein from hepatitis B virus, Mol Cells 34, 165-169.
83. Rucker, A. L., and Creamer, T. P. (2002) Polyproline II helical structure in protein unfolded states: lysine peptides revisited, Protein Sci 11, 980-985.
84. Xu, C., Gagnon, E., Call, M. E., Schnell, J. R., Schwieters, C. D., Carman, C. V., Chou, J. J., and Wucherpfennig, K. W. (2008) Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif, Cell 135, 702-713.
85. Haxholm, G. W., Nikolajsen, L. F., Olsen, J. G., Fredsted, J., Larsen, F. H., Goffin, V., Pedersen, S. F., Brooks, A. J., Waters, M. J., and Kragelund, B. B. (2015) Intrinsically disordered cytoplasmic domains of two cytokine receptors mediate conserved interactions with membranes, Biochem J 468, 495-506.

86. Papayannopoulos, V., Co, C., Prehoda, K. E., Snapper, S., Taunton, J., and Lim, W. A. (2005) A polybasic motif allows N-WASP to act as a sensor of PIP(2) density, Mol Cell 17, 181-191.
87. Santner, A. A., Croy, C. H., Vasanwala, F. H., Uversky, V. N., Van, Y. Y., and Dunker, A. K. (2012) Sweeping away protein aggregation with entropic bristles: intrinsically disordered protein fusions enhance soluble expression, Biochemistry 51, 7250-7262.
88. Li, J., Motlagh, H. N., Chakuroff, C., Thompson, E. B., and Hilser, V. J. (2012) Thermodynamic dissection of the intrinsically disordered N-terminal domain of human glucocorticoid receptor, J Biol Chem 287, 26777-26787.
89. Torrecillas, A., Laynez, J., Menendez, M., Corbalan-Garcia, S., and Gomez-Fernandez, J. C. (2004) Calorimetric study of the interaction of the C2 domains of classical protein kinase C isoenzymes with Ca2+ and phospholipids, Biochemistry 43, 11727-11739.
90. Zhang, X., Rizo, J., and Sudhof, T. C. (1998) Mechanism of phospholipid binding by the C2A-domain of synaptotagmin I, Biochemistry 37, 12395-12403.
91. Tochtrop, G. P., Richter, K., Tang, C., Toner, J. J., Covey, D. F., and Cistola, D. P. (2002) Energetics by NMR: site-specific binding in a positively cooperative system, Proc Natl Acad Sci U S A 99, 1847-1852.
92. Freire, E., Schon, A., and Velazquez-Campoy, A. (2009) Isothermal titration calorimetry: general formalism using binding polynomials, Methods Enzymol 455, 127-155.
93. Herrick, D. Z., Sterbling, S., Rasch, K. A., Hinderliter, A., and Cafiso, D. S. (2006) Position of synaptotagmin I at the membrane interface: cooperative interactions of tandem C2 domains, Biochemistry 45, 9668-9674.
94. Martens, S., and McMahon, H. T. (2008) Mechanisms of membrane fusion: disparate players and common principles, Nat Rev Mol Cell Biol 9, 543-556.
95. Baker, K., Gordon, S. L., Grozeva, D., van Kogelenberg, M., Roberts, N. Y., Pike, M., Blair, E., Hurles, M. E., Chong, W. K., Baldeweg, T., Kurian, M. A., Boyd, S. G., Cousin, M. A., and Raymond, F. L. (2015) Identification of a human synaptotagmin-1 mutation that perturbs synaptic vesicle cycling, J Clin Invest 125, 1670-1678.
96. Bykhovskaia, M. (2015) Calcium binding promotes conformational flexibility of the neuronal ca(2+) sensor synaptotagmin, Biophys J 108, 2507-2520.
97. Fealey, M. E., and Hinderliter, A. (2013) Allostery and instability in the functional plasticity of synaptotagmin I, Commun Integr Biol 6, e22830.
98. Motlagh, H. N., and Hilser, V. J. (2012) Agonism/antagonism switching in allosteric ensembles, Proc Natl Acad Sci U S A 109, 4134-4139.
99. Turner, P. R., Westwood, T., Regen, C. M., and Steinhardt, R. A. (1988) Increased protein degradation results from elevated free calcium levels found in muscle from mdx mice, Nature 335, 735-738.
100. Rief, M., Pascual, J., Saraste, M., and Gaub, H. E. (1999) Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles, J Mol Biol 286, 553-561.
101. Lin, A. Y., Prochniewicz, E., Henderson, D. M., Li, B., Ervasti, J. M., and Thomas, D. D. (2012) Impacts of dystrophin and utrophin domains on actin structural dynamics: implications for therapeutic design, J Mol Biol 420, 87-98.

