

Stability and Strain in Hisactophilin and Mechanism of the Myristoyl Switch

by

Duncan William Scott MacKenzie

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemistry

Waterloo, Ontario, Canada, 2019

© Duncan W.S. MacKenzie 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: G. Andrew Woolley
Professor, Department of Chemistry, University of Toronto

Supervisor(s): Elizabeth Meiring
Professor, Department of Chemistry, University of Waterloo

Internal Member: Thorsten Dieckmann
Associate Professor, Department of Chemistry, University of Waterloo
Michael Palmer
Associate Professor, Department of Chemistry, University of Waterloo

Internal-External Member: Dr. Todd Holyoak
Associate Professor, Department of Biology, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Hisactophilin is a myristoylated, histidine-rich, pH-dependent actin- and membrane-binding protein. In response to cellular changes in pH, this β -trefoil protein reversibly switches between cytosolic and membrane-bound forms. A key feature of the reversible membrane-binding is the covalent acylation of the N-terminal glycine with a C₁₄ myristoyl group. At pH > 6.9, the myristoyl group favours sequestration in the barrel of the β -trefoil, whereas at pH < 6.9, the myristoyl group favours increased solvent accessibility and eventually anchors hisactophilin to the inner leaflet of the cell membrane. In *Dictyostelium discoideum*, membrane-bound hisactophilin also binds and bundles actin, contributing to cell locomotion. Despite widespread myristoylation of eukaryotic proteins, its effects on protein folding, stability, and function are still poorly understood and limits our understanding of a broader set of switches and the ability to design them. Combining equilibrium denaturation, folding kinetics, variable temperature and variable pH nuclear magnetic resonance (NMR), chemical shift perturbations, and reverse micelle encapsulation, we use hisactophilin as a model for characterizing the determinants of finely tuned myristoyl switches.

Equilibrium stability measurements identified hisactophilin mutants with broken switches—in which pH ceases influencing conformational switching—and switches with tuned sensitivities. In a few cases, stunningly small changes to amino acid side chains broke the pH-dependent myristoyl switch. Interestingly, a thermodynamic switch broken by one mutation may be repaired by making additional mutations, illustrating novel synergistic contributions to global stability and switching. Studying the mutants also revealed that a predominant effector of switching appears to be strain from an overpacked core when the myristoyl group is sequestered in the binding pocket. Altering strain through changing the geometry of the myristoyl binding pocket offers a new approach for tuning the sensitivity of hisactophilin and other switch proteins, as well as to inform future design efforts.

The temperature dependence of amide proton chemical shifts localized myristoyl-induced strain to a set of residues in wild-type hisactophilin's myristoyl-binding pocket. The mutants with broken switches (I85L, which favours the accessible state, and F6L/I85L/I93L, which favours the sequestered state), however, no longer show evidence of strain. Nonlinear temperature dependence of chemical shifts indicate that dynamics in residues that report on switching adjacent to the myristoyl group are also attenuated in broken-switch mutants. Thus, the strained residues in the protein core appear to form part of the communication network between the proton binding site(s) and the myristoyl group. Apparent pK_as of backbone amide protons for I85L and LLL obtained by NMR-monitored pH titrations further support the decoupling of residues adjacent to the myristoyl group from switching. While we have achieved a high resolution and unrivalled look at the mechanism of hisactophilin's pH-dependent myristoyl switch, we have also shown the validity and utility of chemical shift temperature dependences for characterizing small, functionally relevant local stability changes

and gaining insight to the near-native energy landscape and its relation to protein function.

Myristoyl switches participate in important cell signalling cascades by forming reversible protein-membrane or protein-protein interactions in response to environmental stimuli. Notwithstanding their prevalence, the high sensitivity and cooperativity of myristoyl switches complicates their study, resulting in poorly understood determinants and mechanisms. By employing thermodynamic measurements and developing a new, general application of an NMR technique, we have developed a detailed picture of the mechanism of a pH-dependent myristoyl switch.

Acknowledgements

Research is sometimes a sprint, but it is mostly a marathon. Long races are made better by the people running with you, even if you're not together the whole race. I have been fortunate in my graduate studies to run with incredible people over the years: Dr. Martin Smith who nudged me toward research; Colleen Doyle and Dalia Naser who tolerated my sarcasm and did not tolerate a curmudgeon; the other grad students, past and present, of the Meiering lab and biochemistry enclave who have made ESC a great place to be and a harder place to leave. The gaggle of undergrads I supervised over the years, Chris Go, Elisa Tran, Maddy Lemke, Lauren Morse, Mikaela Ney, Travis Ko, Puru Shah, Christopher Leo, Jeff Palumbo, and Mike Tarasca left their fingerprints all over my research from helping push it forward.

I would like to thank my supervisor, Dr. Elizabeth Meiering, for her wisdom when I got lost and her generosity when I needed it.

Dedication

To my long-suffering, titanicly patient, and loving Jeff,
My wonderful and supportive parents,
And to the value of a struggle.

Contents

Examining Committee Membership	ii
Author’s Declaration	iii
Abstract	iv
Acknowledgements	vi
Dedication	vii
Abbreviations	xvii
Chapter 1: Introduction	1
1.1 Overview	1
1.2 Protein Stability	1
1.3 Myristoylation	3
1.4 Switching	5
1.5 Folding a Protein	7
1.6 Hisactophilin	8
1.7 Structure of the Thesis	10
Chapter 2: Non-native Interactions in Folding Hisactophilin	11
2.1 Preface	11
2.2 Chapter Abstract	12
2.3 Introduction	12
2.4 Materials and Methods	15
2.4.1 Coarse-Grained Molecular Dynamics Simulations	15
2.4.2 Structure of nonmyristoylated and myristoylated hisactophilin	16
2.4.3 All-atom Molecular Dynamics	16
2.4.4 Chemicals	16
2.4.5 Protein Purification	17
2.4.6 Equilibrium and Kinetic Folding Measurements	17
2.4.7 NMR Experiments	17

2.4.8 Amide Exchange	18
2.5 Results and Discussion	18
2.5.1 Hydrophobic and Electrostatic Nonnative Interactions are Necessary to Accu- rately Simulate the Folding of a Myristoylated Protein	18
2.5.2 Analysis of Myristoyl Interactions in the Transition State and the Native State	22
2.5.3 Energetic Myristoyl Switch with pH	23
2.5.4 Structural Switch of the Myristoyl among Sequestered, Accessible, and Ex- posed States	26
2.5.5 Localized Stability Changes Associated with Myristoyl Switching	29
2.6 Conclusions	31
2.7 Supplemental Information	32
Chapter 3: Thermodynamics of Myristoyl Switching	34
3.1 Chapter Abstract	34
3.2 Introduction	34
3.3 Materials and Methods	36
3.3.1 Expressing Hisactophilin	36
3.3.2 Purification	37
3.3.3 Equilibrium Denaturation	38
3.3.4 Folding Kinetics	39
3.3.5 Mass Spectrometry	40
3.4 Results and Discussion	40
3.4.1 A faster purification protocol optimized for less stable mutants	40
3.4.2 Switching and native state interactions can tune, break, and recover the myris- toyl switch	41
3.4.3 The myristoyl group robustly accelerates folding kinetics in many hisactophilin mutants	49
3.4.4 Folding kinetics at low pH may suggest parallel folding pathways	50
3.5 Conclusions	53
Chapter 4: Local Conformational Stability and Strain in Myristoyl Switching	56
4.1 Chapter Abstract	56
4.2 Introduction	56
4.2.1 Protein Dynamics	56
4.2.2 Strain	58
4.2.3 NMR	60

4.2.3.1 The Chemical Shift	63
4.2.4 Temperature dependence	66
4.3 Materials and Methods	70
4.3.1 Expressing ¹⁵ N-labeled Hisactophilin	70
4.3.2 Purifying ¹⁵ N-labeled Hisactophilin	71
4.3.2.1 Cell Lysis	71
4.3.2.2 Immobilized Metal Affinity Chromatography	71
4.3.2.3 C18 Reverse Phase Chromatography	72
4.3.3 One- Two- and Three-Dimensional Nuclear Magnetic Resonance Experiments	72
4.3.3.1 Sample Preparation	72
4.3.3.2 Experiments for Assignments	73
4.3.3.3 Analyzing and Assigning	73
4.3.3.4 Variable Temperature Nuclear Magnetic Resonance Experiments . .	74
4.3.3.5 Analyzing Variable Temperature Data	74
4.4 Results and Discussion	75
4.4.1 Chemical Shifts	75
4.4.1.1 Backbone Resonance Assignments	75
4.4.1.2 Chemical Shift Perturbations	77
4.4.2 Amide Proton Temperature Coefficients	84
4.4.2.1 Global energetics and WT Temperature coefficients	85
4.4.2.2 WT amide temperature coefficients	87
4.4.2.3 Myristoylation significantly impacts temperature coefficients	94
4.4.2.4 Localizing strain and allostery with thermodynamic cycles	96
4.4.2.5 LLL amide temperature coefficients	100
4.4.2.6 I85L amide temperature coefficients	103
4.4.2.7 Integrating tension into the switching mechanism	105
4.5 Conclusions	108
4.6 Supplementary	109
4.6.1 Additional wild-type temperature coefficients	109
4.6.2 Additional LLL temperature coefficients	110
4.6.3 Additional I85L temperature coefficients	114
4.6.4 Correlating Temperature Coefficients with experiment and simulation	120
4.6.4.1 Secondary chemical shift	121
4.6.4.2 Root mean square fluctuation	121

4.6.4.3 Distance to the Myristoyl Group	126
4.6.4.4 Amide Exchange	129
Chapter 5: Conformational Heterogeneity in Myristoyl Switching	133
5.1 Chapter Abstract	133
5.2 Introduction	133
5.2.1 Dynamics and function	134
5.2.2 Nonlinear Temperature Dependence	137
5.2.3 Titratable residues	141
5.2.3.1 NMR-monitored pH titration	141
5.3 Materials and methods	144
5.3.1 Expressing ¹⁵ N-Labeled Hisactophilin	144
5.3.2 Purifying ¹⁵ N-labeled hisactophilin	144
5.3.3 Measuring nonlinear temperature dependences	144
5.3.4 Analyzing nonlinear temperature dependences	144
5.3.5 NMR-monitored pH titrations	145
5.3.6 Analyzing NMR pH Titrations	146
5.4 Results and Discussion	146
5.4.1 Curvature in Wild-Type	146
5.4.1.1 Curvature in Non-myristoylated Hisactophilin	147
5.4.1.2 Curvature in Myristoylated Hisactophilin	150
5.4.1.3 Myristoylation influences curvature	152
5.4.2 Curvature in F6L/I85L/I93L	156
5.4.3 Curvature in I85L	158
5.4.4 Recapitulating Curvature	163
5.4.5 Macroscopic pK _a s of LLL and I85L	166
5.5 Conclusions	174
5.5.1 Integrating conformational heterogeneity and apparent pK _a	175
5.6 Supplementary	176
Chapter 6: Method Development: automation and bootstraps	180
6.1 Chapter Abstract	180
6.2 Introduction	180
6.3 Evaluating Automation	182
6.4 Peak Picking Algorithms Impact Performance	184
6.5 Excising Experience	186

6.6 Anomalies, Anchors, and Application	188
6.7 Concluding remarks	193
Chapter 7: Proof of concept and future works	195
7.1 Chapter Abstract	195
7.2 Introduction	195
7.2.1 Reverse micelles	196
7.3 Materials and Methods	199
7.3.1 Materials	199
7.3.2 Reverse micelle encapsulation	199
7.3.2.1 Encapsulation in 10MAG/LDAO	199
7.3.2.2 Encapsulation in CTAB/Hexanol	200
7.3.2.3 Titration with cardiolipin	200
7.4 Results and Discussion	201
7.5 Additional future studies on hisactophilin	211
Bibliography	213
Appendices	229
Appendix C	237
1D Excitation Sculpting	237
1H, 15H HSQC	240
15N-Edited NOESY - HSQC	245
15N-Edited TOCSY - HSQC	252
1D 1H Presaturation	259

List of Figures

1	Figure 1.1 Simplified energy diagram two-state folding of a protein and a thermodynamic cycle	2
2	Figure 1.2: Examples of myristoylated proteins.	4
3	Figure 1.3: Basic residues around the opening of hisactophilin	5
4	Figure 1.4: The different states of Hisactophilin	6
5	Figure 1.5: Two conformations of the Ca ²⁺ -binding myristoyl switch protein, recoverin. . .	7
6	Figure 1.6: Sequence and structure of hisactophilin showing symmetry elements and mutants	9
7	Figure 2.1: The effect of nonnative interactions on simulating the folding of non- myristoylated and myristoylated hisactophilin.	14
8	Figure S2.1: The simulated free energy profiles for folding of myristoylated hisactophilin . .	19
9	Figure S2.2: The Δ distance matrices of the transition state and unfolded state ensembles relative to ensembles obtained using the native topology-based models.	20
10	Figure 2.2: Interplay between nonnative interactions on the folding barrier and stability. . .	22
11	Figure S2.3: Compaction of the unfolded and folded states with different native and non- native state interactions	22
12	Figure S2.4: Observed folding rate constants versus urea concentration for variant hisac- tophilins.	23
13	Figure 2.3: Energetics of myristoyl switching.	25
14	Figure 2.4: Structural characterization of the myristoyl switching mechanism.	27
15	Figure S2.5: The probability of populating the myristoyl accessible state	28
16	Figure 2.5: Changes in dynamics upon myristoylation.	29
17	Figure S2.6: Amide H/D exchange protection factors in hisactophilin upon myristoylation. .	30
18	Figure 3.1: Structural models of the β -trefoil protein hisactophilin with the conserved residues of the core at different depths/layers.	35
19	Figure 3.2: Simplified energy diagram and thermodynamic cycle used to measure switch energies.	42
20	Figure 3.4: Shown are bar charts representing the energetics of unfolding at pH 6.2 and pH 7.7	45

21	Figure 3.5: Kinetics of folding and unfolding of myristoylated and non myristoylated hisactophilin mutants.	51
22	Figure 4.1: Funnel model of an energy landscape and the relationship to dynamics.	60
23	Figure 4.2: NMR Spectra of wild-type non-myristoylated hisactophilin	63
24	Table 4.1: Contributing factors to chemical shifts.	65
25	Figure 4.3: Variable temperature 2D NMR, superimposed.	68
26	Figure 4.4: NMR assigned residues in wild-type, LLL, and I85L.	77
27	Figure 4.5: Overlay of the ^1H , ^{15}N HSQC for myristoylated wild-type and mutant LLL at pH 7.7.	78
28	Figure 4.6 Overlay of the 2D ^1H , ^{15}N HSQC of myristoylated I85L and LLL	79
29	Figure 4.7: Residues assigned for I85L, shown on a structure	80
30	Figure 4.8: Chemical shift perturbations upon myristoylation of wild-type hisactophilin	82
31	Figure 4.9: Chemical shift perturbations upon myristoylation of hisactophilin variants at pH 6.8	83
32	Figure 4.10 Bar plots of the amide proton temperature coefficients for wild-type hisactophilin	87
33	Table 4.2: Global thermodynamics and temperature coefficients	88
34	Figure 4.11: Wild-type amide proton temperature coefficients	90
35	Figure 4.12 Change in temperature coefficients induced by the pH change from 6.2 to 7.7 in wild-type	92
36	Figure 4.13: Bar chart of the temperature coefficient differences from the thermodynamic cycle	94
37	Figure 4.14 Change in temperature coefficients caused by myristoylation of wild-type	95
38	Figure 4.15: Temperature coefficient switch energetics	97
39	Figure 4.17: Bar graph showing the values of $\Delta\text{TCoeff}_{\text{switch}}$	99
40	Figure 4.21: Bar chart of the temperature coefficient differences in LLL	102
41	Figure 4.20: Changes in temperature coefficients induced by myristoylation of LLL	102
42	Figure 4.22: Changes in temperature coefficients upon myristoylation in I85L	104
43	Figure 4.23: Structure of hisactophilin illustrating some residues in the myristoyl binding pocket	106
44	Figure 4.19: Amide proton temperature coefficients for F6L/I85L/I93L mapped onto hisactophilin	111
45	Figure 4.24: Change in temperature coefficients induced by changing the pH from 6.2 to 7.7 in LLL	112
46	Figure 4.18: Bar chart of the temperature coefficient differences in LLL	114
47	Figure 4.25: Amide proton temperature coefficients for hisactophilin mutant I85L	115

48	Figure 4.26: Bar chart of the I85L mutant amide proton temperature coefficients	117
49	Figure 4.27: Differences in temperature coefficients resulting from changing from pH 7.7 to pH 6.2	118
50	Figure 4.28: Bar chart of the temperature coefficient differences of I85L	120
51	Figure 4.29: Scatterplots of the correlation between temperature coefficients and secondary chemical shifts	122
52	Figure 4.30: Rank order scatterplots showing the lack of correlation between temperature coefficients and secondary chemical shift.	123
53	Figure 4.31: Correlations between RMSF values obtained from umbrella sampling and tem- perature coefficients.	125
54	Figure 4.32: Structures of hisactophilin resulting from umbrella sampling.	127
55	Figure 4.33: Relating temperature coefficients to proximity of the myristoyl group.	129
56	Figure 4.35: Structure of hisactophilin with the global H/D exchangers indicated	130
57	Figure 4.34: Temperature coefficients versus Hydrogen Deuterium Exchange parameters . .	132
58	Figure 5.1: Timescales and techniques for measuring dynamics and exchange with an excited, or dark state dynamics	135
59	Figure 5.2: the temperature dependence of the amide proton chemical shifts of two core residues in hisactophilin	138
60	Figure 5.3: Structure of wild-type showing residues with curvature	148
61	Figure 5.5 Residues that experience curvature in non-myristoylated hisactophilin	150
62	Figure 5.6: Residues that exhibit curvature in myristoylated hisactophilin at pH 6.2	151
63	Figure 5.8: Structure of myristoylated hisactophilin showing curved residues at 2 pH values	152
64	Figure 5.9: Residues with curvature in myristoylated hisactophilin.	154
65	Figure 5.10: Extent of curvature (a-value) for myristoylated wild-type at pH 6.2.	155
66	Figure 5.11: Structure of LLL showing curved residues	157
67	Figure 5.14: Curved residues in Myr and NM LLL	158
68	Figure 5.15: Top-down view of LLL hisactophilin showing curved residues	159
69	Figure 5.16: Structure of I85L indicating curved residues	161
70	Figure 5.18: Side view of I85L showing curved residues	163
71	Figure 5.19: Top-down view of myristoylated I85L showing curved residues	164
72	Figure 5.20: Side-by-side comparison of curvature in myristoylated WT, LLL, and I85L at pH 7.7.	165
73	Figure 5.21: Free Energy Surface of myristoylated hisactophilin from umbrella sampling . .	165
74	Figure 5.22: Wild-type and mutant hisactophilin apparent pK _a s	167

75	Figure 5.23: The absolute values of changes in $pK_{a,app}$ s upon myristoylation, mapped onto hisactophilin	169
76	Figure 5.24: Scatter plots of the pH dependence of chemical shifts for wild-type and mutants	170
77	Figure 5.25: pH dependence of chemical shifts for unaffected residues	171
78	Supplementary Figure S5.1: Curved residues that depend on pH in non-myristoylated LLL	177
79	Supplementary Figure S5.2: Curved residues that depend on pH in myristoylated LLL	177
80	Supplementary Figure S5.3: pH-dependent changes in curved residues for I85L	178
81	Supplementary Figure S5.4: pH-dependent changes in the residues with curvature in I85L	179
82	Figure 6.1: Manual workflow for analyzing variable temperature data.	181
83	Figure 6.2: Plots of the temperature dependence of amide proton chemical shifts to evaluate automated peak picking without interpolation	183
84	Figure 6.3: Plots of the temperature dependence of amide proton chemical shifts to evaluate automated peak picking with interpolation	185
85	Figure 6.4: Workflow for analyzing variable temperature data with some automation	186
86	Figure 6.5: Scatter plots of residuals and bootstrap residuals	190
87	Figure 6.6: Examples of curvature in the residuals of amide proton temperature dependence data	191
88	Figure 6.7: Residuals from the temperature dependence of reduced pWT SOD1	192
89	Figure 7.1: Models and illustrations of reverse micelles	197
90	Figure 7.2: Representative 1D NMR spectra of reverse micelles.	204
91	Figure 7.3: 2D spectra of reverse micelle encapsulated hisactophilin	206
92	Figure 7.5: Aqueous 2D HSQC spectra of hisactophilin.	207
93	Figure 7.4: Comparisons of encapsulated myristoylated spectra with other species and spectra	209
94	Figure 7.6: Encapsulated myristoylated hisactophilin titrated with cardiolipin	210

Abbreviations

10MAG	Decylmonoacylglycerol
ΔG_{U-F}	Gibbs Free energy of folding
ANS	8-Anilinonaphthalene-1-sulfonic acid
AOT	bis(2-ethylhexyl)sulfosuccinate
ATP	Adenosine triphosphate
cAMP	Cyclic adenosine monophosphate
CARA	Computer Aided Resonance Assignment
CCPN	Collaborative Computing Project for NMR
CCS	(human) copper chaperone for SOD1
CHAPS	3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate
CHESCA	Chemical Shift Covariance Analysis
C_{mid}	Concentration of Urea at the midpoint of unfolding
CoA	Coenzyme A
CPMG	Carr-Purcell-Meiboom-Gill
CTAB	Cetyltrimethylammonium bromide
DEAE	Diethylaminoethyl
DHPC	1,2-Dihexanoyl-sn-Glycero-3-Phosphocholine
DMPC	1,2-Dimyristoyl-sn-Glycero-3-Phosphocholine
DMPG	1,2-Dimyristoyl-sn-Glycero-3-PhosphoGlycerol
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTAB	Dodecyltrimethylammonium bromide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
GCAP	Guanylate cyclase activating protein
GTP	Guanosine triphosphate
H/D Exchange	Hydrogen/Deuterium exchange

HIV(1)	Human immunodeficiency Virus 1
hNMT 1	Human N-myristoyltransferase 1
HSQC	Heteronuclear single quantum correlation
HPLC	High performance liquid chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
k_f^0	Rate of refolding extrapolated to 0 M urea
k_{obs}	the observed rate of folding/unfolding.
k_u^0	Rate of unfolding extrapolated to 0 M urea
LB	Lysogeny Broth
LDAO	Lauryldimethylamine N-oxide
LLL	F6L/I85L/I93L hisactophilin Mutant
LLLG	F6L/I85L/I93L/H90G hisactophilin Mutant
m_2^{ref}	Second order denaturant dependence of the refolding rate.
m_2^{unf}	Second order denaturant dependence of the unfolding rate
MARCKS	Myristoylated alanine-rich C kinase substrate
MBP	Maltose Binding Protein
MES	2-(N-morpholino)ethanesulfonic acid
m_f^\ddagger	Denaturant dependence of the refolding rate
m_u^\ddagger	Denaturant dependence of the unfolding rate
Myr	myristoylated
NM	non-myristoylated
NMR	Nuclear Magnetic Resonance
NMT	N-myristoyltransferase
NOESY	Nuclear Overhauser Effect Spectroscopy
OD	Optical Density
$pK_{a,app}$	Apparent pK_a
PMSF	phenylmethylsulfonyl fluoride
RDC	Residual Dipolar Couplings

RMSF Root mean square fluctuation
SDS-PAGE . . . Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SOD1 (human) superoxide dismutase 1
TOCSY Total Correlation Spectroscopy
TRIS Tris(hydroxymethyl)aminomethane
VMD Visual Molecular Dynamics
WT Wild-Type

Chapter 1: Introduction

1.1 Overview

Proteins effect some remarkable biochemistry — they make processes otherwise unavailable to the cell under biological conditions possible. The folding of these polypeptide chains into beautiful three-dimensional structures, comprising combinations of sheets, helices, and loops, is largely responsible for enzymes' substrate specificity and efficiency, a structural protein's mechanical strength, and/or its lifetime in the cell [1]. Adaptability to different environmental conditions requires that proteins are regulated so that individual proteins or pathways are active only when needed. Switching proteins and sensors respond directly to environmental stimuli, for example, increases in calcium ion concentrations, changes in pH, or illumination by specific wavelengths of light. These stimuli provoke a change to the switching protein's behaviour, such as membrane binding or oligomerization, that may take part in a cascade of other protein activations/deactivations that enable an organism to adapt to its surroundings. How these stimuli manifest and propagate through a single, small sensor protein is an interesting problem that can be addressed by measuring small structural and dynamic changes reflected in subtly shifting energy landscapes.

Functional switching in hisactophilin, a β -trefoil from *Dictyostelium discoideum* and model protein for switching, involves pH-dependent reversible anchoring to a membrane by a myristoyl group. Myristoylation is present in $\sim 0.5\%$ of all eukaryotic proteins [2], however, there is a paucity of high-resolution mechanistic detail of allostery and the general determinants of pH-dependent switching. To address this, I have set out to systematically evaluate the impact of numerous mutations on protein stability, folding, and switching, while also deploying novel methods using nuclear magnetic resonance to identify residues whose interactions govern switching and if those interactions disappear in switch-deficient mutants.

1.2 Protein Stability

One of the themes woven throughout this thesis is protein stability and its tradeoffs with function. The Gibbs free energy of folding (ΔG_{u-f}) is the stability difference between the unfolded and folded states; the larger the difference, the more stable the protein (represented in the left panel of Figure 1.1). The Gibbs free energy is generally described by the equation: $\Delta G = \Delta H - T\Delta S$ (Equation 1.1), where ΔH and ΔS are the enthalpy and entropy changes upon protein folding, respectively. In principle, proteins establish their lowest energy state through the delicate balancing of electrostatic, hydrophobic, and hydrogen bonding interactions to stabilize the folded, native state.

Despite the hundreds, if not thousands, of non-covalent interactions throughout a protein that contribute

to global stability, the Gibbs free energy of folding for a protein is typically in the range of -5- -12 kcal/mol [3]. While the gains in favourable enthalpic interactions in the folded protein help compensate for the concomitant loss of entropy, solvent interactions also contribute to global stability because of the entropy gained by solvent molecules when the hydrophobic groups are packed against one another and thus release their solvent clathrate cages [4]. The stability of a protein is the result of a complex interplay of stabilizing and destabilizing forces. The native state balances maintaining a stably folded protein against allowing flexibility or conformational adjustments necessary for protein function. Changes in the Gibbs free energy of folding may be used to report on the contribution individual residues, post-translational modifications, or ligand-binding make to global protein stability. Furthermore, measuring stability changes upon the stepwise combination of mutations, post-translational modifications, and ligands characterizes interactions and coupling between different parts of a protein (right panel of Figure 1.1)[5,6]. Thus, identifying networks of residues coupled to ligand binding, for example, may reveal the pathway for allosteric communication that effects changes in protein behaviour upon environmental changes [7].

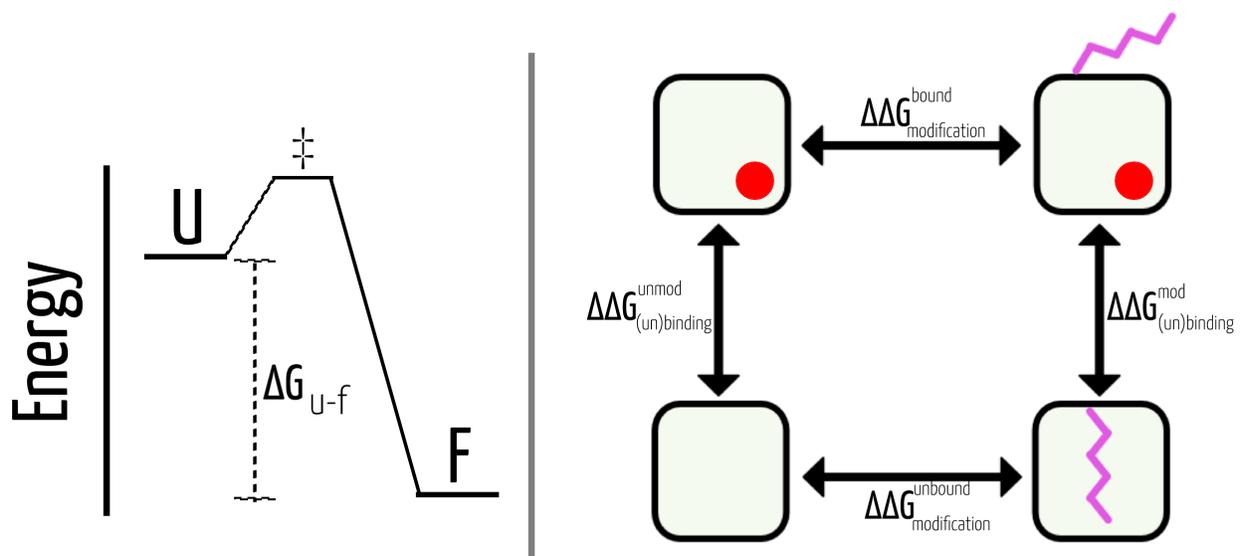


Figure 1.1: (Left panel) Simplified energy diagram of two-state folding of a protein, traversing from the unfolded state (U) to the folded state (F) through a single transition state (\ddagger). The reaction coordinate is along the implied x-axis while the y-axis represents the energy of the system. The gap between U and F is the Gibbs free energy of folding, ΔG_{U-f} . In reality, the energy profile may not be smooth because of numerous energy minima that are part of a rough landscape. (Right panel) Schematic of a thermodynamic cycle as applied to ligand-induced switching to analyze the interaction between binding sites and a modification (e.g. a myristoyl group [8], phosphoryl group [9], or a glycosyl group [10]). The fuchsia lines represent a modification like a myristoyl group, and the red circles represent a bound ligand. The cycle is also discussed in Chapter 3. Each corner represents the ΔG_{U-f} of the **modified** or **unmodified** forms of a protein in the liganded and/or un-liganded states. Differences in these stabilities report on the effect of ligand binding (vertical edges of the cube) or the effect of modification (horizontal edges). The full cycle gives ΔG_{switch} , a measure of how the modification's effects on the stability of hisactophilin interact with ligand binding. Based on [5].

1.3 Myristoylation

In many proteins, often those that participate in signalling, lipidation is an important post-translational modification that facilitates protein-membrane and/or protein-protein interactions [11,12]. Broadly, lipidation includes post-translational modification with fatty acids, isoprenoids, or glycosylphosphatidylinositol anchors, but the principal focus herein will be fatty acids. Namely, myristoylation; the co-translational modification of an N-terminal glycine of a protein with a C₁₄ saturated fatty acid. This is one of the most ubiquitous acyl modifications; it occurs with approximately 0.5% of all eukaryotic proteins in the cell are myristoylated [2,13].

The covalent linkage of the myristoyl group to the N-terminal glycine is effected by N-myristoyltransferase (NMT) [13], which targets the general recognition sequence Gly-X_{neutral}-X_{neutral}-X_{neutral}-Ser/Thr-Lys [14] on the destination protein. NMT is essential for survival—knockout in *Saccharomyces cerevisiae* is recessive lethal [15] and similarly causes embryonic lethality in *Drosophila melanogaster* [16]. There are two variants of NMT in humans, of which hNMT1 (human NMT 1) is essential in early development, cell replication, and, in aberrant cases, tumour cell growth [17,18]. Interestingly, when hisactophilin is expressed in *Escherichia coli*, it was most effectively myristoylated by hNMT1, which has greater sequence homology with the cognate NMT from *Dictyostelium discoideum* than the more commonly used yeast NMT [19,20]. Before encountering NMT, however, fatty acyl-CoA synthetase (FAA1) activates myristic acid with Coenzyme A [21] and methionine amino peptidase removes the N-terminal methionine, exposing glycine on the target protein [2].

The attachment of a myristoyl group has thermodynamic and functional implications. Myristoylation stabilizes the Ca²⁺-sensing protein recoverin (see top left panel of Figure 1.2) [22], Guanylate Cyclase Activating Protein-2 (GCAP-2)[23], GCAP-1 (bottom left panel of Figure 1.2) [24,25], calcineurin (top right of Figure 1.2) [26], HIV-1 Matrix Protein p17 (bottom right of Figure 1.2) [27], and hisactophilin [8]. Increased protein stability likely arises from the thermodynamic advantages of burying additional hydrophobic groups [28]. However, stability gains by myristoylation may not be universal—though no rigorous thermodynamic measurements have yet been undertaken, pulse chase experiments with Lassa virus Z protein indicated that myristoylation was not required for stability in the cell (though it did impact membrane association and infectivity) [29]. Similar measures on the small envelope protein of porcine reproductive and respiratory syndrome virus [30] and the E protein from the envelope of equine arteritis virus [31] agreed with the finding in the Lassa virus protein [29]—myristoylation decreased protein half-life in the cell. When myristoylation of the two viral proteins is disrupted, virus infectivity was attenuated. The mechanism is not entirely elucidated, however, compromised membrane targeting and/or fusion events appear to be a significant factor for the phenotype. These examples aim to illustrate the diversity of proteins that are myristoylated and the varied

effects myristoylation has on the host protein.

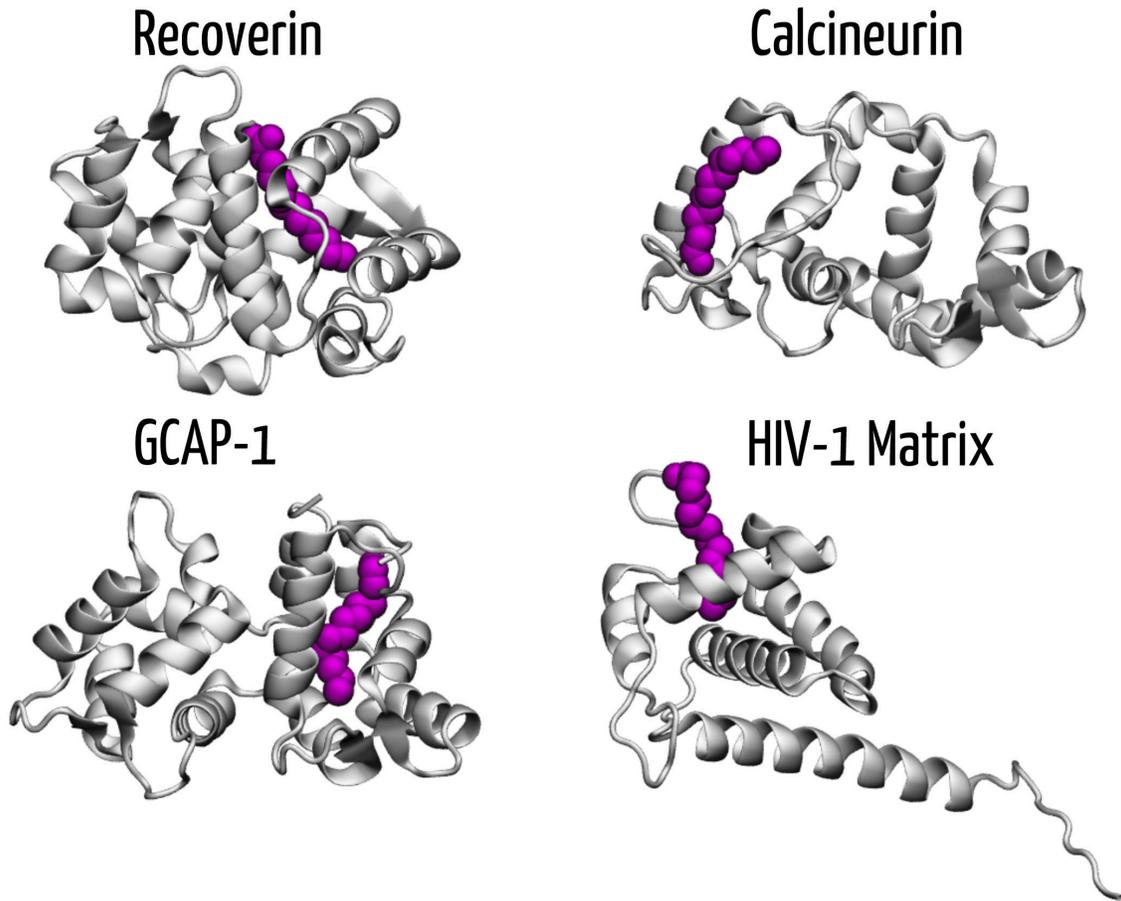


Figure 1.2: Examples of myristoylated proteins. The myristoyl group is shown as magenta spheres covalently linked to the N-terminal glycine of various proteins. The structure of recoverin is adapted from PDB: 1JSA [32], Calcineurin from PDB: 1TCO [33], Guanylate Cyclase Activating Protein-1 (GCAP-1) from PDB: 2R2I [34], and HIV-1 Matrix Protein from PDB: 1UPH [35]. They are all predominantly helical proteins and the myristoyl group tends to be protected from solvent in some binding cleft or pocket. Burial of the myristoyl group is believed to be a contributor to the conferred stabilization. Over 150 mammalian proteins are reported to be myristoylated [36].

Facilitating membrane interactions is one of the core functional contributions of the myristoyl group. However, experiments, demonstrated that the Gibbs free energy change, ΔG_0 , of inserting the myristoyl group into and anchoring a peptide onto a membrane is only approximately -8 kcal/mol, insufficient for stable binding [37,38]. Thus, a second anchoring component is typically involved. Four possible “second signals” are: 1) a polybasic region that interacts with anionic components of a membrane (e.g. phosphatidylserine), 2) hydrophobic residues that interact with the membrane, 3) interactions with another membrane-associated protein, or 4) an additional acyl modification (e.g. palmitoylation, farnesylation) [13]. Hisactophilin, a small pH-dependent myristoyl switch model protein, differs from the example myristoylated proteins in Figure 1.2 in that it folds into an all β -trefoil. The myristoyl group is buried in a pocket down the centre of the β -barrel. Hisactophilin’s second signal for membrane binding is a polybasic region around the opening of the β -barrel

(see Figure 1.3) and together, the two signals yield a ΔG_0 of binding -10.5 kcal/mol [11]. Myristoylated alanine-rich C kinase substrate (MARCKS) [39], c-Src [40], HIV-1 Gag and Nef proteins [37] also rely on a polybasic motif to complement the myristoyl moiety for membrane binding.

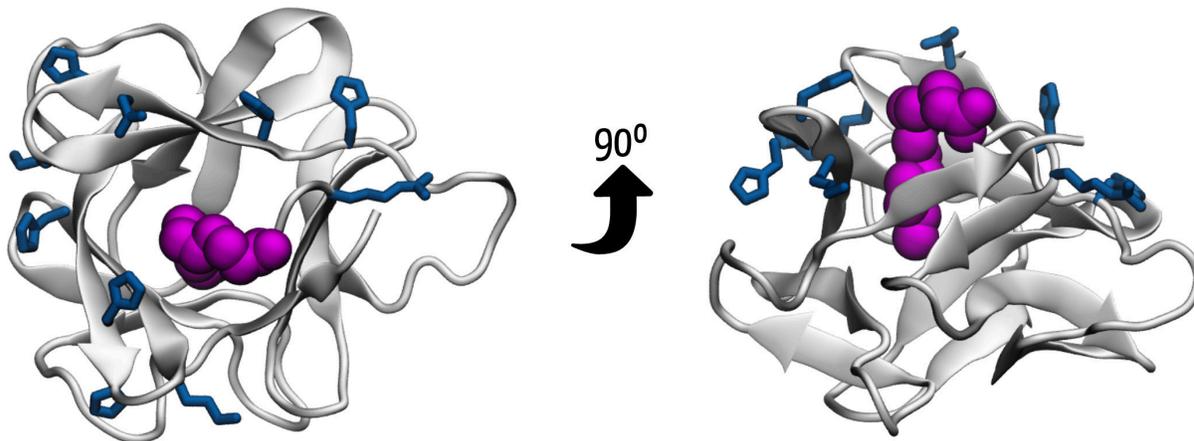


Figure 1.3: Basic residues around the opening of hisactophilin. Hisactophilin is shown in the myristoyl-sequestered state. The myristoyl group is magenta spheres and basic residues near the opening of the β -barrel are blue. They comprise residues R4, H33, H35, N38, H39, K42, H75, H78, H79, and K82. Not shown are four glutamic acids, which could also form electrostatic interactions, albeit not likely with phosphatidylserine.

1.4 Switching

Akin to other myristoylated proteins, membrane binding is reversible. The dual nature of these switch proteins relates to conformations with distinct characteristics: one in which the myristoyl group anchors the protein to the membrane inner leaf; the other where the myristoyl group is sequestered from solvent, inside or on the surface of a cytosolic protein (shown for hisactophilin in Figure 1.4). The reversible switching behaviour creates an environmentally modulated sensor; triggers can vary from Ca^{2+} -binding in recoverin, GTP-binding in ADP ribosylation factor [41], or proton-binding in hisactophilin [42]. In *Dictyostelium discoideum*, hisactophilin bundles actin and switches between cytosolic and membrane-bound conformations in response to pH changes. Pulses of cAMP, a component of chemotactic signalling, cause localized intracellular pH changes [43]. Lowering the pH below ~ 6.9 switches hisactophilin toward bundling actin and membrane binding [44], which enables cell locomotion toward the source of the chemoattractant.

As mentioned above, the myristoyl group stabilizes many proteins. In hisactophilin, increased protein stability by the myristoyl group is greatest at pH values above the midpoint of switching ($\text{pK}_{\text{switch}} = 6.95$), when the acyl group is sequestered and forms favourable interactions with the protein (~ 3.2 kcal/mol more stable than the non-myristoylated form) [8]. *In vitro*, without a membrane present (dashed box in Figure

1.4), crossing to below pK_{switch} results in a net uptake of ~ 1.5 protons and hisactophilin switches toward the accessible state, which has fewer interactions between the myristoyl group and the binding pocket (right side of the dashed box in Figure 1.4). Below pK_{switch} , residual interactions between the myristoyl group and hisactophilin result in global stability increased by by ~ 1.1 kcal/mol.

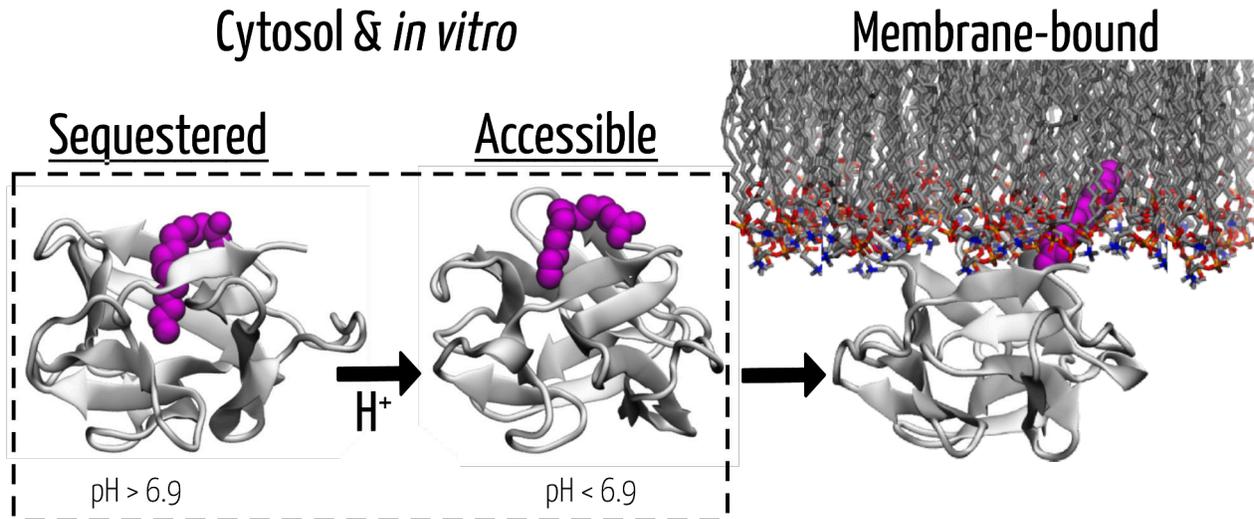


Figure 1.4: Different states of Hisactophilin. Shown here are some of the states of hisactophilin involved in the myristoyl switch. The magenta spheres represent the highly dynamic myristoyl group. The sequestered and accessible states are the two states studied *in vitro* (without a membrane). Hisactophilin favours the sequestered state at $\text{pH} > 6.95$, where the myristoyl group is buried in the core of the β -trefoil and makes favourable interactions that stabilize the protein by ~ 3.2 kcal/mol. Shifting the pH below 6.95 shifts the population toward the accessible state in response to a net uptake of ~ 1.5 protons. Here, residual interactions between the myristoyl group and the protein results in only a 1.1 kcal/mol stabilization [8]. The difference in myristoyl-induced stabilization at different pH values is believed to be a driving force of the switch. The membrane-bound state is meant only to be demonstrative as there is little high resolution structural information available. No interactions with actin—one of hisactophilin’s other binding partners—are shown.

In some proteins, myristoyl switches are associated with large conformational changes. Recoverin’s calcium-binding induced switch of the myristoyl group to an accessible state involves conformational changes around swivel points G42 and G96 (see Figure 1.5). Rotation around G42 changes the interhelical angle of EF-1 toward an open conformation that exposes hydrophobic residues important in target-binding and ejects the myristoyl group. Rotation around G96 changes the angle between domains by 45° [32]. HIV-1 matrix protein also couples large conformational changes to switching; the myristoyl-sequestered monomer converts to a myristoyl-accessible trimer [35]. Some designed light-reactive switches also undergo conformational changes [45,46], for example domain-swapping to control protein-protein interactions [47], regulating protein-DNA binding [48,49], and in engineered photoswitches triggered by high-wavelength light *in vivo* [50]. Switching in hisactophilin, however, does not involve significant rearrangement of secondary structure; so, the conformational changes driven by proton-binding to the all β -protein are likely subtle, but also thermodynamically permissive of the myristoyl accessible state. Another goal of this thesis is to characterize the

communication network between the site of protonation and the myristoyl group that has proven elusive, as well as experimentally defining the molecular mechanism of switching at high resolution.