102. Prochniewicz, E., Henderson, D., Ervasti, J. M., and Thomas, D. D. (2009) Dystrophin and utrophin have distinct effects on the structural dynamics of actin, Proc Natl Acad Sci U S A 106, 7822-7827.
103. Orlova, A., Rybakova, I. N., Prochniewicz, E., Thomas, D. D., Ervasti, J. M., and Egelman, E. H. (2001) Binding of dystrophin's tandem calponin homology domain to F-actin is modulated by actin's structure, Biophys J 80, 1926-1931.
104. McCarthy, M. R., Thompson, A. R., Nitu, F., Moen, R. J., Olenek, M. J., Klein, J. C., and Thomas, D. D. (2015) Impact of methionine oxidation on calmodulin structural dynamics, Biochem Biophys Res Commun 456, 567-572.
105. Norwood, F. L., Sutherland-Smith, A. J., Keep, N. H., and Kendrick-Jones, J. (2000) The structure of the N-terminal actin-binding domain of human dystrophin and how mutations in this domain may cause Duchenne or Becker muscular dystrophy, Structure 8, 481-491.
106. Klauda, J. B., Venable, R. M., Freites, J. A., O'Connor, J. W., Tobias, D. J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A. D., Jr., and Pastor, R. W. (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types, J Phys Chem B 114, 7830-7843.
107. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, J Phys Chem B 102, 3586-3616.
108. Keep, N.H., Winder, S.J., Moores, C.A., Walke, S., Norwood, F.L.M., Kendrick-Jones, J. (1999) Crystal structure of the actin-binding region of utrophin reveals a head-to-tail dimer, Structure Fold. Des. 7,1539-1546.
109. Jorgensen, W. L., Chandrasekhar, J., and Madura, J. D. (1983) Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79, 926-935.
110. Abraham, M. J., Murtola, T., Schulz, R., Páll, R., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2, 19-25.
111. Hess, B., Bekker, H., Berend, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations, J. Comp. Chem. 18, 1463-1472.
112. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method, J. Chem. Phys. 103, 8577-8593.
113. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling, J Chem Phys 126, 014101.
114. Parrinello, M., and Rahman, A. J. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52, 7182-7190.
115. Amadei, A., Linssen, A. B., and Berendsen, H. J. (1993) Essential dynamics of proteins, Proteins 17, 412-425.

116. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics, J Mol Graph 14, 33-38, 27-38.
117. Eisenhaber, F., Lijnzaad, P., Sander, C., and Scharf, M. (1995) The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies, J. Comp. Chem. 16, 273-284.
118. Lin, A. Y., Prochniewicz, E., James, Z. M., Svensson, B., and Thomas, D. D. (2011) Large-scale opening of utrophin's tandem calponin homology (CH) domains upon actin binding by an induced-fit mechanism, Proc Natl Acad Sci U S A 108, 12729-12733.
119. Freire, E. (2001) The thermodynamic linkage between protein structure, stability, and function, Methods Mol Biol 168, 37-68.
120. Spolar, R. S., Livingstone, J. R., and Record, M. T., Jr. (1992) Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water, Biochemistry 31, 3947-3955.
121. Kumar, S., Tsai, C. J., and Nussinov, R. (2002) Maximal stabilities of reversible two-state proteins, Biochemistry 41, 5359-5374.
122. Privalov, P. L., and Khechinashvili, N. N. (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study, J Mol Biol 86, 665-684.
123. Fuson, K., Rice, A., Mahling, R., Snow, A., Nayak, K., Shanbhogue, P., Meyer, A. G., Redpath, G. M., Hinderliter, A., Cooper, S. T., and Sutton, R. B. (2014) Alternate splicing of dysferlin C2A confers Ca(2)(+)-dependent and Ca(2)(+)-independent binding for membrane repair, Structure 22, 104-115.
124. Herrick, D. Z., Kuo, W., Huang, H., Schwieters, C. D., Ellena, J. F., and Cafiso, D. S. (2009) Solution and membrane-bound conformations of the tandem C2A and C2B domains of synaptotagmin 1: Evidence for bilayer bridging, J Mol Biol 390, 913-923.
125. Avery, A. W., Crain, J., Thomas, D. D., and Hays, T. S. (2016) A human beta-III-spectrin spinocerebellar ataxia type 5 mutation causes high-affinity F-actin binding, Sci Rep 6, 21375.