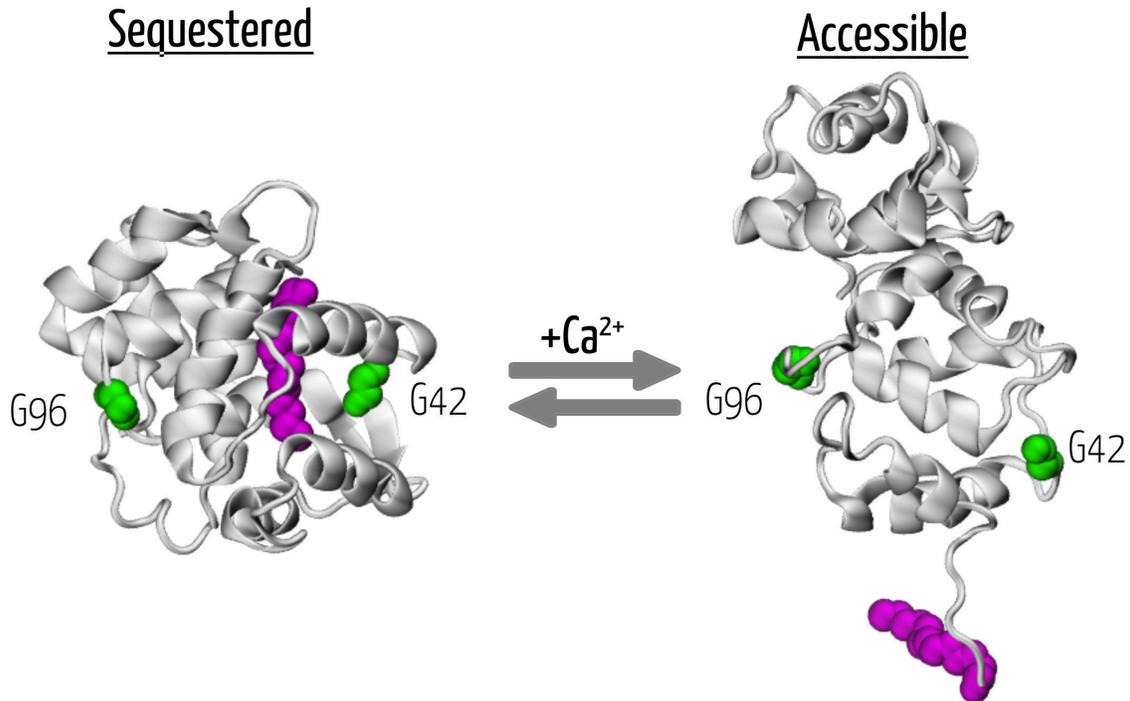


Figure 1.5: Two conformations of the Ca²⁺-binding myristoyl switch protein, recoverin. The myristoyl-sequestered state (left) lacks calcium ions and the accessible state (right) is calcium-bound. The two hinge glycines are shown as green spheres—G42 and G96. The myristoyl group is shown as magenta spheres. The structure of the accessible state was taken from PDB: 1JSA [32] and the sequestered state from PDB: 1IKU [51], and visualized in VMD [52]. Switching here involves a relatively large conformational shift.

1.5 Folding a Protein

Folding a protein into a functional three-dimensional shape is a fundamentally complex problem because an unfolded/denatured protein can explore a significant conformational space before finding a functional conformation [53]. A random search of the conformational space is not tractable on a biological timescale and it is thought that various pathways guide folding [54]. The classical model of folding is depicted in Figure 1.1 (left panel)—folding into the native state involves passage through a high-energy transition state without the accumulation of any meta-stable intermediate structures along the pathway [55]. The denatured state is an ensemble of rapidly fluctuating conformations with transient elements of structure that may initiate folding [56]. The transition state of folding can be thought of as an ensemble where native-like non-covalent interactions are forming [57]. The native state emerges from the native-like interactions that form in the

transition state but non-native interactions—interactions that are not represented in the folded protein[58]—may compete with them and frustrate folding [53]. While some non-native interactions frustrate and slow folding [59], they have also been shown to accelerate or help folding [8,60,61], and could further aid and abet function [62]. Determining the roles of post/co-translational modifications and prosthetic groups on the energy landscape is still under investigation by other groups [63–65] and it is also one of the goals of this thesis.

1.6 *Hisactophilin*

Hisactophilin is amenable for use as a model system for switching because it captures many dimensions of interactions in a small (13.5 kDa) and relatively well-behaved protein. It is regulated by phosphorylation [11], an osmoprotectant [66], an N-terminally myristoylated pH-dependent switch, reversibly binds to membranes and bundles actin. Its small size makes it feasible to obtain atomic resolution structural and dynamic information accessible by nuclear magnetic resonance (NMR) experiments. Furthermore, it folds by an apparent two-state mechanism without evidence of an intermediate, so stability measurements to analyze the energetics of switching are relatively straightforward. The primary sequence and structure (determined by NMR [42]) are given in Figure 1.6.

Hisactophilin comprises 118 amino acids, however methionine aminopeptidase cleaves the N-terminal methionine to expose glycine so N-myristoyltransferase can covalently attach myristic acid, a C₁₄ saturated fatty acid. Named for its high proportion of **histidine** and affinity for **actin**, thirty-one of hisactophilin's 118 amino acids are histidine, which likely relates to hisactophilin's osmoprotectant properties and the pH-dependent behaviour centered at pH ~6.95 [8]—within the typical range of titrating histidine residues. The primary sequence folds into a β -trefoil of 12 β -strands with elements of three-fold pseudosymmetry.

The β -barrel component is defined by 6 β -strands (two per trefoil element) while the remaining 6 β -strands close off one end with a hairpin triplet. Other members of the β -trefoil family are acidic and basic fibroblast growth factors [67], ricin B-like toxin, and interleukin 1 β and interleukin 33 [68]. Despite the similarities in structure, the other members of the β -trefoil family share less than 10% sequence identity with hisactophilin [69,70]. Furthermore, those β -trefoils are not myristoylated. In hisactophilin, when sequestered (pH > 6.95), the terminal methyl of the myristoyl moiety is within 5 Å of residues F6, I85, I93, and F113 [8].

The global thermodynamics of non-myristoylated hisactophilin have been thoroughly characterized, including folding/unfolding kinetics [8,71–73], as well as its global stability via fluorescence- or circular dichroism-monitored urea equilibrium denaturation at pH 5.7, 6.7, 7.7, 8.7, and 9.7. Below pH 7.7, global stability is strongly pH-dependent and decreases to as low as ~2.4 kcal/mol at pH 5.7 (from ~10.1 kcal/mol at pH 9.7). The thermodynamics of myristoylated hisactophilin were similarly interrogated and illustrated the

Hisactophilin

MGNRAFKS**HH**GHFLSAEGEAVK**THHGHHDHHTHFH**VENHGGKVALK**TH**CGKYLSIGD**HK**
 QVYLS**HH**LHGDHSLF**H**LE**HH**GGKVS**IKGHHHH**YISAD**HHGH**VSTKE**HH**DHDTTFEE**III**

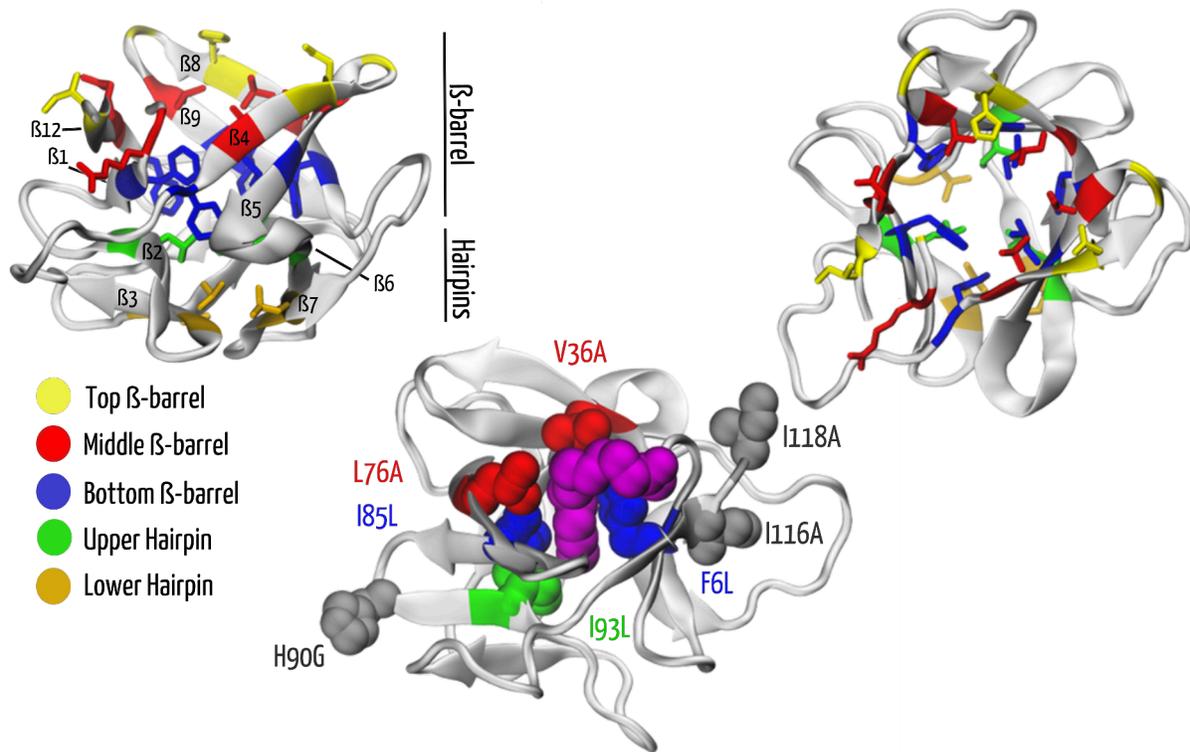


Figure 1.6: The primary sequence and structure of Hisactophilin I [42]. All 118 amino acids are shown, though the N-terminal methionine is cleaved to expose glycine before myristoylation by N-myristoyltransferase [2]. Noteworthy as well is that 31 of 118 amino acids are histidine (bolded), a feature believed to be important in pH-depending switching and osmoprotection. The left and right structures show a side and top view of hisactophilin without the myristoyl group. Hisactophilin folds into a β -trefoil comprising 12 β -strands; four per element of pseudosymmetry, six form a β -barrel and the other six close one end of the barrel. The core is made up of semi-conserved hydrophobic residues (with other β -trefoils) at various depths; the top, middle, and bottom of the β -barrel, and then the upper and lower hairpins. The bottom panel is an off-axis top view where the N-terminal myristoyl group is shown as magenta spheres, and positions of studied mutations are shown as spheres following the colour scheme of the preceding two structures; F6L, V36A, L76A, I85L, H90G, I93L, I116A, and I118A.

strong pH-dependence of its disposition [8]. Furthermore, myristoylated and non-myristoylated hisactophilin NMR spectra were assigned previously and subjected to pH titration to measure the apparent pK_a s of backbone amide protons, which eventually suggested candidate residues for driving the myristoyl-switch [8,74]. There is a strong basis of structural and thermodynamic knowledge of hisactophilin, which forms a sturdy foundation from which the molecular mechanism and determinants of pH-dependent myristoyl switching and allostery may be investigated at high resolution and in great detail.

1.7 Structure of the Thesis

The thesis that follows is written so that chapters are semi-modular and can be read fairly independently, each generally comprising an introduction to the theory of the experimental methods and context of the research, methods, results, and a discussion to contextualize the findings in the broader field. After Chapter 1, the introduction, Chapter 2 presents work done in collaboration with Dr. Yaakov Levy at the Weizmann institute [73] that uses a combined experimental and computational approach to describe the unique roles of the myristoyl group and non-native interactions (that differ from other post-translational modifications) in the folding and function of hisactophilin. Chapter 3 investigates the native-state stability of mutations throughout the protein to determine the role of specific and highly tuned interactions in folding and switching. Chapters 4 and 5 show that variable temperature nuclear magnetic resonance experiments are a robust and highly sensitive technique for obtaining high resolution detail of the near-native energy landscape and the molecular mechanism of switching. The combination of temperature-dependence experiments with the pH-dependence of chemical shifts also used in Chapter 5 offer a novel approach to identify components of the allosteric network linking pH-sensing to the myristoyl group. The development and evaluation of a generalizable and widely applicable method for analyzing linear and non-linear temperature dependences of chemical shifts is presented in Chapter 6. Lastly, Chapter 7 highlights some preliminary work on high resolution detail of membrane-bound hisactophilin in a reverse micelle and other outstanding questions and future work.

Chapter 2: Non-native Interactions in Folding Hisactophilin

2.1 Preface

The following work was the result of a collaboration with Dr. Yaakov Levy to delve deeper into questions asked and discoveries made in an earlier publication on hisactophilin from our research group [8]. That work presented quantitative measurements of the thermodynamic implications of myristoylation on global stability and folding kinetics. There, a formalism was developed that extracted the energetics of switching—later published in *Protein Engineering Design and Selection* [5]—that indicated important native-state interactions between the myristoyl group and the rest of the protein important in pH-dependent myristoyl switching. Beyond native-state interactions, folding kinetics experiments revealed that non-native interactions between protein and myristoyl group accelerate folding, that the myristoyl group may induce strain in the native-state, and that the aforementioned non-native interactions relieve that strain. The following work attempts to localize the interactions between myristoyl group and protein that are responsible for the observed thermodynamic effects with combined efforts in experimental work, coarse-grained and atomistic simulations—identifying strain and its manifestations will be an ongoing theme in this thesis.

Dr. Yaakov Levy's group carried out the coarse-grained simulation (Dalit Shental-Bechor) that were designed to model the native and non-native interactions between the protein and the myristoyl group (e.g. Figure 2.1, 2.2A,B,C, Figure 2.4A). Dr Aron Broom from the Meiering Lab was responsible for some of the atomistic simulations that recapitulated States I, II, III (e.g. Figure 2.4C, D, E). Experimental work was shared between Dr. Martin Smith and Duncan MacKenzie. Martin Smith designed mutants, acquired hydrogen-deuterium exchange experiments, some of the equilibrium thermodynamic measurements, and analyzed a portion of the folding kinetics data. Duncan MacKenzie developed and executed the strategy for purifying the new mutants (with the help of undergraduate Chris Go), acquired and analyzed the folding kinetics experiments, and acquired the remaining equilibrium thermodynamic experiments. Writing and editing were a collaborative effort. The objective of the experiments were to measure alterations to the native- and transition-states upon mutation of residues simulations predicted to form interactions (non-native and native) with the myristoyl group.

Note that the supplementary figures are presented here in the main text for ease of reading, however some additional results and analysis reside in the 2.7 Supplemental Information.

2.2 Chapter Abstract

We present an integrated experimental and computational study of the molecular mechanisms by which myristoylation affects protein folding and function, which has been little characterized to date. Myristoylation, the covalent linkage of a hydrophobic C₁₄ fatty acyl chain to the N-terminal glycine in a protein, is a common modification that plays a critical role in vital regulated cellular processes by undergoing reversible energetic and conformational switching. Coarse-grained folding simulations for the model pH-dependent actin- and membrane-binding protein hisactophilin reveal that nonnative hydrophobic interactions of the myristoyl with the protein as well as nonnative electrostatic interactions have a pronounced effect on folding rates and thermodynamic stability. Folding measurements for hydrophobic residue mutations of hisactophilin and atomistic simulations indicate that the nonnative interactions of the myristoyl group in the folding transition state are nonspecific and robust, and so smooth the energy landscape for folding. In contrast, myristoyl interactions in the native state are highly specific and tuned for sensitive control of switching functionality. Simulations and amide hydrogen exchange measurements provide evidence for increases as well as decreases in stability localized on one side of the myristoyl binding pocket in the protein, implicating strain and altered dynamics in switching. The effects of folding and function arising from myristoylation are profoundly different from the effects of other post-translational modifications.

2.3 Introduction

Protein folding is governed by various physicochemical forces that bias the native state, which for many proteins is also the functional state, over the many alternative nonnative states. The network of native interactions has been found in many cases to be sufficient to capture the folding mechanism and kinetics of proteins [75,76]. The discrimination between native and nonnative interactions is the foundation of the principle of minimal frustration [53] and explains the power of native topology-based models in studying folding biophysics [77]. The dominant role of native interactions is manifested by the funnel-shaped energy landscape for folding that suggests folding is robust and an efficient process. The information stored in the native topology may, however, be tuned by various factors such as confining the protein in a small space, crowding agents, or conjugating the protein to other biomolecules [e.g., oligosaccharides [63] or fatty acyl chains such as myristoyl [8]]. In addition to manipulating folding characteristics by modifications or environmental conditions, nonnative interactions, which are by definition in conflict with the native state, may decorate the folding funnel [59,78] by increasing energetic frustration [79] between interactions and therefore landscape roughness. The degree of roughness, which affects the trapping of the protein in nonnative states, depends on the particular sequence of the protein and can be tuned by mutations.

Investigations of several proteins have reported evidence for nonnative interactions that assist, rather than hinder, folding [60,61,64,80], and more importantly they may also support function [8,62,81] by assisting conformational changes. Residues in functional sites in proteins have been implicated in causing geometric frustration [82] or increasing localized energetic frustration [62]. Nonnative interactions that result with localized frustration can transiently be formed, for example, between hydrophobic residues [61] or between oppositely charged residues [83–85]. This is akin to frustration in RNA folding that arises from negatively charged groups [86]. The formation of nonnative interactions may affect folding in various ways. For example, nonnative interactions in the unfolded state may affect its entropy and therefore the overall stability of the protein. Also, nonnative interactions in the transition state that support the critical nucleus may speed up the folding process.

In the current study, we investigate the effect of myristoylation on protein folding and in particular the involvement of nonnative interactions. Myristoylation is a common modification, where a saturated C_{14} fatty acyl chain is covalently linked to the N-terminal glycine in a protein [13]. In many proteins the myristoyl interconverts between a sequestered state, where it is located in a hydrophobic binding pocket, and an accessible state, where it is available to bind to membrane or other proteins. Often this interconversion, or switching, is reversible and controlled by the binding of ligand. Myristoyl switching is associated with diverse and vital regulated signaling pathways in cells [13].

We investigate here the effect of myristoylation on the folding kinetics and thermodynamics of the β -trefoil protein, hisactophilin, a small (118 residues) protein from *D. discoideum* (Figure 2.1A). The function of hisactophilin is to reversibly recruit actin filaments to membranes during chemotaxis and osmotic stress. Previous studies revealed that increasing hisactophilin charge with decreasing pH favours actin and membrane binding but decreases protein stability and folding kinetics [71,87]. The myristoyl increases hisactophilin stability with an apparent pK_a of 6.95 as it switches between a “sequestered” state at high pH, where the myristoyl is buried in the protein core (Figure 2.1A), and an “accessible” state at low pH for membrane binding [8]. Concomitantly, the myristoyl markedly accelerates both folding and unfolding kinetics, and undergoes rapid native-state switching, implicating the long and flexible hydrophobic myristoyl chain in creating strain in the native state and forming nonnative interactions during folding and switching. Here we report an integrated approach combining various computational and experimental methods to analyze myristoylated hisactophilin, including characterization of hydrophobic mutations distributed in the primary and tertiary structure of the protein (Figure 2.1B). Variants include a triple myristoyl binding pocket mutant (F6L/I85L/I93L) in which pH-dependent switching is abolished [5], and single mutations that alter stereochemistry (I85L) or truncate side chains inside (V36A and L76A) or outside (I118A) the binding pocket. The results provide unique insight into the role of nonnative interactions in folding as well as the atomistic mechanism of myristoyl switching.

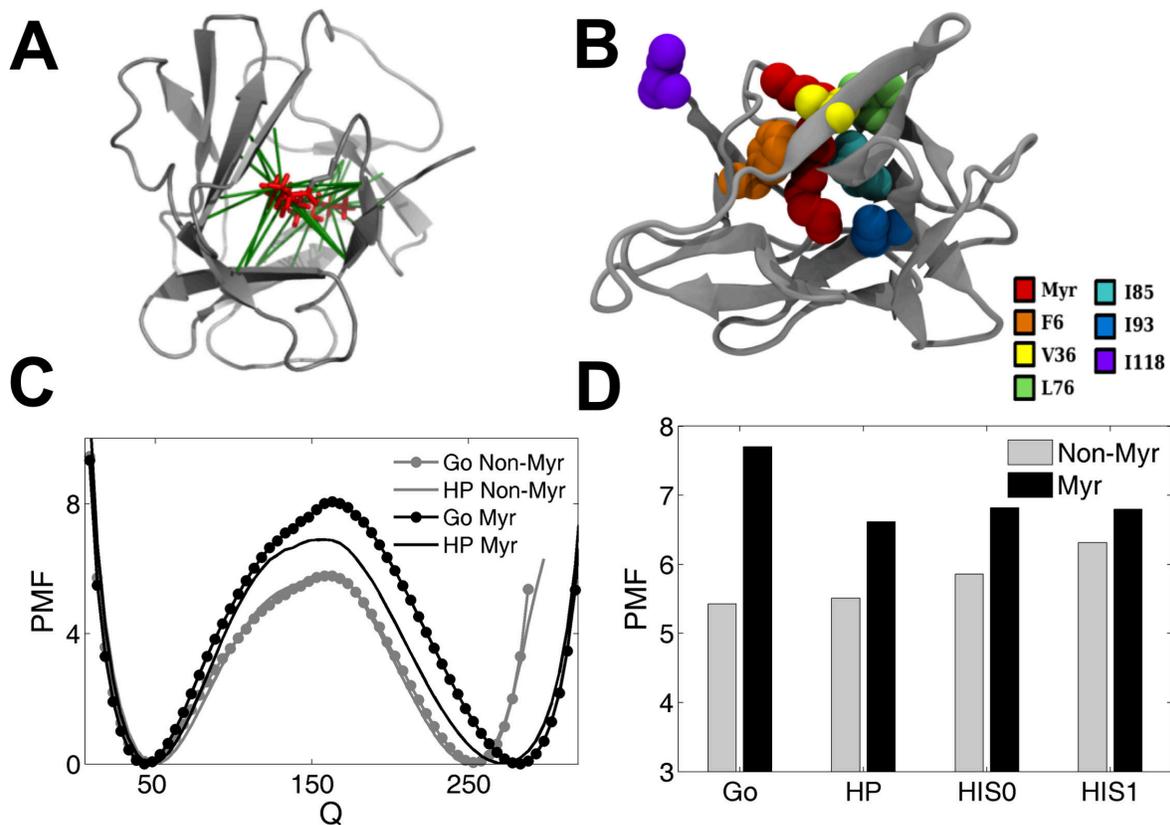


Figure 2.1: The effect of nonnative interactions on simulating the folding of non-myristoylated and myristoylated hisactophilin. (A) Top view of myristoylated hisactophilin (backbone in grey cartoon) with native contacts (green) between the myristoyl group (red) and the protein. (B) Side view of hisactophilin with space-filling representations for mutated residues. (C) Potential of mean force vs. Q , the number of native contacts within hisactophilin, for folding of myristoylated (black) and non-myristoylated (grey) hisactophilin calculated for four different simulation models: the native topology-based model (Go), nonnative hydrophobic model (HP), charged residue model at high pH (His0), and positively charged histidine at low pH electrostatic model (His1). The His0 and His1 electrostatic models also include the HP model. Folding barriers are calculated at the folding temperature, T_F . (D) The folding barrier for non-myristoylated (grey) and myristoylated (black) hisactophilin at the T_F of each system.

2.4 Materials and Methods

Two coarse-grained models of nonnative interactions were used: nonnative hydrophobic interactions modeled by a Gaussian potential [61] in which the κ is the hydrophobicity strength, (HPGo model) and nonnative electrostatic interactions modeled by the Debye–Huckel potential (His0 and His1, for high and low pH) [85]. Each native interaction is modeled using Lennard–Jones potential with strength of ϵ . To model non-myristoylated hisactophilin all the contacts of the myristoyl beads were removed except for the virtual bonds that connect the beads to each other (SI Materials and Methods). All-atom molecular dynamics were based on the refined structure of myristoylated hisactophilin and run using the AMBER simulation package (SI Materials and Methods). Preparation of hisactophilin variants and measurements of protein stability by equilibrium denaturation curves [87], folding kinetics by stopped flow [71], and amide H/D exchange by NMR [88] were performed as described previously (SI Materials and Methods).

2.4.1 Coarse-Grained Molecular Dynamics Simulations

We used simplified coarse-grained models in which each residue is represented by a single bead centered at the alpha carbon. The model is described in detail in [89]. This is a topology-based model and the energy function is based on the structure of the native state of the protein. Each native interaction (either residue-residue or residue-myristoyl) is modeled using a Lennard-Jones term,

$$E_{ij}^{Nativecontact} = \epsilon \left(5 \frac{A_{ij}^{12}}{r_{ij}^{12}} - 6 \frac{A_{ij}^{10}}{r_{ij}^{10}} \right)$$

where A_{ij} is the optimal distance in Å between beads i and j and r_{ij} is the distance (in angstroms) between beads i and j in a given conformation along the trajectory. ϵ is the strength of the native interaction. The nonnative interactions are all the possible interactions between hydrophobic beads in the protein (including the myristoyl group), which are not in contact in the native structure and are not close in sequence ($i < j - 3$). In this work, we followed the procedure introduced by Chan and coworkers [61]. Shortly, we added an attractive potential between each pair of hydrophobic residues (pairs were defined as described above) in addition to the repulsive interactions between pairs that are not in contact in the native state. Any nonnative hydrophobic interaction is therefore modeled as,

$$E_{HP} = -K_{HP} \kappa_i \kappa_j \exp\left(-\frac{(r_{ij} - 5)^2}{2}\right)$$

where κ_i and κ_j is the hydrophobicity strength of beads i and j that represents an amino acid or a carbon of the myristoyl group and its value ranges from 0 to 1. K_{HP} is the overall (i.e., native and nonnative)

strength of the hydrophobic forces that was selected to be 0.8. In the present study, alanine, valine, leucine, isoleucine, methionine, tryptophan, phenylalanine, and tyrosine were considered to be hydrophobic. The nonnative electrostatic interactions were modeled using the Debye-Huckel potential [85]. These interactions were defined between each pair of charged residues (lysine, arginine, and aspartic and glutamic acids). In this model, called His0, histidine residues had neutral charge and mimic high pH. Modeling low pH was achieved by adding a positive charge to all the histidine residues (called His1 model). When nonnative electrostatic interactions were modeled, the nonnative HP interactions were included as well.

2.4.2 Structure of nonmyristoylated and myristoylated hisactophilin

We did some refinements on the model structure that was provided from the NMR measurements [PDB: 1HCD and [8]]. We ran long all-atom molecular dynamics simulations of the myristoylated protein and clustered the conformations that were generated during the simulation to find the representative conformation to be treated as the native structure. We used this conformation to construct the native topology-based model. In the coarse-grained simulations, the protein was represented only by its C_{α} atoms and the myristoyl was represented by six beads (each represents two aliphatic carbons of the myristoyl chain). We used this conformation of the myristoylated protein to model the non-myristoylated variant. To this end, we removed all the contacts of the myristoyl beads except for the virtual bonds that connect the beads to each other.

2.4.3 All-atom Molecular Dynamics

All-atom molecular dynamics were performed using the AMBER simulation package [90]. The TIP3P water model was used and all hydrogens were constrained to have rigid bonds, with a 2 fs timestep being used. The initial wild-type (WT) model was built from the NMR structure (PDB 1HCD), and the two binding pocket mutants (I85L, F6L/I85L/I93L) were initially constructed using PyMol (www.pymol.org). AMBER parameters for the myristoyl group were determined using the RED RESP charge-fitting server package to be consistent with the AMBER03 forcefield used for the remainder of the protein [91]. The models were initially refined by conjugate gradient energy minimization until no further decreases in energy were found, and then equilibrated in an NPT ensemble (Langevin thermostat, Berdnensen barostat) for 20 ns. Final production was done for 80 ns, with snapshots taken every 20 ps, and analyzed using visual molecular dynamics [52].

2.4.4 Chemicals

Chemicals were analytical grade and obtained from Bioshops Inc. in Canada, unless otherwise specified. Urea concentrations were confirmed by refractometry.

2.4.5 Protein Purification

Site-directed mutagenesis for mutant hisactophilin were performed by GenScript Inc. (Piscataway, USA). Myristoylated hisactophilin was purified as described previously in Chapter 3 and [8]. Briefly, cells containing myristoylated hisactophilin were resuspended in Tris buffer pH 8 and lysed using an Emulsiflex C-5 emulsifier (Avestin Inc.). Next, 0.5% (w/v) [3-(3-Chloramidopropyl)] dimethylammonio-1-propanesulfonate (CHAPS) was added to the crude cell lysate, which was then incubated for 2 h at 4 °C to facilitate solubilization of membrane-bound hisactophilin. Hisactophilin was purified using DEAE (Biorad Inc.) anion exchange chromatography, followed by gel filtration using a HiLoad 26/60 Superdex 75 column (GE Healthcare). Separation of myristoylated and non-myristoylated hisactophilin was obtained using RP-HPLC with an acetonitrile gradient using a C₁₈ column (Waters Inc.). Purified protein was exchanged into 25 mM ammonium carbonate, lyophilized, and stored at -80 °C.

2.4.6 Equilibrium and Kinetic Folding Measurements

Equilibrium denaturation curves were measured as described previously and in greater detail in Chapter 3 [8]. Stock protein solution was prepared by dissolving lyophilized hisactophilin to a concentration of 2 mg/mL in either 500 mM 2-morpholino-ethanesulfonic acid (MES) or 500 mM potassium phosphate. Protein stock was diluted tenfold in water and stock urea to the desired final concentration of urea. Samples were equilibrated at 25.0 °C in a water bath for at least ten half-lives as measured for kinetic folding/unfolding transitions. Samples were monitored by fluorescence using a Fluorolog3-11 spectrofluorometer (Horiba-Jobin-Yvon Spex Inc.) as described previously with excitation and emission wavelengths of 277 nm and 306 nm, respectively [8]. Equilibrium denaturation was also monitored by circular dichroism at 227 nm using a J715 spectropolarimeter (Jasco Inc.) as described previously [8]. Folding rates were measured using the Fluorolog3-11 interfaced with a SFM4/Q (Molecular Kinetics Inc.). Kinetic data were fit using the Biokine 2.1 software (Molecular Kinetics Inc.). Data were then fit to a two-state unfolding model using the binomial extrapolation method (BEM) as described previously [8].

2.4.7 NMR Experiments

¹⁵N-labeled hisactophilin was prepared by growing *E. coli* in M9 minimal media with ¹⁵NH₄Cl (Cambridge Isotopes, MA) as the sole nitrogen source as described previously [8]. Resonance assignments were made using 2D homonuclear NOESY and TOCSY as well as 3D ¹⁵N-edited HSQC-NOESY and HSQC-TOCSY spectra of myristoylated hisactophilin at pH 6.8. Assignments for non-myristoylated hisactophilin were used as a starting point for making assignments of the myristoylated protein, followed by confirma-

tion and extension using standard procedures. Assignments were obtained for the α -proton and amide nitrogen/proton resonances of 96 of 118 residues; the majority of unassigned residues are in the loop consisting of residues 25–32.

2.4.8 Amide Exchange

Amide hydrogen exchange samples were prepared containing approximately 1 mM ^{15}N -enriched myristoylated hisactophilin in 50 mM phosphate buffer at pH 8.1 or pH 5.9 and then lyophilized. Exchange was initiated by dissolving lyophilized protein in D_2O , the sample was thermally equilibrated in the NMR spectrometer (Bruker Avance TCI 700 cryoprobe), and acquisition began after a dead time of approximately 22 min. Amide exchange decays were monitored using successive ^{15}N - ^1H HSQC spectra using gradients for artifacts and water suppression as described previously (8)[88]. A spectrum acquired every 12 min over the first 72 h and rechecked periodically thereafter for 120 days. NMR data were processed using Bruker NMRSuite. Amide exchange decay rates, k_{obs} , were monitored by integrating amide cross-peaks and fitting successive areas to a single exponential decay. Intrinsic exchange rates, k_{int} , were calculated using the SPHERE server (<http://www.fccc.edu/research/labs/roder/sphere/>). Protection factors, $P = \frac{k_{\text{obs}}}{k_{\text{int}}}$ (where k_{obs} is the observed rate constant for exchange and k_{int} is the intrinsic exchange rate in a random coil) for myristoylated hisactophilin were calculated for comparison with simulated values.

2.5 Results and Discussion

2.5.1 Hydrophobic and Electrostatic Nonnative Interactions are Necessary to Accurately Simulate the Folding of a Myristoylated Protein

We first simulated the folding of hisactophilin in its non-myristoylated form using the native topology-based (Go) model that considers only native interactions found in the high-resolution structure of the protein (Figure 2.1C, SI Results). For hisactophilin, we observed two-state folding with a relatively high energy barrier ($5.9 \times T$) (Figure 2.1C) consistent with experimentally observed relatively slow folding [71]. While the native topology-based model captures many features of the folding energy landscape, it neglects the roughness of the landscape due to nonnative interactions, which surely exist to some extent. Nonnative interactions may transiently form in either the unfolded state or the transition state and influence both folding thermodynamics and kinetics. We then modeled potential nonnative interactions by the formation of nonspecific interactions between hydrophobic residues (HP model) as well as electrostatic nonnative interactions among Asp, Glu, Lys, and Arg residues (model His0) and with positively charged histidines (model His1) (Figure 2.1D, SI Results).

The modeling is consistent with experimentally observed large decreases in hisactophilin folding rate with decreasing pH [71,84].

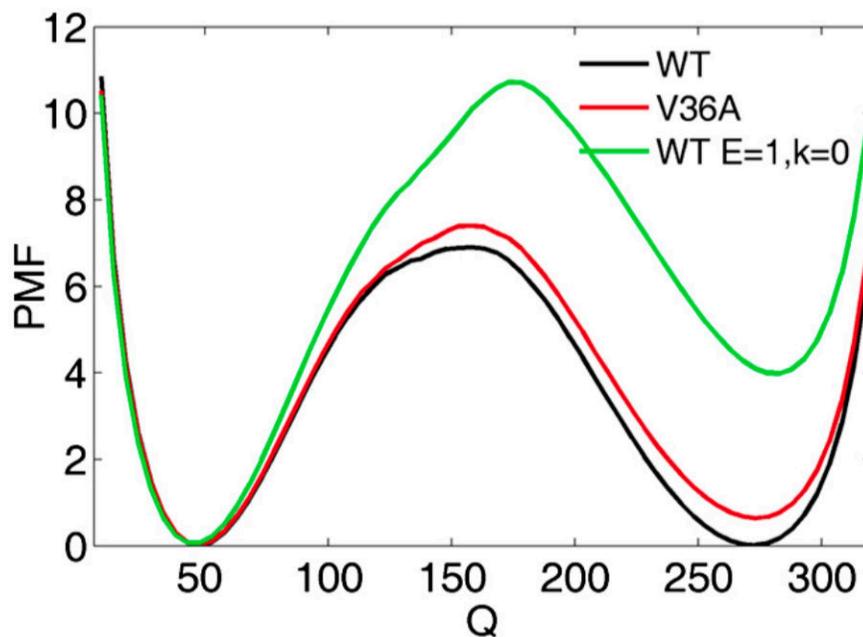


Figure S2.1: The free energy profiles for folding of myristoylated hisactophilin. Profiles are shown with (black) and without (green) non-native interactions between the myristoyl and the protein. All PMFs are shown at T_F of WT myristoylated hisactophilin. Free energy profile for V36 is shown in red where all the nonnative hydrophobic interactions between V36 and the myristoyl group and the protein have been removed. The mutation V36A results with higher free energy barrier. The role of nonnative interactions is illustrated by the folding barrier for the system in which the myristoyl has only native interactions with the protein but no nonnative hydrophobic interactions.

We then studied the effect of the myristoyl group on folding kinetics using a coarse-grained simulation. The energy barrier at T_F was increased from 5.4 kT to 7.7 kT due to myristoylation. Adding the hydrophobic nonnative interactions between the myristoyl group and the protein resulted in a dramatic acceleration of folding and a decrease in the energy barrier to 6.6 kT (Figure 2.1D and Figure S2.1). Note that the energy barrier of the myristoylated protein is higher than that of the non-myristoylated protein (even when the hydrophobic nonnative interactions are included), because the barrier heights were estimated at T_F of each system. When, however, the folding barrier heights were measured at the same temperature for all the systems, the experimentally observed accelerated folding upon myristoylation is reproduced. Integration of electrostatic nonnative interactions (at low and high pH) within the myristoylated protein did not have an additional effect on the energy barrier. From examination of the distances between residues and the change in distances upon integration of the various nonnative interactions we conclude that the major effect on the folding is in the transition state ensemble. Hydrophobic nonnative interactions position the myristoyl group 15 Å closer to the other hydrophobic residues in the protein with respect to the native topology-based simulations that included only native interactions (Figure S2.2). The electrostatic interactions, especially at high pH, had a dual effect

on the distances; the nonpolar residues were still closer but the His residues were slightly farther from one another. It is possible that simulations at different ionic strength would also affect electrostatic interactions and therefore folding, as shown for RNA folding [92]. The incorporation of the nonnative interactions (both hydrophobic and electrostatic) also influences the thermal stability of the protein as measured by the simulation, with a pH-dependence that is consistent with experiment [8,87] (SI Results).

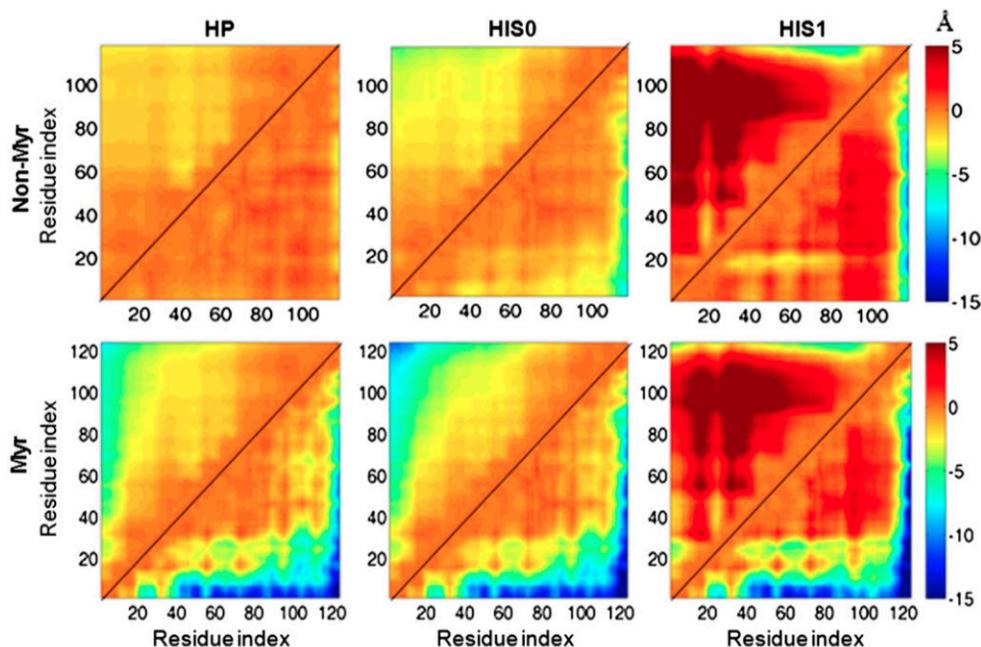


Figure S2.2: The Δ distance matrices of the transition state (lower) and unfolded state (upper) ensembles calculated relative to the corresponding ensemble obtained using the native topology-based models. Pair-wise distances between residues i and j , $R_{ij} = R_{ij}^{\text{nonnative}} - R_{ij}^{\text{Go}}$. Red colour indicates larger distance while blue colour indicates contraction due to the nonnative interactions.

To better understand the effect of the myristoylation on the kinetics and stability of the protein, we gradually changed the strength of the native and nonnative interactions between the myristoyl group and the protein. To this end, we gradually changed the value of ϵ and χ —which correspond to the strength of native and nonnative contacts, respectively—formed between the myristoyl group and the relevant amino acids. When ϵ equals unity, the contacts between the myristoyl group and the protein are equivalent in their energy contribution to the rest of the contacts within the protein. On the other hand, when ϵ equals zero, these contacts do not make any enthalpic contribution. The contacts between the myristoyl and the protein are shown in Figure 2.1A. The matrix in Figure 2.2A presents variation in the T_F of myristoylated hisactophilin for different values of χ and ϵ . In general, the T_F is increased (i.e., the protein is more thermostable) as ϵ and χ are increased, probably because more interactions are formed between the myristoyl and the protein and because the enthalpy is larger. Calculation of the average radius of gyration for the protein, $\langle R_g \rangle$, shows that the folded state is mostly affected by variations in the strength of ϵ and χ while the $\langle R_g \rangle$ of the unfolded state is robust (Figure

S2.3). We therefore conclude that the major effect on the free energy is on the folded state. The changes in the dimensions of the folded state are in agreement with changes in the enthalpy and entropy of the folded state. The more compact folded state with increased ϵ and χ has correspondingly lower enthalpy and lower entropy, which results in an overall decrease in the free energy of the folded state. It is notable that the strength of nonnative interactions in the simulations seems to have an effect on both the native and the transition state. The nonnative interactions in the native state may represent interactions involved in switching.

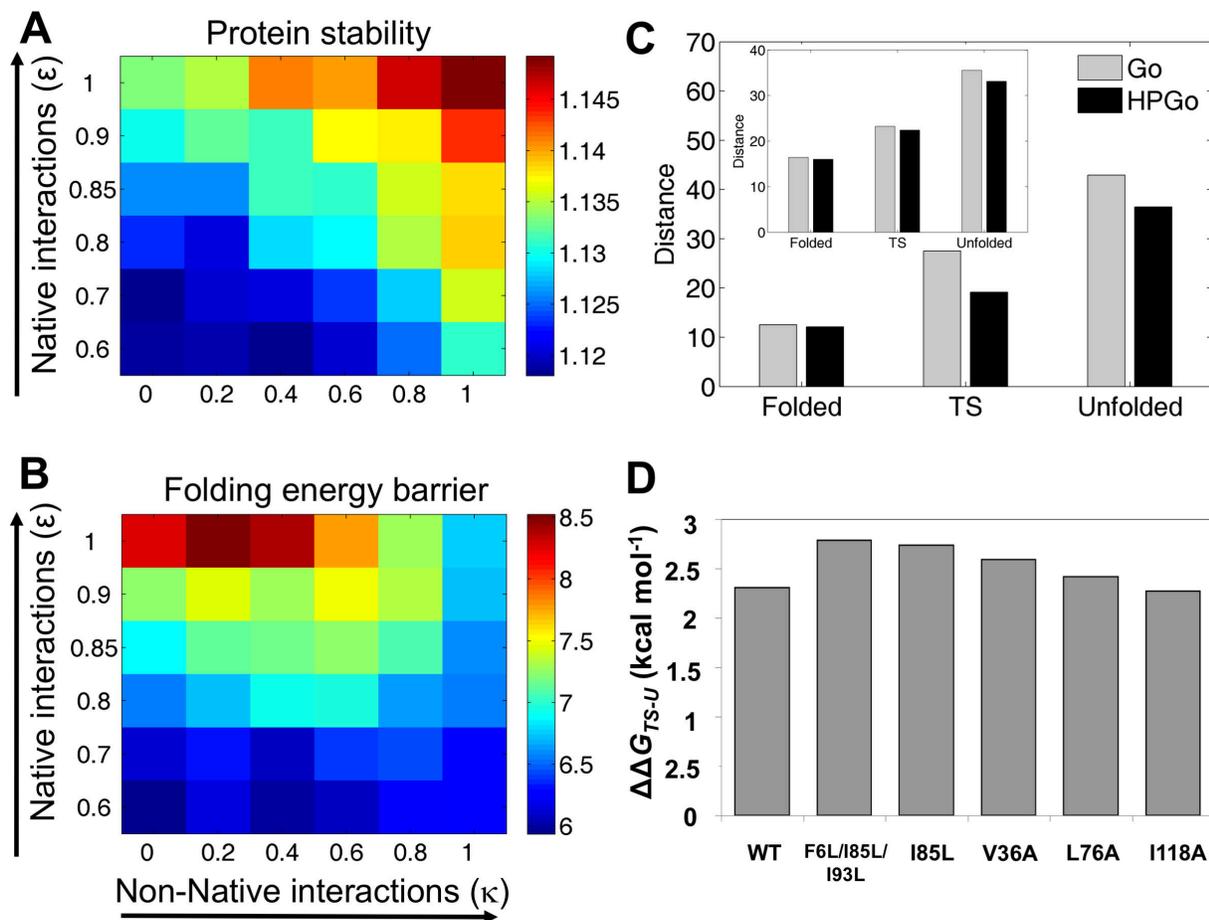


Figure 2.2: Interplay between nonnative interactions on the folding barrier and stability. (A) The value of the folding temperature, T_F , at different strengths of native (ϵ) and nonnative (χ) interactions between the myristoyl and the protein. (B) The value of the free energy barrier for folding at different values of ϵ and χ (at T_F). In these simulations, the ϵ of the protein contacts equals 1. $\chi=0$ corresponds to the pure native topology-based model. (C) The median of the distances of all the hydrophobic nonnative pairs between the protein and the myristoyl in the folded-, unfolded-, and transition-state ensembles for the native topology-based model (Go, grey) and the model supplemented by nonnative hydrophobic interactions (HP_{Go}, black). The inset shows the median of all hydrophobic nonnative interactions within the protein. (D) The decrease in the energy barrier for folding of WT and mutant hisactophilin upon myristoylation, $\Delta\Delta G_{TS-U} = m_{f,avg} \cdot [(urea)_{ln(kf=0),myr} - (urea)_{ln(kf=0),nonmyr}]$, calculated using the average denaturant dependence for \ln_{kf} , $m_{f,avg}$, for all the variants (Figure S2.4).

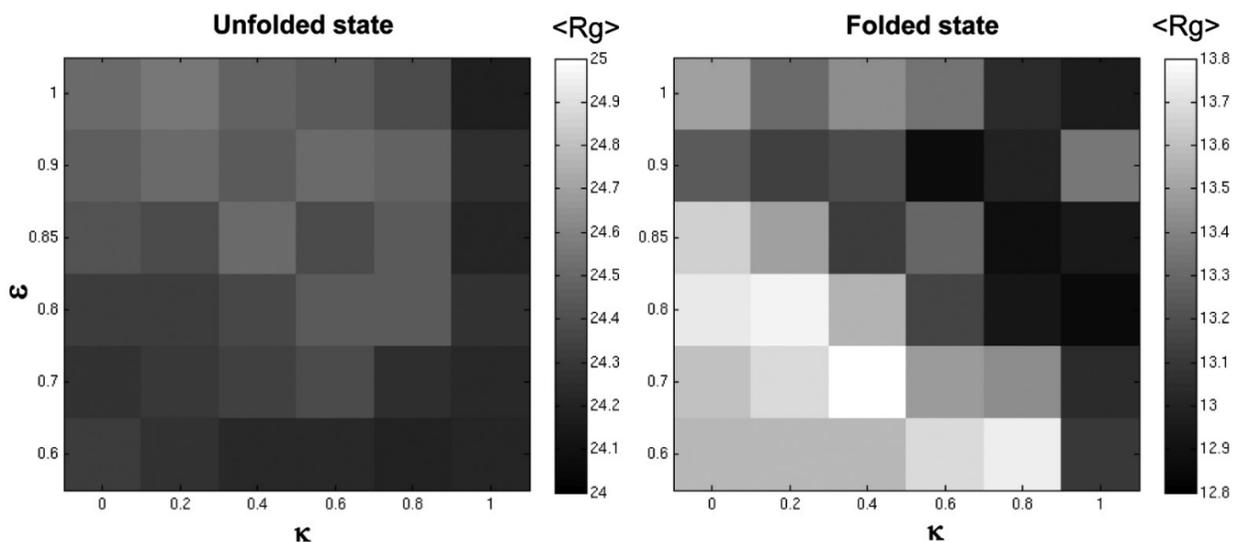


Figure S2.3: Compaction of the unfolded and folded states as a function of the strength of native and nonnative interactions. The compactness of the hisactophilin structure illustrated by the value of the mean radius of gyration, $\langle Rg \rangle$, of the unfolded state (Left) and the folded state (Right) of myristoylated hisactophilin modeled with different strengths of ϵ (native) and χ (nonnative) contributions. In the calculation of $\langle Rg \rangle$ only the beads of the protein are considered.

2.5.2 Analysis of Myristoyl Interactions in the Transition State and the Native State

Figure 2.2B presents a summary of the folding energy barrier of the myristoylated protein for differing strengths of native and nonnative interactions. The matrix illustrates that the energy barrier for folding increases with the value of ϵ —i.e., when the myristoyl is forced to be sequestered in the protein (and interacts strongly with its pocket). When hydrophobic nonnative interactions are included, the energy barrier may be significantly decreased, in accordance with the experimental finding that the nonnative interactions play a role in accelerating folding kinetics of myristoylated hisactophilin [8]. Notably, the increase in the energy barrier is significantly moderated when the nonnative interactions are included ($\chi=1$) compared to when they are omitted ($\chi=0$). The formation of the hydrophobic nonnative interactions at the folding transition state is illustrated by shorter pairwise distances between hydrophobic residues. These distances in the transition state

ensemble are generally shorter for nonnative interactions between the myristoyl and hydrophobic residues than between two hydrophobic residues (Figure 2.2C).

Experimental measurement of folding kinetics for hisactophilin variants can provide information on the formation of interactions for specific groups (e.g., myristoyl and/or amino acid sidechains) during protein folding. As mentioned above, such measurements have shown that the myristoyl group greatly accelerates the folding of wild-type (WT) hisactophilin and that this involves the formation of nonnative interactions in the transition state [8]. The effects of mutations (Figure 2.1B) that abolish switching (e.g., F6L/I85L/I93L (25) and I85L, *vide infra*) and truncate hydrophobic residues inside and outside the myristoyl binding pocket (V36A, L76A, and I118A) range from slightly increasing to considerably decreasing the folding rate (Figure S2.4). Remarkably, however, the large increase in folding rate conferred by myristoylation remains largely unchanged for all mutants (Figure 2.2D). In other words, none of the mutated residues appear to make critical, specific interactions with the myristoyl group in the transition state ensemble. This suggests that the myristoyl group accelerates the rate of folding by making nonspecific, including nonnative, interactions in the transition state. These interactions may be highly robust to mutation and have the effect of smoothing the folding energy landscape and reducing frustration.

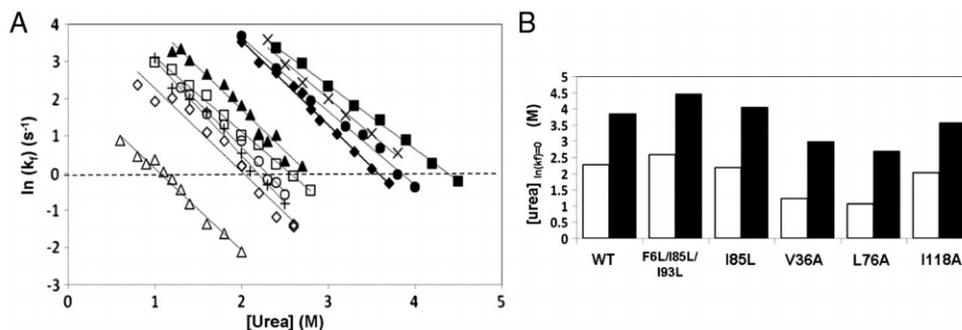


Figure S2.4: Observed folding rate constants versus urea concentration for variant hisactophilins. (A) Rate constants are shown for myristoylated proteins: L76A (closed triangles), WT (closed circles), I85L (Xs), I118A (closed diamonds) and F6L/I85L/I93L (closed squares); and for the corresponding non-myristoylated proteins: L76A (open triangles), WT (open circles), I85L (crosses), I118A (open diamonds) and F6L/I85L/I93L (open squares). The dashed line illustrates where $\ln(k_f) = 0$. (B) Urea concentration where the rate constant for folding of variant hisactophilin, k_f , is $1/\text{s}$, i.e. $[\text{urea}]_{\ln(k_f)=0}$ where $\ln(k_f) = 0$, for non-myristoylated (white) and myristoylated (black) proteins. Proteins with higher values of $[\text{urea}]_{\ln(k_f)=0}$ fold faster.

2.5.3 Energetic Myristoyl Switch with pH

Experimental studies of hisactophilin show a clear energetic switch in the myristoylated relative to non-myristoylated form upon changing from low to high pH [8]. At low pH, in excess ligand (H^+) the stabilization upon myristoylation is much smaller than at high pH, where the absence of bound H^+ differentially increases the stability of both the myristoylated and non-myristoylated protein (Figure 2.3A). Assuming that the change

in pH mostly affects the protonation state of histidine residues, we simulated the protein using the His0 and His1 models (which model the protein at high and low pH, respectively). In the simulations the myristoylated protein is more stable than the non-myristoylated protein both at low and high pH (Figure 2.3B), similar to experiment. However, the modulation of the stabilization with pH could not be simply observed in the simulations, but only by assuming that the strength of the interactions of the myristoyl with the protein (either native or nonnative) is indirectly affected by the pH due to the change in the histidine protonation state. This energetic switch is accompanied by a conformational switch: at low pH the myristoyl group is in the accessible state and at high pH it is sequestered. We observe this dual effect of energetic and structural switching when we manipulate the strength of the native or nonnative interactions between the myristoyl group and the protein. In this way, we can decrease the amount of extra stability that was achieved by adding the myristoyl, implying that changes in pH not only affect the protonation state of the histidine residues, but also change the strength of the interactions between the protein and the myristoyl. We conjecture, therefore, that the switch originates from a thermodynamic effect. In the absence of ligand (at high pH), the protein is thermodynamically stable and undergoes minor structural fluctuations; therefore the probability of the switching between the sequestered and accessible states is low (high free energy). In excess of ligand (at low pH), the protein is less stable, the probability for the switching is higher, and the accessible state is more populated. This leads to a lower thermodynamic stability due to the loss of enthalpic contributions from the direct interaction of the myristoyl and the hydrophobic pocket. Thus, simulations provide evidence that the experimentally observed energetic switch is governed by the balance of native and nonnative interactions. This highlights the importance of modelling nonnative interactions for understanding the mechanism of myristoyl switching.

We then analyzed the contributions of individual residues to switching, using a combination of experimental measurements and atomistic simulations for mutant proteins. In previous experiments, we found that the switch in hisactophilin is broken in the triple myristoyl binding pocket mutant, F6L/I85L/I93L, such that the myristoyl group remains in the sequestered state and does not switch to the accessible state with decreasing pH [5] (Figure 2.3C). We now report that the single mutation, I85L, breaks the switch in the opposite way—i.e., by weakening interactions of the myristoyl with the protein such that the mutant protein remains in the accessible state. It is noteworthy that this highly conservative mutation, changing only the stereochemistry of a single sidechain, essentially abolishes switching. We speculate that moving the branch point in the sidechain from the β - to the γ -carbon may create steric clashes with the myristoyl, interfering with it adopting its fully sequestered conformation, and concomitantly increasing protein dynamics such that coupling between the myristoyl binding pocket and sites of protonation is disrupted.

The F6L/I85L/I93L and I85L mutations have dramatically different effects on the transition state compared to the native state. Rather counterintuitively, despite the broken switching, both F6L/I85L/I93L and I85L fold slightly faster than WT in the myristoylated form (Figure S2.4A). This suggests that the folding

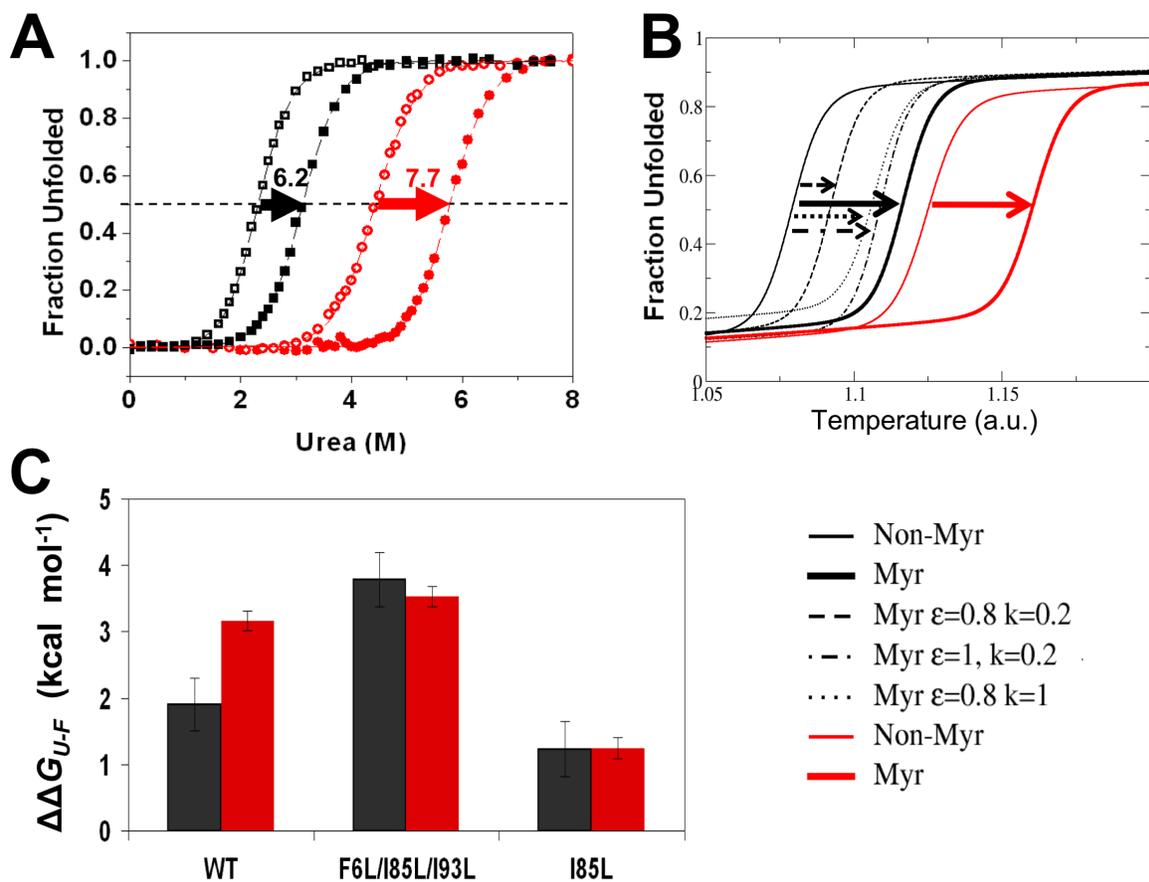


Figure 2.3: Energetics of myristoyl switching. (A) Equilibrium urea denaturation curve measurements of stability for non-myristoylated (open symbols) and myristoylated (closed symbols) hisactophilin at pH 6.2 (black squares) and 7.7 (red circles). Stabilization upon myristoylation, $\Delta\Delta G_{u-f} = \Delta C_{mid} \times m_{avg}$, calculated from the difference in the midpoint of denaturation, ΔC_{mid} , for myristoylated relative to non-myristoylated hisactophilin multiplied by m_{avg} , the average denaturant dependence of ΔG_{U-F} for the 2 forms of hisactophilin [8]. (B) Simulated fraction unfolded vs. temperature for myristoylated (solid bold line) and non-myristoylated (solid line) hisactophilin at low pH (black) and high pH (red) under different strengths of ϵ (native) and χ (nonnative) myristoyl interactions. (C) Stabilization upon myristoylation, $\Delta\Delta G_{U-F}$, for WT, F6L/I85L/I93L, and I85L hisactophilin at pH 6.2 (black) and pH 7.7 (red)

of the WT protein is slightly slowed—i.e., frustrated—by residues required for switching functionality. Nevertheless, as mentioned above, the energetics of the transition state relative to the unfolded state are robust (Figure 2.2D). In contrast, the energetics of the native state and protein function are dramatically affected by the mutations (Figure 2.3C). This suggests that I85L introduces too much strain into the native state to allow for switching, while F6L/I85L/I93L causes too large a reduction in strain, consistent with altered dynamics observed in atomistic simulations (*vide infra*).

2.5.4 Structural Switch of the Myristoyl among Sequestered, Accessible, and Exposed States

From our folding simulations, we identified the conditions where myristoyl switching occurs. Figure 2.4A shows a two-dimensional free energy surface for the coupling between folding of hisactophilin (depicted by Q_{Folding}) and switching of the myristoyl group (depicted by $Q_{\text{Prot-Myristoyl}}$) for three sets of ϵ - and χ -parameters for native and nonnative interactions, respectively, of the myristoyl group. At low values of Q_{Folding} , the protein is unfolded and the myristoyl group is highly exposed to solvent (i.e., $Q_{\text{Prot-Myristoyl}}$ is low but some sporadic interactions between the myristoyl and the protein are found). At high values of Q_{Folding} , the protein is folded and the myristoyl forms many more contacts with the protein as it fits in the hydrophobic pocket. When $\epsilon=0.6$ and $\chi=0$, the enthalpic stabilization is insufficient for populating the sequestered state. However, increasing the strength of the native and nonnative interactions ($\epsilon=0.8$ and $\chi=0.2$) results in the full insertion of the myristoyl into its pocket. The switching of the myristoyl from the exposed to the sequestered state follows the folding of the protein (Figure 2.4A). Only when the strength of the interactions of the myristoyl with the protein are sufficiently strong (close to unity), a coupling between folding and switching emerges (i.e., the folding follows more two-state rather than three-state behaviour). In this scenario, the strong coupling between folding and switching can be reduced when the strength of nonnative interactions is increased, as they may allow the protein to populate states other than the fully sequestered state while the rest of the protein is folded.

To further elucidate the switching mechanism, we analyzed the position of the tip of the myristoyl with respect to the bottom of its hydrophobic pocket; specifically, we measured the distance between the bead representing C_{13} and C_{14} of the myristoyl group and the centre of mass of the beads representing the alpha carbons of three residues at the bottom of the barrel (V21, V61, and V101). This analysis revealed three typical structural states of the myristoyl group relative to the protein. The two limiting states correspond to the myristoyl being fully sequestered in the protein binding pocket versus the fully solvent exposed state of the myristoyl (states I and III in Figure 2.4). In the remaining state, the myristoyl is partially accessible to the solvent (state II) and is stabilized by nonnative hydrophobic interactions with residues located at the rim of the

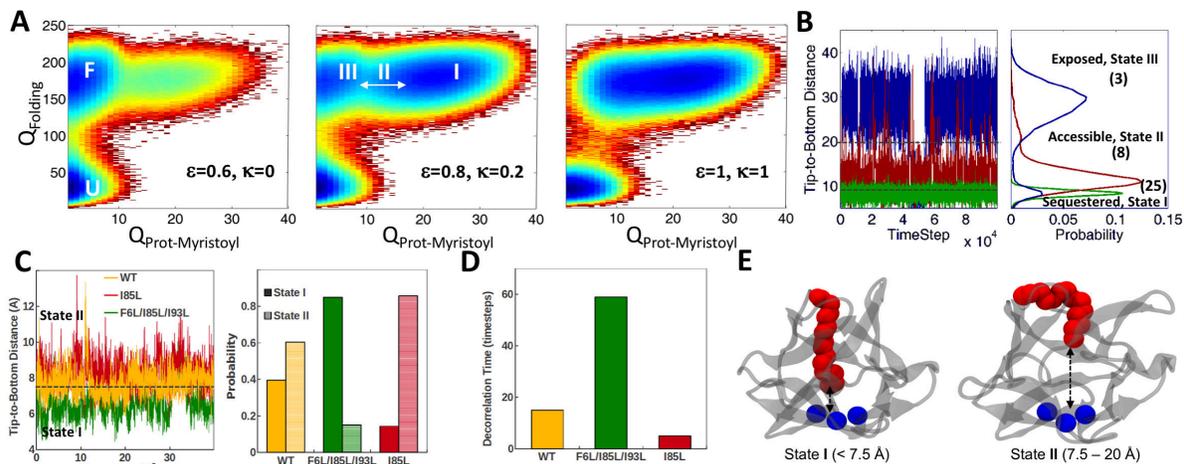


Figure 2.4: Structural characterization of the myristoyl switching mechanism. (A) The coupling between folding and switching is shown by projecting the free energy onto two reactions coordinates: Q_{Folding} (number of native contacts within hisactophilin) and $Q_{\text{Prot-Myristoyl}}$ (number of native contacts between the myristoyl and hisactophilin) for $\epsilon=0.6$, $\chi=0$; $\epsilon=0.8$, $\chi=0.2$; and $\epsilon=1$, $\chi=1$. Blue and red areas correspond to highly and poorly populated states, respectively (free energy in units of kT). (B) Definition of the three states of the myristoyl relative to the bottom of the barrel: fully sequestered (state I, $<7.5 \text{ \AA}$), accessible (state II, $7.5\text{--}20 \text{ \AA}$), and exposed (state III, $>20 \text{ \AA}$). The numbers in brackets refer to the average $Q_{\text{Prot-Myristoyl}}$ in each state. The blue, red, and green trajectories were simulated using different values of (ϵ ; χ) which equal (0.6; 0), (0.8; 0.2), and (1.0; 1.0), respectively. (C) Analysis of atomistic molecular dynamics simulations for WT, F6L/I85L/I93L, and I85L showing effects of mutations on the probability of populating states I and II. (D) Decorrelation time of the tip- to-bottom distance. (E) Representative conformations of the myristoyl for states I and II that were found in atomistic simulations.

pocket (Figure 2.4E). This intermediate state (state II, in which the myristoyl forms about 8 native interactions with the protein compared to 25 in the fully sequestered state, Figure 2.4B) resembles the accessible state that was observed in experiments as pH was lowered. In this accessible state, the myristoyl group is less buried; however, there remain significant interactions between the myristoyl and the protein, based on measurements of energetics and NMR data [8].

In the simulations, the strength of native and nonnative interactions between the protein and the myristoyl has a major impact on the position of the myristoyl with respect to the protein. In general, we see that when the native interactions are strong ($\epsilon \geq 0.9$), the protein is mostly in the sequestered state (state I) regardless of the strength of the hydrophobic nonnative interactions. The exposed state (state III) is common in simulations in which both the native and nonnative interactions are weak (χ and $\epsilon < 0.6$). The partially accessible state (state II) is most common when the nonnative interactions are strong ($\chi = 1$), but native contacts between the myristoyl and the pocket are weak ($\epsilon < 0.7$, Figure 2.4A and Figure S2.5). Based on previous analyses of the pH dependence of stability and NMR data, switching from the sequestered state to the accessible state is caused by the binding of approximately 1.5 protons by a small number of histidine residues localized on one side of the protein [5,8]. Also, the protonation of the histidines is coupled to the state of nearby hydrophobic residues including F6, I85 and I93, which sensitively communicate pH changes to the myristoyl binding pocket as evidenced by switching being abolished in the F6L/I85L/I93L [5] and I85L (Figure 2.3C) mutants.

Therefore, we speculate that in hisactophilin at low pH, where the accessible state is favoured, the nonnative interactions between the protein and the myristoyl are strong while histidine protonation weakens the native protein-myristoyl interactions with respect to the rest of the native interactions in the protein, and that this combination leads to the intermediate position of the myristoyl group relative to the protein.

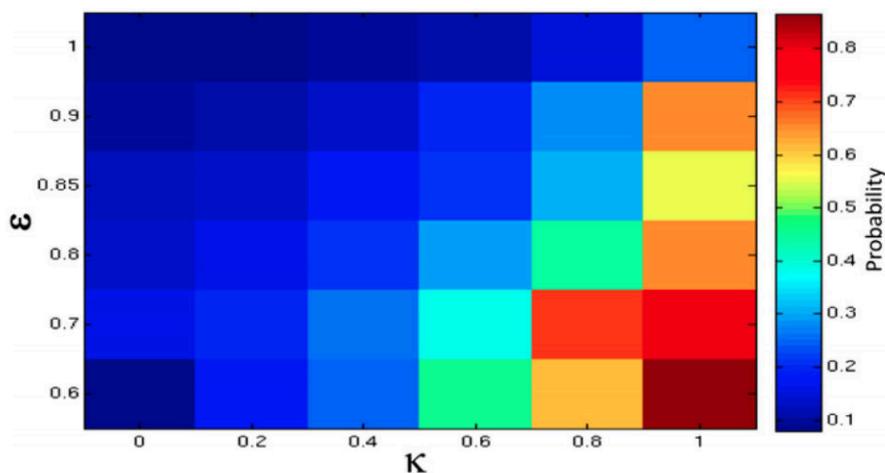


Figure S2.5: The probability of populating state II (the myristoyl accessible state) at different strengths of native, ϵ , and non-native, χ , protein-myristoyl interactions.

We used molecular-dynamics simulations of the folded state for WT, F6L/I85L/I93L, and I85L to assess the extent of burial of the myristoyl group in the protein binding pocket by measuring the distance from the tip of the myristoyl to the bottom of the barrel. We also calculated the time for the position of the myristoyl group in the protein to become decorrelated, which provides a measure of myristoyl dynamics. The simulations suggest that WT accesses both the sequestered and partially accessible states (states I and II). In contrast, F6L/I85L/I93L is predominantly in a sequestered state (state I), with the myristoyl inserted further into the protein than in WT and little fluctuation of the myristoyl into the accessible state (state II). On the other hand, in I85L the myristoyl may be slightly less buried than in WT, with increased population of the accessible state (state II). In addition, the triple mutant transitions between states more slowly than WT based on increased decorrelation time for the position of the myristoyl group compared to WT, whereas the single mutant shows the opposite behaviour, transitioning more rapidly with a shorter decorrelation time (Figure 2.4D). Thus, both the location and dynamics of the myristoyl group in the atomistic simulations are consistent with the experimental energetic data (Figure 2.3C) and folding simulations (Figure 2.4A–C), which taken together indicate that the myristoyl: (i) makes stronger interactions with the protein in F6L/I85L/I93L, causing it to remain in the sequestered state; (ii) makes weaker interactions in I85L so that it stays in the accessible state II; and (iii) is poised to switch between the two states in WT.

2.5.5 Localized Stability Changes Associated with Myristoyl Switching

We obtained higher-resolution insights into the energetics of switching by further analysis of simulations combined with experimental measurements of amide H/D exchange rates. Full stabilization of the protein occurs only after the myristoyl is switched from the accessible to the sequestered state (transition from state II to state I). We identified 12 residues that become significantly more ordered after the myristoyl is fully inserted into the pocket (Figure 1.5A, blue). Some of these residues are close in sequence to the myristoyl and may be expected to behave in this way, whereas others are far in sequence but close in structure. Notably, the N- and C-termini of the protein exhibit the largest increase in order when the myristoyl becomes sequestered.

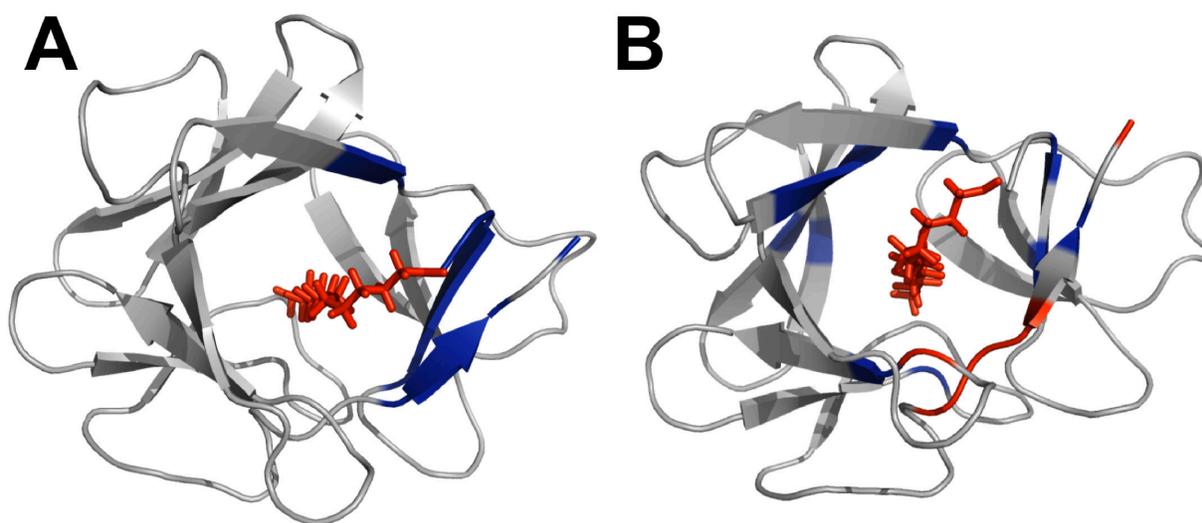


Figure 2.5: Changes in dynamics upon myristoylation. (A) Top view of hisactophilin with residues predicted to become ordered (blue, G2, N3, R4, A5, F6, K7, H35, V36, E114, E115, I116, I118) upon sequestering of the myristoyl group as observed by coarse-grained simulations. (B) Measured change in amide exchange protection factor for myristoylated compared to non-myristoylated WT hisactophilin, $\Delta P_{pH8.1} = \log(P_{myr} \cdot P_{non-my}^{-1})$ (Figure S2.6), at pH 8.1, where the myristoyl group is sequestered. R4, V43, A95, and I116 (blue, Figure S2.6) show the largest increase in protection upon myristoylation while other residues are slightly protected (blue); V83, T112, F113, E114, and I118 (red) show decreased protection. The myristoyl group is shown in red stick representations.

We obtained an experimental measure of changes in protein flexibility upon myristoylation using amide H/D exchange rate measurements for approximately 35 amides located throughout the protein structure (Figure 2.5B and Figure S2.6A). Increased amide H/D protection factors upon myristoylation are observed for many amides throughout the protein, at both pH 5.9 where the protein is predominantly in the accessible myristoyl conformation (state II) and at pH 8.1 where the protein is predominantly in the sequestered conformation (state I) (Figure S2.6A). This protection is consistent with myristoylation increasing the global protein stability, as has also been observed in chemical denaturation experiments [8,88]. Furthermore, the

largest increases in amide protection arising from the myristoyl group from pH 5.9 to 8.1 are observed for amides close to the N- and C-termini (Figure S2.6B). This is consistent with the increase in structure upon full insertion of the myristoyl group observed by simulation (Figure 2.5A). The amide protection factors also reveal that a small group of amides (V83, T112, F113, E114, and I118) show small decreases in protection upon myristoylation (Figure S2.6A). These residues are clustered at one side of the myristoyl binding pocket, suggesting that mobility in this region may be linked to and facilitate switching between accessible and sequestered states.

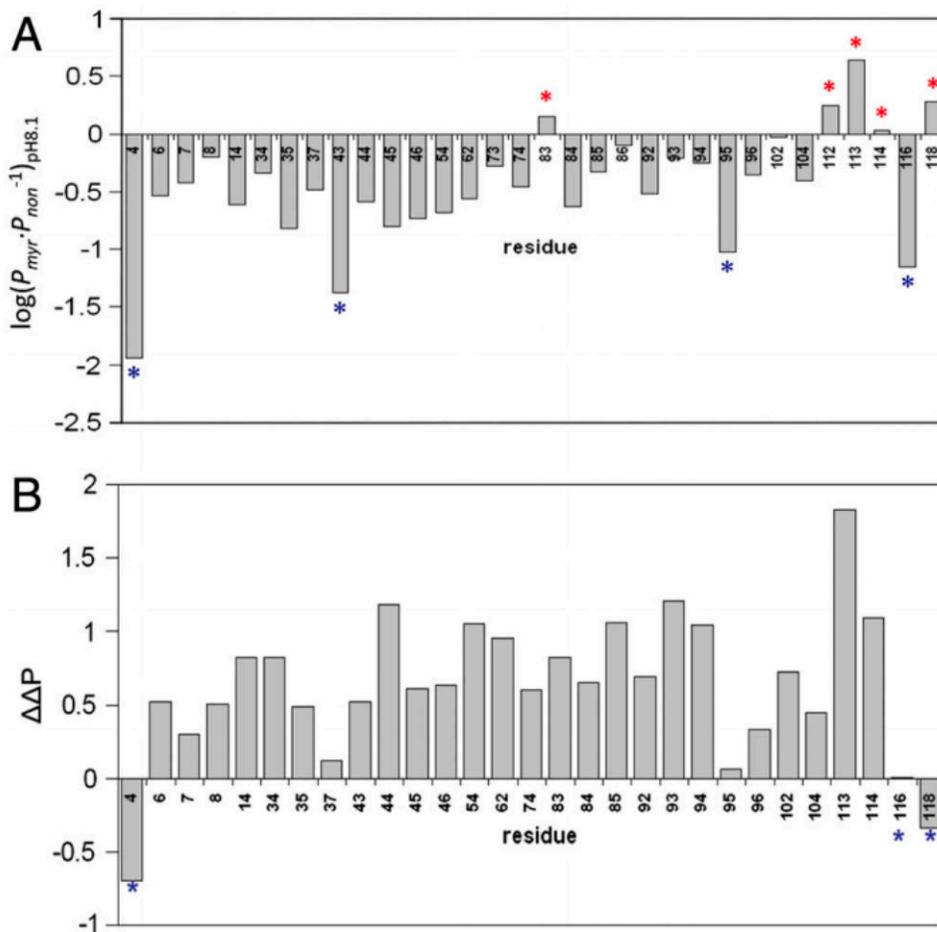


Figure S2.6: Amide H/D exchange protection factors in myristoylated hisactophilin relative to non-myristoylated hisactophilin. (A) Changes in amide protection factors, P , are plotted versus residue number. Change in protection upon myristoylation, $\Delta P_{pH 8.1} = \log(P_{myr} \cdot P_{non}^{-1})_{pH 8.1}$, for individual NH groups in hisactophilin at pH 8.1. The overall negative values of $\Delta P_{pH 8.1}$ reflect increased global stability upon myristoylation. Residues that are relatively more protected ($\Delta P_{pH 8.1} < 0$) and less protected ($\Delta P_{pH 8.1} > 0$) are marked with blue and red asterisks, respectively. (B) Change in protection upon switching, $\Delta\Delta P = \log(\Delta P_{pH 8.1} \cdot \Delta P_{pH 5.9}^{-1})$, for individual amide groups in hisactophilin. $\Delta\Delta P$ is a measure of the relative change in protection, and hence change in dynamics, associated with the myristoyl group switching from state II at pH 5.9 to state I at pH 8.1. Due to the very high sensitivity of hisactophilin stability to pH, there is a systematic offset for all $\Delta\Delta P$ values due to small differences in pH between myristoylated and non-myristoylated samples. Residues, marked with a blue asterisk, at the N- and C-termini show the most pronounced decrease in dynamics (indicated by relatively more negative values of $\Delta\Delta P$) upon myristoyl sequestration.

2.6 Conclusions

Post-translational modifications are a common means to regulate protein function. In many cases the regulation is achieved by the modification directly modulating the protein biophysical properties. Our study reveals molecular details of significant changes in protein stability and kinetics caused by myristoylation. The effects of myristoylation are more pronounced than those observed for glycosylation [63] and ubiquitination [93] and can be simply rationalized by the close interactions of the hydrophobic myristoyl group with its binding pocket, while in the case of glycan and ubiquitin conjugates, the interface with the protein is smaller and dominated by excluded-volume effects. The extensive interface the myristoyl forms with the protein may therefore change the folding enthalpy as well as the protein internal dynamics.

Using a range of complementary experimental and computational approaches, we characterized the pronounced effects of myristoylation on the folding of hisactophilin. The effects on folding cannot be explained solely by the native interactions defined in the NMR structure of hisactophilin with the myristoyl sequestered in its binding pocket [8]; additional nonnative hydrophobic and electrostatic interactions also have significant roles. The coarse-grained simulation models highlight the importance of nonnative hydrophobic interactions in reducing the free energy barrier for folding and stabilizing the protein. The simulations reveal an intermediate state, distinct from the sequestered one observed by NMR and a state with fully exposed myristoyl, where the myristoyl is partially accessible and makes stabilizing nonnative interactions with the protein. In simulations and experiments, hisactophilin has maximal global stability only after the myristoyl switches to the sequestered state. Furthermore, coarse-grained simulations and amide H/D exchange measurements reveal molecular details of switching mechanism involving changes in local stability clustered at one side of the binding pocket and including the N- and C-termini of the protein. The effects of mutating hydrophobic residues distributed around and outside the myristoyl binding pocket reveal the robust effect of myristoylation on folding kinetics, which contrasts with the relative ease of breaking native-state switching. These results suggest that the long and flexible nature of the myristoyl alkyl chain may enable sampling of a relatively broad, nonspecific ensemble of hydrophobic interactions in the transition state that is much more restricted and finely tuned in the native state.

In summary, a combination of computational and experimental approaches provides unique insights on how a common fatty acyl protein modification can tune folding through both native and nonnative interactions, and how changes in stability and dynamics control function. The effects of myristoylation in hisactophilin are similar to results for other proteins where regions of increased local stability are counterbalanced by regions of decreased local stability implicated in function regulation by ligand binding [94–96]. The folding energetics and mechanism for myristoylated hisactophilin support the notion that local energetic frustration and the accumulation of strain during switching can be linked to function as observed in allosteric proteins

[62] and for functional and/or hydrophobic residues in other proteins [59,60,97,98].

2.7 Supplemental Information

Hydrophobic and Electrostatic Nonnative Interactions are Necessary to Accurately Simulate the Folding of a Myristoylated Protein. We simulated the folding of hisactophilin in its non-myristoylated form using the native topology-based model that considers only native interactions found in the high-resolution structure of the protein. For hisactophilin, we observed two-state folding with a relatively high energy barrier (5.9 kT) (Figure 2.1C) consistent with experimentally observed relatively slow folding [87]. While the native topology-based model captures many features of the folding energy landscape, it neglects the roughness of the landscape due to nonnative interactions, which surely exist to some extent. Nonnative interactions may transiently form in either the unfolded state or the transition state and influence both folding thermodynamics and kinetics. We first modeled potential nonnative interactions by the formation of nonspecific interactions between hydrophobic residues (HP model). The integration of hydrophobic nonnative interactions did not change the folding energy barrier, implying that hydrophobic nonnative interactions do not participate in the folding and are likely to have a minor effect on folding of non-myristoylated hisactophilin. A comparison of all the pairwise distances in the unfolded state indicates more compact conformations in the ensemble of the unfolded state when hydrophobic nonnative interactions are included in addition to the native interactions. In the ensembles of transition and folded states, the differences due to the hydrophobic interactions are smaller.

In addition to nonnative interactions between hydrophobic residues, electrostatic interactions may also contribute to the roughness of the energy landscape. The effects of nonnative electrostatic interactions compared to hydrophobic interactions obviously depend on the protein sequences as well as the fact that the electrostatic interactions are long-range while the hydrophobic forces are short-range by nature. Transient interactions between hydrophobic residues and between charged residues can both increase the ruggedness of the landscape and change the folding thermodynamics and kinetics. Integration of electrostatic nonnative interactions among Asp, Glu, Lys, and Arg residues together with the hydrophobic nonnative interactions (model His0) results in a more compact unfolded state compared to the unfolded state generated from the model with hydrophobic nonnative interactions only. The electrostatic nonnative interactions result in a higher folding barrier (an increase of approximately 5% in comparison to the HP model, Figure 2.1D). Hisactophilin contains 31 histidine residues (average pK_a value of approximately 6.8 [74]), which may also participate in electrostatic nonnative interactions. When we modeled the histidine residues with positive charge (model His1), we observed a significantly increased folding barrier (approximately 13%) owing to the repulsion among the numerous positive charges in the protein. The repulsive interactions are reflected by the increased distances between specific areas in the protein in the unfolded and transition states. The model-

ing is consistent with experimentally observed large decreases in hisactophilin folding rate with decreasing pH [8,71].

The incorporation of the nonnative interactions (both hydrophobic and electrostatic) also influences the thermal stability of the protein as measured by the simulation. Hydrophobic nonnative interactions slightly increased the protein stability by 0.1%. The inclusion of the electrostatic interactions between the charged residues increased the stability by 0.8%. When we then added charges to the histidine residues the stability dropped by 3.3%. From this we conclude that the hydrophobic and electrostatic nonnative interactions contribute to the thermodynamic stability of the protein, but when too many positive charges are present in the protein stability is decreased, as has also been observed experimentally [8,74]. In terms of stability, the hydrophobic nonnative interactions increased the T_F by 0.5% relative to the T_F when nonnative interactions were excluded. The electrostatic nonnative model at high pH increased the T_F by 1.1% and at low pH decreased the T_F by 2.7%.

Chapter 3: Thermodynamics of Myristoyl Switching

3.1 Chapter Abstract

Myristoylation, the covalent linkage of a C₁₄ fatty acid group to the N-terminal glycine of a protein, is a key feature of pH-dependent myristoyl switching. However, the effect of the myristoyl group on protein folding and its specific interactions with the native-state protein that enable switching are little understood, limiting the design of novel switching proteins and a general understanding of reversible membrane-binding proteins. We systematically characterized the equilibrium thermodynamics and folding kinetics of the myristoyl switch protein, hisactophilin, to tease out the pH-dependent contributions of interactions between sites of mutation in the core of the protein and the acyl group to protein folding and function. We show that, although the myristoyl group favours folding, it may also unfavourably interact with some residues to create strain in the native state that is relieved by switching or mutation. The myristoyl group consistently accelerates folding at pH 7.7 in the studied mutants, further supporting previous work indicating interactions between the myristoyl group and the rest of the protein in the folding transition state are less specific than in the native state. For folding kinetics on the other side of the myristoyl switch at pH 6.2, the impact of histidine ionization and the myristoyl group are complex and may suggest non two-state folding, proceeding, for example, via parallel pathways. The amount of native-state strain and the impetus to relieve it is a possible mechanism for driving or tuning the sensitivity of the pH-dependent switch, an understanding of which may be crucial to future efforts in characterizing and designing functional sensors and sensitive switches.

3.2 Introduction

Myristoylation is one of the most common lipid modifications of proteins in nature and facilitates interactions with other proteins or membranes. The effects on the energy landscape of adding a saturated C₁₄ fatty acid and its intramolecular interactions with the protein have, however, been explored only sparingly. Such interactions are particularly interesting in hisactophilin because the conformation of the myristoyl group is coupled to changes in cytosolic pH. At pH > ~7, the myristoyl group is predominantly sequestered in hisactophilin's core, but it switches to a more solvent-accessible conformation at pH values below ~7 [8,11]. This pH-dependent switching behaviour allows its host organism, *Dictyostelium discoideum*, to organize and control actin bundling for locomotion and stability in response to environmental stimuli. Being a relatively small protein of 117 amino acids (~13.5 kDa), hisactophilin is an accessible model for teasing out the network of interactions that link proton binding to myristoyl switching.

Hisactophilin, a β -trefoil, has three-fold pseudosymmetry with numerous structurally conserved residues

with other β -trefoil proteins, which constitute its major hydrophobic core (see Figure 3.1). The structurally conserved residues can be subdivided into layers of residues that occur at symmetry-related positions of the β -trefoil at different depths within the β -barrel (coloured residues in Figure 3.1). Some of these core residues diverge from a symmetrical sequence so, because the myristoyl group interacts with the core, is it possible asymmetry is functionally important for folding and function? Threefoil is an engineered β -trefoil with perfect threefold symmetry that exhibits extraordinary kinetic stability and is also a functional trivalent sugar binding protein [99,100]. Meanwhile, another engineered symmetric protein Symfoil, based on Fibroblast Growth Factor 1, has lost FGF-1-like activity, including receptor binding, heparin binding, localization, etc. [101]. So, in one case, symmetry permitted function while ablating function in the other. One of the potential difficulties arising from symmetry is that it may contribute to non-native interactions during folding that form traps, frustrating the folding landscape [102,103]. Threefoil, though, overcomes these traps in part due to its significant folding cooperativity [99].

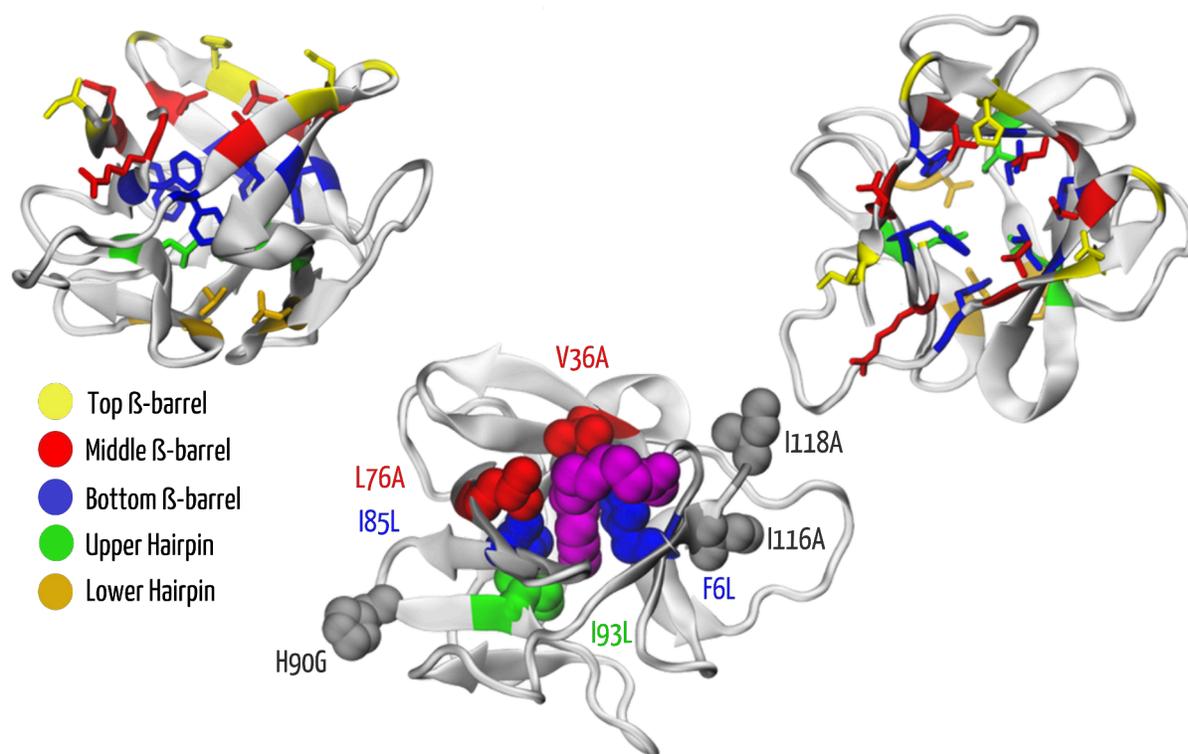


Figure 3.1: Structural models of the β -trefoil protein hisactophilin with the conserved residues of the core at different depths/layers. Mutations of interest are also shown. The layers in the upper panels are coloured as follows: yellow is the top of the β -barrel (comprising residues G2, N38, G41, H78, G81, and I117), red is the middle of the β -barrel (R4, V36, V43, L76, V83, and E115), blue is the bottom of the β -barrel (F6, F34, L45, F74, I85, and F113), green is the upper hairpin layer (L14, L53, I93), and gold is the lower hairpin layer (V21, L63, and V101). The 18 conserved core residues are those in the middle and bottom β -barrel and upper and lower hairpin layers. The lower panel indicates the locations of the mutations studied herein, coloured according to the layer to which they belong, and grey if they are elsewhere in the protein. The sidechains are represented by space-filling spheres and the N-terminal myristoyl group is shown as magenta spheres.

Hisactophilin is only pseudosymmetric, so, while it is not expected to suffer the same degree of unfavourable non-native interactions that may plague highly symmetric proteins [104], co/post-translational modifications might be expected to frustrate folding as they do for other proteins [63]. Interestingly, though myristoylation in hisactophilin appears to strain the native state, folding kinetics are enhanced by orders of magnitude [8,73], even in variants where partial symmetry is restored in the core (Chapter 2) [73]. So, the effects of the myristoyl group were contrary to expectation and also relatively unaffected by mutation.

These same mutations, spread over the protein, are also an avenue to characterize native state interactions between the myristoyl group, proton binding, and how the disruption of those interactions hampers coupling, allostery, and ultimately function. Some of the mutations came about as a result of earlier work that looked to systematically restore symmetry around the hisactophilin core at different depths/layers (represented by the colours in the top two panels of Figure 3.1) that are generally conserved in β -trefoils [105]. The symmetry-related mutants under scrutiny here are single, double, triple, and quadruple mutant combinations of core residues F6L, I85L, I93L, as well as H90G which is external to the core, in the β 9- β 10 loop. I93L restores 3-fold leucine symmetry in the upper hairpin layer, and F6L and I85L complete symmetrical elements in the lower β -barrel, close to the tip of the myristoyl group. We have already shown that the I85L mutation, and the triple mutant F6L/I85L/I93L, can each break the myristoyl switch and decouple proton binding from the state of the myristoyl group [5,73], and now we look to further dissect these interactions by considering the single mutations and pairs of mutations en route to the triple mutant. Chapter 2 introduced additional mutants that were informed by coarse-grained simulations to predict residues that were integral to stabilizing the transition state of folding with the myristoyl group. However, we suspect that mutations in these locations (all positions for mutations are shown in the lower panel of 3.1), V36A, L76A, I116A, and I118A may also be integral for balancing the energetics of switching.

Herein, we present an expansion of some of the equilibrium thermodynamic experiments explored in Chapter 2 to a wider array of mutations, as well as the folding kinetics at pH values on either side of pK_{switch} to characterize the interactions between folding and function, pH and the myristoyl group, and the bearing that core symmetry has on tuning the pH-dependent myristoyl switch.

3.3 Materials and Methods

3.3.1 Expressing Hisactophilin

Purifying hisactophilin has followed two tracks; the original method implemented by a previous graduate student, Chengsong Liu [106] and expanded by Joe Meissner to include handling the myristoylated variant [19], and then a newer protocol I developed that was inspired by the poorly behaved mutants first studied in Chapter 2 that resisted purification. The goal was to optimize yield and reduce the opportunity for the

protein to aggregate. The original method lysed cells with a homogenizer, followed by a combination of anion exchange chromatography, size exclusion chromatography, and reverse phase high performance liquid chromatography with a C₁₈ column. Requisite buffer exchanges and concentration steps were carried out with an amicon stirred cell. The newer method (described in greater detail in Chapter 4 : Purifying ¹⁵N Labeled Hisactophilin, though here with ¹⁴N protein) represented a change in tack that, following cell lysis with a homogenizer, leveraged hisactophilin's high histidine content by using nickel affinity chromatography with low pH-induced elution that also unfolded hisactophilin. Buffers were exchanged by overnight dialysis and then protein was concentrated in the amicon stirred cell prior to reverse phase high performance liquid chromatography with a C₁₈ column.

Growing and harvesting unlabelled protein from rich media for the earlier purification technique started with the same vectors and transformation protocol described in Chapter 4: Expressing ¹⁵N Labeled Hisactophilin. However, the 10 mL overnight cultures were used as the inoculum (1:100) for 1L rich medium (Lysogeny Broth) flasks containing 100 µg/mL ampicillin and 30 µg/mL kanamycin. The flasks were incubated at 37 °C at 200 rpm until an OD₆₀₀ = 0.25, at which time 10 mL of 20 mM sodium myristate was added to each flask and growth continued until OD₆₀₀ = 0.7. Protein expression was induced with 1 mM IPTG and allowed to continue undisturbed for 6 hours. Cells were then harvested in the centrifuge at 5,000 x g for 12 minutes at 4 °C, and cell pellets were stored at -80 °C if they were not immediately processed for purification.

3.3.2 Purification

Cell lysis was executed as described in Chapter 4, except that the next step was anion exchange rather than immobilized metal affinity chromatography.

Anion exchange chromatography used the BioRad Macro-prep DEAE weak anion exchange resin packed into a glass econo-column and connected to a BioRad low-pressure chromatography system (BioRAD Bio-Logic LP, BioRad Laboratories Inc, Hercules, CA) with conductivity meter and UV absorbance at 280 nm. The binding buffer was 20 mM Tris at pH 8.0 with 1 mM EDTA, the elution buffer comprised a gradient of sodium chloride up to 400 mM with fractions collected by the Model 2110 Fraction Collector. Following the overnight run, hisactophilin-containing fractions were located 15% SDS-PAGE.

Fractions containing hisactophilin were pooled and buffers were exchanged for 150 mM potassium phosphate at pH 7.7 with 0.25 mM EDTA and 1 mM DTT using the amicon stirred cell (Millipore Sigma, Burlington, MA) with a 3 kDa molecular weight cutoff regenerated cellulose filter in preparation for size exclusion chromatography. The size exclusion column was the HiLoad Superdex 75 from GE Health Sciences (Pittsburgh, PA) connected to the BioRad Duoflow medium pressure chromatography system and BioFrac fraction collector. Hisactophilin, at this stage, is the majority species and elutes after ~100 minutes. The fractions containing hisactophilin were collected and had their buffer exchanged again in the amicon stirred cell for

water and were then concentrated to 1 - 2 mL in anticipation of reverse phase chromatography.

At this point, the hisactophilin fraction contains a mixture of myristoylated and non-myristoylated hisactophilin that are resolvable by reverse phase HPLC/FPLC with a C₁₈ column (Waters Inc, Milford, MA). Trifluoroacetic acid (0.1%) was added to the acetonitrile gradient as a weak ion pairing agent and to ensure protein was unfolded. The BioRad DuoFlow was again used as the chromatography system. Myristoylated and non-myristoylated-containing fractions were pooled separately, the pH was quickly adjusted to the physiological range with 25 mM ammonium carbonate, and then more thoroughly exchanged from acetonitrile against ammonium carbonate in the amicon stirred cell prior to storage in the -80 °C freezer or lyophilization.

3.3.3 Equilibrium Denaturation

Equilibrium denaturation curves were obtained as previously published [72]. In brief, lyophilized hisactophilin (myristoylated or non-myristoylated) was dissolved to a final concentration of 2 mg/mL in 500 mM potassium phosphate at pH 7.7 or 500 mM MES (2-morpholino-ethanesulfonic acid). The protein stocks were then diluted by a factor of 10 in water and denaturant to reach the final urea concentration target—usually close to twenty different urea concentrations that spanned from 0 M to ~9 M. These samples equilibrated at 25 °C for at least 10 half-lives of unfolding (as measured from folding kinetics experiments), in practice at least 3 hours. Unfolding was quantified by measuring tyrosine fluorescence at 306 nm with excitation at 277 nm [72] on a Fluorolog3-11 or Fluorolog3-22 spectrofluorometer (Horiba, Kyoto, Japan). The resulting curves were fit to either a linear extrapolation model,

$$Y = \frac{(Y_N + S_N[\text{urea}]) - ((Y_N + S_N[\text{urea}]) - (Y_U + S_U[\text{urea}]))(e^{\frac{\Delta G_{U-F} - m_{eq}[\text{urea}]}{RT}})}{1 + e^{\frac{\Delta G_{U-F} - m_{eq}[\text{urea}]}{RT}}}$$

Equation 3.1, or to the binomial extrapolation model,

$$Y = \frac{(Y_N + S_N[\text{urea}]) - ((Y_N + S_N[\text{urea}]) - (Y_U + S_U[\text{urea}]))(e^{\frac{\Delta G_{U-F} - m_{eq}[\text{urea}] + m_2[\text{urea}]^2}{RT}})}{1 + e^{\frac{\Delta G_{U-F} - m_{eq}[\text{urea}] + m_2[\text{urea}]^2}{RT}}}$$

, Equation 3.2. Here, ΔG_{U-F} is the free energy of unfolding in water, m_{eq} is the [urea]-dependence of ΔG_{U-F} , Y is the optical signal of the native (N) or unfolded (U) states, while S is the [urea]-dependence of each of those signals. In the binomial extrapolation method, m_2 is an additional term to describe deviation from linearity of urea concentration and folding energy.

3.3.4 Folding Kinetics

Kinetic folding and unfolding experiments were carried out by manual mixing or with the SFM4/Q (Bio-Logic SAS, Seyssinet-Pariset, France) stopped flow instrument connected by fiberoptics to the Fluorolog3-11. Manual mixing experiments were favoured when the folding or unfolding half lives exceeded ~15 seconds. They involved mixing folded protein (for unfolding experiments) or acid-unfolded protein (for refolding experiments) into a potassium phosphate buffer for pH 7.7 or MES for pH 6.2 with varying amounts of urea. Dead times were on the order of 8-12 seconds. Like with the equilibrium denaturation experiments, folding or unfolding events were monitored by tyrosine fluorescence (excited at 277 nm, emission at 306 nm). In the stopped flow experiments, rapid mixing was handled by the SFM4/Q system and Biokine 2.1 software with dead times on the order of milliseconds. Folding/unfolding reactions for each urea concentration were measured in quadruplicate and monitored for at least 10 half-lives of the reaction (or 10 half-lives of the slowest component, when possible); the first trace was often discarded as it flushed the system and the remaining three traces were averaged and then fit to a single- or multi-exponential process. (Un)folding rates were then fit to a 2-state or 3-state model as previously described [71], the former given by the equation,

$$\ln(k_{obs}) = \ln(k_f^0 \cdot e^{m_f^\ddagger[urea]-m_{2ref}[urea]^2} + k_u^0 \cdot e^{m_u^\ddagger[urea]-m_{2ref}[urea]^2})$$

Equation 3.3.

Variables k_f^0 and k_u^0 are the folding and unfolding rate constants extrapolated to 0 M urea (i.e. in water), m_u^\ddagger and m_{2unf} are the linear and quadratic dependence of the unfolding rates on urea concentration, and m_f^\ddagger and m_{2ref} are the same but for the urea dependence of the refolding rate.

Folding kinetics experiments above pH 7 are generally well-described by the 2-state model given the general agreement of the kinetics-derived thermodynamic parameters with those obtained by equilibrium denaturation. For example, when the natural logarithm of the observed folding/unfolding rates are plotted against [urea], the data often form a chevron (e.g. panel D of Figure 3.5) and the minimum of the chevron corresponded well with the midpoint of a denaturation curve [8,107]. However below pH 7 (i.e. below pK_{switch}), the fit to a 2-state model is somewhat more fraught and demanded a less straight-forward interpretation (vide infra).

The pH-dependence of refolding of wild-type was measured essentially as the above, with the principal difference being that the buffer was switched to potassium phosphate at pH (7.2, and 6.7). ANS-binding experiments (8-Anilino-naphthalene-1-sulfonic acid) were executed on the stopped flow instrument with a setup identical to the refolding experiments, except that 1 mM ANS was included in the syringe that contained 500 mM pH 6.2 MES buffer. There were attempts to include the ANS in the protein stock but the kinetics were 3-phase, noisy, and otherwise unnecessarily difficult to interpret.

3.3.5 Mass Spectrometry

Preparing samples for mass spectrometry varied depending on need and disposition/state of the protein. Generally, they were thoroughly exchanged against water in an amicon centrifugal filter with a 3 kDa molecular weight cutoff and concentrated to approximately 10 μ M, and then diluted fold into 0.4% formic acid and acetonitrile for electrospray ionization in the University of Waterloo mass spectrometry facility's Micromass Q-TOF Ultima global instrument. Occasionally, samples were sent to the mass-spec directly from the C₁₈ reverse phase column as they were already in a volatile buffer.

3.4 Results and Discussion

3.4.1 A faster purification protocol optimized for less stable mutants

The five hydrophobic truncation mutants that we studied as a result of the collaboration with Dr. Yaakov Levy in Chapter 2 [73], V36A, L73A, L76A, I116A, and I118A, were selected for further study because they were predicted to impact folding and/or switching. However, they were not well-behaved and there was great difficulty in purifying sufficient quantities for thermodynamics experiments using established protocols. Myristoylated V36A tended to co-purify with an alternatively acylated population of lauroylated hisactophilin that required further purification steps to resolve, L73A expressed exceedingly poorly, and L76A, I116A, I118A, and also V36A would succumb to an apparent instability that manifested as aggregation and final yields on the order of milligrams or less per litre rather than tens of milligrams per litre for wild-type. These problems suggested that the purification be overhauled. SDS PAGE of the whole cell lysate gave all indications that expression levels of the recombinant proteins (except L73A) were comparable to wild-type and the symmetry-derived (e.g. I85L, I93L, F6L, etc.) mutants. Thus, the purification protocol was revamped to limit the time elapsed between the beginning and end of purification in hopes of reducing the opportunity for hisactophilin to aggregate or otherwise degrade.

As it was, purification involved three chromatography steps: anion exchange, size exclusion, and a C₁₈ reverse phase column with buffer exchanges and sample concentration as necessary. Removing the requirement for one of these chromatography steps would represent an appreciable improvement in the process. Anion exchange chromatography only marginally increased the purity of hisactophilin despite the overnight run time. At \sim pH 8, above its pI, hisactophilin is estimated to have only a slight negative charge (\sim -2.9), not nearly so exceptional as to distinguish it from other proteins. Despite hisactophilin's glut of histidine, the inherited prevailing wisdom was that a nickel affinity column does not work. Instead, we thought to harness the positive charge on hisactophilin below \sim pH 7 using a cation exchange resin. In the context of the cell lysate, crossing hisactophilin's pI results in catastrophic aggregation. However, these aggregates are

nearly exclusively hisactophilin, so we attempted to resolubilize them to continue with purification. Lauren Morse, an undergraduate student, and I systematically tried combinations of heat, chaotropes (urea, guanidine hydrochloride, guanidine thiocyanate), and detergents (Triton X-100, deoxycholate, octyl- β -glucoside, CHAPS, SDS), and time. Unfortunately, none of these resolubilized enough hisactophilin from the insoluble aggregates for further pursuit to be worthwhile.

On the suggestion of my committee, nickel-charged immobilized metal affinity column was revisited because 31 of 117 residues in mature hisactophilin are histidine. Binding is reversible, and elution is possible by a pH-drop or an imidazole gradient. Qualitatively, we observed that hisactophilin is well-behaved when unfolded and quantitatively, it is amenable to refolding as it is a fast, two-state folder [8,71]. So, unfolding the protein early in the purification may mitigate problems with time-dependent aggregation. To that end, hisactophilin was eluted from the nickel affinity column with a pH \sim 3.5 buffer. The hisactophilin-containing fractions were very nearly free of other proteins. However, separating the myristoylated and non-myristoylated variants still required the reverse phase C₁₈ column chromatography. Where buffer exchanges were now executed by dialysis overnight (otherwise unused downtime), the time elapsed between cell lysis and pure acylated and non-acylated hisactophilin was reduced from six days to three. Yields per litre were also approximately doubled for wild-type hisactophilin. But, more importantly, the aforementioned well-expressed but poorly behaved mutants emerged from the purification with double-digit mg yields that, though still lower than other variants, were sufficient for equilibrium and kinetic folding experiments. Some of the other mutants from the symmetry set also resisted purification and they, like the hydrophobic truncation mutants, benefitted from the new methodology.

3.4.2 Switching and native state interactions can tune, break, and recover the myristoyl switch

The equilibrium experiments were a joint effort between Dr. Martin Smith, a former graduate student in the laboratory, and myself. Analyzing interaction energies by fluorescence-monitored equilibrium denaturation involved a set of four experiments for each mutant; myristoylated and non-myristoylated forms each at high (7.7) and low (6.2) pH [5,8]. We were each responsible for acylated and non-acylated forms of hisactophilin mutants at a single pH, high pH by Dr. Smith, and low pH by me. The resulting denaturation curves were fit to a 2-state unfolding model comprising only the folded (F) and unfolded (U) states to extract the free energy of unfolding, ΔG_{U-F} (left panel of Figure 3.2). Combined, the completed thermodynamic cycle allows measurement of interactions between protonation events and the myristoyl group and how they vary by mutation, or the measurement epistasis [108,109] (right panel of Figure 3.2).

Previous work with hisactophilin showed that myristoylation increases the global stability of his-

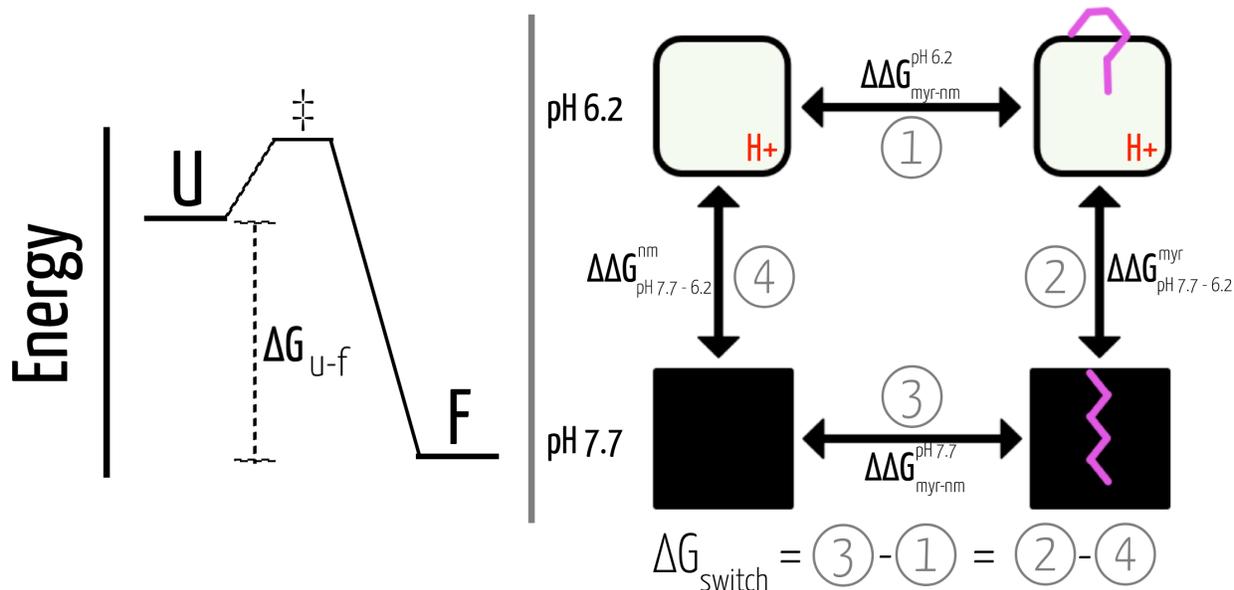


Figure 3.2: Simplified energy diagram and thermodynamic cycle used to measure switch energies. This figure illustrates some of the thermodynamic paradigms employed in the analyses to follow. Left panel is a simplified energy diagram of a 2-state transition between the unfolded state (U) and the folded state (F) with a single rate-limiting step. The energy difference between the unfolded and folded states, as determined by fitting equilibrium chemical denaturation data to Equation 3.1, is denoted as ΔG_{U-f} . The right panel is a schematic of the thermodynamic cycle as applied to analyzing switching in hisactophilin [5,8]. Each corner of the square represents the ΔG_{U-f} of hisactophilin at either pH 6.2, 7.7, and myristoylated (magenta) or non-myristoylated. The difference in the stability measurements, $\Delta\Delta G$, reports on the effect of myristoylation (edges 1 and 3) or pH changes (edges 2 and 4) on the protein, depending which edge of the cycle is considered. ΔG_{switch} is given by the difference between either set of opposing edges (i.e. edges 3-1, or 2-4). This same formalism is also applicable to studying other interactions in proteins, for example mutating residues [6].

actophilin [8], which, depending on pH, is ~ 3 kcal/mol. The pH dependence of this stabilization by myristoylation ($\Delta\Delta G_{U-F} = \Delta G_{\text{myr},U-F} - \Delta G_{\text{nm},U-F}$) is given by Equation 3.4). Horizontal edges in right panel in Figure 3.2) that drives the myristoyl switch. Of mutants studied herein, H90G is the only one that increases global stability for myristoylated and non-myristoylated forms at pH values on either side of the midpoint of switching (relative to wild-type. See panels A and B in Figure 3.4). Similar stability increases for each of the four forms of H90G supports that interaction between residue H90 and the myristoyl group is slight (but not non-existent). Increased stability is likely due to the structural predilection of the inter- β -strand loops to form a favourable tight turn if there is a glycine in the $i+3$ position, as there is in the analogous position in the other two trefoils of hisactophilin, and in position 93 of the structurally related FGF-1 (human acidic fibroblast growth factor-1) [110]. ΔG_{switch} (the difference between two edges in right panel of Figure 3.2) decreases, and thus stabilization of the $\beta 9$ - $\beta 10$ turn may affect the relative energetics of the switch and slightly disfavour the sequestered state.

Generally, the remaining point mutations are destabilizing for most or all pH values and acylation states. The most destabilized are F6L, V36A, L76A, I116A, and I118A. F6L is destabilizing at high and low pH in myristoylated and non-myristoylated forms, which is broadly consistent with a mutation to a smaller residue in the hydrophobic core. Changes there may be expected to manifest in all forms of the mutant because of the large number of altered interactions. The pH dependence of stabilization by myristoylation (panel D of Figure 3.4) is less than it is in wild-type, meaning that the stability change upon adding the myristoyl group is more similar at both pH values than in wild-type—here they are both large (panel C of Figure 3.4). The stability of the non-myristoylated form is more negatively impacted by the mutation than that of the myristoylated form, which suggests that the change in the hydrophobic core is somewhat compensated for by the myristoyl group. However, the balance of energetics between the myristoyl sequestered and accessible states is sufficiently perturbed that the former appears to be more favoured (larger bars in panel C), as losing stabilizing interactions with the myristoyl group in the core through switching would be especially unfavourable because of the mutation. The balance is not so perturbed that the thermodynamic switch is fully broken, however.

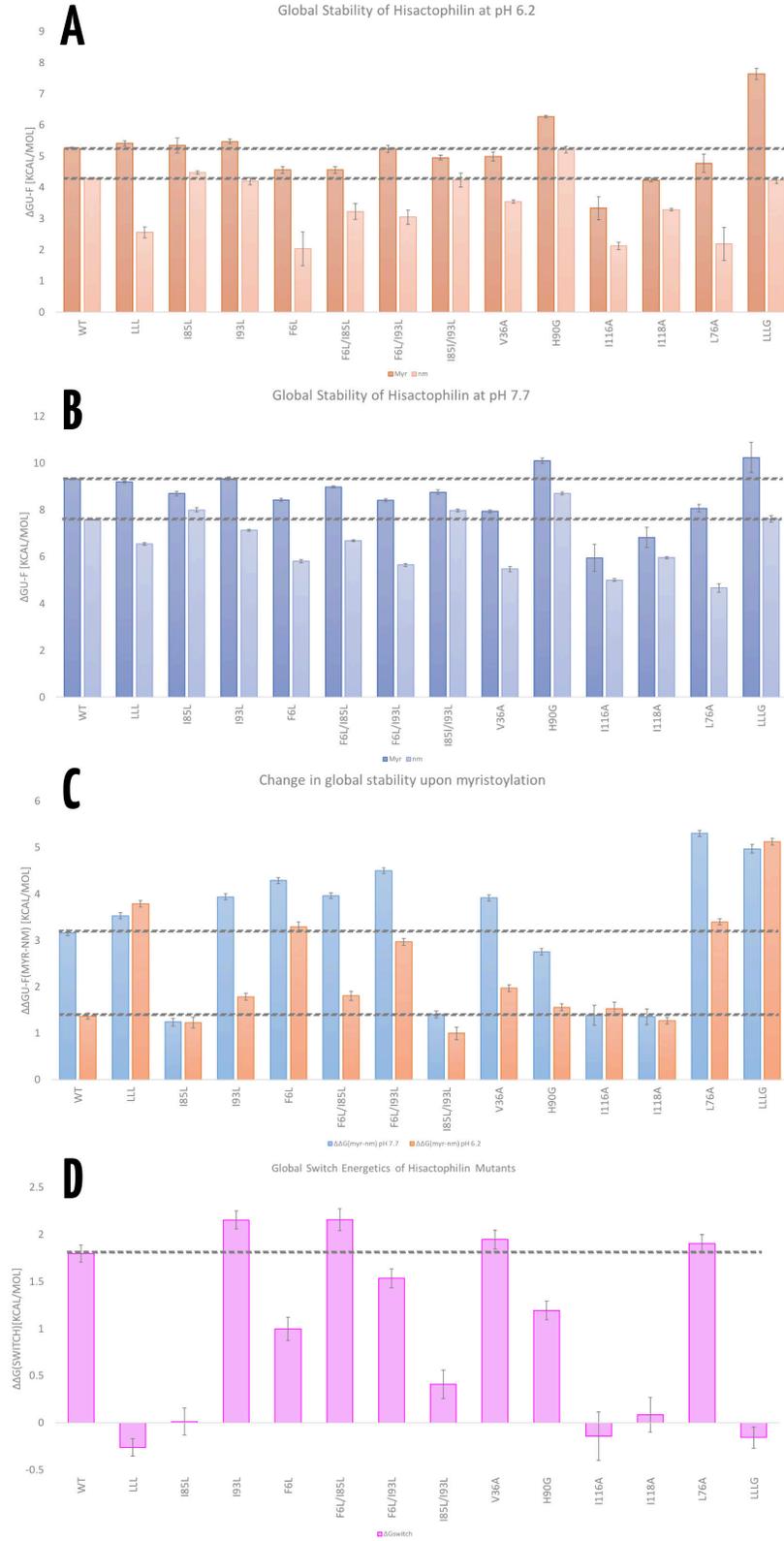


Figure 3.4: Bar charts representing the energetics of unfolding at pH 6.2 (A) and pH 7.7 (B), as well as the difference energies corresponding to the edges labelled 1 and 3 of the thermodynamic cycle shown in the right panel of Figure 3.2, $\Delta\Delta G_{\text{myr-nm}}$ (C), or ΔG_{switch} (D). On each panel, the dotted line represents the wild-type values for that particular plot. $\Delta G_{\text{u-f}}$ was determined by fitting equilibrium urea denaturation curves to the binomial extrapolation method described earlier. ΔG_{switch} represents the interaction energy between sites involved in protonation and the myristoyl group. Mutants with a value near zero have decoupled the two and are therefore interpreted to have a broken switch (e.g. I85L, LLL, LLLG, I116A, I118A).

V36A, another core-facing single mutation, follows a similar pattern as F6L; all four forms are destabilized relative to wild-type, but the non-myristoylated form at both pH values is more strongly affected by the mutation than the myristoylated one. This behaviour is consistent with the expected destabilization by a hydrophobic truncation mutation. Whereas in the myristoylated form, interactions with the acyl may compensate for losses in favourable contacts in the non-myristoylated form arising from the mutation. However, the pH dependence of stabilization by the myristoyl group (ΔG_{switch} , panel D) remains wild-type-like, so overall, the balance of the switch is maintained. L76A is positioned similarly to V36A in the structure, but in an adjacent trefoil component. L76A has almost the same ΔG_{switch} as V36A (1.90 kcal/mol and 1.94 kcal/mol, respectively), and so pH has similar effects on stability as changes conferred by myristoylation in wild-type. However, the additional stability gained by adding (and/or burying) the myristoyl group is much larger than in wild-type, highlighting a potential avenue for tuning the pH-dependent switch. Adding the myristoyl group appreciably stabilizes these mutants, suggesting that burying the myristoyl group is favoured because the increased stability for doing so (3-5 kcal/mol) would relatively disfavour the accessible state. Manipulating the global stability of the protein and the energy difference of sequestering the myristoyl group is a potential avenue to tune the populations of the sequestered and accessible states at a given pH. The effects of the mutations also point to the importance of the specific geometry in the core for the myristoyl switch. Truncating V36 and L76 to alanine destabilizes the non-myristoylated form, which can be taken to mean that interactions usually formed to stabilize the accessible state are altered by the mutations. On the other hand, an alternate interpretation is that relative to the non-myristoylated form, truncated residues relieve internal strain that the wild-type amino acids create by removing specific, unfavourable interactions from the sequestered state. In the context of more general switch design, manipulating the energy difference between two states could be the difference between a functional switch, one that leaks, or an opportunity to start again.

The C-terminal residues I116A and I118A behave differently. Relative to wild-type, both mutants are significantly destabilized at both pH values and acylation states. Surprisingly, the myristoylated forms of both mutants are the most destabilized and at pH 7.7 relative to wild-type—the expectation from wild-type is that this ought to be the most stable condition. Stabilization by myristoylation ($\Delta\Delta G_{\text{myr-nm,U-F}}$, panel C) is only 1.3 - 1.5 kcal/mol at pH 6.2 and 7.7. Wild-type, by comparison, is stabilized by 3.2 kcal/mol by the myristoyl group at pH 7.7 and 1.4 kcal/mol at pH 6.2. Furthermore, the extent of stabilization by myristoylation is not appreciably pH-dependent ($\Delta G_{\text{switch}} \sim 0$) which we interpret to mean coupling between proton binding and

the myristoyl group is ablated, i.e. the switch is broken [5,8]. The similarity between the mutants' (slight) gain in stability with the myristoyl group and the myristoyl-induced wild-type stabilization at pH 6.2 gives intriguing insight into the possible impact of these mutations on switching. If the hydrophobic c-terminus is involved in interactions with the rest of the protein and/or myristoyl group that favour the sequestered state, then a disruption of those interactions may be sufficient to tip the balance of switching toward the accessible state.

In the structural model of hisactophilin, I118 is proximal to the N-terminus and the point of attachment of the myristoyl group to the protein backbone (see Figure 3.1), so, I118 may help to latch the N-terminus in place when the myristoyl group is sequestered. Mutation of I to A may disrupt the N-terminal latch and the communication network between the myristoyl group and the site of protonation, tipping the balance of states away from sequestered, thus breaking the switch. I116 also appears to break the switch and effects a significant destabilization of the protein (which likely contributed to the difficulties in isolating this mutant). Unlike I118, interactions with the N-terminus are not as clear. However, disruptions at I116 may destabilize the final β -strand (β 12) of the protein, which forms part of the β -barrel with β -strand 1. Moreover, another potential cause of decreased local stability by these mutations is that isoleucine has higher propensity to form β -structure than alanine [111]. This could account for the overall destabilization of hisactophilin while indirectly disrupting the I118/C-terminal latch.

V36A, L76A, I116A, and I118A are the variants most destabilized by mutation, however interesting results are not limited to this set. I93L, found in the upper hairpin layer of hisactophilin's hydrophobic core (see the green residues in Figure 3.1), creates a 3-fold leucine element of symmetry. It is an interesting residue because it has NMR-derived NOEs to the terminal methyl of the myristoyl moiety, so it is possible that changes here could alter interactions with the myristoyl group. Although a seemingly minor mutation, there is precedent for small mutations to have large effects on β -trefoils, as in the I85L mutation (vide infra and [73]), or an analogous mutation in FGF-1 (V109L) that affected stability through a four-fold increase in the unfolding rate [112]. Equilibrium denaturation experiments reveal modest global stability changes (< 0.5 kcal/mol) relative to wild-type for the myristoylated form at pH 7.7 and 6.2, and non-myristoylated at pH 6.2. On the other hand, the non-myristoylated variant at pH 7.7 was destabilized by 0.7 kcal/mol. It was previously posited that this could be the result of forming a new or slightly larger cavity at the bottom of the myristoyl-binding pocket [105]. The myristoyl group could fill this cavity and otherwise compensate for some of the attenuated interactions at pH 7.7.

I85L is a mutant we have extensively studied because simply moving the branch point in the sidechain apparently disrupts the energetics of hisactophilin enough to break the myristoyl switch [73]. The fragility of the switch highlights the finely tuned interactions and energetics that drive switching. $\Delta\Delta G_{\text{myr-nm}}$ for I85L shows that the myristoyl group grants relatively small additional stability at both pH values, which suggests

fewer favourable interactions between the protein and the myristoyl group. Simulations also support the hypothesis that I85L favours the accessible state [73]. At pH 7.7, ΔG_{U-F} of the myristoylated form is less than wild-type and the non-myristoylated form is more stable than wild-type. The stabilization of the non-myristoylated form (at both pH 7.7 and 6.2) may offer structural insight as to why sequestering the myristoyl group appears unfavourable. The void in the centre of non-myristoylated hisactophilin in the absence of the myristoyl group likely contributes to its lower global stability relative to when that space is filled by the myristoyl group. However, the steric change in I85L may have two consequences: conflict between leucine and the myristoyl group when the myristoyl group samples the sequestered state and over-strains the system; leucine could also imitate the tip of the myristoyl group by forming favourable interactions with nearby residues in the bottom of the β -barrel and occlude the pocket. Stabilizing the non-myristoylated form suggests the myristoyl-accessible state is relatively favoured. In terms of coupling, the loss of the pH dependence of stabilization of the protein by the myristoyl group suggests that proton binding is insufficient to tilt the balance of populations in I85L.

The combination of three mutations, F6L/I85L/I93L (LLL), also has a broken switch [5,73]. Different from I85L, $\Delta\Delta G_{\text{myr-nm}}$ (panel C in Figure 3.4) indicates the myristoyl group confers significant stabilization to hisactophilin, independent of pH. Stabilization can be explained by a myristoyl group that prefers sequestration, even at pH values where wild-type tilts toward the accessible state. The myristoyl group's proclivity for sequestration simply arises from its hydrophobicity. With wild-type, this preference is countered at low pH by further overpacking the core; this effect, in turn, can be abolished by relieving steric conflicts through mutation. Without the strategic, unfavourable, and strained interactions with the myristoyl group, the energy gap between the accessible and sequestered states is too great to be overcome by proton-binding.

Relieving strain could be the result of an enlarged binding pocket, replacing slightly unfavourable interactions between the protein and the sequestered myristoyl group with more permissive ones where the binding pocket is effectively larger (similar in F6L). Creating a larger, solvent-accessible void in the protein is also unfavourable, so this may further favour the sequestered state where the void is occupied by the myristoyl group. If this is the case, the non-myristoylated form would likely be destabilized relative to wild-type. The ΔG_{U-F} of non-myristoylated LLL is destabilized relative to wild-type; at pH 7.7, ΔG_{U-F} are 6.6 kcal/mol and 7.6 kcal/mol for LLL and wild-type, respectively, and 2.6 kcal/mol compared to 4.3 kcal/mol at pH 6.2. The myristoylated forms, on the other hand, have stabilities very similar to wild-type (within 0.15 kcal/mol), so, interaction between the myristoyl group and the two proteins are consistent with sequestration. These findings match with simulations of LLL that show it more extensively samples the myristoyl sequestered state [73].

Many interactions in the core are changed between wild-type and LLL, and using double mutant cycles [6] offers a methodology for dissecting the interactions between two sites of mutation, an established method

for measuring the pairwise interaction between two positions in a protein and any synergistic interactions that may exist. In light of the thermodynamic cycles used to analyze switching interactions (right panel of Figure 3.2)[5,8], numerous cycles for dissecting switch energies in conjunction with differences between mutants are possible. Focusing on the switch energies (ΔG_{switch}) of the pairwise combinations of mutations that comprise LLL may illuminate the contributions of those residues to switching; e.g. considering I85L and I93L separately and then in a double mutant I85L/I93L; F6L and I85L, then F6L/I85L; F6L and I93L, then F6L/I93L. Above, we showed that I85L broke the myristoyl switch ($\Delta G_{\text{switch}} = 0.01$ kcal/mol) in favour of the accessible state and I93L tuned the switch ($\Delta G_{\text{switch}} = 2.15$ kcal/mol vs 1.80 kcal/mol for wild-type) by slightly favouring the sequestered state over the accessible (the non-myristoylated form was relatively destabilized). Adding I93L to the broken I85L recovers some of the switching energy ($\Delta G_{\text{switch}} = 0.41$ kcal/mol), which may indicate that the expected increase in space in the core for the myristoyl group resulting from I93L offsets some of the unfavourable interactions caused by I85L.

Interestingly, combining F6L with I85L's broken switch restores the switch energetics to near wild-type levels with a $\Delta G_{\text{switch}} = 2.16$ kcal/mol (similar to I93L alone). It is possible that increased unfavourable interactions between the buried myristoyl group and L85 are compensated for by a larger pocket near the myristoyl-terminus due to the mutation of phenylalanine to leucine. Notable also is that beyond the agreement with I93L's switch energies, in F6L/I85L's the stabilization upon myristoylation ($\Delta\Delta G_{\text{myr-nm,U-F}}$) also matches I93L, so this may lend further support to the importance of interactions with—and geometry of—the pocket around the terminal methyl of the myristoyl group to switching.

Neither of the individual mutations for F6L/I93L break the switch, however, they each tune it; F6L reduces ΔG_{switch} while I93L increases it. The combination of the mutations moves the switch energy midway between the individual mutations, albeit not additively. This finding generally supports that the different dimensions of the thermodynamic cube: pH, myristoylation, and mutation, all interact to some degree and impact switching. Furthermore, as observed in Chapter 4 and Chapter 5, switching may be sensed over the whole protein.

Interestingly and finally, adding H90G to the LLL mutation maintains the broken switch and stabilizes both the myristoylated and non-myristoylated forms at pH 7.7, likely by a similar mechanism to the stabilization of the wild-type background by H90G. LLLG is more stable than LLL by ~ 2 kcal/mol for all variants/-conditions with the myristoylated forms slightly more stabilized than non-myristoylated, which supports that stabilizing the $\beta 9$ - $\beta 10$ turn slightly favours the sequestered state. Since previous work [8] singled out this turn as a potential region for the binding of switch proton(s), it is not surprising that disturbing the energetics here affects the coupling between proton binding and the myristoyl group. However, the single mutation of H90G would be expected to break the switch if it were predominantly responsible for proton binding. H90 may not be the actual proton-binding site but adjacent to it (there are four histidines nearby: H88, H89, H90, H91), or

as other studies with haemoglobin and envelope protein of West Nile virus have shown, multiple ionizable residues drive or modulate switching [113,114] so a lone gunman may not be the only responsible party. The NMR-monitored pH titrations in Chapter 4 indicate the $pK_{a,app}$ s of many residues respond to switching and thus, there may be fractional proton binding at numerous sites (including H90).

The above results reveal that hisactophilin is a finely tuned switch where the relative stabilities and interactions of the accessible and sequestered states governs the sensitivity of the switch to pH changes. A potential mechanism of keeping hisactophilin on a razor's edge so that small pH changes drive the switch is by straining the core of hisactophilin through strategic unfavourable interactions with the myristoyl group. The results presented here suggest that mutations F6L, V36A, L76A, I93L, F6L/I93L, and F6L/I85L relieve these strained interactions and increase the relative stability of the myristoylated form. pH-dependent switching is maintained, however, when the amount of strain varies with pH such that proton-binding effects changes in hisactophilin that further destabilize the sequestered state and drive toward the accessible state. But, if strain is relieved in both states and decoupled from proton binding, then situations like LLL and LLLG arise where the myristoyl group appears to favour the sequestered state, independent of pH. Mutations may also ratchet up the strain, as in I85L, I85L/I93L, and, to a lesser extent, H90G. Here, myristoylation confers little additional stability to hisactophilin, and one explanation could be that the accessible state is preferred because sequestering the acyl chain could over-strain the protein and ablate any stability gains from burying it. The switch energetics of I116A and I118A present similarly to the over-strained mutants, and while these mutations may also mediate strain, their positions at the C-terminus and toward the top of the β -barrel instead suggest they play different structural roles in stabilizing/interacting with the different states, whether ensuring β -structure formation or latching the N-terminus in place to secure the sequestered state. What these different mutations demonstrate, however, is that many interactions are responsible for balancing the energetics of the sequestered and accessible states and that many roads may lead to switching.

3.4.3 The myristoyl group robustly accelerates folding kinetics in many hisactophilin mutants

While equilibrium denaturation experiments measure the global stability of the native state, we can further quantify how the myristoyl group and pH impact the energy landscape of hisactophilin, in particular the transition states of folding (and unfolding) with folding kinetics experiments. The folding of hisactophilin has already been subjected to considerable scrutiny in the past [8,71,72], and most recently in Chapter 2 [73]. Myristoylation affects the folding of hisactophilin by significantly accelerating the folding rate and, interestingly, also unfolding. The effect on unfolding informed the hypothesis of a strained native state as a driver of switching. Furthermore, the effect of myristoylation on the folding of LLL, I85L, V36A, L76A,

and I118A at pH 7.7 evinces that its effects are fairly robust and that, unlike the native state, the mutations do not appreciably hamper the myristoyl group from making favourable interactions in the transition state.

Here, folding experiments were expanded to include mutants F6L, I93L, H90G, and I116A. Conducting a similar analysis to the published work [73] by comparing the concentration of urea where the folding rate is 1 s^{-1} (to reduce errors from extrapolating to 0 M urea), the previously observed trend that the myristoyl group robustly stabilizes the folding transition state and accelerates folding is maintained (see panel A in Figure 3.5). Myristoylation in wild-type stabilizes the transition state ($\Delta\Delta G_{\ddagger-U, \text{myr-nm}}$) by 2.0 kcal/mol. In comparison, F6L was also stabilized by 2.0 kcal/mol which suggests that interaction between position 6 and the myristoyl group are similar for mutant and wild-type. However, the stabilities of the myristoylated and non-myristoylated folding transition states for F6L are decreased which may reflect the loss of significant hydrophobic interactions in the core [115]. The transition state of I93L is stabilized by myristoylation by 2.3 kcal/mol, H90G by 2.7 kcal/mol, and I116A by 2.8 kcal/mol, which are all suggestive of maintained and slightly more favourable interactions between the sites of mutation and the myristoyl group in the folding transition state.

Native state equilibrium stability measurements indicated that I93L and H90G reduced strain in the native state, so, if some native-state interactions are formed in the transition state of folding, then a similar stabilization of the transition state might be expected. In FGF-1, H93G is analogous to H90G and the mutation speeds folding, which is attributed to an increased propensity to form a tight turn [110,112]. However, these mutations and those previously measured at pH 7.7 have only a small effect on the transition state, so the myristoyl group appears to make many and varied favourable interactions that lower the energy of the transition state and the change of one (or three) residues does not hamper the myristoyl group's positive impact on folding.

3.4.4 Folding kinetics at low pH may suggest parallel folding pathways

Inasmuch as we were curious to measure the interactions between the myristoyl group and the sites of protonation in the native state, the same sites of potential interactions may be examined in the transition state, i.e. how do protonation and electrostatics modulate the effects of myristoylation on the folding of hisactophilin and is there any evidence of switch-like behaviour in the transition state? Methods for measuring folding kinetics at high pH were already well established in the laboratory, so expanding to include pH 6.2 was expected to be a straightforward evolution of the project. It was not.

With the goal of complementing the folding kinetics measurements at pH 7.7, refolding (and some unfolding) kinetics were enthusiastically measured at pH 6.2 for wild-type, I85L, I93L, H90G, LLL, LLLG, L76A, I116A, and I118A. Those data for the myristoylated form, however, diverge from the character of the pH 7.7 sets in a way that complicates interpretation. Kinetic traces from fluorescence-monitored pH-jump

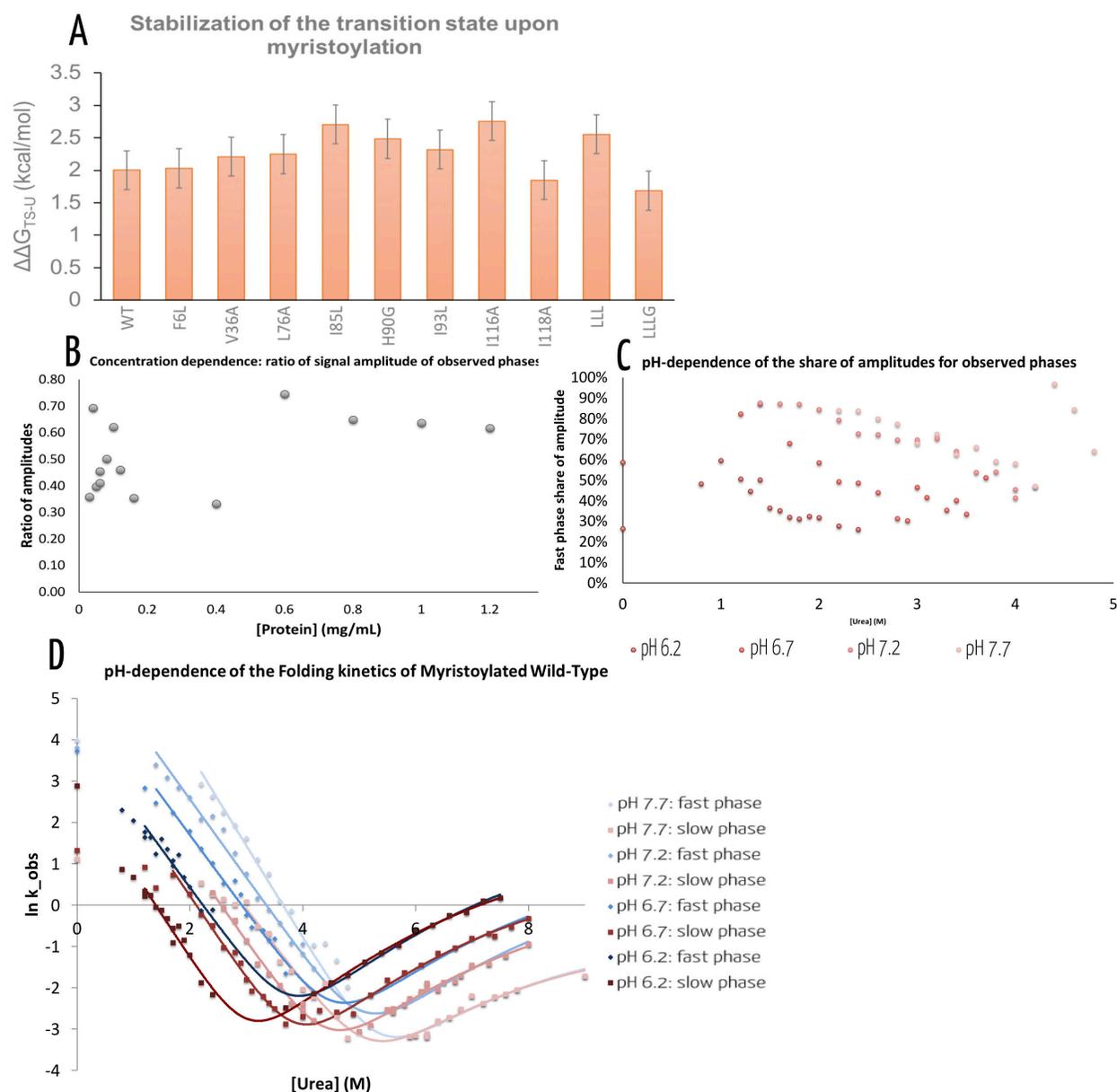


Figure 3.5: Kinetics of folding and unfolding of myristoylated and non myristoylated hisactophilin mutants. (A) Bar chart showing the robust enhancement of folding rate upon myristoylation at pH 7.7. These values were calculated by determining the [urea] where $k_{obs}=1$ for the major(fast) phase, and then taking the difference of myr-nm and multiplying by an average m_f -value of 1.38 to determine the $\Delta G_{u-\ddagger}$ [73]. (B) The ratio of amplitudes for the two observed refolding phases at 1.4 M urea with a range of protein concentration from 0.01 mg/mL - 1.2 mg/mL to test for protein concentration-dependent aggregation of myristoylated wild-type that might account for the additional kinetic phase at pH 6.2. There is no conclusive trend. (C) The pH dependence of the share of the total amplitude from the fast phase from refolding kinetics of myristoylated wild-type. Above pH 6.7, the faster phase is responsible for the majority of the signal amplitude. At pH 6.7 and below, a slower phase is increasingly responsible for the signal amplitude. (D) pH dependence of the refolding rates of the two phases (fast and slow) for myristoylated wild-type. Darker colours represent lower pH values, the blues are the fast phase and reds are the slow phase.

refolding experiments as a function of denaturant concentration were better described by the sum of two exponential functions at pH 6.2, indicating that the system may no longer be best described by a 2-state model of folding. A 2-state model implies a rapidly equilibrating unfolded state that refolds without reaching a metastable state along the way [116,117]; however, multi-exponential kinetics may arise for reasons relating to a more complicated folding mechanism that involve one or more intermediates (which may be on- or off-pathway) or parallel folding pathways, for example. Depending on the folding mechanism, the observed phases could correspond to any of the steps (e.g. U-F, and the second phase is folding from an intermediate on a parallel pathway). Alternatively, multi-exponential kinetics may result from a contaminating protein (like a population of non-myristoylated protein), proline cis-trans bond isomerization (hisactophilin is proline-free, though), reversible transient protein aggregation that occurs during refolding, or aggregation in the protein stock, to name a few causes.

Eliminating erroneous sources of the second phase of the folding kinetics began with checking the myristoylated protein preparations destined for kinetics experiments for contaminants (e.g. non-myristoylated protein) by mass spectrometry. However, after verifying that both myristoylated and non-myristoylated forms have similar ionization behaviour in the mass spectrometer, it was concluded that there were no contaminants that accounted for the amplitude of signal in the additional observed refolding phase. Protein aggregation was tested for by dissolving hisactophilin in the experimental buffers, measuring the refolding kinetics arm of the chevron on half of the protein and reserving the remainder for experiments twenty-four hours later. Aggregation may manifest as a time-dependent change of the relative signal amplitudes of the components of the observed kinetics. However, no time-dependent amplitude changes were observed, so, such aggregation is an unlikely culprit for the additional folding phase. Transient aggregation during kinetic experiments can also be protein concentration dependent but the ratio of amplitudes was inconclusive (see panel B in Figure 3.5); though, there was some indication that the faster phase has a protein concentration dependence that plateaus, but this is inconsistent with *m*-values and apparent midpoints of denaturation. Thus, combined with a multitude of instrument tests for mixing artefacts and temperature control, evidence was mounting that the emergence of second folding phase may report on a real folding process.

At pH 7.7, myristoylated hisactophilin's folding kinetics are dominated by a single phase whose change in signal (tyrosine fluorescence) over time is consistent with a two-state transition from the unfolded to folded state (left panel of Figure 3.2) [8]. Under those conditions, myristoylation accelerates folding by orders of magnitude relative to non-myristoylated hisactophilin. Changes were observed for the folding kinetics measured at 0.5 pH unit increments from pH 7.7 down to 6.2 for myristoylated wild-type (see panel D in Figure 3.5). The results show that a second, slower refolding phase increasingly contributes to the observed kinetics with decreasing pH (see panel C of Figure 3.5). At pH values below the midpoint of switching (~pH 6.9), the equilibrium stability measurements correspond more closely to the slower of the two apparent phases

(in terms of m -values and C_{mid}), which may suggest that the slower phase reflects the folding transition of the unfolded to the native state. If true, then the myristoyl group accelerates folding by only $\sim 5x$ at low pH, much less than at high pH ($\sim 55x$)[8]. The increase in native-state stability conferred by myristoylation is much less at low pH, so, that may account for some of the attenuation of the acyl group's effect on the rate of folding if the myristoyl group also forms fewer favourable interactions in the transition state. The explanation might be that increased charge in the unfolded state at low pH hampers the non-native interactions between the myristoyl group and the protein that help folding [73]. The impact of electrostatic interactions in the unfolded and transition states can be significant [85,118–121].

If interactions with the myristoyl group in the transition state change the pK_a of ionizable residues, the shifts in ionization may be responsible for myristoylation-induced changes to folding at the lower pH. So, the extra, faster folding phase could arise from a transition state where the myristoyl group still forms helpful non-native interactions, though why this faster pathway to the native state is not dominant is not easily explained and suggests that this simplistic model is not sufficient. Alternatively, interleukin-33, another β -trefoil, shares some kinetic folding features (e.g. appearance of a new phase on the same order of rates as the major phase) with hisactophilin and they proposed a folding mechanism involving parallel pathways [122,123]. Conclusive analysis of the folding pathway of myristoylated hisactophilin at low pH requires more complex folding experiments, such as double jump interrupted refolding and unfolding experiments [124–126] that can identify which phases are associated with forming native molecules, for example. The complexity of these experiments, however, were beyond the scope of my scientific curiosities and intended project. And so, low pH folding kinetics were bravely abandoned until another student dons the mantle and analyzes the trove of folding data that were acquired for the hisactophilin mutants.

3.5 Conclusions

The systematic measurements of the impact of mutations on global stability, protein-myristoyl interactions, and switching described here served to deepen our understanding of hisactophilin's mechanism and determinants of switching. Already established was that the I85L and F6L/I85L/I93L mutations could break the thermodynamic switch resulting in limiting cases where the myristoyl group appears to favour the accessible state or the sequestered state, regardless of pH, respectively. These data were complemented by studying a broader set of mutants designed to increase the three-fold symmetry of the core [105], and others suggested by coarse grained simulations to capture non-native interactions in the folding transition state [73]. The effects of many of the mutations on switching and on native state energetics can be rationalized in terms of modulating the balance of favourable and unfavourable interactions with the myristoyl group in the accessible or sequestered state. Single mutations may tune the switch by making small adjustments to these interactions,

and interestingly, combinations of core mutations may have non-additive effects that culminate in rescuing a broken switch (F6L/I85L), or failing to rescue the switch (I85L/I93L). These findings have broader implications for switch design, showing that the energy difference between functional conformations may be tuned by making small steric changes that strain or relieve strain in the different states of the protein.

That the effects of mutations depend on the context (e.g. sequence, structure) in which they occur is a growing area of study in proteins. In fields focused on protein evolution, it is called epistasis [127,128], a theorized mechanism by which new evolutionary pathways are made available to a protein. For example, a permissive mutation that increases protein stability may allow for a subsequent functional mutation that previously would have destabilized the protein [127,129,130]. For example, suppose the evolutionary ancestor of modern hisactophilin is the more symmetrical LLL. An L85I mutation (F6L/I93L mutation in Figure 3.4) creates the myristoyl switch, but destabilizes the protein which could negatively impact survivability. The L6F mutation (I85L/I93L in wild-type), however, slightly improves protein stability but with only a marginal switch. With epistasis, both mutations together, L6F/L85I (I93L in wild-type) increase protein stability and function synergistically, having a greater effect where the whole is greater than the sum of the mutations.

In wild-type hisactophilin, single mutations made in different contexts have very different effects. For example, I85L breaks the switch to favour the accessible state, whereas when I85L is made in the F6L/I93L background, the switch is again broken, but instead to favour the sequestered state. This might be an epistatic effect that also highlights the importance of specific, finely tuned interactions in the core and with the myristoyl group in balancing the switch energetics. Furthermore, there is perhaps something to be said for the necessary divergence from symmetry (e.g. LLL) to allow for function [101]. Beyond just breaking the switch, mutations tuned it and further illuminated how some of the favourable and unfavourable interactions formed in the core exert their influence on hisactophilin by creating and/or relieving native-state strain. Startlingly few and/or small changes to the protein can disrupt the coupling between proton binding and the myristoyl group. This effect was also seen in recent work with the ancestral reconstruction of the mollusk ortholog of the steroid hormone receptor, where two mutations were sufficient to abolish allostery. The two mutations were either H524F/I536F or A415W/I536F, which affected hydrophobic packing by filling or creating cavities and tuning interactions that selectively stabilize the active vs inactive states [131].

In additional hisactophilin mutants, the myristoyl group continued to robustly smooth the folding funnel and generally speed folding at high pH, contrary to the more drastic effects these mutations had on the native state. Protein modifications that form non-native interactions that help folding are still uncommon, so some of these findings may be used to improve the folding landscape and/or increase the global stability of slow-folding, designed proteins. The folding kinetics at low pH, however, increase in complexity, possibly owing to the additional frustration from having to reconcile more electrostatic repulsion upon folding [119]. In any case, hisactophilin—a highly dynamic and responsive myristoyl switch protein—continues to make an

excellent subject for studying fundamental principles in biochemistry and biophysics, from allostery to ligand- and membrane-binding and beyond.

Chapter 4: Local Conformational Stability and Strain in Myristoyl Switching

4.1 Chapter Abstract

Conformational switching is a central feature of protein function, such as ligand binding-induced shifts between an active and inactive state. The molecular mechanisms underlying switching remain a challenge to understand at high resolution. Describing switching mechanisms involves the confluence of thermodynamic stability, allosteric communication, and protein dynamics in and between different functionally relevant conformations and states. Here, we apply measurements of the temperature dependence of amide proton chemical shifts as a novel means to define how changes in local stability at individual residues contribute to switching, which involves taking measurements of myristoylated and non-myristoylated hisactophilin at pH values on either side of the switch and applying a thermodynamic cycle-like analysis. We show that the method can go beyond the established application of detecting hydrogen bonds and identify the allosteric communication network between ligand binding and the myristoyl group, in addition to locating sources of conformational tension and strain that drive switching. When the myristoyl group is sequestered in the hydrophobic core of the protein, resulting in increased global stability, numerous residues are unexpectedly destabilized which indicates conformational strain. In two mutants with a broken switch, evidence of conformational strain/tension is greatly diminished, which further supports the importance of specific unfavourable interactions to balance the thermodynamics of switching. Thus, in addition to identifying strained residues important to allostery and switching, we have demonstrated the power of applying the temperature dependence of amide proton chemical shifts to variants of a protein to detect the nuanced and difficult-to-measure changes that may accompany mutation, ligand binding, post-translational modification, or more general conformational changes.

4.2 Introduction

4.2.1 Protein Dynamics

The prevalence and dominance of x-ray crystallography for determining the structure of a protein has perhaps contributed to the misapprehension that the folded, native state of a protein is static and that its function springs from that single, most stable conformation. The modern view of protein folding where nascent polypeptides explore a number of different conformations and pathways on an energy landscape, shedding entropy until they converge at an energy minimum may only reinforce this view [53]. The native

state, however, is not likely to be a smooth basin on the landscape, but comprises numerous microstates close in energy to the minimum and accessible from the thermally induced fluctuations and wanderings [96,132–136]. A more rigorous understanding of the mechanisms and determinants of some protein function requires invoking conformational fluctuations to proximate conformations on the energy landscape, this applies to catalysis [137–139], molecular recognition [140], and allostery [141–143]. These microstates and excited states may arise from changes in side chain rotamers, the movement of elements of secondary structure, or changes in loop conformation [144]. However, almost by definition, the states are only sparingly populated, which makes them inherently difficult to measure and detect because if a probe for such a state exists, its signal is likely to be overpowered by the more dominant native state. Recent advances in solution-state nuclear magnetic resonance (NMR) proffer some new and exciting tools for interrogating such functionally relevant and fleeting states of the protein and tracing the sorts of excursions and motions a protein can undergo in the line of duty [145].

The motions and dynamics of different parts of a protein can range in timescale from small bond vibrations and rotamer interconversions on the sub-picosecond timescale to catalysis and protein folding over microseconds to greater than seconds. The interconnectedness and hierarchy of dynamics, whether fast motions facilitate large-scale, slower motions, is still up for debate [146,147]. There isn't a single experiment that measures dynamics across all time regimes, so a biochemist with an interest in protein motions requires a diverse toolkit: amide exchange, Carr-Purcell-Meiboom-Gill (CPMG), chemical shifts, and order parameters measure motions on the slow, intermediate, fast, and very fast timescales, respectively. The finely tuned relationships between stability, structure, and flexibility are integral to protein function, and sometimes function and flexibility come at the cost of stability. The protein SlyD has two domains, the less stable of which functions as a chaperone, but is also highly flexible to facilitate binding to a wide array of targets. Interestingly, despite the difference in stability of the two domains, unfolding is still apparently cooperative [148,149]. A loss of local structural stability and increased flexibility has also been implicated in misbehaved and aggregation-prone proteins, like the disease-associated superoxide dismutase 1. An edge β -strand with a higher than average propensity to explore conformations with a loss of secondary structure on a fast timescale, or lowly populated conformers that form non-native dimers stabilized by deviant intermolecular interactions are suggestive of possible pathways to pathogenesis [150,151].

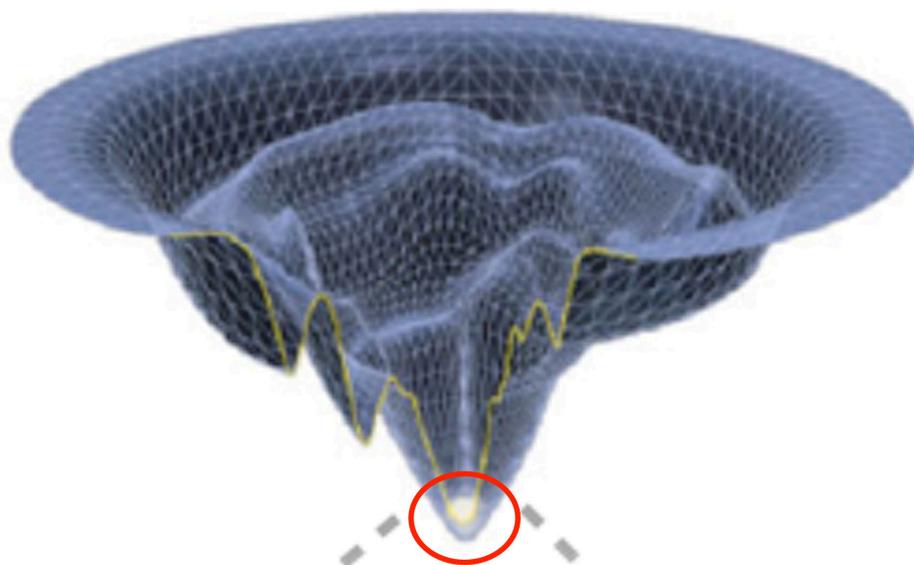
Fortunately, many timescales are accessible by NMR (with varying degrees of difficulty) and since myristoyl switching in hisactophilin is expected to be in the microsecond regime (estimated from NMR peak line-shape analysis in [8]), chemical shifts and their perturbations seem promising to obtain structural and dynamic information on the motions that may drive hisactophilin's pH-dependent switch.

4.2.2 Strain

Canonically, strain is a phenomenon that refers to the energetic penalty imposed on the native state in exchange for favourable energetics for function that manifests in the kinetics of folding, be the function catalysis or conformational rearrangements [152,153]. Residues destabilizing the native state by introducing strain countermands the notion that proteins are strictly under evolutionary pressure for stability when, instead, there is a nuanced negotiation between the conformational and energetic requirements for function, stability and solubility of the protein [153–156]. The buildup of innate strain in the switching protein adenylylate kinase is sufficient to unfold the entire protein during some simulations, so it is believed that some strain unfolds localized areas of the protein to allow the conformational change from open to closed states integral to efficient catalysis. This strain-powered unfolding has been called cracking [157]. Mutating the catalytic serine in β -lactamases to glycine increased the melting temperature by as much as 13 °C in the KPC-2 variant; however, the tradeoff for stability was a 500 fold decrease in the catalytic rate [158]. Strain is present in proteins beyond enzymes; the allosteric proteins CDC42 and L-lactate dehydrogenase also show a similar buildup of energetic frustration in functionally relevant areas such as hinges. This localized and selective destabilization enables some large-scale conformational changes by increasing localized flexibility and unfolding/cracking [62]. So, although strain may appear thermodynamically disadvantageous, it appears necessary for the function, catalytic or otherwise, of numerous proteins.

In fast-folding hisactophilin, the myristoyl group robustly smoothes the protein's folding landscape (at high pH) while simultaneously and paradoxically inducing native state strain that might have functional implications [8,73]. In this chapter, I have slightly subverted the usage of strain—as discussed above, strain typically involves global or sub-global unfolding events (e.g. loss of secondary or tertiary structure, generally represented by the funnel in the upper panel of Figure 4.1) that tend to manifest in folding kinetics. Strain in the native state is important in functional myristoyl switching and this chapter aims to identify and dissect the sources of strain in the native and near-native states (lower panel of Figure 4.1). However, it may be a misnomer to refer to changes in strain while characterizing only near-native states. For example, reporting that strain is relieved could erroneously imply that a near-global unfolding event is involved which is beyond the scope of this chapter. Instead, here, the result of strain in the native state will be referred to as conformational tension. The relaxation of tension may be associated with a conformational or dynamic change consistent with access to a state nearby in energy, but still ostensibly in the native basin (lower panel of Figure 4.1).

Global Folding



Native Basin

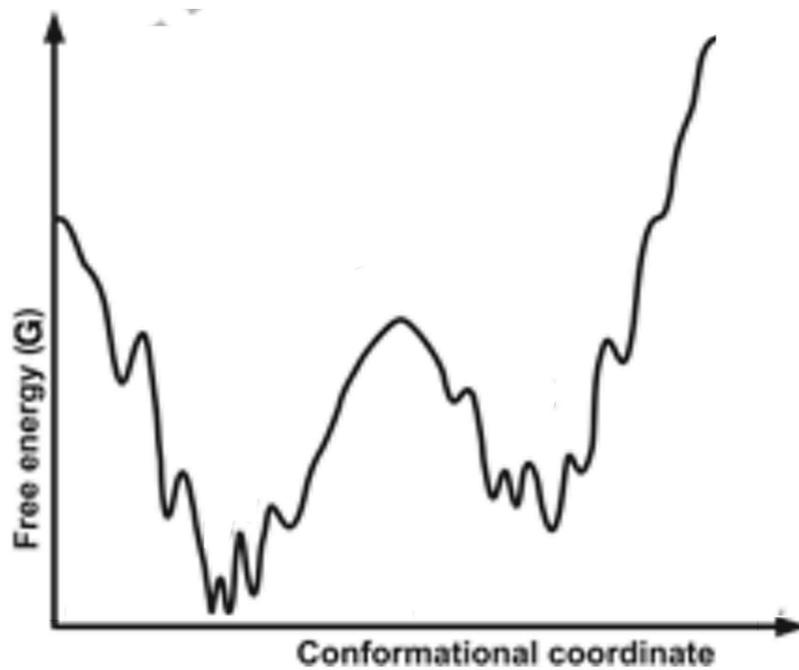


Figure 4.1 Adapted from [159,160]. The energy landscape of a protein. (Upper Panel) a folding energy landscape where the top, wide mouth of the funnel represents the unfolded state where entropy is maximized. As the funnel narrows, entropy is reduced as folding proceeds along any number of pathways that might include local minima that might favour forming productive folding intermediates or traps that hinder folding, until the global minimum and native state is reached. (Lower Panel) The native state basin may not be perfectly smooth and instead comprise a number of microstates nearby in energy, that might be critical to the functioning of the protein. This microstates account for some of the dynamic personalities of proteins across many timescales [161].

4.2.3 NMR

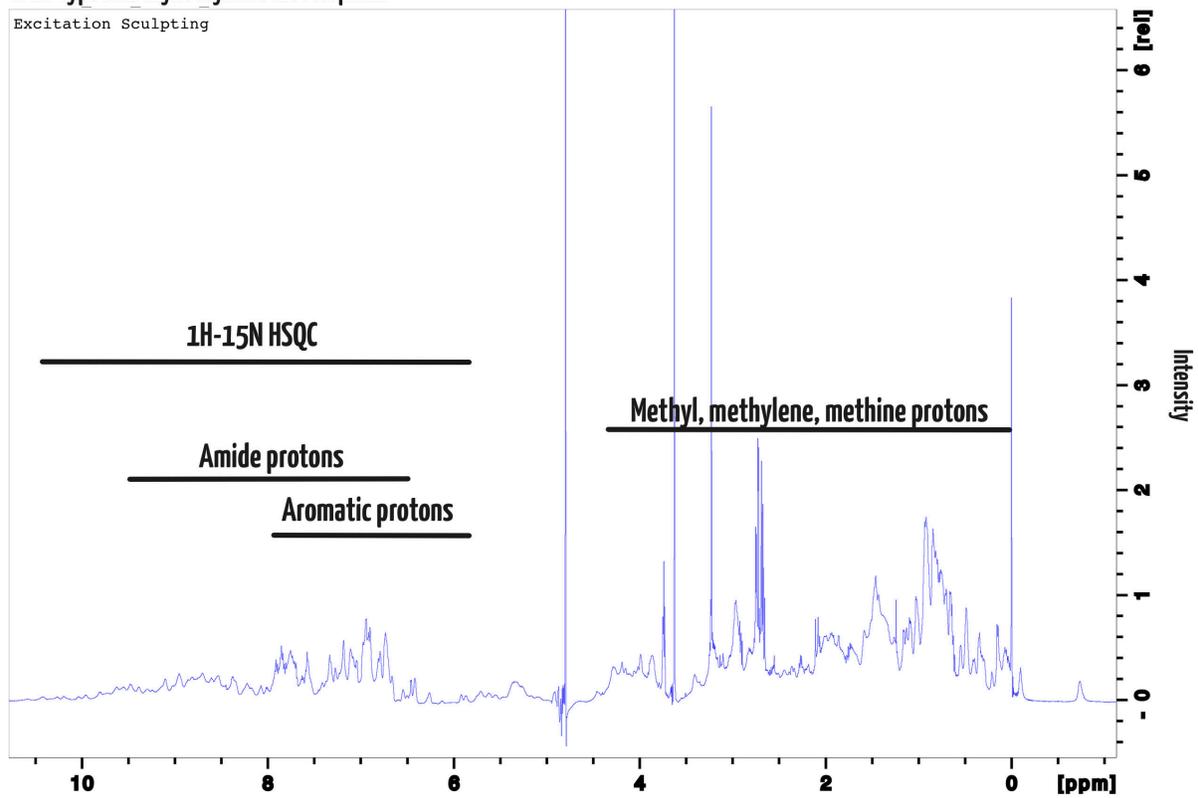
Nuclear Magnetic Resonance (NMR) has emerged as a powerful means to extract structural information from biomolecules with atomic resolution as an alternative to x-ray crystallography [162]. Beyond that, however, NMR is an invaluable tool for measuring the motions and dynamics of proteins, including those that are slow, which may encompass folding events and catalysis [163–165]; to the intermediate timescales comprising domain rearrangements, molecular tumbling, and catalysis [[166]; [167]; [168]; [169]]; to fast loop motions and side-chain rearrangements [170–173]. Nuclear magnetic resonance uses a strong external magnetic field, modernly produced by a liquid helium- and nitrogen-cooled superconducting magnet, combined with radio frequency pulse generators, a bevy of electronic controllers and detectors, a concentrated protein sample, and an enthusiastic (under)graduate student or victim.

When a nucleus experiences the strong magnetic field, its resonance frequency (ω , in rad s^{-1}) is proportionate to the strength of the field (B_0) and the gyromagnetic ratio (γ), as given in equation 1, $\omega = \gamma B_0$, where the gyromagnetic ratio is an innate property of an atom. However, not all nuclei interact with a strong magnetic field and therefore may not be appropriate for NMR—only those with a spin of $1/2$ are suitable for the NMR experiments that are going to be described and used throughout this work. An atom with spin $1/2$ is one with an odd mass number, regardless of atomic number [174]. Commonly studied nuclei in proteins are ^1H , ^{13}C , ^{15}N , ^{19}F , and ^{31}P , and proton, carbon, and nitrogen are the common nuclei of interest in proteins. Unfortunately, the most abundant isotopes of carbon (^{12}C) and nitrogen (^{14}N) are not spin $1/2$, so proteins have to be made in hosts fed a strict diet of ^{13}C glucose and ^{15}N NH_4Cl to ensure uniform labelling so they satisfy the spin requirement.

The resonance frequency of an atom in the NMR is also affected by its local magnetic environment, which may include the identity of nearby nuclei, covalent bonds, non-covalent interactions with adjacent atoms (e.g. nearby aromatic groups), and their relative orientation in space of the atoms, for example (*vide infra*) [162]. As a result, two protons may have different resonance frequencies because they experience different chemical environments. Modern NMR spectrometers report the resonance frequencies of atoms as chemical shifts, which is the observed resonance frequency relative to a standard reference compound. It is necessary to standardize resonance frequencies to a standard field so that chemical shifts are comparable from one

spectrometer to another [175]. Example reference compounds are tetramethylsilane, trimethylsilylpropanoic acid, or 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). A one-dimensional (1D) NMR spectrum reports the chemical shifts of a single type of nucleus (^1H , for example) by measuring the decay of observable magnetization following an excitation radio-frequency pulse. The resulting intensity versus time plot (called a free induction decay, or FID) is processed by a Fourier Transformation of the time domain data to extract the resonance frequencies of the atoms that contribute to the FID. Each resonance (converted to chemical shift in ppm as discussed above) is plotted as peaks on a horizontal axis vs intensity. In small molecules, the peaks are narrow, well-defined, and usually sufficiently well-resolved from other nuclei with similar chemical shifts as to extract significant structural information relating to the covalent structure of the molecule. Larger molecules, like polypeptides and proteins, are orders of magnitude larger and more complex. A 1D spectrum of a protein is composed of many overlapped and relatively broad peaks because, rather than tens of protons, there are hundreds or thousands. A limiting challenge of NMR is peak broadening that arises from a protein's large size, which increases how quickly magnetization decays. The large number of peaks in a protein 1D spectrum leads to a crowded spectrum that complicates extracting precise and meaningful structural information (as seen in the upper panel of Figure 4.2).

Wild-type Non-myristoylated Hisactophilin



2D 1H-15N HSQC Wild-type Non-myristoylated Hisactophilin

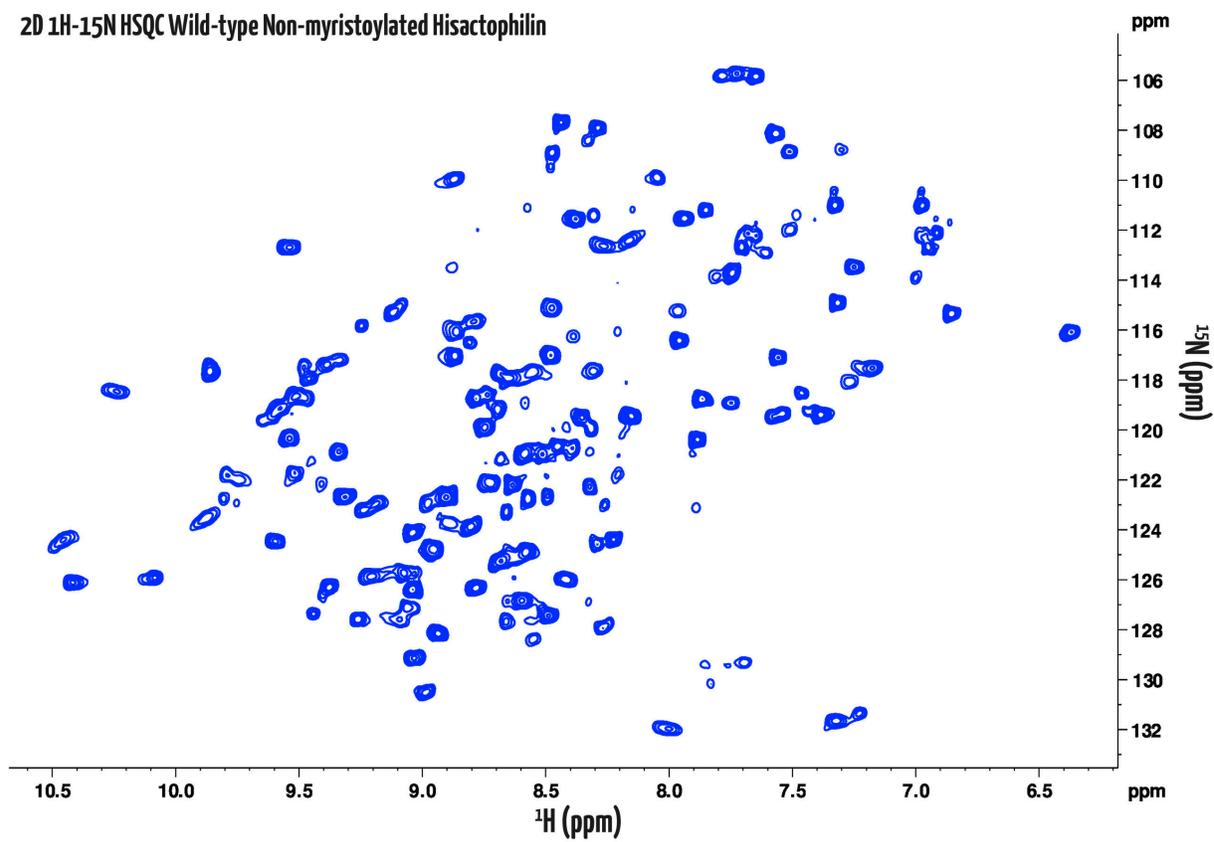


Figure 4.2 NMR Spectra of wild-type non-myristoylated hisactophilin acquired on a Bruker 600 MHz Avance NMR spectrometer. The sample contained 50 mM MES at pH 6.2, 1 mM EDTA, 10% D₂O, 1mM DSS, 2 mM protein. (Upper panel) 1D ¹H spectrum of hisactophilin, acquired with double pulse field gradients for water suppression. The approximate range of chemical shifts of different types of protons are labelled. The principal point, however, is the large number of peaks that make up a 1D protein spectrum. (Lower panel) A ¹H, ¹⁵N HSQC spectrum of the same protein as in the upper panel, but the second dimension resolves individual peaks and partially solves the overlap issue.

Multidimensional NMR is one method for overcoming the cruel reality of a crowded spectrum by correlating the proton resonances with another nucleus, like ¹⁵N. This correlation occurs through bonds by J-coupling, for example, or through space by the Nuclear Overhauser Effect (NOE). A commonly used multidimensional experiment in biomolecular NMR is the ¹H, ¹⁵N-heteronuclear single quantum correlation (HSQC) experiment which correlates the mercurial and sensitive amide protons to directly bonded ¹⁵N atoms (lower panel of Figure 4.2). Compared to the 1D experiment described above, the 2D HSQC spectrum is immediately simplified because protons not directly bonded to a nitrogen are not observed. As the name implies, data are plotted in an extra dimension corresponding to the resonance frequency of the ¹⁵N atom which, like a proton, is very sensitive to its local magnetic environment [176,177]. In the two-dimensional spectrum, individual peaks are resolved according to the resonances of their correlated hydrogen and nitrogen atoms with minimal overlap in a cooperating and well-behaved protein. In practical terms, a protein HSQC produces one peak per amino acid except for proline (which lacks an amide proton), and peaks for nitrogen-containing side chains, with the caveats that the protein is smaller than ~30 kDa, that no slow exchange processes that could create a second peak for a subset of amino acids, and no intermediate exchange processes that might broaden a peak to the point of undetectability [178].

To further resolve peaks and glean more structural information from a protein, a third dimension can be added to correlate amide groups to other entities, such as covalently linked ¹³C atoms or non-amide protons, or other protons that are nearby in space but not necessarily in primary structure (using NOEs). For example, a ¹H, ¹⁵N NOESY-HSQC records cross-peaks for any proton within ~5 Å of an amide proton; and a ¹H, ¹⁵N TOCSY-HSQC (TOCSY = Total Correlation Spectroscopy) records a cross-peak for protons, including in side-chains, in the same amino acid as a given amide group. Together, these two 3D experiments solve one of the limitations of the unaccompanied HSQC, that although each peak corresponds to an amino acid we don't know which amino acid. The information from a TOCSY-HSQC can help identify the type of amino acid while the NOESY-HSQC fills in information about the connectivity of the amino acids to build the sequence and infer and assign the exact identity of each HSQC cross-peak [162].

4.2.3.1 The Chemical Shift

The chemical shift is one of the more accessible measures obtained from protein NMR and it has been somewhat overshadowed of late by relaxation and dynamics experiments [179]}. Recent work has led to

a resurgence in chemical shifts for obtaining highly informative and precise structural and dynamic data [143,176,180–186]. The exquisite sensitivity of chemical shifts to the molecular geometry, bonding partners, and electronegativity of a nucleus has been both a boon (because of the uniqueness of the environment around a given nucleus) and a bane (because the effect of these environments on chemical shifts is challenging to model). Traditionally, chemical shifts as a purveyor of approximate structural features was a matter of course; however extracting precise, three dimensional structural information depends on accurately predicting chemical shifts from structural information or vice versa, which is an ongoing challenge [180,184,187,188]. The momentum of progress has only decreased the challenge extracting precise structural information posed, with new methods that are sequence-based, and others that are structure-based. The former relies on predicting chemical shifts based on structural alignment with existing NMR data, and the latter calculates chemical shifts directly from atomic coordinates [189]. Some of the tools developed for these calculations, albeit not an exhaustive list, include ShiftX[190], ShiftX2[180]}, PROSHIFT[191]}, GeNMR[184], CHESHIRE [192], CheShift-2 [193], SHIFTCALC[194], SPARTA+[195], SHIFTY[196], and E-Thrifty[181].

Applying these algorithms and other methodologies to chemical shifts have yielded numerous successes in predicting secondary structure [176,197–201], residue solvent accessible surface area [176], protein flexibility [202,203], structure [204–206], and torsion angles [207–209]. Part of the difficulty of accurately predicting chemical shifts is the significant impact of non-covalent interactions and dynamics that dominate protein structure, in addition to the dependence on covalent bonds. One model for chemical shifts is the sum of non-covalent terms given by,

$$\Delta\delta = \delta_{total} - \delta_{rc} = \delta_{anis} + \delta_{ring} + \delta_{HB} + \delta_e + \delta_{side} + \delta_{misc}$$

(4.2) where δ_{rc} is the random coil chemical shift, δ_{anis} is the contribution of backbone torsion, δ_{ring} is the ring current contribution, δ_{HB} is the contribution from hydrogen bonding or close contacts, δ_e is the local electric field contribution, δ_{side} is the side chain torsion contribution, and δ_{misc} is a catch-all that includes contributions from solvent, temperature, and motional averaging [177,190].

Defining determinants of the chemical shift is ongoing. In designing SHIFTX2, the Wishart group evaluated the contribution of numerous factors derived from machine learning techniques, sequence-based, and structure-based prediction methods to chemical shifts (Table 4.1). The table lists twenty of the most influential factors to the chemical shifts of backbone atoms, of which the amide proton (^1HN) is the most germane to the discussions to follow. The factors in Table 4.1 do not coincide directly with the general formula given by Equation 4. 2, however the torsion angles psi and phi contribute to anisotropy, chi to side chain torsion, and ring currents, electric field, and hydrogen bonding map onto the model. Deviations from the random coil chemical shift are most responsive to changes in the psi angle of the preceding residue (ψ_{i-1}), ring current effects, and the presence and disposition of any H-bonding to the amide or neighbouring oxygen [180,210].

Table. 4.1: Factors that contribute to chemical shifts. Table adapted from [180], illustrating the relative contribution (in percent of total) of the 20 most significant features to chemical shift prediction by the SHIFTX+ module of SHIFTX2. Contributions were determined by systematically omitting one factor at a time from the SHIFTX+ model and quantitating its effect on the RMS error which was further refined with tenfold cross validation. R. coil shift refers to the random coil chemical shift from [211]. AA is short for amino acid identity. ϕ and ψ are the backbone phi and psi angles, respectively. X are side chain torsion angles. θ is the angle between consecutive C_{α} , and SS refers to secondary structure. Electric fields contributions were calculated as in [212]. Ring currents were calculated as described in [213]. Hydrogen bond effects were calculated as discussed in [214]. Subscripts i , $i-1$, $i+1$ refer to an amino, the previous, and next amino acids, respectively.

Feature	$^{13}C'$	$^{13}C_{\alpha}$	$^{13}C_{\beta}$	1HN	$^1H_{\alpha}$	^{15}N
R. coil shift	22.5	50.0	58.5	3.0	21.3	35.9
AA_i	0.6	11.6	15.4	0.5	0.8	3.4
AA_{i-1}	0.4	0.1	0.1	0.4	0.2	2.9
AA_{i+1}	2.3	1.0	0.3	0.3	0.7	0.3
φ_i	5.8	11.0	8.1	4.4	29.9	4.5
φ_{i-1}	0.6	0.4	0.3	2.1	1.0	2.1
φ_{i+1}	3.6	1.1	0.6	0.9	1.3	0.9
ψ_i	13.9	10.4	5.7	5.3	3.8	7.1
ψ_{i-1}	1.4	0.3	0.2	15.3	0.4	18.7
ψ_{i+1}	8.6	0.9	0.3	0.6	0.6	0.5
ψ_{i-2}	0.4	0.2	0.2	5.9	0.4	0.5
χ_{1i}	4.1	2.6	1.3	0.8	1.3	5.9
χ_{2i}	3.1	2.2	1.4	0.5	0.4	1.6
θ_i	2.3	0.6	0.3	5.3	0.8	0.5
κ_i	2.5	0.3	0.2	3.1	0.4	0.4
SS_i	8.1	0.1	0.1	0.1	0.6	0.0
Electric field	0.0	0.3	0.0	2.7	12.9	0.0
Ring current	0.0	0.5	0.9	11.5	11.2	0.6
Surface area	4.2	0.3	0.2	1.2	0.6	0.5
Hbond effect	0.0	0.0	0.0	18.4	0.3	0.0

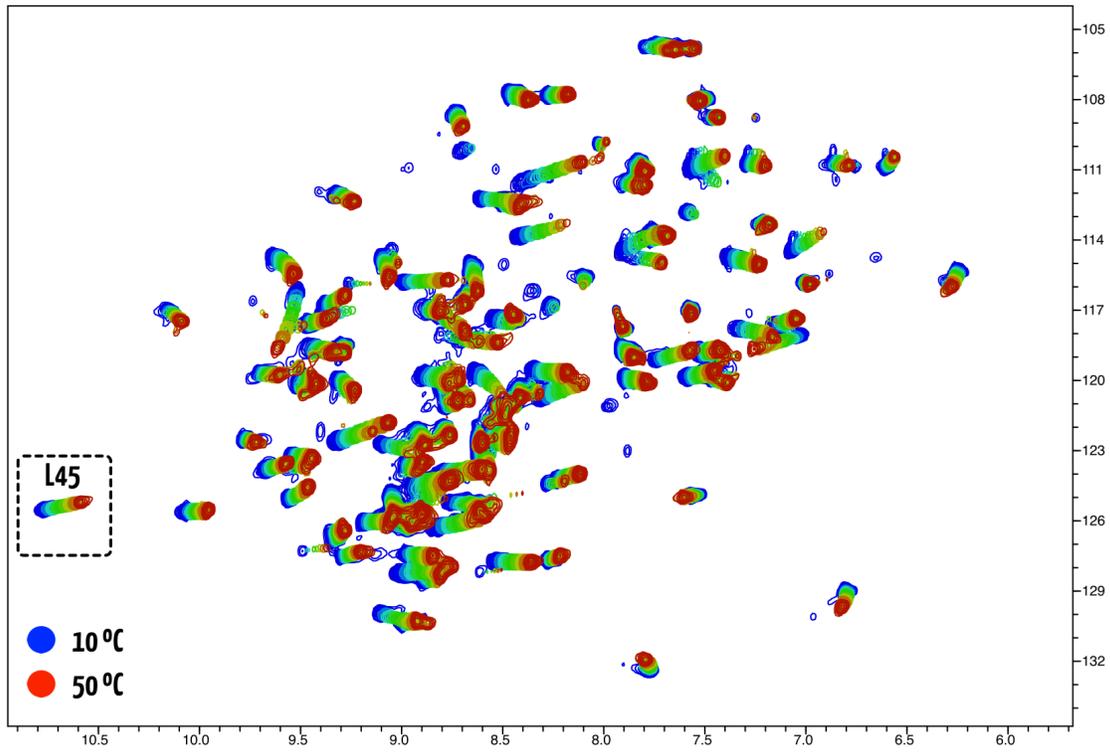
Chemical shifts require further discussion here because temperature coefficients (*vide infra*) are determined by changes in chemical shifts as a function of temperature, and the extent of these changes offers some structural information about the studied protein [150,183,215]. The exceptional sensitivity of amide protons to hydrogen bonding, the psi torsion angle of the preceding residue, and ring current effects is fortuitous because all three depend to some extent on secondary and tertiary structure and thus offers a framework for interpreting changes in chemical shifts. It is also consequential to note that chemical shifts represent the population-weighted average of the contributing terms (e.g. hydrogen bonding, anisotropy, ring current effects, etc.), so any protein motions that are fast on the NMR timescale may be captured by the chemical shift. What is meant by fast on the NMR timescale is if you have conformations X and Y, two slightly different conformations with distinct proton chemical shifts, and the speed of interconversion between X and Y is much faster than the difference between the two chemical shifts, then the chemical shift will reflect the population weighted average of conformations X and Y.

Despite progress in predicting, calculating or rationalizing carbon chemical shifts [216], amide protons are still a significant challenge because what makes them a useful and sensitive reporter is also what makes them difficult to model—they are sensitive to even small structural changes. In practical terms, even small discrepancies between the structural model used to parameterize predictors and the ensemble, the average structure that amide proton chemical shifts report on can give rise to aberrant results and a poorly performing predictor [210,217]. Quantum mechanical calculations have successfully captured some of amide proton sensitivity to small structural changes; however, on the whole they have not achieved the accuracy of some empirical models [210,218]. The progress of and ongoing interest in predictors and modelling recapitulates that the sensitivity of amide proton chemical shifts to small structural perturbations makes them an invaluable tool to interrogate small and subtle structural changes.

4.2.4 Temperature dependence

With chemical shifts well established as sensitive to local magnetic environments, which in proteins generally arise from the local conformation and motional averaging, considering how chemical shifts change as a function of a perturbation from ligand binding, pH or temperature changes, for example, can offer insight into the nature of the interactions and their effects on a protein. Analyzing chemical shift perturbations can help identify binding sites [177,219–222], the dielectric constant of the interior of a protein [223], or the $pK_{a,apparent}$ of backbone amide groups [8,224]. In essence, monitoring chemical shifts as they respond to changes in environment can offer insight into the structural and/or dynamic response of the protein. Acquiring NMR spectra at incremental temperatures and tracking the linear movement of individual resonances (Figure 4.3) provides residue specific temperature coefficients (sometimes referred to elsewhere as temperature gradients) that offer similar insights and structural information as other chemical shift-based methods.

Wild-Type Myristoylated Hisactophilin at pH 6.2



Residue L45

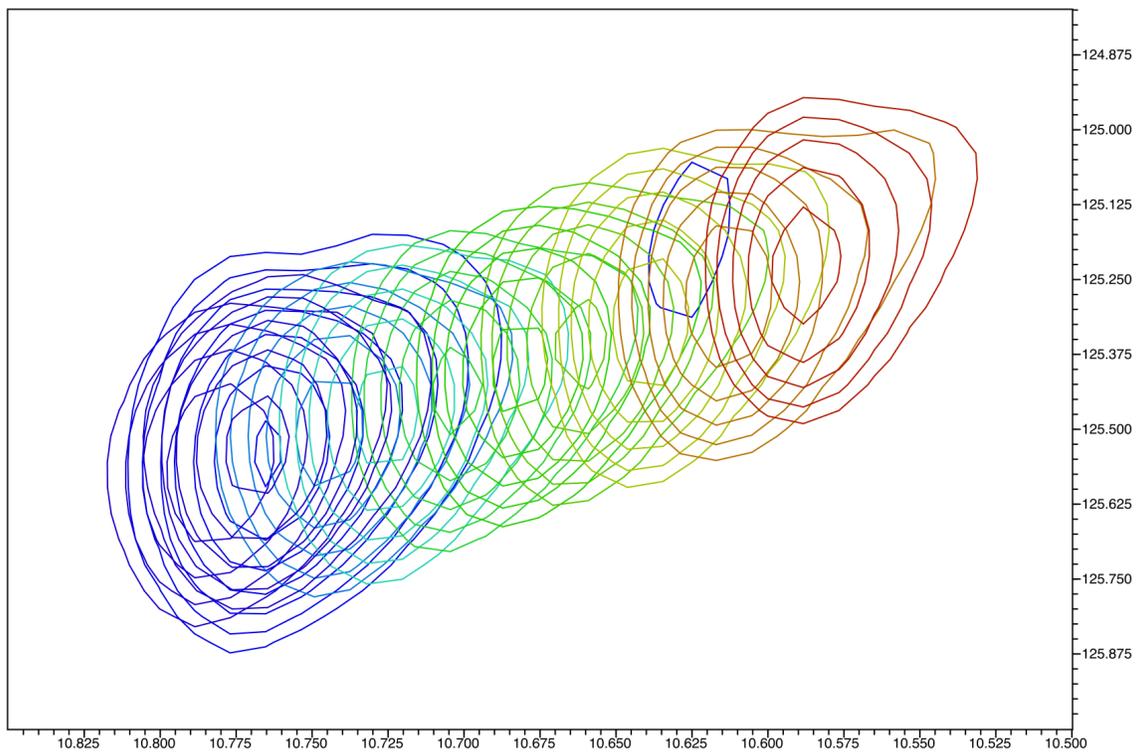


Figure 4.3: ^1H - ^{15}N HSQC recorded as a function of temperature. Experimentally, spectra were acquired at 2.5 °C increments, however for clarity and legibility, only select temperatures between 10 - 50 °C are shown for illustrative purposes. (Upper Panel) ^1H - ^{15}N HSQC of a uniformly labelled myristoylated, wild-type hisactophilin at pH 6.2 acquired on a Bruker Avance 600 MHz spectrometer with a triple resonance TXI 5 mm probe. The spectrometer temperature controller was set to 10 - 50 °C, however the difference in chemical shift between DSS and water was used as an internal calibration of the actual temperature. Overlays were made in CARRA NMR [225]. Of importance is that different residues are differentially responsive to changes in temperature. (Lower Panel) Enlarged area of the ^1H - ^{15}N HSQC spectrum centred on residue L45 to demonstrate the regular movement of its resonance with increasing temperature.

Temperature coefficients were first described in 1969 by Ohnishi and Urry [226,227] for peptides gramicidin S and Valinomycin, where they were pitched as an alternative to hydrogen/deuterium exchange experiments to define hydrogen-bonded amides, and their use in this capacity continued intermittently into the 1990s [228–230]. Their adaptation to proteins, which have relatively more defined structures than most peptides, however, did not occur until later in the 1990s and early 2000s [136,231–234]. Initially, applying temperature coefficients to proteins was limited to a sort of binary scale, where temperature coefficients on one side of an empirical (but somewhat arbitrary) cut-off was likely to indicate intramolecular hydrogen bonding, while those on the other side were likely to be intermolecularly hydrogen-bonded, say to solvent [231,233]. Based on 793 amide protons in 14 globular proteins, a cut-off was defined as -4.6 ppb/K and more positive temperature coefficients were 85% predictive of intramolecular hydrogen-bonds, increasing to 93% when further constrained to between -4 and -1 ppb/K to exclude those with neighbouring aromatic rings. Conversely, 80% of intermolecularly hydrogen-bonded amide protons have temperature coefficients more negative than -5 ppb/K [231,232].

This simple view of temperature coefficients to conveniently add additional constraints to structure calculations with acceptable accuracy is now standard practice. Temperature coefficients with this inter-/intramolecular cutoff have helped to define catalytically-relevant thermal plasticity in NADH oxidase [235], the impact of charge mutations on the stability and function of the Pin1 WW domain [236], interrogated changes in the hydration shell of γS -crystallin and crowding [237], offered insight into a collagen model peptide and the role of hydrogen-bonding with water [238], attributed the disfunction of a multitopic receptor to changes in intramolecular hydrogen-bonding networks [239], characterized a sparsely populated cold-denatured state of a Cro protein [240], and determined an intact β -barrel in immature SOD1 [151]. Researching and using temperature coefficients has not stagnated, however, as our understanding of temperature coefficients and their determinants continues to evolve, opening avenues for their application to problems that involve higher resolution data. Temperature coefficients could do more—if they are only 85% predictive of intramolecular hydrogen bonds and 80% predictive of intermolecular bonds [231,232], what other factors might temperature coefficients report on?

At their heart, temperature coefficients are chemical shift-based measures where the information obtained from changes in chemical shifts report on changes to their determinants (as discussed in 4.2.3.1 The Chem-

ical Shift). The sensitivity of chemical shifts, and thus temperature coefficients, to hydrogen bonding bears comparison to NMR hydrogen/deuterium exchange experiments. In H/D exchange, the base-catalyzed exchange of protein protons for de facto NMR-invisible deuterons results in a time-dependent decrease in signal intensity for each proton resonance [241–243]. The rate of signal loss depends on the pH of the environment, temperature, and how protected a proton is from the solvent. Protection most intuitively comes from involvement of the amide in hydrogen-bonds that anchor secondary structure, but also from tertiary structure that could exclude solvent, and any dynamics and motions that might transiently undermine secondary or tertiary structure. Though amide H/D exchange is widely used and powerful, some of the limitations of H/D exchange do not apply to temperature coefficients — temperature coefficients are largely pH-independent (barring pH-induced conformational changes [136,183,232,232,244]), and surface-exposed amides that exchange too fast for H/D exchange are readily measured by temperature coefficients, and so also reporting on local stability of the local magnetic environment as a proxy for local conformational stability [183]. Applying temperature coefficients to measure conformational stability, however, is a recent application.

A systematic characterization of the temperature coefficients of the B1 domain of protein G with varying pH, and a more recent mutational study of human SOD1, demonstrated that they are a palimpsest of more meaningful structural information than whether or not an amide is hydrogen-bonded [150,183]. These works expanded temperature coefficients to encompass a more general measure of structural stability, contributors to which may include the details of hydrogen bonding, the relative orientation of nearby aromatic groups, and the temperature-dependent propensity to lose structure. In SOD1, the global stability of wild-type and mutants measured by differential scanning calorimetry have less negative average temperature coefficients (indicative of greater structural stability). Immature forms of SOD are less globally stable and correlated with more negative temperature coefficients (supporting reduced global stability). As a reporter of local structural stability, temperature coefficients localized unstable areas at the edges and periphery of the structure for the otherwise highly globally stable mature SOD1. Given SOD1's propensity to aggregate, this could have meaningful implications for disease [150].

Broadly, a more negative a temperature coefficient corresponds to lower structural stability, which could arise from weakly participating in secondary structure, a neighbouring unpaired charged residue, or an aromatic side chain with a high propensity for movement and rearrangement[183]. Qualitatively, hydrogen bonds polarize the electron density of the peptide bond perpendicular to the N-H and deshield it; however, as the temperature increases and structures expand and explore more of the conformational landscape, deshielding is reduced and peaks generally move upfield [136,215]. That said, understanding the precise mechanism and origin of temperature coefficients is a work in progress. In an independent set of experiments, amide proton temperature coefficients correlated well with measures of temperature-dependence of through-hydrogen-bond coupling, $^3J_{\text{NC}}$, and were further supported by parallel tempering molecular dynamics simulations that

demonstrated thermal expansion of the hydrogen bond and protein structure are important components of the amide proton temperature coefficient [215,234]. Furthermore, it is well established that hydrogen bond length and strength are correlated [245,246]. Taken together with the findings that temperature coefficients are determined by the propensity of a region to become disordered with increasing temperature [183], there is compelling evidence that temperature coefficients expound the structural stability of the local environment of an amide proton, allowing us to determine the relative energetic contributions of individual residues to global stability [150].

In the work that follows, I expanded the use of temperature coefficients by applying them over a wider range of temperatures and with more increments than previously reported. Temperature coefficients have been used to study a single protein [136] with varying pH [183], and to compare different maturation states and mutations in SOD1 [150]. Here, I apply temperature coefficients to hisactophilin as a new tool to describe its near-native energy landscape and to develop a mechanism of pH-dependent myristoyl switching by localizing the sources of conformational strain that drive it. This was accomplished by systematically measuring temperature coefficients at pH values on either side of the myristoyl switch for acylated and non-acylated wild-type protein. When the same methodology was applied to two switching-impaired mutants, indications of strain in residues identified in wild-type disappeared, which demonstrates the utility of temperature coefficients in future studies of allostery and structure-function.

4.3 Materials and Methods

4.3.1 Expressing ¹⁵N-labeled Hisactophilin

Hisactophilin was expressed using previously published methods [8,87] and subsequently purified using a newly developed purification protocol (discussed in Chapter 3). *E. coli* strain BL21 was doubly transformed with plasmids pHV738 and pHW. PHV738 encoded the IPTG-inducible gene for human N-myristoyl transferase 1 (NMT) under the control of the Ptac promoter to handle the N-terminal co-translational acylation of hisactophilin by myristic acid, the *E. coli* methionine aminopeptidase gene to expose the requisite N-terminal glycine for NMT activity, and kanamycin resistance for selection. The pHW plasmid encoded the IPTG-inducible hisactophilin mutant, similarly under the control of the Ptac promoter, and ampicillin resistance for selection.

Transformation was effected by electroporation [247] on cells that were made competent by harvesting mid-log phase by centrifugation at 4000 x g for 15 minutes then chilled and kept on ice for the remainder. The resulting cell pellet was resuspended and then again pelleted as above in 500 mL, 250 mL, 20 mL, and then 2 mL of ice-cold 10% glycerol before being stored at -80 °C. 40 µL to 80 µL of the now-electrocompetent cells were mixed with 2 µL of plasmid DNA in a low ionic strength buffer in a 0.1 cm electroporation cuvette

(BioRad Laboratories Inc., Hercules, CA) and subjected to a 1.8 kV electric field for approximately 1 s (Eppendorf Eporator, Eppendorf, Hamburg, Germany). Cells were immediately transferred to 1 mL of sterile SOC liquid medium (2% Bacto tryptone, 0.5% bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose) for recovery and to allow for expression of ampicillin- and kanamycin-resistance genes before being plated on ampicillin- and kanamycin-containing selective lysogeny (LB) agar plates.

A single viable colony was selected to inoculate 10 mL of LB (100 µg/mL ampicillin, 30 µg/mL kanamycin) culture for growth overnight (~16 hours) at 37 °C and 200 rpm in a shaker/incubator. These overnight cultures were used to inoculate (1:100) larger, 1L M9 minimal media containing Na₂HPO₄ (6 g/L), KH₂PO₄ (3 g/L), NaCl (0.5 g/L), ¹⁵NH₄Cl (0.5 g/L), MgSO₄ (2mM), glucose (4 g/L), thiamine (0.0005 g/L), CaCl₂ (0.1 mM), ampicillin (100 µg/mL) and kanamycin (30 µg/mL), fortified with micronutrients ((NH₄)₆(MO₇)₂₄ (124 µg/L), H₃BO₃ (25 µg/L), CoCl₂ (24 µg/L), CuSO₄ (25 µg/L), MnCl₂ (20 µg/L), and ZnCl₂ (14 µg/L)) to ease the stresses of growing in minimal media. Growth continued under the same shaker/incubator parameters as the overnight cultures. Once an optical density at 600 nm (OD₆₀₀) of 0.25 was read by a UV/Vis spectrophotometer (Cary 300 UV-Vis, Agilent, Santa Clara, CA), 10 mL of 20 mM sodium myristate was added to each culture flask and then, once OD₆₀₀ = 0.7 was reached, cells were induced with 1 mM IPTG. Optimization experiments determined that yields were greatest for Wild-type and LLL when incubation temperatures were adjusted to 25 °C and allowed to continue for 18 - 24 hours before harvesting. Cells expressing I85L, however, responded with small cell pellets and less hisactophilin expression to growth at 25 °C, so growth conditions were adjusted to a more canonical 37 °C for 6 hours [8]. Cells were harvested by centrifugation at 5,000 x g for 12 minutes at 4 °C and stored at -80 °C.

4.3.2 Purifying ¹⁵N-labeled Hisactophilin

4.3.2.1 Cell Lysis

Purification began with thawing 2 L or 3 L of cultures worth of cell pellet and resuspending in 50 mM Tris (pH 8.0, 200 mM NaCl) and adding 1 mM MgCl₂, 0.1 mM of the protease inhibitor PMSF, and then lysing the cells with three passes through the emulsifier (Emulsiflex C5, pressure sweeps from 20,000 psi to 8,000 psi). 12 mM 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate (CHAPS) and DNase I were added to the lysate to help solubilize any hisactophilin that might be membrane-bound and reduce the viscosity, respectively, and incubated at 4 °C for one hour. Two rounds of centrifugation (48,000 x g, 22 minutes, 4 °C) resulted in a soluble fraction that contained nearly all the hisactophilin, and this was further clarified with a 0.4 µm syringe filter before nickel immobilized metal affinity chromatography.

4.3.2.2 Immobilized Metal Affinity Chromatography

Traditionally, nickel affinity chromatography takes advantage of an engineered N- or C-terminal histidine-tag, however hisactophilin's natural abundance of 31 histidines of 117 residues is sufficient for binding the nickel resin (Profinity IMAC, BioRad Laboratories Inc, Hercules, CA). We used a low-pressure chromatography system (BioRAD BioLogic LP, BioRad Laboratories Inc, Hercules, CA) with integrated conductivity metre and UV-absorbance at 280 nm with the nickel resin, a binding buffer of 50 mM Tris (pH 8.0, 200 mM NaCl), a 50 mM potassium phosphate wash buffer (pH 6.3, 300 mM NaCl), and a 50 mM formate elution buffer (pH 3.5, 300 mM NaCl). The presence of hisactophilin was verified by 15% SDS-PAGE.

The hisactophilin-containing fraction was dialyzed in 3.5 kDa molecular weight cut-off dialysis tubing (Spectrum Laboratories Inc., Rancho Dominguez, CA) against three to four exchanges of four litres of Mili-Q water. This was followed by concentrating to 3 mL in an Amicon stirred cell (Millipore Sigma, Burlington, MA) with a 3 kDa molecular weight cutoff regenerated cellulose filter (Millipore Sigma, Burlington, MA) in preparation for reverse-phase chromatography.

4.3.2.3 C18 Reverse Phase Chromatography

At this stage of the purification, owing to imperfect efficiency of NMT in an *E. coli* host, hisactophilin exists as a mixture of myristoylated and non-myristoylated forms that are indistinguishable by prep-grade size-exclusion chromatography or SDS-PAGE because the change in mass is only 210 Da. Resolving the two forms of hisactophilin was achieved by a C₁₈ chromatography column (Waters Inc., Milford, MA) and a BioRad DuoFlow medium-pressure chromatography system (BioRad Laboratories Inc, Hercules, CA). Hisactophilin was unfolded with 0.1% trifluoroacetic acid and eluted from the column with a gradient of acetonitrile from 2% to 100%. At this stage, myristoylated and non-myristoylated hisactophilin were well-resolved and >95% pure. The acetonitrile was dialyzed against an acidic but weak buffer over two exchanges, and then hisactophilin was refolded into 20 mM ammonium carbonate, concentrated using an amicon stirred cell (Millipore Sigma, Burlington, MA), and lyophilized (Labconco Freezone 4.5, Labconco Corporation, Kansas City, MO).

4.3.3 One- Two- and Three-Dimensional Nuclear Magnetic Resonance Experiments

4.3.3.1 Sample Preparation

Preparing hisactophilin for NMR, fortunately, is not especially arduous because it is amenable to lyophilization, and both myristoylated and non-myristoylated forms are soluble to concentrations in excess of our target of 25 - 30 mg/mL (1.8 - 2.2 mM), albeit unsurprisingly slightly pH-dependent. Wild-type, I85L, and LLL were all prepared in the same fashion; 13-15 mg of lyophilized protein was dissolved directly in a buffer containing 50 mM potassium phosphate (final pH 7.72 or pH 6.76, 1 mM EDTA), 1 mM DTT, 10 % D₂O, and 1 mM DSS to a final volume of 500 μ L. The set up is identical for experiments at pH 6.2, however

rather than potassium phosphate, we used 50 mM 2-(N-morpholino)ethanesulfonic acid (MES) with a final pH of 6.18. The lyophilized protein cake often entirely redissolved for non-myristoylated hisactophilin and about 85-90% for myristoylated hisactophilin. However, to overcome this shortcoming of solvation, the protein can be first dissolved and unfolded in a small volume of 5-10 mM acetic acid, followed up dilution with the NMR buffer to the desired volume. This tends to result in 100% dissolved and cheerfully refolded protein.

4.3.3.2 Experiments for Assignments

The spectrometer used for the following experiments was a Bruker Avance spectrometer operating at 600 MHz with a 5 mm triple resonance probe (TXI) with x, y, and z-axis gradients (Bruker, Billerica, MA) and BCU 05 variable temperature control unit. Control of the NMR and subsequent basic spectrum processing for 1D and 2D experiments was carried out in Bruker XWIN-NMR, and processing 3D datasets by Bruker TopSpin 3.5.

Three variants of hisactophilin in varied states of acylation were assigned, resulting in five endeavours in assigning resonances: myristoylated WT at pH 6.18, myristoylated and non-myristoylated LLL at pH 6.76, and myristoylated and non-myristoylated I85L at pH 6.76. The suite of experiments used for each assignment exercise comprised $^{15}\text{N}, ^1\text{H}$ HSQC [248–251], ^{15}N -edited NOESY - HSQC [249–252] for WT, LLL, and I85L (pulse sequence can be found in Appendix C, and ^{15}N -edited TOCSY-HSQC [249–251] for WT, and LLL (pulse sequence can be found in Appendix C). The $^{15}\text{N}, ^1\text{H}$ HSQC that preceded and followed each 3D experiment was executed at a set temperature of 25 °C, used INEPT magnetization transfers [253], ^{15}N decoupling by GARP [254], and water suppression using double pulse field gradients. The ^{15}N -edited NOESY - HSQC had a 150 ms mixing time to allow for NOE transfer and to allow some dubious but occasionally very helpful spin diffusion, and the ^{15}N -edited TOCSY-HSQC had a 60 msec mixing time.

4.3.3.3 Analyzing and Assigning

3D data were processed with some amount of zero filling, linear prediction in the indirect dimensions, a sine shifted bell window function, and baseline correction to strike a balance between signal intensity and apparent resolution, particularly in the indirect ^1H dimensions. Further analysis of the assigning-oriented 2D and 3D spectra employed the software Computer Aided Resonance Assignment (CARA)[225] which provided a friendly but fraught cross-platform graphical user interface for calibrating spectra, building databases of crosspeaks, and most importantly finding tracts of sequential resonances. We used the open source software Visual Molecular Dynamics (VMD)[52,255–257] to map assignments onto existing solution NMR structures [42]. The assignments for LLL and I85L (myristoylated and non-myristoylated) were transposed from pH 6.8 where they were determined to pH 6.2 and 7.7 with the help of an NMR $^{15}\text{N}-^1\text{H}$ HSQC-monitored pH titration experiment. The titration experiment will be explained in greater detail and treated more quantitatively

in Chapter 5.

4.3.3.4 Variable Temperature Nuclear Magnetic Resonance Experiments

Variable temperature NMR experiments were employed to measure the temperature dependence of amide proton chemical shifts of ^{15}N -labeled protein on temperature. These experiments involved preparing a myristoylated or non-myristoylated NMR sample (as described in the 4.3.3.1 Sample Preparation section) at either pH 7.7 or pH 6.2. The sample was then loaded into a Bruker ceramic spinner, suitable for temperatures up to 423 K (rather than the standard blue POM spinner), at a depth of 18 mm. The sample equilibrated for ten minutes in the NMR, followed by matching, tuning, and automatic shimming (by gradshim) at a set temperature of 298 K (controlled by the BCU 05 with 400L/hr of nitrogen gas flow). Once the 90° pulse was measured and the optimal power for water suppression by excitation sculpting [258] calculated, a 1D spectrum centred on water was measured to determine the chemical shift of DSS, followed by an 8 scan ^{15}N - ^1H HSQC. After this first HSQC, the temperature was lowered to a set temperature of 283 K, allowed to equilibrate for 10 minutes, matched, tuned, shimmed, had its pulse lengths measured, and optimized power levels for water suppression. Following the 1D experiment as above and ^{15}N - ^1H HSQC, the temperature was incremented by 2.5 K and this process repeated until spectra were measured at 323 K. The temperature was then returned to 298 K where a final spectrum was measured to ensure reversibility and quantify the amount of protein, if any, lost to aggregation.

4.3.3.5 Analyzing Variable Temperature Data

Methods and workflows for analyzing the temperature dependence data to determine temperature coefficients and non-linearities are covered in greater depth in Chapter 6. However, in brief, 2D spectra were processed using XWIN NMR/TopSpin and referenced to the chemical shift of DSS. TopSpin's automated peak-picking algorithm generated peak lists ($\sim 2\times$ the number of expected peaks) from parabolically interpolated peak shapes. The well-defined relationship between separation of the water and DSS peaks was used as an internal thermometer to determine the actual temperature for each increment [259].

The generated peaks lists, actual temperatures, and any assignments were given as input to the script Shift-T written by Kyle Trainor, a PhD student in the Meiering Group, written in Python using components from the SciPy ecosystem [260–262]. Shift-T grouped and tracked the temperature dependent movement of peaks based on the given assignments and calculated temperature coefficients from the slope (by linear least squares regression) of amide proton chemical shift vs temperature. Residues for which Shift-T could not find a solution had their peaks manually and laboriously tracked and recorded in CARA [225] and Microsoft Excel. Temperature coefficients were then written into PDB files (with a straightforward python script) in the beta-factor field and mapped onto hisactophilin's structure using VMD [52,255–257] and rendered using the tachyon ray tracer.

The earliest variable temperature experiments were analyzed without Python-assisted automation. Each set of spectra was imported into CARRA where peaks were manually picked, centred, and assigned at the starting temperature of ~298K. That assigned peak list was duplicated and each peak was adjusted to its new position for the next temperature increment. Repeated ad nauseam for each temperature for wild-type myristoylated and non-myristoylated, pH 7.7 and 6.2. Later reprocessing of the data by Shift-T produced, on average, less than a 5% difference in the resulting chemical shifts, so processing entire data sets manually has been discontinued.

4.4 Results and Discussion

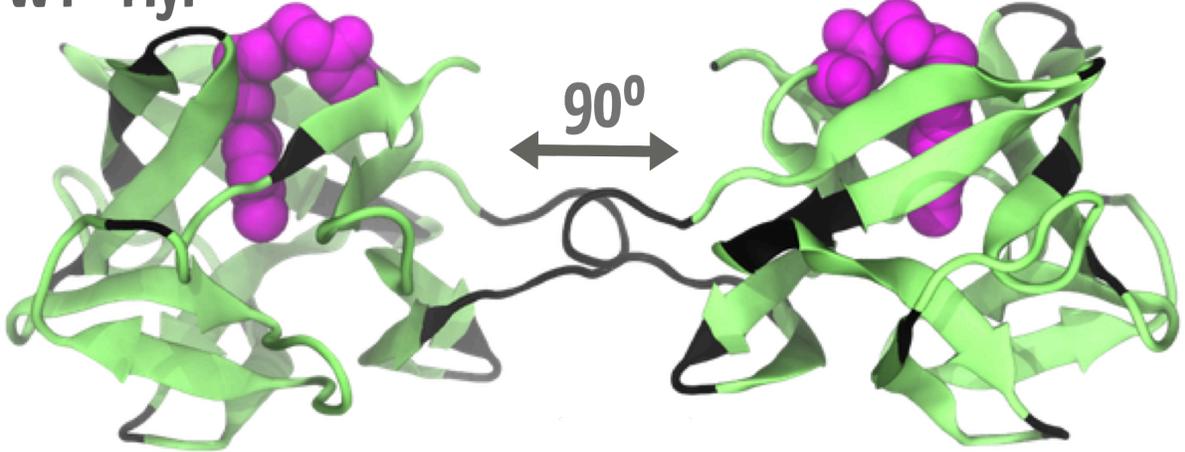
4.4.1 Chemical Shifts

4.4.1.1 Backbone Resonance Assignments

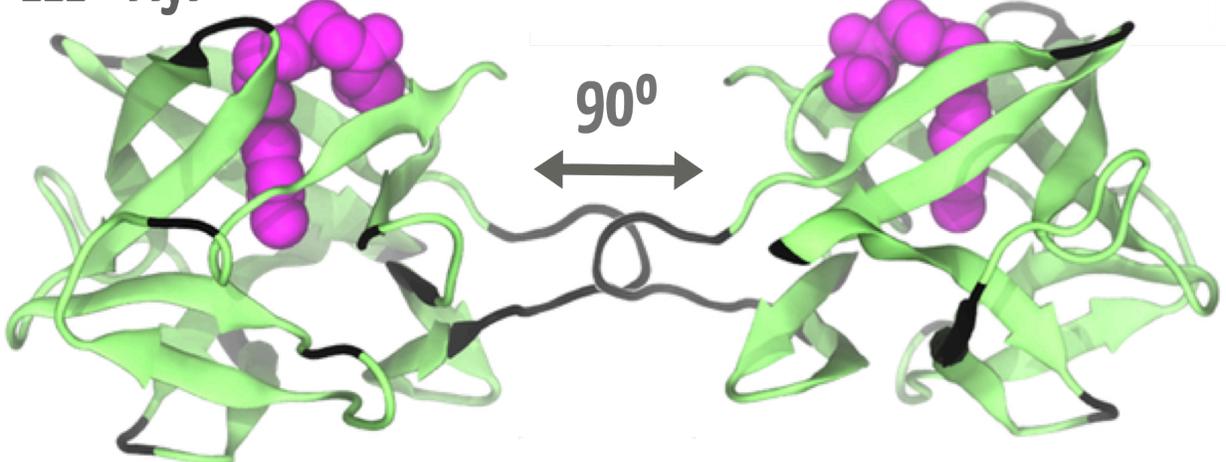
One of the first challenges faced by anyone with aspirations of nuclear magnetic resonance experiments is assigning the resonances that appear in a ^{15}N - ^1H HSQC to backbone amide protons belonging to specific amino acids in the protein of interest—here, hisactophilin. Although not a trivial exercise, assigning amino acid residues to resonances de novo can be achieved with a robust suite of three-dimensional NMR experiments that require ^{13}C - and ^{15}N -labelled protein. If aspirations end at assigning the backbone resonances and do not extend to determining a structure, ^{15}N -labelled protein and ^{15}N -edited NOESY - HSQC and ^{15}N -edited TOCSY - HSQC may be sufficient for the task, especially if there are existing assigned and similar spectra from which to work. Assignments for wild-type non-myristoylated hisactophilin are known and pH-titration experiments allow for fairly straightforward transposition of the assignments to a host of different experimental conditions [42,74]. While we do have assignments for wild-type myristoylated hisactophilin at pH 6.8 [8], it was reassigned from new NMR experiments at pH 6.2 to confirm the assignments transferred using the pH-titration.

Equipped with a ^{15}N -edited NOESY - HSQC and ^{15}N -edited TOCSY - HSQC and the deft eyes of Mikaela Ney, an undergraduate student, we successfully assigned 97 of 117 possible backbone amide resonances to their respective residues, as mapped onto hisactophilin in the upper panel of Figure 4.4. Nine of the missing residue assignments are located in a single loop between β -strands three and four. The remaining residues are randomly distributed through the structure, though they are more often in loop structural elements whose sometimes dynamic or poorly structured nature complicates the assigning process.

WT - Myr



LLL - Myr



LLL - Non myr

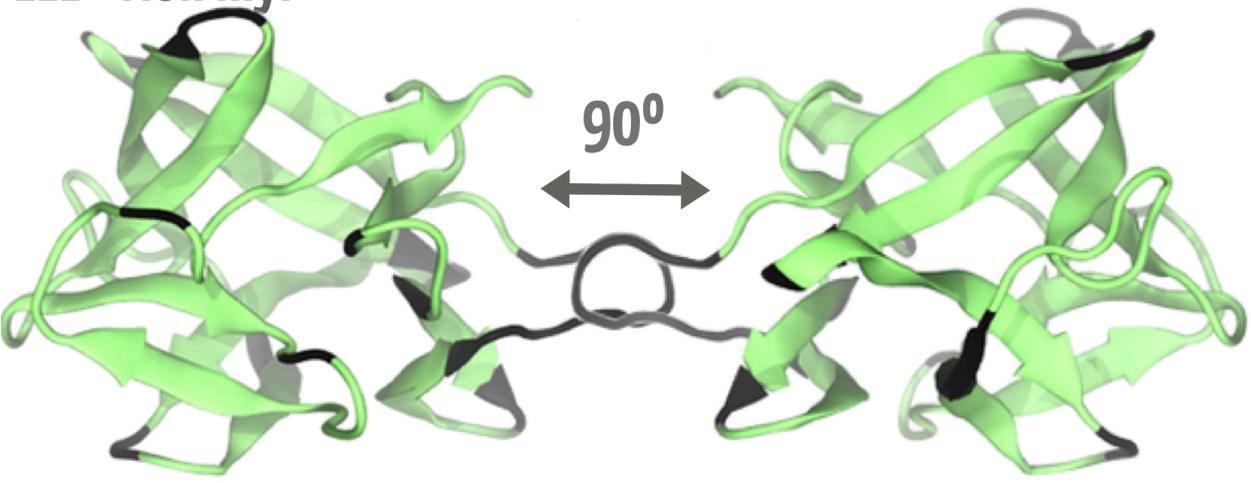


Figure 4.4: Illustration of assigned residues mapped onto the structure of hisactophilin for wild-type and F6L/I85L/I93L (LLL). This structure is based on models built for representing the position of the myristoyl group based on NMR-based NOE constraints [8] in the previously determined solution structure [42]. Assigned residues are shown in green, unassigned residues are in black, and the myristoyl group is in magenta. (Upper panel) Myristoylated wild-type hisactophilin with successful assignments at pH 6.2, determined with the valuable help of Mikaela Ney. Many of the unassigned residues are in loops, the loop spanning residues 25-33 between β -strands three and four being the longest unassigned stretch. The difficulty in assigning this loop could be because it is highly mobile and dynamic. Some overlap between peaks in the spectra also accounts for some of the incomplete assignments. Overall, 97/117 were assigned in a short period of time. (Middle panel) Assignments of for the myristoylated form of (LLL) at pH 6.8. 97 residues were assigned with the help of Purnank Shah. Many of the missing assignments mirror those missing in wild-type. (Lower Panel) Assignments for the non-myristoylated form of LLL determined with the help of Travis Ko and Christopher Leo. 97 residues were assigned.

Interest in studying two broken-switch hisactophilin mutants, the triple mutant F6L/I85L/I93L (LLL, broken toward the sequestered state) and the single mutant I85L (broken toward the accessible state), necessitated assigning projects for the myristoylated and non-myristoylated variants. The mutations composing LLL were all in the core of the β -trefoil and included mutating an aromatic residue (F6L), so although there are similarities in its spectrum to wild-type (illustrated in Figure 4.5), differences were sufficiently pronounced that we acquired ^{15}N -edited NOESY-HSQC and ^{15}N -edited TOCSY-HSQC spectra for assigning. The non-myristoylated LLL was assigned with the help of two consecutive undergraduate students, Travis Ko and Christopher Leo. Together, we confidently assigned 97 of 117 residues. Like in wild-type, missing assignments are, for the most part, in loops, particularly concentrated between β -strands three and four (see the lower panel in Figure 4.4). Assigning myristoylated LLL was first undertaken by another undergraduate student, Purnank Shah whose work I expanded on to assign (coincidentally) 97 of 117 residues with the same problematic areas noted in the non-myristoylated variant (middle panel of Figure 4.4).

After wild-type and LLL were assigned, it was predicted that I85L's 2D spectrum would resemble wild-type and LLL closely enough that a ^{15}N -edited NOESY - HSQC would provide information enough to confirm assignments (see Figure 4.6). Largely, this was true; however, I85L's spectrum suffered from more overlap than the other mutant and the HSQC-NOESY were less helpful in determining sequential assignments (suggesting perhaps that I85L is more dynamic), resulting in approximately 80 confident assignments (by Christopher Leo) for myristoylated and non-myristoylated I85L (mapped onto the hisactophilin structure in Figure 4.7). Fortunately, the assignments we do have are distributed somewhat evenly throughout the protein and are many of the same for LLL and wild-type, so the results of our temperature dependence experiments still readily compare to one another and report on similar sections of the protein.

4.4.1.2 Chemical Shift Perturbations

Having NMR assignments available for hisactophilin as we perturb it with a myristoyl group, changes in pH, and by mutation can clarify the impact each of these changes to the structure has by straightforwardly measuring the deflection of chemical shifts in response to these changes to the protein. This may be referred to

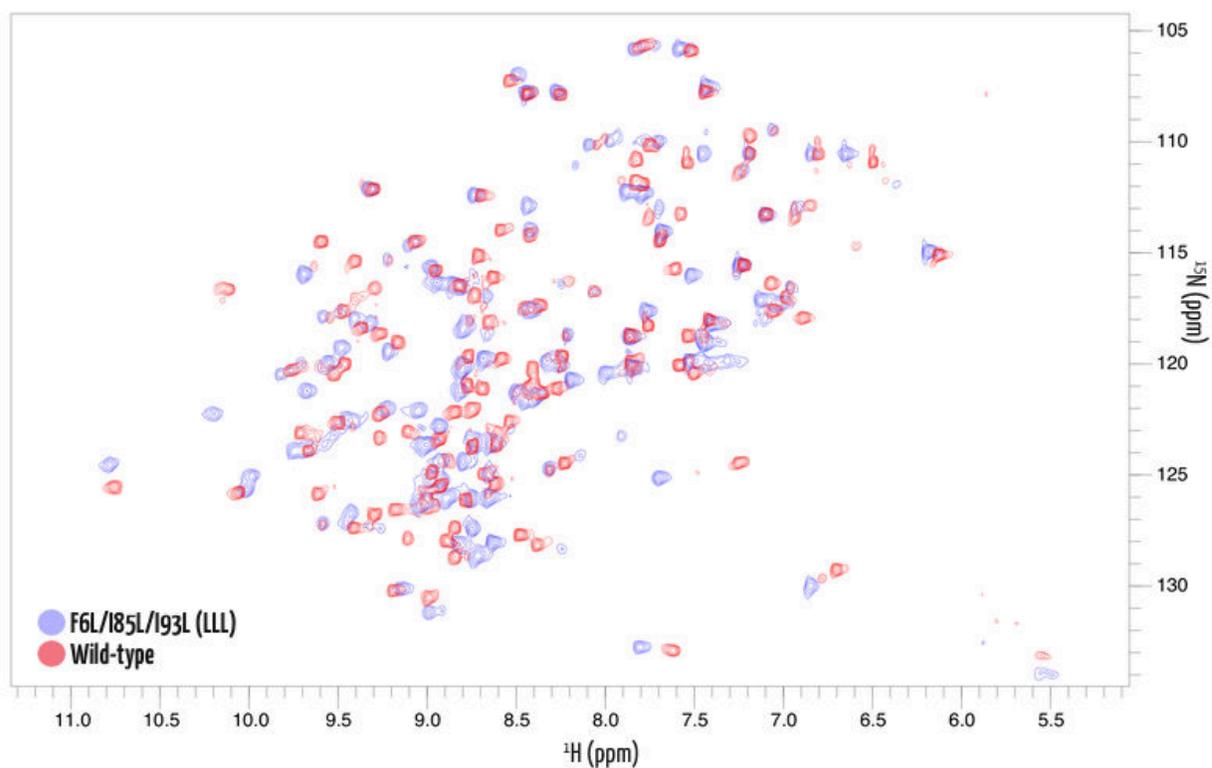


Figure 4.5: Overlay of the ^1H , ^{15}N HSQC for myristoylated wild-type and mutant LLL at pH 7.7. Both spectra were acquired on a Bruker 600 MHz spectrometer at 298K on samples prepared nearly identically — 50 mM potassium phosphate buffer, 1 mM EDTA, 1 mM DTT, 1 mM DSS, ~ 2 mM protein. Overlay was made using CCPNMR software version 2.4.2 [263]. The plot highlights that although there are similarities between the wild-type and LLL spectra, where some peaks remain in roughly the same location, a sufficient number are significantly perturbed that it was prudent to acquire a ^{15}N -edited TOCSY-HSQC as well as a ^{15}N -edited NOESY-HSQC to aid assigning.

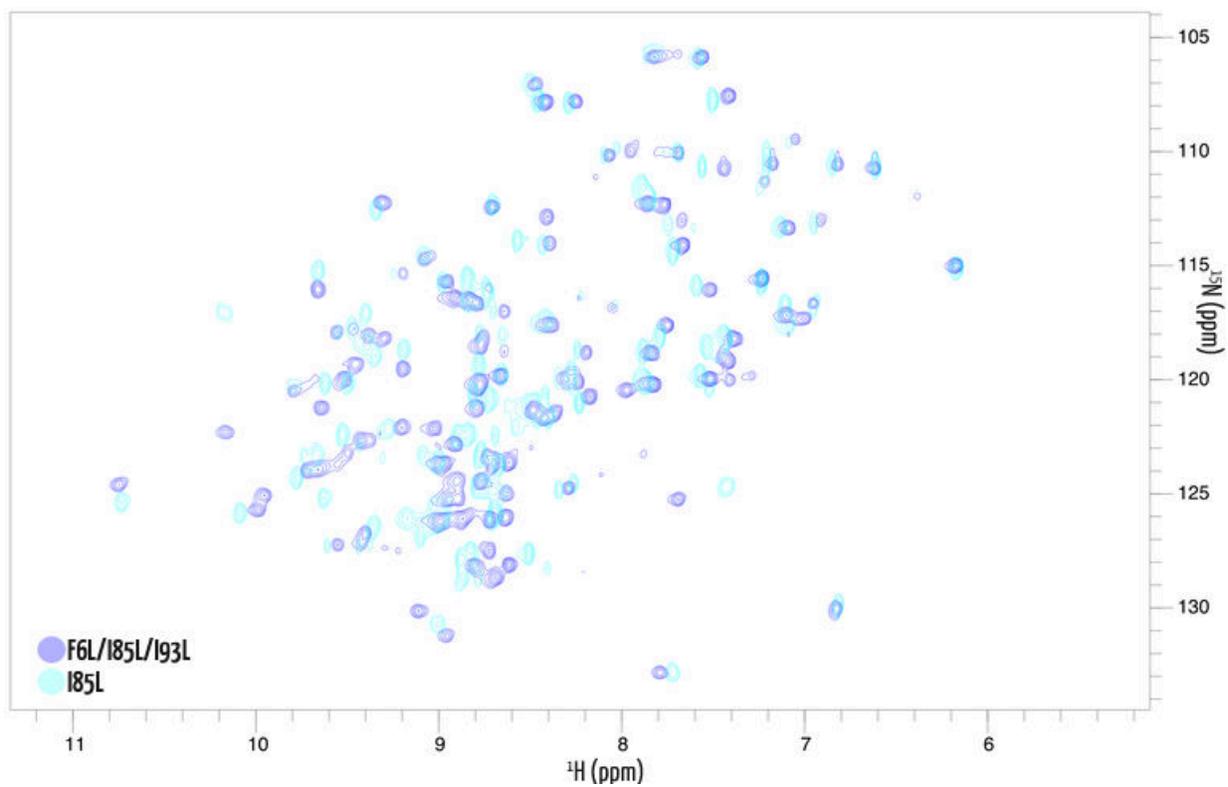


Figure 4.6: Overlay of the 2D ^1H , ^{15}N HSQC of myristoylated I85L and LLL to illustrate some of the similarities between the spectra of LLL and I85L that allowed assignments to be made using only a 3D ^{15}N -edited NOESY-HSQC. Both spectra were acquired under the same conditions listed in Figure 4.5. The overlay was made using CCPNMR version 2.4.2 [263].

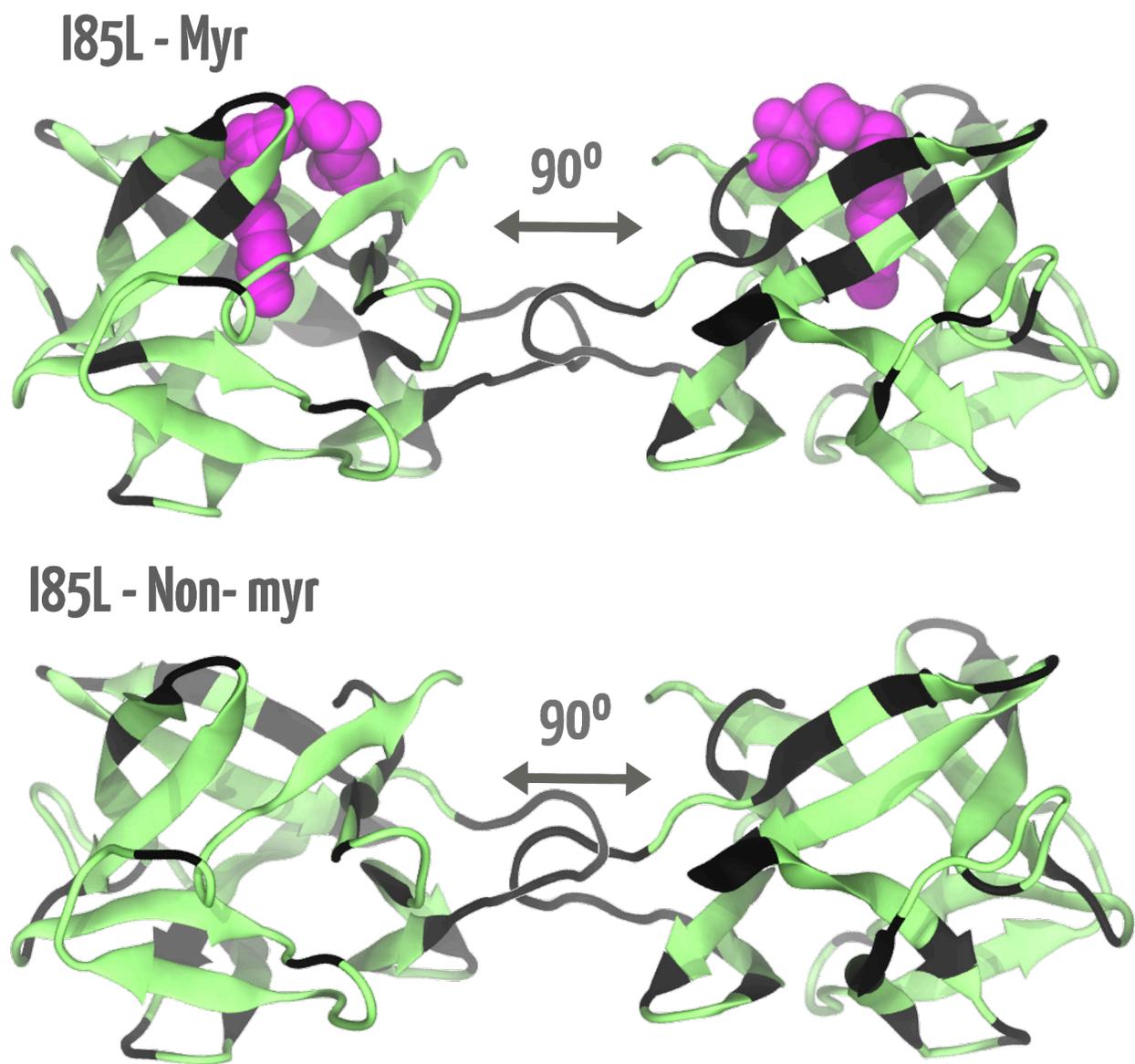


Figure 4.7: Illustration of assigned residues for myristoylated (upper panel) and non-myristoylated (lower panel) I85L. They are mapped onto the original structure that had the myristoyl group modelled in based on NMR-derived NOE constraints [8,42]. Assigned residues are green, unassigned residues are black, and the myristoyl group is magenta. Similar to wild-type and LLL, the loop between β -sheets three and four is the longest contiguous stretch of unassigned residues, but there are also assignments missing in secondary structure, for example V36, and H75 at the top of the barrel. This could be the result of overlapped peaks, the loss of NOE resonances due to intermediate exchange, or just bad luck.

as chemical shift perturbation (CSP), chemical shift mapping, or in specific cases involving ligand-binding, complexation-induced changes in chemical shifts (CIS). Regardless of the moniker, CSPs can quantify chemical shift changes (which can be construed as reflecting structural changes) in response to titration of a ligand (in more conventional experiments), or the changes in local environment arising from point mutations or post-translational modification [177]. The crux of the methodology is the chemical shift, so perturbations thereof are sensitive to the ring currents, backbone bond angles, and hydrogen-bonds that contribute to amide proton and nitrogen chemical shifts.

With refreshed assignments for wild-type at pH 6.2, existing assignments for other pH values and the non-myristoylated variant, and similar data for I85L and LLL, there are numerous dimensions of comparisons for understanding how mutations in the core modulate the structural changes associated with myristoylation and proton binding. First published elsewhere [8], adding the myristoyl moiety to wild-type (at pH 8.7) induced minor amide proton chemical shift changes throughout the protein, however some of the most significant (> 0.75 ppm) are in residues that pack against the sequestered fatty acid, F6, I85, I93, F113. These same residues experienced little change in NOEs, a readout of through-space interactions, compared to the non-myristoylated variant. Thus, myristoyl sequestration does not appear to be associated with any major conformational restructuring. Results from the chemical shift perturbation by myristoylation at pH 6.2 support these findings, that aside from the protein termini, some of the greatest chemical shift changes from myristoylation occur for myristoyl-adjacent residues the bottom of the β -barrel and top of the hairpins, like I85, H89, I93, and in the loop comprising D70 - S72 (Figure 4.8). The trend of perturbation does not present as strongly pH-dependent (though the magnitude of the shift changes have a minor pH dependence), which supports that switching is fast on the chemical shift timescale and that the myristoyl group sufficiently samples the sequestered state at low pH to continue perturbing the chemical shifts of core residues.

Mutations modulate the effect myristoylation has on the chemical shifts of these residues. The average chemical shift perturbation of WT is greater than either of I85L and LLL (Upper panel of Figure 4.9). Residues of particular interest, like L85, I/L93, and others such as D70-S72, no longer experience large perturbations; for wild-type, I85 shows a combined chemical shift perturbation, d , of 1.07 ppm and I93 of 1.27 ppm. In LLL and I85L, the Leu in position 85 is perturbed by 0.18 ppm and 0.29 ppm, respectively, and in position 93 the perturbation is 0.05 ppm and 0.36 ppm—much smaller by comparison. This could speak to fewer interactions between these areas of the protein and the myristoyl group, which is consistent with our hypotheses for why LLL and I85L have broken switches — that the geometry of LLL's core and channel for the myristoyl group is more accommodating resulting in fewer interactions, or that the increased space allows the myristoyl group to adopt a greater number of conformations and the chemical shifts represent the average of more disparate conformations whose chemical shifts could approach those of the non-myristoylated form. I85L, on the other hand, may have a myristoyl pocket that is less sterically accommodating to the myristoyl

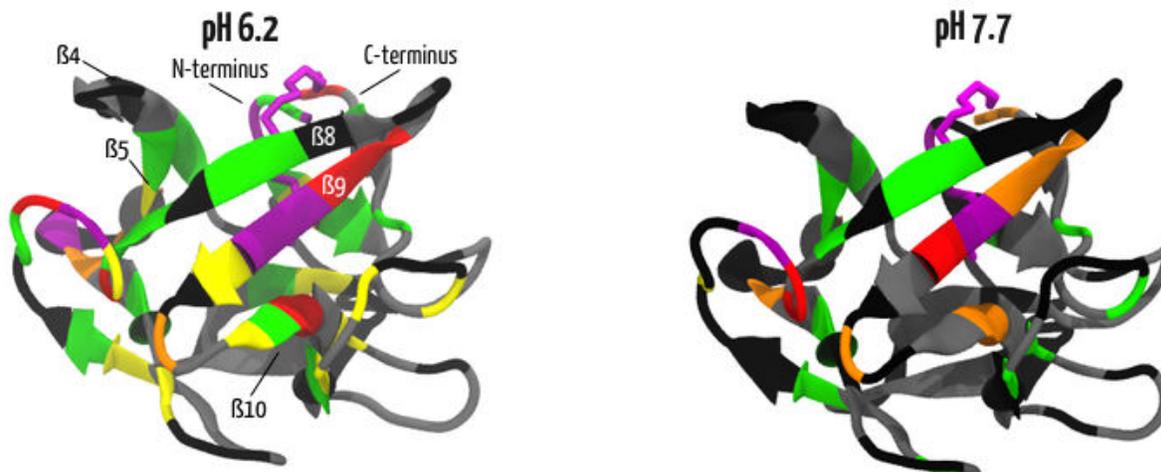


Figure 4.8: Chemical shift perturbations upon myristoylation of wild-type hisactophilin at pH 6.2 (left panel) and pH 7.7 (right panel) mapped onto the structure. Chemical shift perturbations were calculated for assigned residues shared between the myristoylated and non-myristoylated forms. Changes in the nitrogen and proton dimensions are reflected in this figure— the distance moved by a chemical shift was calculated by $d = \sqrt{\frac{\delta_H^2 + \alpha \cdot \delta_N^2}{2}}$, where δ represent the chemical shift of each nucleus, and α is a scaling factor so that the inherently large chemical shifts of amide nitrogen do not dominate the equation. 0.14 was used as a scaling factor [177], though there is not a clear consensus as to the ‘correct’ value [221,264–267]. Black residues are missing assignments in either the myristoylation or non-myristoylated forms, and grey, green, yellow, orange, and red, represent a relative scale that increments by a fifth the range (excluding outliers) of perturbations for the variant. Grey show very little perturbation, red are a significant perturbation and purple are off scale.

moiety, encouraging it to sample the accessible state where chemical shifts at the bottom of the β -barrel will look non-myristoyl-like (Lower panels in Figure 4.9). The myristoyl group may still try to sequester itself and those fleeting interactions may be the cause of chemical shift perturbations that are slightly larger than for LLL (while still much less than wild-type). Interestingly, the chemical shifts toward the top of the barrel in LLL and I85L are more perturbed than they are for wild-type. The greater perturbation could indicate that the myristoyl moiety has a greater impact on those residues in the mutants than for wild-type because of the broken switch. Particularly in I85L, the myristoyl group may more frequently sample the accessible state where it could interact with residues showing the large perturbations. The chemical shift perturbations support aspects of the broken switch model in Chapters 2 & 3, that with LLL the myristoyl group could favour the sequestered state while I85L favours the accessible state. These CHESCA-like results may helpfully and semi-independently be corroborated by measuring the temperature dependence of chemical shifts and enable an argument for the energetic basis of the pH-dependent myristoyl switch and highlight the residues most attuned to the link between proton binding and the myristoyl group.

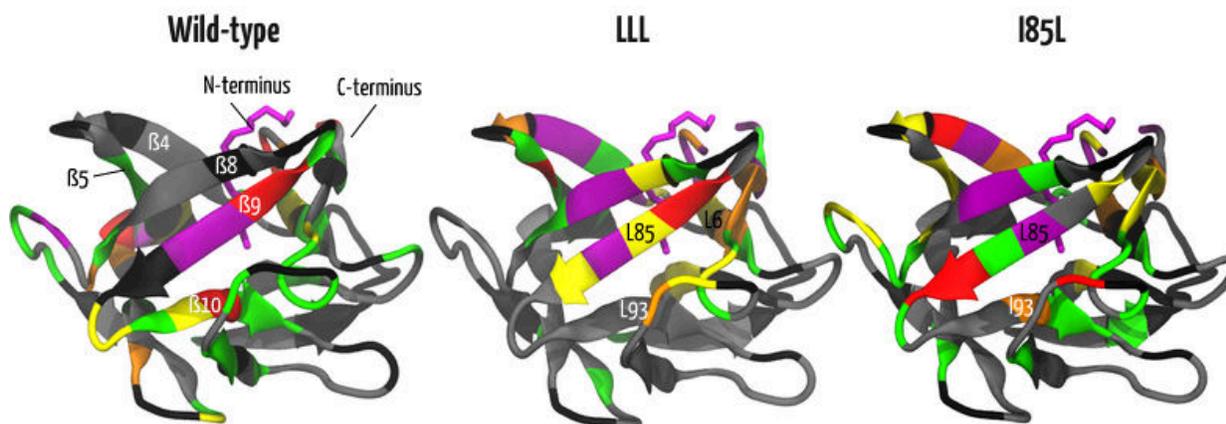
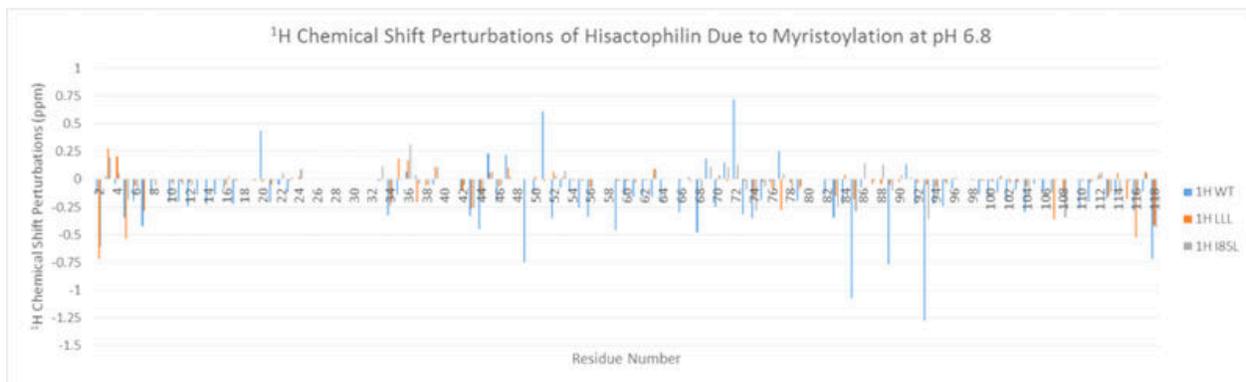


Figure 4.9: Chemical shift perturbations upon myristoylation at pH 6.8 of hisactophilin variants. (Upper panel) Bar plot to show the sizes of the amide proton chemical shift perturbations upon myristoylation for wild-type (blue), LLL (orange), and I85L (grey). Of note is that the proton chemical shifts of wild-type are more sensitive to the myristoyl group than either of the mutants. (Lower panels) Chemical shift perturbations upon myristoylation at pH 6.8 for the three variants studied. Black residues are missing assignments in either the myristoylation and/or non-myristoylated forms, and grey, green, yellow, orange, and red, represent a relative scale that increments by one fifth the range (excluding statistical outliers) of perturbations for the variant. Grey show very little perturbation, red are a significant perturbation and purple are off scale. The calculated perturbations are the euclidian distance moved of a residue in both the proton and nitrogen dimensions, given by $\sqrt{\delta^2 + \alpha^2}$, where δ represent the chemical shift of each nucleus, and α is a scaling factor set to 0.14 [177].

4.4.2 Amide Proton Temperature Coefficients

Amide proton temperature coefficients (or just ‘temperature coefficients, going forward) are a simple yet robust way to interrogate the local stability of the backbone of hisactophilin for which we have, or could realistically obtain, assignments. Because of our extensive characterization of hisactophilin in its acylated and non-acylated forms (and numerous mutants thereof) by bulk measurements like global chemical equilibrium denaturation (Chapter 2 & 3, and [5]), temperature coefficients are a tantalizing complementary experiment to ascribe changes in global stability to individual amino acids or structural elements. More than just rationalizing the global stability, however, the goal is to use temperature coefficients to elucidate the mechanism of the pH-dependent myristoyl switch by identifying where the protein might be strategically destabilized and map the communication network between the putative proton binding site(s) and the myristoyl pocket.

In principle, using the systematic perturbation of chemical shifts to map allostery in a protein is spiritually similar to the CHESCA methodology developed by the Melacini group [186]. CHESCA, or CHEmical Shift Covariance Analysis, is built on the notion that accurately measured variations in chemical shifts are sufficiently sensitive to highlight functionally relevant conformational changes that are relevant to allosteric communication, and that perturbing a component of a communication pathway will be sensed by the chemical shifts of other residues of the same communication pathway. One of the first steps of implementing CHESCA is to settle on how to perturb the protein; options include carefully considered mutations to the protein or chemically modifying the allosteric effector to attenuate its effectiveness to create a range of agonists and antagonists. Pairs of residues affected by the same exchange processes in response to this perturbation are, in an ideal case, linearly correlated and subsequently mapped onto a correlation matrix. The results of this correlation and mapping are subject to hierarchical clustering algorithms and verified by singular value decomposition [186,268,269]. CHESCA has been implemented to identify allostery in the exchange protein activated by cAMP (EPAC) [186], allosteric control of tryptophan synthase[270,271], open to closed transitions of PKA [272], ligand binding [273], and allostery through flexible linkers [143]. However, the methodology leans on having a library of available perturbations (mutations/agonists/antagonists) with well characterized functional profiles which may not be so trivial a thing. Temperature coefficient analysis requires only native ligands/‘perturbations’ and offers a straightforward alternative to (or corroborator of) CHESCA to verify models of allostery but with additional energetic arguments as well. Subsequent statistical and/or covariance analysis on temperature coefficients is not yet as sophisticated, however that may soon change...

Accessing the energetics of the myristoyl switch, however, requires that we consider not just how the myristoyl group interacts with the protein, but also how it is affected by pH changes. We quantified this interaction by measuring the change in stability on account of myristoylation on one side of the switch, say when the myristoyl group is accessible at pH 7.7, and again on the other side of the switch at pH 6.2. This

analysis is a thermodynamic cycle analogous to that used in double mutant cycles to assess pairwise interactions between residues [5,6,8,73], but here we use the temperature coefficients as a measure of local stability rather than the free energy of global unfolding ΔG_{D-N} (Left panel of Figure 1.1). In essence, analysis of the switch energetics comprises measuring the temperature dependence of amide proton chemical shifts in the myristoylated and non-myristoylated forms, each at high and low pH (pH 7.7 and 6.2, respectively) to construct each side of the thermodynamic cycle. Each difference in temperature coefficient upon myristoylation or change of pH reflects how the local stability of each amide is impacted. Together, the cycle describes the switch energy, the interaction energy and coupling between the amide, proton binding, and the myristoyl group—those residues that show an interaction are likely to contribute to the mechanism of the switch.

4.4.2.1 Global energetics and WT Temperature coefficients

Of the available assignments for myristoylated wild-type hisactophilin, 95 backbone probes were amenable to being tracked for most, if not the entire temperature range to calculate temperature coefficients. The temperature dependent behaviour of amides is illustrated in Figure 4.3, and the trends of the movement of amides is varied; some move relatively far upfield, corresponding to a very negative temperature coefficient; others move hardly at all, and their temperature coefficients have a smaller magnitude. Work in peptides first ascribed temperature coefficients to whether or not that particular probe was hydrogen bonding to its neighbour [226], however more contemporary studies have broadened the scope of temperature coefficients to reporting on local structural stability, or more specifically, the temperature dependent loss of structure [150,183].

WT Temperature Coefficients

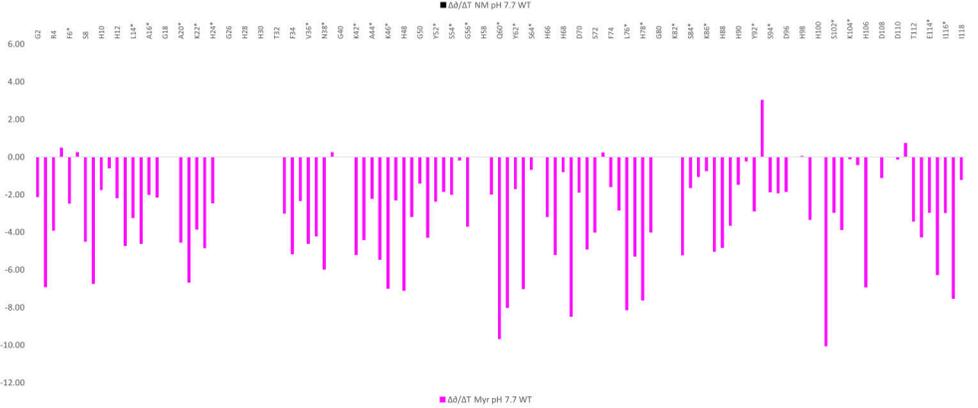
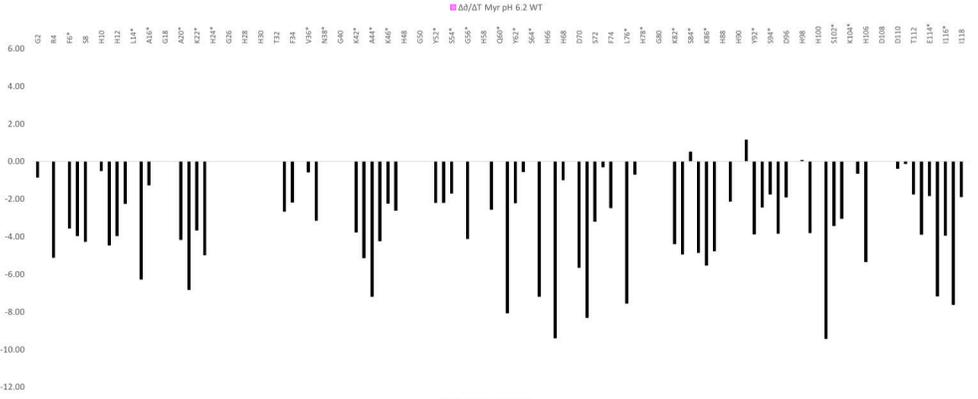
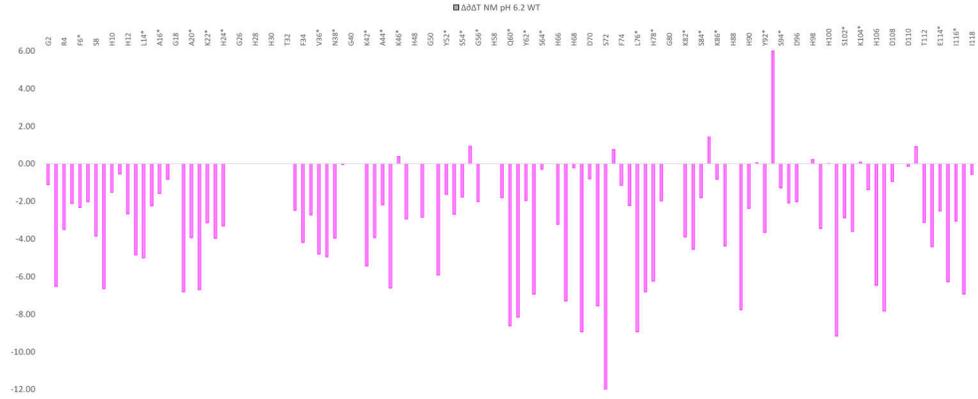
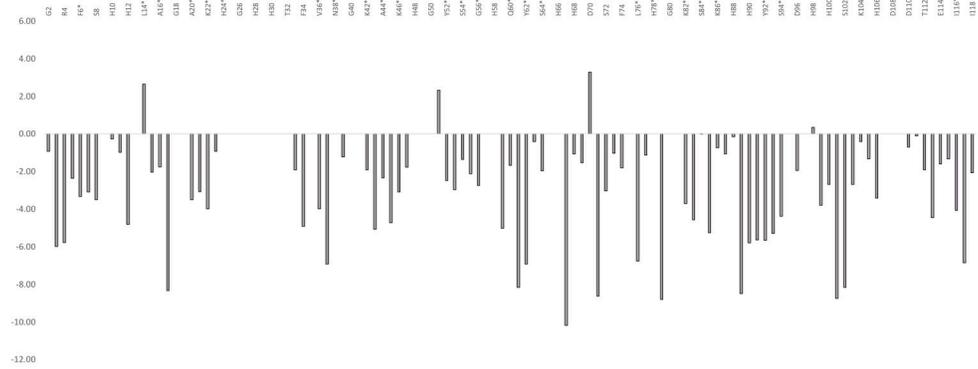


Figure 4.10: Bar plots of the amide proton temperature coefficients for wild-type hisactophilin at pH 6.2: non-myristoylated (top panel) and myristoylated (second panel). Temperature coefficients are also shown for wild-type at pH 7.7: non-myristoylated (third panel) and myristoylated (bottom). These represent 2D ^1H , ^{15}N HSQC spectra measured over the temperature range of 10 - 50 °C at 2.5 °C increments. The movement of individual peaks were tracked on spectra referenced to DSS and fit, versus temperature, to a least squares regression line. The slope of that line is the temperature coefficient. The values correspond to those mapped onto the backbone in Figure 4.11. Residue labels with an asterisk (*) form β -structures.

The temperature coefficients for non-myristoylated wild-type (WT NM) at pH 7.7 are shown in the upper panel of Figure 4.10; the values are generally within the typical range of -11 - +1 ppb/K [232] with an average value of -3.49 ppb/K. In principle, averaging the local structural stability across the whole of the protein should reflect its global stability, in this case 10.20 kcal/mol (Table 4.2 and [8]). Human SOD1, another protein extensively characterized using temperature coefficients, is a dimeric metalloprotein whose thermodynamic stability is on the order of 33 kcal/mol and its average temperature coefficient reflects the global stability by being less negative than hisactophilin at -3.2 ppb/K [150]. Comparing local and global stabilities of non-myristoylated wild-type and LLL hisactophilin with the myristoylated form at pH 7.7 and 6.2 (Table 4.2), a similar trend holds, where variants with greater thermodynamic stability have less negative average temperature coefficients consistent greater local structural stability. This trend, however, is not a rule as there are exceptions when considering all variants and experimental conditions. That said, the r^2 correlation of equilibrium stability and average temperature coefficient is 0.35 so there is a weak linear relationship. That the relationship is imperfect could arise from assignments that are not randomly dispersed through the protein, i.e. the loop comprising residues 25-33 is wholly missing, or the slight upfield shift of temperature coefficients because of hisactophilin's many aromatic residues (five phenylalanines, three tyrosines, and thirty-one histidines) contributing ring current effects to chemical shifts.

Looking just at residues that hydrogen/deuterium amide exchange experiments indicated only lose structure upon global unfolding [73,88], the correlation is somewhat stronger. As predicted by global energetics, myristoylated hisactophilin has a less negative average temperature coefficient (and thus a greater structural stability) than the non-myristoylated form. Though this trend weakens with mutants LLL and I85L, it is not unexpected given what we know about the state of their thermodynamic switches. For example, myristoylated I85L is posited to access the sequestered state less (the myristoyl group buried in the core) than wild-type, and this is further supported by only a small difference in average temperature coefficient with the non-myristoylated form. The average temperature coefficients for myristoylated and non-myristoylated LLL, the other broken switch mutant, are also similar, so the average temperature coefficients may be recapitulating the changing interactions between the myristoyl group and the core that are captured by global thermodynamics measurements [8].

4.4.2.2 WT amide temperature coefficients

Table 4.2: ΔG_{U-F} and average temperature coefficients of hisactophilin variants. The ΔG_{U-F} of unfolding is shown, as calculated using the binomial extrapolation method (Chapter 3) for myristoylated and non-myristoylated wild-type, LLL (F6L/I85L/I93L), and I85L at pH 6.2 (below pK_{switch} , accessible state in wild-type) and 7.7 (above pK_{switch} , the sequestered state). The average temperature coefficients are a mean value excluding residues without assignments or undefined temperature dependences. The average of the global exchangers is the mean of residues that hydrogen deuterium exchange identified as exchanging upon global unfolding [88]: 44, 45, 46, 54, 62, 74, 85, 86, 93, 94, 113, 114, shown in Figure 4.35). Appendix A is complete list of temperature coefficients for the studied variants.

	pH 7.7			pH 6.2		
	Equilibrium ΔG_{D-N} (kcal/mol)	Average Temperature Coefficient (ppb/K)	Average Temperature Coefficient of global exchangers (ppb/K)	Equilibrium ΔG_{D-N} (kcal/mol)	Average Temperature Coefficient (ppb/K)	Average Temperature Coefficient of global exchangers (ppb/K)
WT Myristoylated	9.33	-3.4	-2.3	5.27	-3.3	-1.3
WT Non-myristoylated	7.58	-3.5	-3.4	4.29	-3.3	-3.5
LLL Myristoylated	9.20	-3.4	-2.8	5.41	-3.6	-3.0
LLL Non-myristoylated	6.56	-3.7	-2.7	2.55	-3.8	-2.8
I85L Myristoylated	8.70	-3.5	-2.8	5.35	-3.3	-3.0
I85L Non-myristoylated	8.00	-3.5	-2.5	4.47	-3.6	-2.7

In non-myristoylated hisactophilin at pH 7.7, the measured temperature coefficients range between -9.43 ppb/K for V101, and +1.15 ppb/K for H91 (Figure 4.10 and Appendix A). Positive temperature coefficients have been reported for other protein and can arise from ring current contributions to chemical shifts from adjacent aromatic residues. Because ring current effects are orientation dependent, temperature coefficients can report very loudly on any temperature-induced changes to the local environment. Thus, the temperature coefficients for hisactophilin provide a complex and nuanced view of local stabilities, as observed in other proteins [150,231,232]. The nuance and complexity is demonstrated by a structurally destabilized residue (e.g. H75 in red in strand β 8, upper right panel of Figure 4.11) adjacent to one that is in a stable regime (L76 in cyan). An explanation for this could involve the positioning of the residues in the context of the whole protein, the nature and flexibility of the side chains, and inter- versus intramolecular hydrogen-bonding. The N- and C-termini, for example, tend to have relatively low structural stability in accordance with their likely relatively high flexibility, amides in the middle of β -strands at the bottom of β -barrel have more positive temperature coefficients and thus appear structurally stable, the residues at the edges of a strand can be less stable than their centrally-located counterparts, and residues in loops are more variable in their structural stabilities.

The sometimes-alternating pattern of temperature coefficients along edge β -strands illustrates some of the origins of temperature coefficients. Structural stability [150] encompasses varied mechanisms that can account for temperature coefficients; changes in inter- and intramolecular hydrogen bond distances or angles [180,190,234,274], movement of aromatic residues relative to a probe (or vice versa), proximity of charged residues, peptide bond anisotropies, or fast chemical exchange with an alternate conformation [190,275]. Traditionally, temperature coefficients were invoked mainly as a reporter of intramolecular hydrogen bonding using an empirically determined cut-off of \sim -4.6 ppb/K, residues with more positive values had an 85% chance of having an intramolecular hydrogen bond, those more negative likely hydrogen-bond with solvent. However, its performance as a predictor for the hydrogen-bonding of any single probe is not robust enough to be fully diagnostic, and it could be in the cases where the simple empirical cutoff fails might be the most interesting [235].

If temperature coefficients are used as an indicator of intramolecular hydrogen bonding in hisactophilin, whole stretches of residues would be mischaracterized. For example, the existing solution structure indicates residues 10, 11, 12, 33, 34, 72, 73, 74, 98, 99, 100, 110, 111, 112, 113 form hydrogen-bonds with solvent (as expected for loop residues) but temperature coefficients suggest the formation of intramolecular hydrogen bonds. On the other hand, edge-strands β 4 and β 8 have an alternating pattern of more positive and more negative temperature coefficients, the more negative temperature coefficients might be explained by the expected correlation with hydrogen-bonds to bulk solvent. However, destabilized residues in the middle of a β sheet, like A44 (β 5) are not so readily explained only by differences in hydrogen-bonding. The temperature

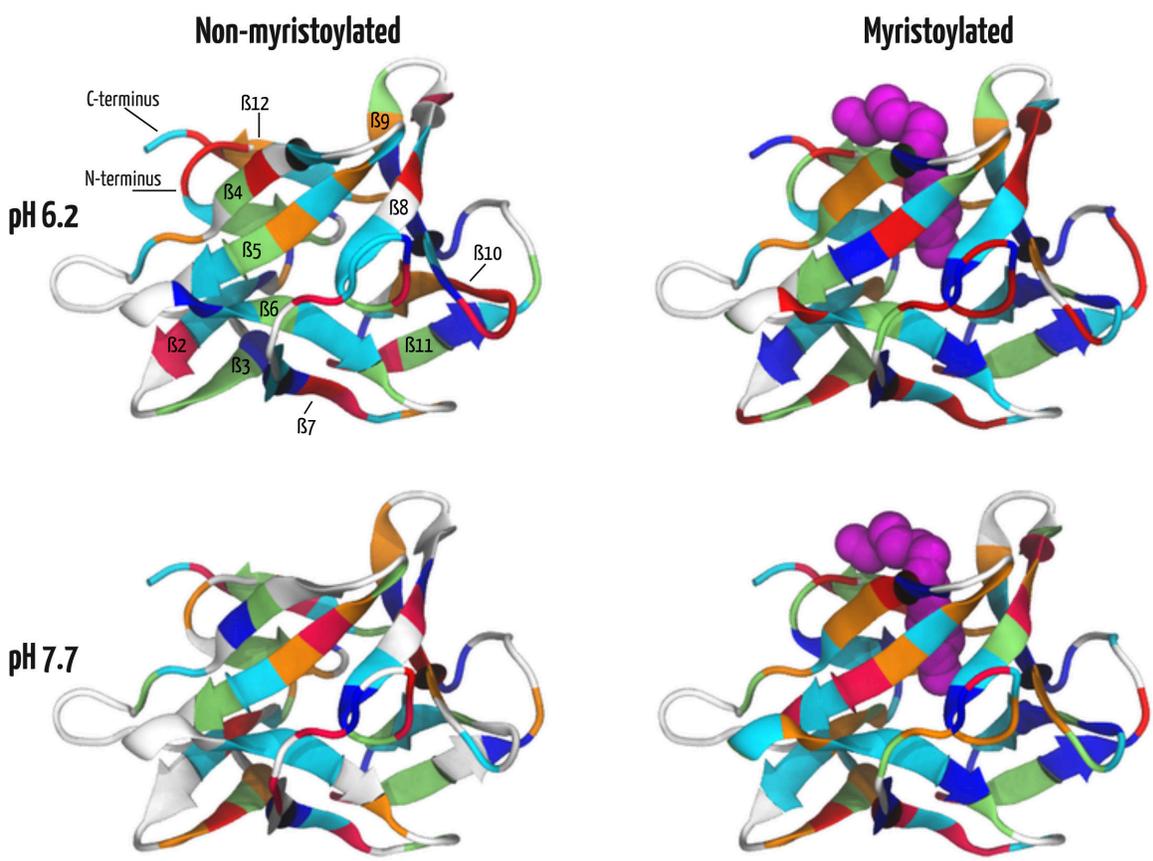


Figure 4.11: Amide proton temperature coefficients for wild-type hisactophilin mapped onto ribbon representation of the protein structure. Upper left and Lower left are non-myristoylated hisactophilin at pH 6.2 and 7.7, respectively. Upper right and lower right are myristoylated hisactophilin at pH 6.2 and 7.7, respectively. The myristoyl group, when applicable, is depicted as magenta spheres. Residues without assignments are white, dark red are temperature coefficients more negative than -7 ppb/K, red are between -5.5 and -7 ppb/K, orange are between -4 and -5.5 ppb/K, green are between -2.5 and -4 ppb/K, cyan are between -1 and -2.5 ppb/K, and blue are temperature coefficients more positive than -1 ppb/K. More negative temperature coefficients indicate less local structural stability.

coefficient of A44 is -7.2 ppb/K (non-myristoylated at pH 7.7) despite a hydrogen bond to the amide proton. So, although hydrogen-bonding may be a substantial determinant of temperature coefficients, within hisactophilin, there are many residues for which cut-off might be wrong. The effects of changes in the relative orientation of ring currents [232], charged residues [183], or backbone angles may be responsible for some of these discrepancies. As a result of the possibility of the mixed mechanisms that contribute to temperature coefficients, they may be best considered to report on the umbrella of structural stability, which is in line with studies on SOD1 [150] and the B1 domain of protein G [183].

For myristoylated hisactophilin at pH 7.7, the temperature coefficients range from -10.05 ppb/K to +3.02 ppb/K and follow the same general trends as non-myristoylated concerning the location of stable and unstable residues, though there is an increasing number of exceptions where the middle of a β -barrel strand is unstable, for example strands 4, 5, and 8 (see the lower right panel of Figure 4.11). Lowering the pH affects numerous temperature coefficients (they range from -12.54 ppb/K to +6.00 ppb/K), however the trends for the localization of the relatively less stable and more stable residues are similar to pH 7.7. Some residues opposite to the N- and C-termini on strands $\beta 8$, $\beta 9$, $\beta 10$ affected by the pH change are more stable at pH 6.2 than at pH 7.7 (orange residues in the right panel of Figure 4.12). Notably, this may indicate that this stretch of residues paradoxically destabilized at high pH are responsible for slightly raising the energy and destabilizing the myristoyl-sequestered state.

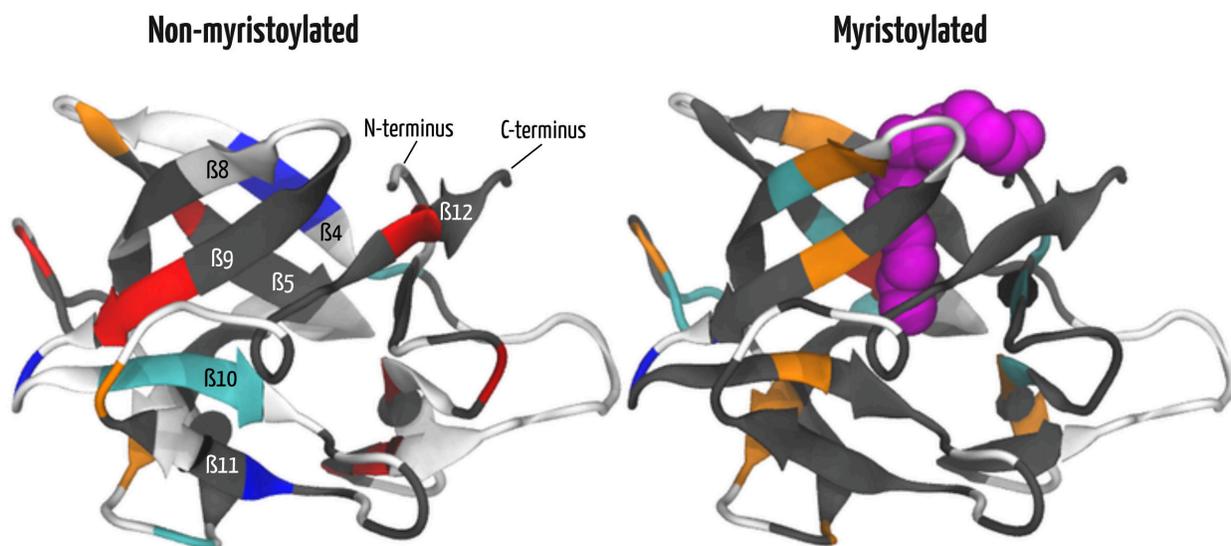


Figure 4.12: Change in temperature coefficients induced by the pH change from 6.2 to 7.7 mapped onto the hisactophilin structure for non-myristoylated (left panel) and myristoylated (right panel) wild-type. The myristoyl group is in magenta. Residues shown in white are lacking an assignment for that form at either pH 6.2 and/or 7.7. The difference in temperature coefficient was calculated as the coefficients at pH 7.7 minus the coefficients at pH 6.2. Positive values indicate that a residue has greater structural stability at pH 7.7 than pH 6.2, and negative values indicate that a residue has greater structural stability at pH 6.2 than pH 7.7. Grey residues show a small change in temperature coefficient of between -1 and 1 ppb/K, orange residues have a change of temperature coefficient of between -1 and -3 ppb/K, and red residues had a change in temperature coefficient more negative than -3 ppb/K. The temperature coefficient of cyan residues increased by between 1 and 3 ppb/K, and blue residues increased by more than 3 ppb/K. Negative values indicate less stability at pH 7.7 and positive values are more stable at pH 7.7. Particularly interesting are the orange residues on strands $\beta 8$, $\beta 9$, and $\beta 10$ on the myristoylated form that are destabilized at pH 7.7—these strands are the location of some mutations of interest: $\beta 8$, L76; $\beta 9$, I85; $\beta 10$, I93; and the $\beta 9/\beta 10$ loop, H90.

WT Temperature Coefficients Differences

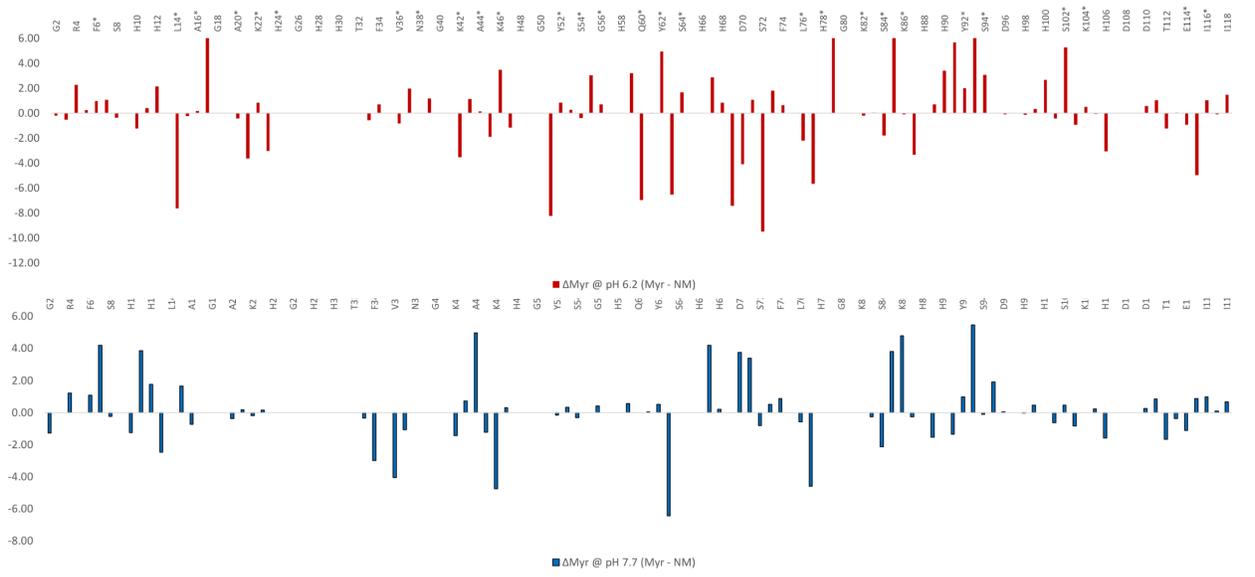
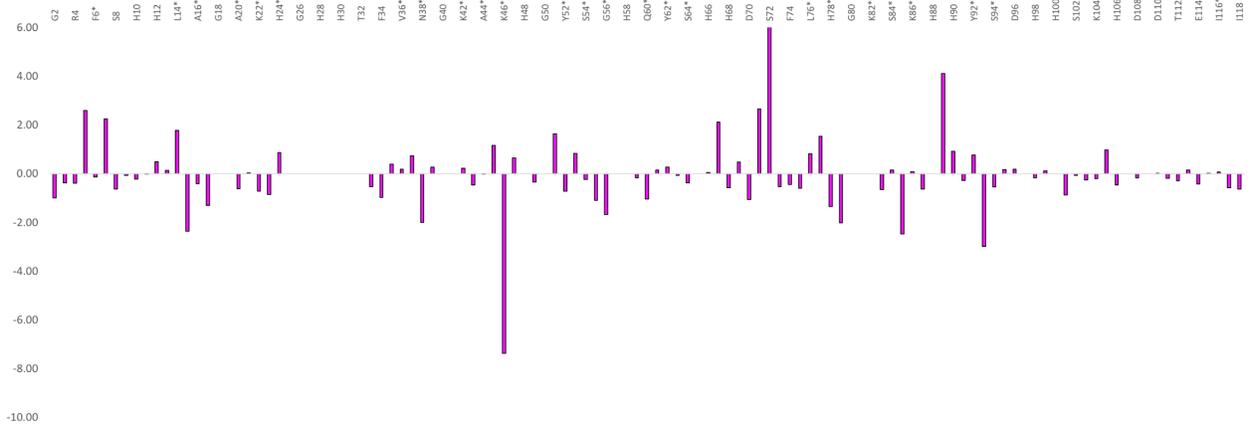
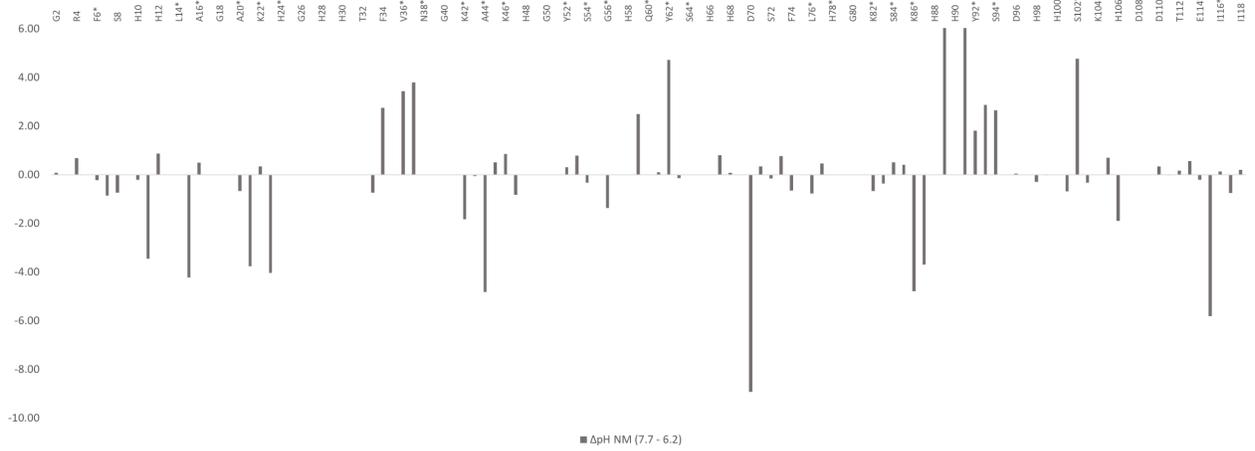


Figure 4.13: Bar chart of the temperature coefficient differences corresponding to each of the edges in the thermodynamic cycle in Figure 3.2 for wild-type hisactophilin. (Top & second panels) Difference in temperature coefficients upon pH change for the non-myristoylated (top) and myristoylated forms (second) by subtracting pH 6.2 from 7.7. The values are mapped on hisactophilin's structure in Figure 4.12. (Third & bottom panels) Difference in temperature coefficients upon myristoylation (myr - non-myr) at pH 6.2 (third) and pH 7.7 (bottom), as in Figure 4.14.

Comparing the temperature dependence data at two pH values for the myristoylated form does not allow us to distinguish whether these energetic changes arise from switching or just changes in ionization of the many histidine residues. In non-myristoylated wild-type at pH 7.7, residue I85, which packs against the myristoyl group has a temperature coefficient of -1.05 ppb/K; this value falls to -4.85 ppb/K when the myristoyl group is removed, an apparent decrease in structural stability. The trend is similar at pH 6.2 (-5.26 ppb/K to 1.43 ppb/K upon myristoylation, see text in the supplementary). The decrease likely reflects hydrogen bonding to water molecules that now fill the central cavity. Other residues in the middle of strands, like L45 in $\beta 5$, also have amide protons that might interact with buried solvent molecules or other hydrogen-bond acceptors. This is true for the non-myristoylated variant at pH 6.2 as well, though it is expectedly even more structurally destabilized than at the higher pH.

Considering the range of temperature coefficients for the non-myristoylated and myristoylated variants at pH 7.7 and 6.2 and considering that we know myristoylation increases global stability [8] and that decreasing pH decreases global stability [8,87], it is consistent that myristoylated and non myristoylated variants have more negative (destabilized) temperature coefficients at pH 6.2 than at pH 7.7. While many residues decrease in stability when pH is shifted to 6.2, those that increase in stability could have functional importance.

4.4.2.3 Myristoylation significantly impacts temperature coefficients

Uncovering which residues experience the strain that drives switching can be achieved by finding those that couple the myristoyl group to the site(s) of proton binding. With the non-myristoylated variant as a reference point at each pH, comparing the myristoylated variant highlights how interactions with the myristoyl group alter the local structural stability of the residues (Figure 4.14, bar chart in Figure 4.13). First, the straightforward expectation informed by global equilibrium thermodynamics is that the stability increases conferred by myristoylation when it is sequestered will be reflected by temperature coefficients that are less negative. Deviation from this behaviour may highlight residues that are strategically strained for function.

It is worth noting that analyzing the differences in temperature coefficients circumvents some of the caveats of analyzing temperature coefficients directly by allowing us to cancel some of the terms that contribute to chemical shift, for example the influence of the amino acid neighbours in the $i+1$ and $i-1$ positions, the surface area, and electric field effects (which should be practically invariant between the myristoylated and non-myristoylated forms at constant pH[180]). This allows us to isolate the non-covalent structural determinants of chemical shifts and temperature coefficients that are likely to vary most strongly when we perturb

the system by mutation, myristoylation, or pH.

Myristoyl-induced changes to temperature coefficients at pH 7.7 (right panel in Figure 4.14 and Figure 4.13) show negligible changes for 39 (grey) of 70 probes (NB the average of the negligibly affected residues is slightly stabilizing), which is to say that when the myristoyl group is sequestered in hisactophilin's core, those residues' stability is not impacted by interactions with the myristoyl moiety. Note, this analysis is contingent on having assignments for both myristoylated and non-myristoylated variants at a given pH, so, there are fewer probes informing this analysis compared to the individual proteins in Figure 4.11 because of a mismatch in the assignments between forms. Fourteen residues displayed a change in temperature coefficient consistent with increased structural stability upon myristoylation: R4, F6, K7, G11, H12, S15, A44, L67, D70, H71, I85, K86, I93, and A95 (cool colours, right panel of Figure 4.14). Many of these residues are located in the lower β -barrel and upper hairpin—approximately the depth of the terminal methyl of the myristoyl group in the sequestered state [8,73], and thus could be the result of favourable interactions of core with the myristoyl group. On the other hand, 17 residues are destabilized relative to the non-myristoylated variant (warm colours in Figure 4.14), which suggests unfavourable interactions induced by the myristoyl group. This destabilization informs the hypothesis that local, selective destabilization may be part of the switching mechanism, and it will have more merit if the extent of (de)stabilization is different at pH 6.2.

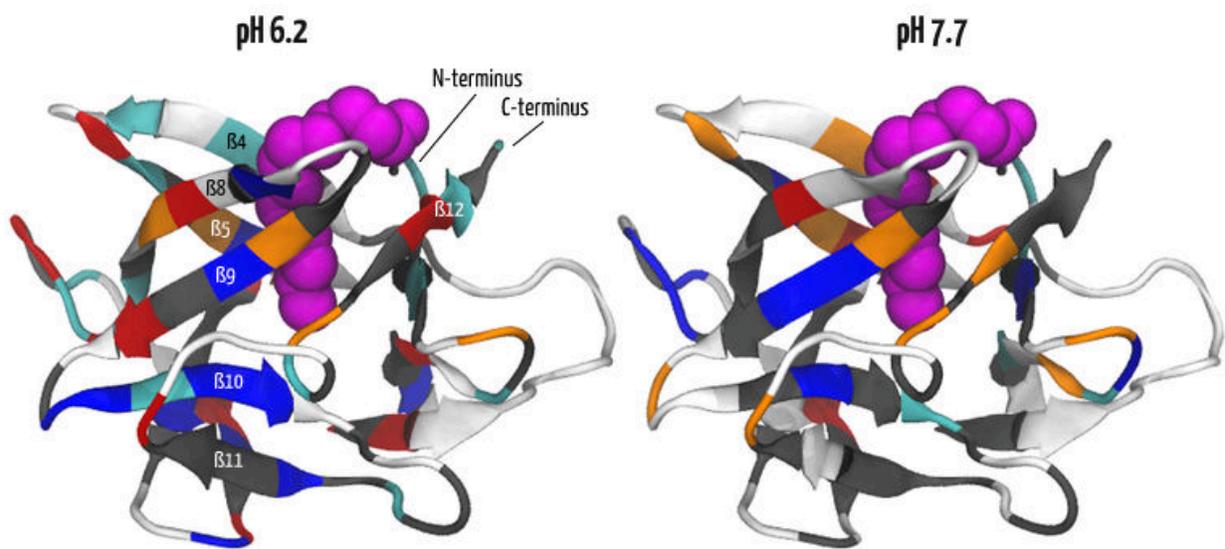


Figure 4.14: Change in temperature coefficients caused by myristoylation of wild-type at pH 6.2 (left panel) and pH 7.7 (right panel) mapped onto the structure of hisactophilin. The myristoyl group is in magenta. Differences in temperature coefficients were calculated by subtracting the values obtained for non-myristoylated from myristoylated. Residues in white are residues lacking assignments in either form of the protein for each pH. Grey residues are residues whose temperature coefficients changed by less than 1 ppb/K in either direction. Orange and red residues' temperature coefficients changed by between -1 and -3 ppb/K, or by more than -3 ppb/K, respectively. The temperature coefficients of the residues coloured cyan or blue changed between 1 and 3 ppb/K, or by more than 3 ppb/K, respectively. A positive value (cool colours) mean that myristoylation induces greater structural stability at that position, and negative values (warm colours) suggest that myristoylation decreases structural stability.

At pH 6.2, differences in temperature coefficient upon myristoylated are determined for 83 residues, and 36 of them show negligible changes at low pH (grey residues, left panel in Figure 4.14). Although we expect the myristoyl group to be in the accessible state and make fewer interactions with the protein at pH 6.2 than at pH 7.7 [73], equilibrium stability measurements show that at pH 6.2, the myristoylated form is still ~ 1 kcal/mol more stable than the non-myristoylated form [8]. The 27 residues that show increased conformational stability upon myristoylation may reflect that residual stabilization (cyan and blue residues in the left panel of Figure 4.14). The distribution of such stabilized residues is similar to pH 7.7, but with more of them appearing at the top of the trefoil barrel, which supports the model that in the accessible state—the myristoyl moiety climbs the barrel and interacts more favourably with residues around hisactophilin’s opening (the accessible state in Figure 1.4). Concurrently, 20 residues show decreased structural stability upon myristoylation at pH 6.2 — many are localized in the hairpin triplets where actin is thought to bind and the residues might be involved in structural plasticity during binding.

4.4.2.4 Localizing strain and allostery with thermodynamic cycles

Mechanistic detail of hisactophilin’s pH-dependent switch from favouring the myristoyl-sequestered state to favouring the myristoyl-accessible state may be revealed by comparing myristoyl-induced changes in structural stability at the two pH values — $\Delta T\text{Coeff}_{\text{switch}}$, an analogous analysis to ΔG_{switch} (Figure 3.2) [5,8]. The upper left panel of Figure 4.15 maps these switch coefficients onto hisactophilin’s backbone (graph of values in Figure 4.17). The complex response to myristoylation and pH change, in which myristoylation stabilizes some residues more at high pH (blues) and others are more stabilized at low pH (reds), could be an integral mechanism of switching.

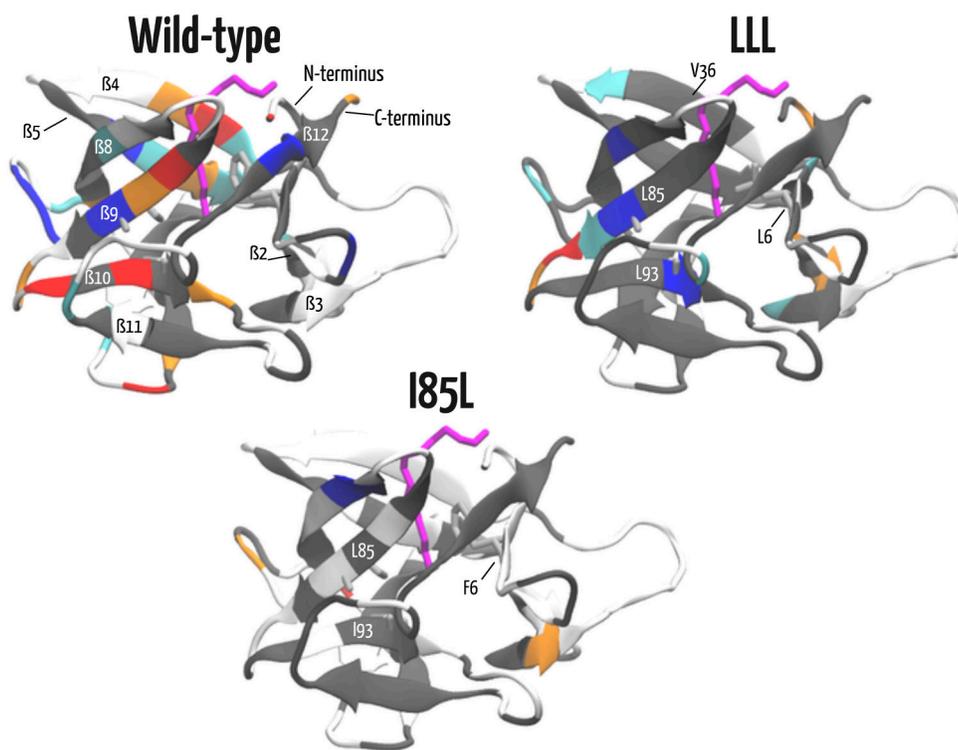


Figure 4.15: Switch energetics for hisactophilin determined from temperature coefficients by subtracting the temperature coefficients of myristoylated from non-myristoylated at pH 7.7, and subtracting from that the difference of temperature coefficients of myristoylated from non-myristoylated at pH 6.2. See the thermodynamic cycle in Figure 3.2: switch energetics is the top edge subtracted from the lower edge. The energetics are illustrated for wild-type (top left), LLL (top right), I85L (bottom). Residues shown in white are those without a full complement of assignments for all four states (myristoylated and non-myristoylated each at high pH and low pH). Grey residues are those whose $\Delta\text{TCoeff}_{\text{switch}}$ are between -1 and +1 ppb/K. Orange and red residues are those who $\Delta\text{TCoeff}_{\text{switch}}$ are between -1 and -3 ppb/K and beyond -3 ppb/K, respectively. Cyan and blue residues represent $\Delta\text{TCoeff}_{\text{switch}}$ on residues between 1 and 3 ppb/K, and beyond 3 ppb/K, respectively. Here negative values (warm colours) represent residues for which myristoylation is more stabilizing at pH 6.2 than at pH 7.7. Positive values (cool colours) are those that behave more predictably, where myristoylation confers greater structural stability at pH 7.7 than at pH 6.2. Of note is that for wild-type, there are many residues that are coupled to myristoylation and proton binding, particularly V36, I85, H89, H91, I93, and I118 (C-terminus). LLL has far fewer residues that exhibit coupling of the myristoyl and site(s) of protonation, and I85L has fewer still. The mutants also do not have warm coloured residues that might contribute to functional destabilization of residues in the sequestered state relative to the accessible state.

Switch Temperature Coefficients

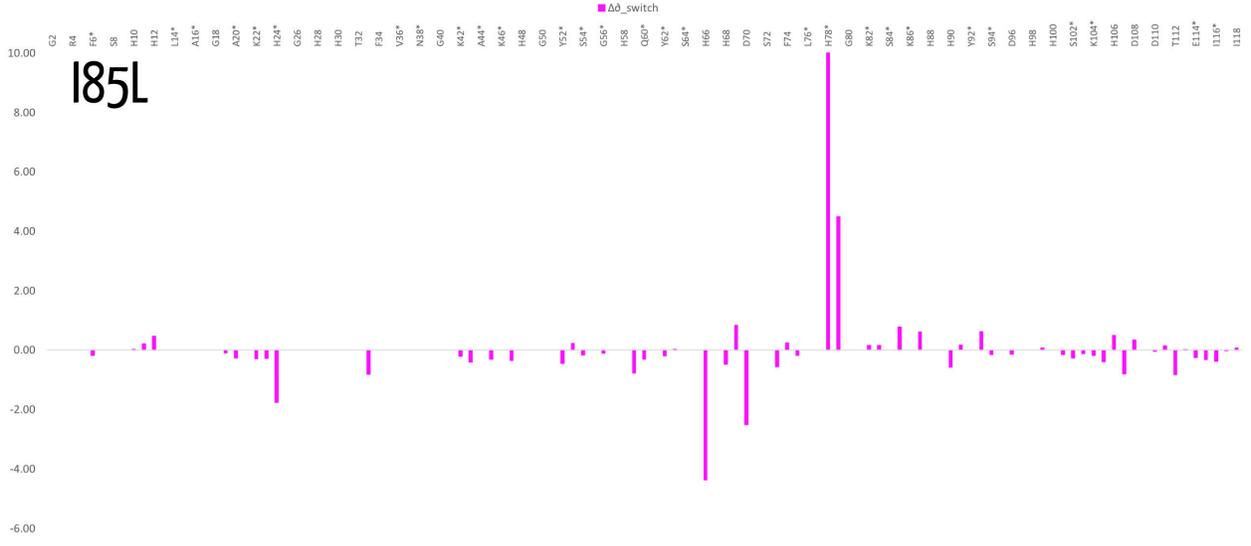
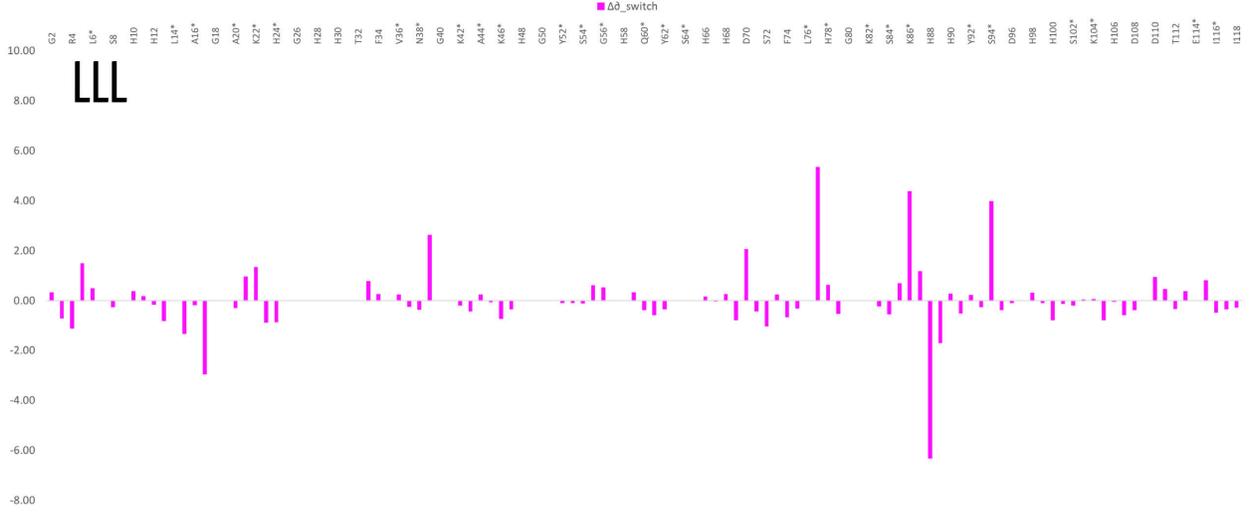
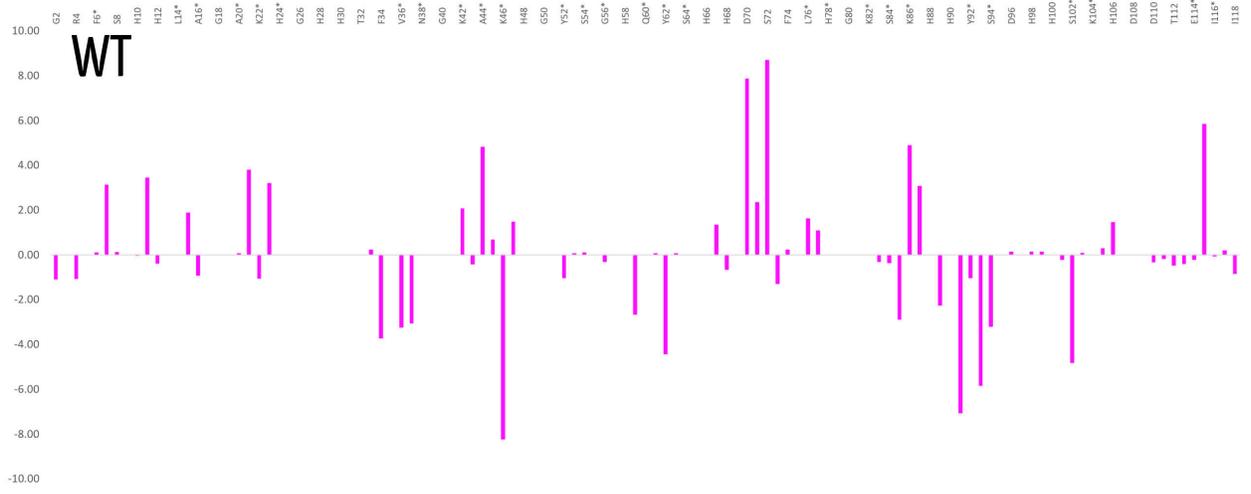


Figure 4.17: Bar graph showing the values of $\Delta\text{TCoeff}_{\text{switch}}$, calculated by subtracting the top horizontal edge from the bottom edge in the thermodynamic cycle shown in 3.2. Negative values are red/orange in Figure 4.15, which represent residues where coupling is favourable for switching to the accessible state. Positive values (blue/cyan in Figure 4.15) couple favourably to the sequestered state. Residues throughout the protein show signs of positive and negative coupling for Wild-Type (top panel). However, the $\Delta\text{TCoeff}_{\text{switch}}$ coupling in LLL (F6L/I85L/I93L; middle panel) and I85L (lower panel) is muted. Residues labelled with an asterisk (*) are part of β -structures.

The simple prediction informed by global stability measurements is that the protein, and thus its residues, should have greater structural stability at pH 7.7 than at pH 6.2. In the upper left panel of Figure 4.15, the residues with negative $\Delta\text{TCoeff}_{\text{switch}}$ in red and orange experience increased structural stability by myristoylation at low pH relative to high pH, and the blue/cyan coloured residues are those that are more stabilized at pH 7.7 than 6.2. Blue residues may represent unfavourable coupling energy that favours the sequestered state and the red residues represent a favourable coupling that encourages switching to the accessible state. For example, I93 is located in strand β 10 and points into the myristoyl pocket (a red residue in the upper left panel of Figure 4.15), shows more signs of strain when the myristoyl group favours the sequestered state (high pH) compared to low pH. When the myristoyl group switches to the accessible state, I93 is relatively stabilized because some of the strained interactions have been relieved. On a local level, the difference in strain between the sequestered and accessible states could be one of the drivers of the switch. The relative destabilization afforded by select residues in hisactophilin's sequestered state may serve to keep the thermodynamic balance between the sequestered and accessible states close enough that a small perturbation (i.e. proton binding) is sufficient to tip the preference from the sequestered to the accessible state.

Strain is sometimes invoked in the context of protein stability/function tradeoff, i.e. that functionally important residues are not evolutionarily selected for their contribution to thermodynamic stability, but for their functional role [155,158,276–278], which may be inherently destabilizing. Thus, strain is a thermodynamic penalty paid by the protein upon folding in exchange for function—catalysis or structural rearrangement—and the penalty can manifest in lower global stability of the native state. The increased unfolding rate upon myristoylation, despite increased global stability, suggests native state-strain is relieved in the folding transition state [8]. Furthermore, the dynamics of switching measured by NMR line shape analysis at pH 6.1 indicated that both states rapidly interconvert, which may further support that strain is present at pH 7.7. Sampling the sequestered state under presumably 'unfavourable conditions' (pH 6.2) speaks to the delicate balance of energetics that governs switching and this balance is likely responsible for its exquisite sensitivity.

Double mutant cycles were originally conceived to measure pairwise interactions between residues in proteins by measuring stability changes in response to individual and then combinations of mutations [6], and adaptation of the cycle here allows us to make similar measurements between the sites of proton binding and the myristoyl group [5]). Beyond strain and tension, we are also measuring the interaction and allostery between protonation and the myristoyl group [8]. All the residues with a non-zero (non grey/white) interaction

energy in Figures 4.15 and 4.17, have some interaction with the proton-binding site(s) and the myristoyl group, so they may illuminate the communication pathway that carries signals from proton-binding to the myristoyl pocket and flip the switch. In wild-type, the implicated residues are many and distributed widely throughout the protein (upper left panel of Figure 4.15), particularly interesting are residues that experience greater relative local structural stability when the myristoyl group is accessible (F34, V36, E37, K46, K59, Q60, Y62, S84, I85, H89, H91, I93, S94, and I118). Spatially clustered residues whose side chains could interact with one another are even more compelling candidates for forming an allosteric communication pathway, like S84 ($\beta 9$), I85 ($\beta 9$), H89 ($\beta 9\beta 10$), H91 ($\beta 9\beta 10$), and I93 ($\beta 10$) that are toward the bottom of the myristoyl binding pocket in trefoil 2, or F34 ($\beta 4$), V36 ($\beta 4$), E37($\beta 4$), and K46 ($\beta 5$) which are in trefoil 1 and nearer the top of the barrel.

The change in $pK_{a,apparent}$ (as previously determined by NMR titration experiments) for residues H91, L45, and I85 implicated them as candidates for allosteric communication [8], and some of the same residues are implicated by temperature coefficients too. That temperature coefficients incriminate residues based on tension that were singled out by other experiments is further evidence of the importance of numerous residues in allostery and the myristoyl switch.

If the tension that we observe in select residues in wild-type contributes to switching, then the switch should be broken if we remove the strain at that residue by mutation...

4.4.2.5 LLL amide temperature coefficients

The LLL mutant (F6L/I85L/I93L) was previously established to have a broken thermodynamic switch (Chapter 3) [5,73], meaning that the thermodynamic stabilization conferred by myristoylation is effectively the same at pH 7.7 as it is at pH 6.2. Based on the global energetics, myristoylation increases stability at both pH values by approximately 4 kcal/mol in LLL compared to 3.2 kcal/mol and 1.38 kcal/mol at pH 7.7 and pH 6.2 for wild-type, so, the model is that because the myristoyl group appears to make favourable interactions at both pH values, hisactophilin prefers to stay in the sequestered state. The model is further supported by simulations (Figure 2.4[73]). Why the switch is broken—by what mechanism has proton binding been decoupled from myristoyl switching—remains an open question that the level of detail obtained from NMR and variable temperature experiments can address (additional plots for LLL temperature coefficients may be found in the supplementary).

LLL Temperature Coefficients Differences

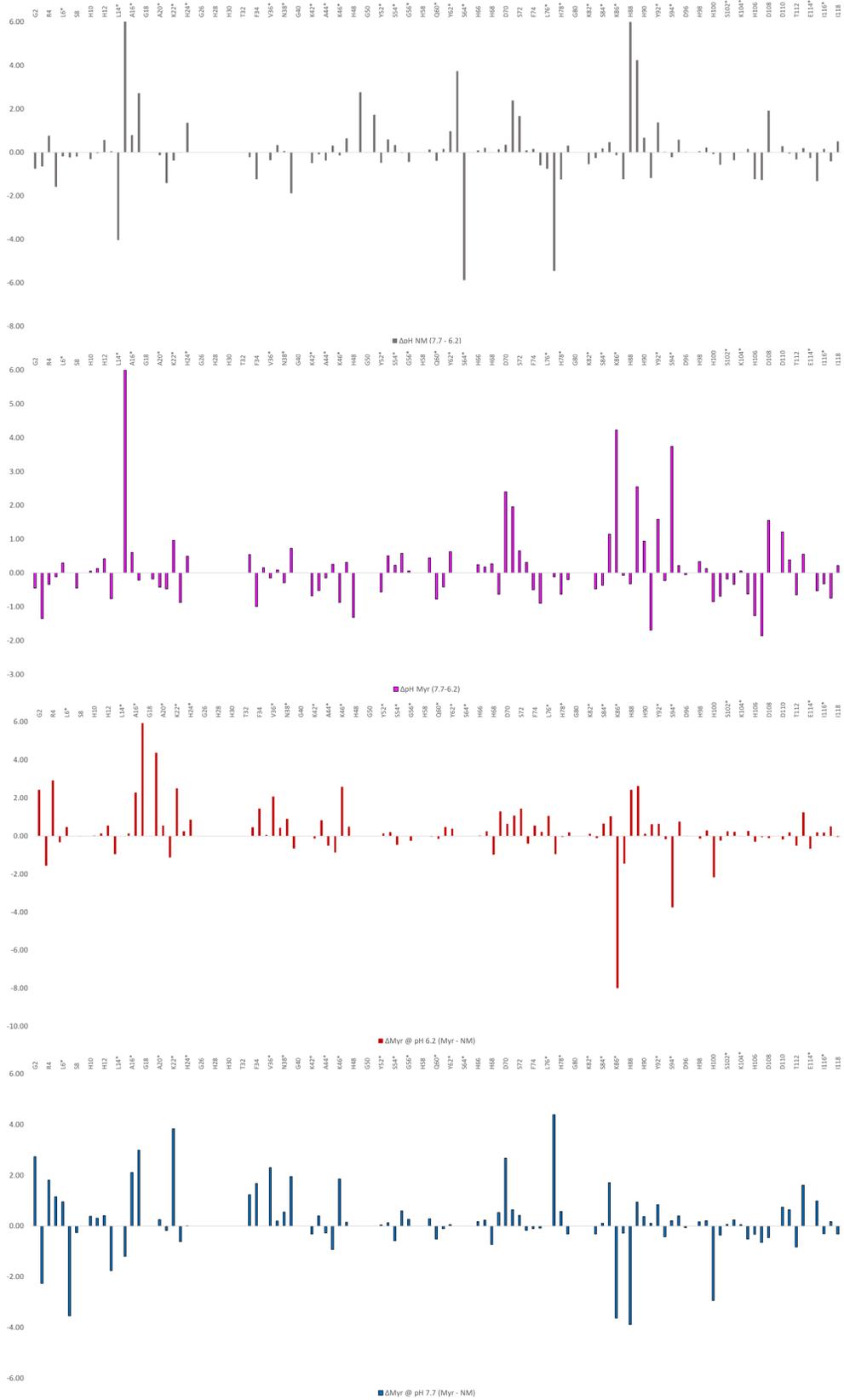


Figure 4.21: Bar chart of the temperature coefficient differences corresponding to each of the edges in the thermodynamic cycle in 3.2 for LLL hisactophilin. (Top & second panels) Difference in temperature coefficients upon pH change for the non-myristoylated (top) and myristoylated forms (second) by subtracting pH 6.2 from 7.7. The values are mapped on hisactophilin's structure in Figure 4.24. (Third & bottom panels) Difference in temperature coefficients upon myristoylation (myr - non-myr) at pH 6.2 (third) and pH 7.7 (bottom), as in Figure 4.20.

Considering first pH 7.7 to examine the effect of myristoylation on structural stability, the prevalence of the colour grey on the structure indicates that the temperature coefficients of fewer residues are sensitive to myristoylation in LLL than in wild-type (right panel of Figure 4.20 vs Figure 4.14; and charts in Figure 4.21). 22 of 86 amide protons show a change in temperature coefficient upon myristoylation, of which only 7 are structurally destabilized by the acyl group. This is in stark contrast to the 31 (of 70) residues in wild-type that have an appreciable change in temperature coefficient with myristoylation, with 17 destabilized by the myristoyl moiety (Figure 4.14). At pH 6.2, 23 (of 90) residues report on myristoylation for LLL, with a similarly small number (6) for pH 7.7 of those with increased structural stability upon myristoylation. Comparatively, at pH 6.2, wild-type has 47 (of 83) residues that are sensitive to myristoylation, of which 20 show increased structural stability at the lower pH. Thus, the temperature coefficients of the triple mutant, LLL, are not particularly responsive to myristoylation in comparison to wild-type.

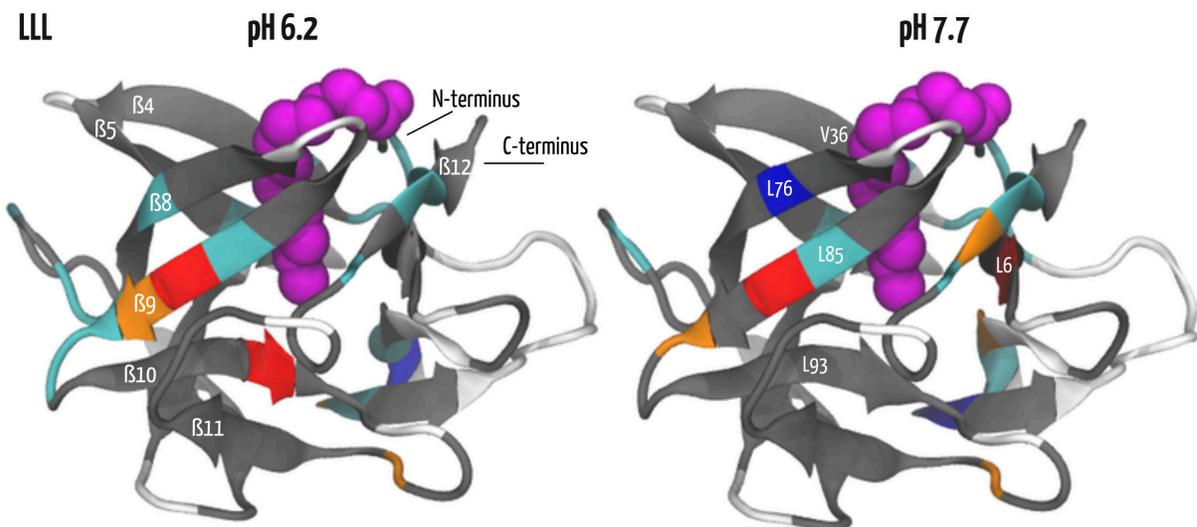


Figure 4.20: Changes in temperature coefficients induced by myristoylation of LLL at pH 6.2 (left panel) and pH 7.7 (right panel) mapped onto the structure of hisactophilin. The myristoyl group is shown in magenta. Differences in temperature coefficients were determined by subtracting the coefficients from the non-myristoylated form from the myristoylated form at each of the two pH values. Residues in white represent those for whom temperature coefficients could not be calculated for either the myristoylated or non-myristoylated form. Grey residues are residues whose temperature coefficients changed by less than 1 ppb/K in either direction. Orange and red residues' temperature coefficients changed by between -1 and -3 ppb/K, or by more than -3 ppb/K, respectively. The temperature coefficients of the residues coloured cyan or blue changed between 1 and 3 ppb/K, or by more than 3 ppb/K, respectively. A positive value (cool colours) mean that myristoylation induces greater structural stability at that position, and negative values (warm colours) suggest that myristoylation decreases structural stability.

Comparing the effect of myristoylation on LLL at pH values 6.2 and 7.7 allows for the quantitation of the interaction energy between proton binding sites and the myristoyl group. Reduced interactions between proton binding and the myristoyl group in residues that form the communication network should be the hallmark of a broken switch. Indeed, the $\Delta\text{TCoeff}_{\text{switch}}$ shown in the upper right panel of Figure 4.15 bear this out; changes in local stability conferred by the myristoyl group are similar at both pH values (grey residues) for 71 residues of the measured 85. Eight of the remaining 14 residues are stabilized more by the myristoyl group at pH 7.7 than at pH 6.2. The other 6 residues show the opposite effect, they are more stabilized by the myristoyl group at pH 6.2 than at pH 7.7. These changes for LLL are much more tepid than for wild-type.

The experiments with wild-type had identified numerous residues that were strained by the myristoyl group at pH 7.7, and therefore possibly involved in switching. In the broken-switch mutant LLL, we find that these same residues (e.g. H35, V36, S84, I85, H89, H91, I93, and A95) are no longer strained. This supports the model of a disrupted communication pathway between the putative proton-binding site(s) and the myristoyl pocket. Without energy and/or stability differences impacted by protonation or the state of the myristoyl group, one cannot drive the behaviour of the other.

4.4.2.6 I85L amide temperature coefficients

I85L is another mutant with a broken thermodynamic switch (Chapter 2)[73], however with a different character than LLL. Global thermodynamics (Chapter 3) show that the degree of stabilization conferred by myristoylation at high and low pH is small (1.2 kcal/mol), similar to the stability increase when wild-type is in the myristoyl-accessible state at pH 6.2. Atomistic simulations support the interpretation that it is less favourable for the myristoyl group to sequester at pH 7.7 and 6.2 [73]. We hypothesize that the preference for the accessible state is because the bottom of the binding pocket is partially occluded by mutation. The resulting changes in structural stability ought to be captured by temperature coefficients and reveal a lack of strain. See the supplementary section for additional discussion of I85L temperature coefficients.

In I85L, myristoylation causes temperature coefficient changes in 19 (of 62) residues at pH 7.7 (right panel of Figure 4.22), and 20 (of 62) residues at pH 6.2 (left panel of Figure 4.22). In wild-type, 31 (of 70) residues report on myristoylation at pH 7.7 and 47 (of 83) at pH 6.2 (Figure 4.14). Fewer residues that are affected by myristoylation in I85L could support the hypothesis that the myristoyl moiety is more prone to the accessible state where it interacts less with core residues (e.g. $\beta 9$, $\beta 10$). However, that there are fewer probes of temperature coefficients for I85L warrants some caution in interpreting an otherwise straightforward consideration of the data.

The structure in the lower panel of Figure 4.15 shows $\Delta\text{TCoeff}_{\text{switch}}$, the difference between the effects of myristoylation on temperature coefficients at pH 7.7 and pH 6.2, and highlights residues whose changes in structural stability in response to myristoylation and proton binding may drive pH-dependent switching.

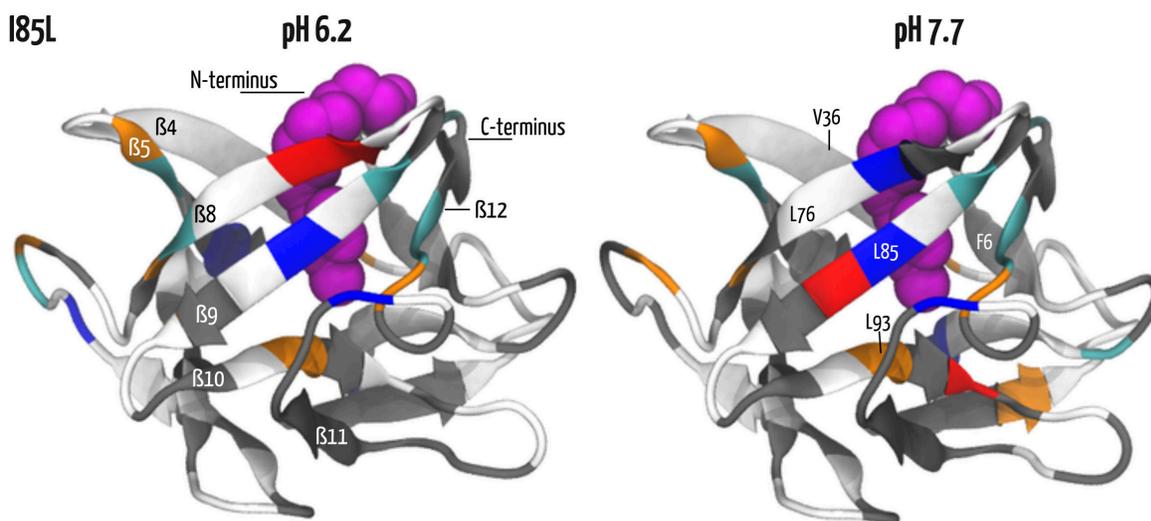


Figure 4.22: Changes in temperature coefficients the result of myristoylation in hisactophilin mutant I85L at pH 6.2 (left panel) and pH 7.7 (right panel). The myristoyl group is shown in magenta. Differences in temperature coefficients were determined by subtracting the coefficients from the non-myristoylated form from the myristoylated form at each of the two pH values. Residues in white represent those without temperature coefficients either the myristoylated or non-myristoylated form. Grey residues have temperature coefficients change by less than 1 ppb/K in either direction. Orange and red residues' temperature coefficients change by between -1 and -3 ppb/K, or by more than -3 ppb/K, respectively. The temperature coefficients of the cyan or blue residues change by between 1 and 3 ppb/K, or by more than 3 ppb/K, respectively. A positive value (cool colours) mean that myristoylation induces greater structural stability at that position, and negative values (warm colours) suggest that myristoylation decreases structural stability.

Strikingly, the residues from wild-type that felt strain are colourless in I85L (I85, I93, V36, L45, in Figure 4.15, for example). All but 5 residues (of 58 probes) from I85L have a $\Delta T\text{Coeff}_{\text{switch}}$ of less than 1 ppb/K. Furthermore, 36 (of 68) wild-type residues are sensitive to the coupling of proton binding and myristoyl switching. In I85L, the switch is no longer influenced by proton-binding and its temperature coefficients recapitulate the broken switch we observe in global energetics experiments (Chapter 3). In combination with the minor effects to temperature coefficients/structural stability due to myristoylation, it supports that differences in temperature coefficients may be reporting on the reduced interactions between the myristoyl group and residues in the binding pocket.

4.4.2.7 Integrating tension into the switching mechanism

We hypothesized that switching is driven by strain in the native state that arises from an over-packed core when the myristoyl group is buried [8]. Strain, here, is meant as the energetic cost paid and stored in the structure as conformational tension upon folding/burying the myristoyl group into too small a cavity, and then redeemed to overcome the forces that favour the sequestered state/disfavour the accessible state to drive the switch. The destabilization of the native state reduces the difference in energy between the myristoyl accessible and sequestered states so that small changes in pH (crossing pH 6.9) are sufficient to tip the balance from the accessible to sequestered state or vice versa with high sensitivity. A consequence of these precariously and finely balanced energetics is that under high pH conditions, which favour myristoyl sequestration, the accessible state is sufficiently close in energy for hisactophilin and the myristoyl group to explore.

Integrating the results of temperature coefficients with the global energetics measurements proffers further evidence that strain and its relief are an effector of the switch and that the switch may be broken by modulating strain and upsetting the thermodynamic balance. A mental schema and model for the shifting energy landscapes of the switch is a seesaw. On one end of the seesaw is the accessible state, on the other end is the sequestered state, and pulling down on either side are the thermodynamics favouring that state. Strain keeps the two sides near-equilibrium, allowing them to readily oscillate. Being nearly balanced also allows small perturbations (by mutation or proton-binding) to readily shift favour toward one state or the other, as might be required for function. The apparent energy-landscape of hisactophilin's switch is consistent with a dynamic energy landscape view of a population shift model of coupled ligand binding and conformational change [279], which suggests hisactophilin could be a model for characterizing allostery and switching in such systems, or that temperature coefficients are an approach to test them directly [280].

The mutations investigated with temperature coefficients, LLL and I85L, upset the careful balance of the sequestered and accessible states. In LLL, for example, the sequestered state is more favoured such that the small protonation doesn't affect the oscillation of the seesaw. Oscillation has not necessarily stopped, but the

seesaw is no longer influenced by proton binding.

These results integrate well with earlier work indicates that the switch is highly dynamic [8,73], that residues that interact with the myristoyl group in the core of hisactophilin exhibit the largest chemical shift perturbations upon myristoylation [8], and the magnitude of the perturbation in the broken switch mutants is markedly decreased (see Figure 4.9) which supports fewer interactions between the binding pocket and the myristoyl group in LLL and in I85L (possibly owing to the larger pocket or the myristoyl group favouring the accessible state, respectively).

In the broken switch mutants, the loss of coupling between proton binding and the state of the myristoyl groups agrees with that conformational tension caused by sequestering the fatty acid is an impetus for switching. Switches in other mutants (Chapter 3) supports that one of the mechanisms of tuning switching is through adjusting core packing. Many residues with an apparent decrease in local structural stability when the protein was myristoylated and in the sequestered state form the bottom of the β -barrel (Figure 4.15), close to the terminal methyl of the myristoyl group, which is in agreement with some of the residues that have the greatest chemical shift perturbations upon myristoylation (Figure 4.8). The seemingly overpacked core reduces the local structural stability of select residues, then, a pH-induced adjustment of the core may further increase tension so that the accessible state, where tension is relieved, is more favourable.

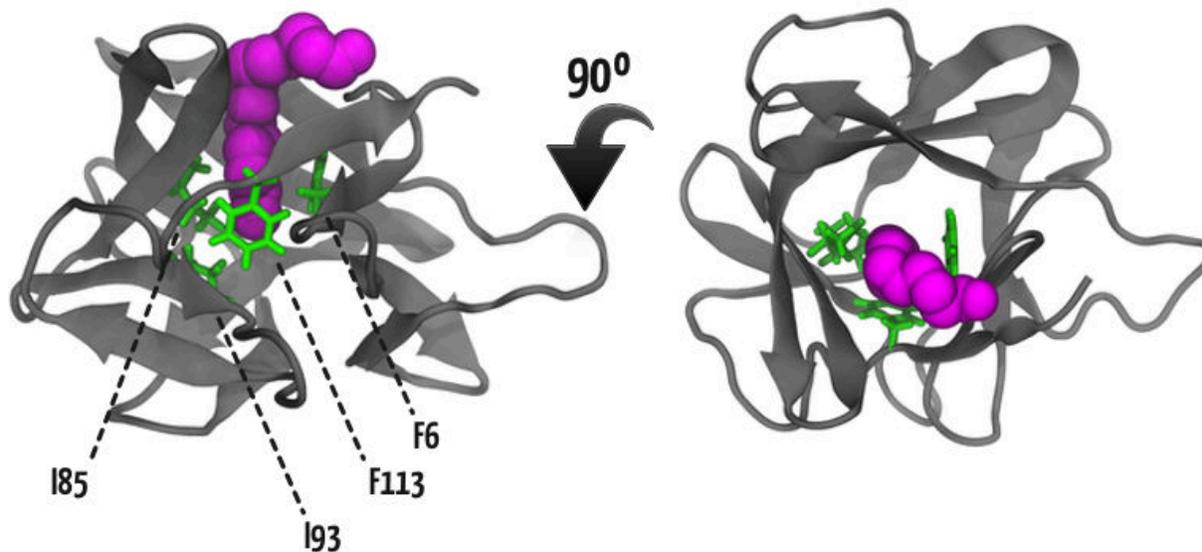


Figure 4.23: Structure of hisactophilin with the myristoyl group shown as magenta spheres and the residues F6, I85, I93, and F113 shown as green sticks and labelled accordingly. Those four residues had measured NOEs to the terminal methyl of the the myristoyl group, which is to say that they are within ~ 5 Å of one another. (Left panel) is a side-on view and the (right panel) is a 90° rotation out of the plane of the page to look down the β -barrel.

Comparing wild-type to the LLL mutant at pH 7.7 offers insight on how these mutations affect the core.

Earlier work that generated the myristoyl-sequestered model of hisactophilin identified NOEs between the terminal methyl group of the myristoyl moiety and residues F6, I85, I93, and F113 [8], so these residues should report on the interaction between the myristoyl group and the bottom of the binding pocket (see green residues in Figure 4.23). At pH 7.7 in myristoylated hisactophilin, the LLL mutation changes the temperature coefficient from +3.05 in wild-type to -1.82 for residue I/L93, from -1.05 to -2.19 for I/L85, from -4.26 to -2.97 for F113, and -2.47 to -2.76 for F/L6 (Appendix A). Residues I/L85 and I/L93 (and slightly for F/L6) are structurally destabilized by the mutations which a myristoyl group that is more mobile in the pocket could explain. Interestingly, increased stability in residues distant from the sites of mutation may compensate for the loss of stability. Changes to temperature coefficients and thus structural stability by mutation appear to support that switching is broken in LLL by altering sterics in the core and position of the myristoyl group so that existing allosteric pathways linking proton binding and the myristoyl group that include I85, I93, and F6 are disrupted.

The impact on the core of moving the γ -methyl in I85L appears somewhat more straightforward. Comparing to wild-type at pH 7.7, residue I93 in the myristoylated mutant is destabilized by -6.91 ppb/K, and yet more pronounced at pH 6.2 where the residue is destabilized by -10.17 ppb/K. This may result from clashes of the altered binding pocket with the myristoyl group when it attempts to sequester, creating large amounts of conformational tension and tipping equilibrium toward the accessible state. In non-myristoylated I85L, the same residue is more stable than the myristoylated form (temperature coefficient of \sim -2.3 ppb/K), as expected for a residue in secondary structure in the core of a protein.

Myristoylation of wild-type and I85L at pH 7.7 shows a decrease in local structural stability at the bottom of the β -barrel, adjacent to the myristoyl group (compare Figures 4.14 and 4.22). This mutation-induced increase in tension over and above what was measured in wild-type may disrupt the finely balanced energetics of the thermodynamic switch.

Analyzing temperature coefficients with our thermodynamic cycle identified residue-level changes in structural stability sensitive to pH, myristoylation, and mutation to a network of connectivity between the putative but elusive proton binding site(s) and the myristoyl pocket. Previous pH titration experiments implicated residues H75, H91, H78, and H107 as likely candidates for binding switching protons [8], and temperature coefficients further support the involvement of residue H91 as its structural stability is strongly responsive to both the state of myristoylation and the pH of the experiment. Allosteric signalling may emanate from H91 in turn β 9 β 10 and increase strain around residues I85 and I93, which favours switching to the accessible state. On the other side of and farther up the β -barrel, residues H35, V36 (β 4), L45 and K46 (β 5) show a relief of strain when the myristoyl group switches to accessible, so they may also contribute to shifting the balance of conformations toward accessible by stabilizing the accessible state.

4.5 Conclusions

Interest in switching proteins is long-established; as a class of protein, their diversity of sensitivity to effectors and triggers like light [48,49,281], calcium [95,221,282,283], phosphorylation and magnesium binding [284], nucleotides [157,280,285,286], RNA [287], other proteins [288,289], and protons [8,11,290–292] makes them subjects of great interest and the target of ongoing research to understand how they work and how to manipulate them into custom sensors for biotechnological applications, for example. Switching and allostery are tightly related concepts because a switch couples the detection/binding of an effector to some change, whether conformational or dynamic, that then influences protein function. Unsurprisingly, the coupling of ligand binding to a functional change does not follow a universal mechanism across switches sensitive to different effectors (or even the same effector), and it is still somewhat of an open question whether allostery tends to propagate via population-shift, where a switching protein may transiently sample the ‘go’ conformation in the absence of the ‘go’ signal (which prefers to bind to the active form of the protein), ergo the ‘go’ conformation seems to precede the signal, or via an induced fit mechanism where ‘go’ signal binding is the impetus for conformational/dynamic change to the active state. While the goal of this research was not expressly to determine hisactophilin’s mechanism of allostery, the data suggest population-shift given the apparent sampling of the myristoyl-accessible state shown by NMR lineshape analysis [8] and by molecular dynamics [73,293]. This is consistent with the work of Okazaki & Takada suggesting the population-shift mechanism is likely favoured by systems whose ligand is small and that require weaker and shorter-ranged interactions for switching [279]. Hisactophilin’s ligand is a proton and the highly dynamic nature of the myristoyl group supports a population-shift mechanism for the myristoyl-sequestered to -accessible switch.

Further reasons why switches are an interesting and active area of research are that there is no single tool for teasing out pathways of interacting residues that might be responsible for the coupling through the switching protein, or a single tool for designing a switch. One tool for measuring residue interactions relied on evolutionary data from the large PDZ protein family to statistically identify positions in the protein that were coupled. A long tract of residues was predicted as important in allosteric communication across the protein and strategic mutation of residues in the pathway supported their claim and validated the method [288]. Another method is the previously discussed CHESCA [186], which relies on a library of well-characterized functionalized ligands to map covarying chemical shift perturbations in the target protein upon binding. Unfortunately, neither of these approaches are especially amenable to hisactophilin; first, although there are structural homologs to hisactophilin, there are no other β -trefoils with pH-dependent myristoyl switches, so there is no large family from which to glean evolutionary data; and second, using a set of different mutations throughout the protein is perhaps fraught in lieu of having ligand alternatives for proton binding (or perhaps other N-terminal fatty acids) for covariance analysis by CHESCA. Instead, we employed a complementary

and robust thermodynamic cycle-like analysis to the temperature dependence of chemical shifts to identify residues coupled to the myristoyl group and proton binding. Beyond relying on chemical shifts which are straightforward to precisely determine, this method requires only knowledge of the approximate melting point of the target protein to define the temperature range of experiments and the (perhaps not so trivial) NMR assignments.

Myristoyl-induced strain in the native state of hisactophilin and the energetic impetus to relieve that strain was proposed as a driver of the myristoyl switch [8], and here temperature coefficients have enabled us to localize regions of conformational tension or localized frustration [62,294] throughout the β -barrel that appears to be relieved at low pH when the myristoyl group switches to the accessible state. Areas involved in switching/strain are residues 36, 45, 76, 85, and 93. Taking a step further, we subjected two mutants whose pH-dependent myristoyl switches were broken (F6L/I85L/I93L and I85L) to the same cycle analysis and the evidence of strain/tension/local frustration was missing from most of the protein, including those regions implicated in switching for wild-type. So, this further supports that an energetic penalty incurred by burying the myristoyl group and creating an overpacked core is an important effector of switching [281]. Energetically, the conformational tension appears to ensure local minima in the native basin are sufficiently close in energy that ionizing some incognito histidines could alter the relative populations [291,295–298].

Temperature coefficients have demonstrated that they are exquisitely sensitive to the small conformational and dynamic changes involved in the myristoyl switch because, unlike some other well-studied switches [283,286,299,300], there is little conformational change in the backbone that accompanies switching. Applying variable temperature NMR and temperature coefficients to hisactophilin provides insight into the molecular mechanism of the pH-dependent myristoyl switch and highlights the importance of myristoyl binding pocket architecture to creating a delicate balance of both favourable and unfavourable interactions that allow and govern switching.

4.6 Supplementary

4.6.1 Additional wild-type temperature coefficients

Reducing the pH from 7.7 to 6.2 affects many temperature coefficients throughout hisactophilin (which now range from -10.18 ppb/K to +3.28 ppb/K; Figure 4.10 and Figure 4.12), and with some exceptions, the above trends still generally hold true; the regions of the termini tend towards less structural stability, where the bottom of the β -barrel meets the hairpin triplet is highly stable, and the ends of strands tends toward a relative decrease in structural stability. Despite being at the interface between the bottom of the β -barrel and the hairpin triplet, the I93-containing strand β 10 is curiously and noticeably unstable in non-myristoylated wild-type hisactophilin. Areas that show significant changes in stability when the pH changes from 6.2 to

7.7 highlight areas that are responsive to proton-binding (see the left panel of Figure 4.12, bar chart in top & second panels of Figure 4.13), and while a few show increased stability at low pH (e.g. C-terminal end of strand β 9), the majority of the areas that respond to pH are stabilized by increasing pH, consistent with the pH dependence of global stability—hisactophilin is more globally stable at high pH (Table 4.2 and [87]).

4.6.2 Additional LLL temperature coefficients

At pH 7.7, the temperature coefficients for the non-myristoylated variant of LLL range from -9.80 ppb/K to +0.51 ppb/L, values comparable to those of WT (-9.43 ppb/K to +1.15 ppbK; see Appendix A for complete a complete list of temperature coefficients). At lower pH, the range expands (-12.93 ppb/K to +0.89 ppb/K) and predictably indicates overall destabilization (Table 4.2). The ranges of temperature coefficients for myristoylated LLL are -10.31 ppb/K to +1.15 ppb/K and -12.80 ppb/K to +1.43 ppb/K for pH 7.7 and 6.2, respectively. At first glance, the state of myristoylation or changes in pH do not appear to dramatically affect the range of temperature coefficients, however this is less surprising in the context of a broken switch. Temperature coefficients are mapped onto the structure in Figure 4.19 (and in chart form Figure 4.18).

In non-myristoylated LLL, 25 residues of the 90 analyzed temperature coefficients show appreciable changes in local structural stability (i.e. changes in temperature coefficients) when the pH is increased from 6.2 to 7.7 (left panel in Figure 4.24, top panel of Figure 4.21). 12 of the 25 are residues that are more structurally stabilized at high pH (blue/cyan), consistent with increased global stability, and 13 that are more structurally stabilized at pH 6.2 (reds). The number of and distribution of these pH sensitive residues is akin to non-myristoylated wild-type, highlighting that proton binding events are felt throughout the protein and that local destabilizations may be compensated for by increased structural stability elsewhere, similar to the propagating effects of mutations seen in other proteins [150,301]. The similarities between non-myristoylated LLL and wild-type here may not be surprising because without the myristoyl group there is no switch, water likely has comparable access to the core, and the incognito proton-binding histidine(s) have not had their interactions interfered with because the sites of mutation are distant and in the core. Interestingly, turn β 9 β 10 and strand β 10—which contain H89 and H91—are pH-sensitive (vide infra). With the myristoyl group present, however, the number of residues whose structural stability is pH-sensitive is depressed to 16, and just 5 are relatively destabilized at pH 7.7 compared to the 12 (of 22) in wild-type (right panel in Figure 4.24 and 4.12). In terms of an abrogated switch, this is another bit of evidence in support of the model that in LLL the myristoyl group makes similar favourable interactions with the rest of the protein at pH 6.2 as it does at pH 7.7. A corresponding physical explanation is that the myristoyl group is sequestered to the same extent at both pH values. Fewer residues report increased structural stability at pH 6.2 relative to pH 7.7 in LLL (compared to WT), which may be the first inklings that strain in the native state is the effector of change because its attenuation correlates with an ablated switch.

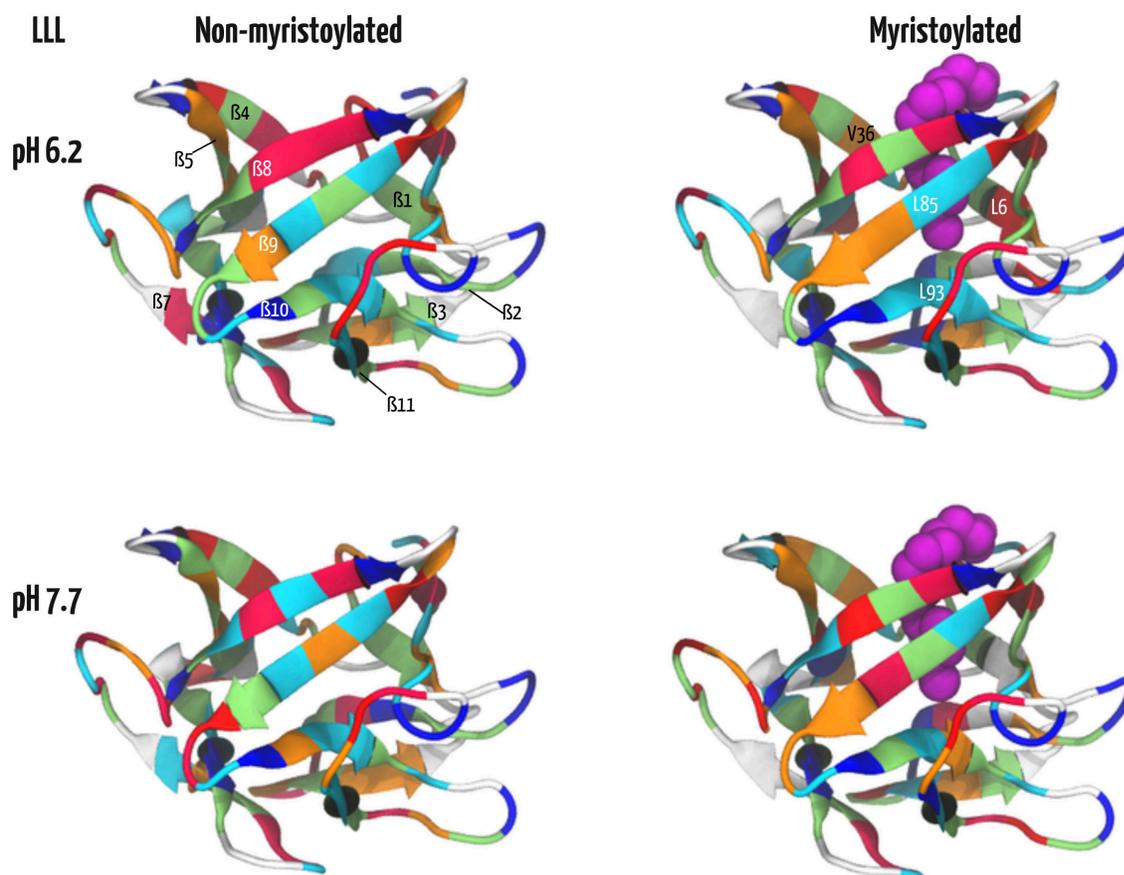


Figure 4.19: Amide proton temperature coefficients for F6L/I85L/I93L mapped onto the backbone of hisactophillin. Upper left and lower left are non-myristoylated forms at pH 6.2 and pH 7.7 respectively. Upper and lower right are the myristoylated forms at pH 6.2 and pH 7.7. The myristoyl group is shown as magenta spheres. Residues shown in white lack assignments or could not be reliably tracked with temperature. Dark red residues have temperature coefficients less than -7 ppb/K, red are between -5.5 and -7 ppb/K, orange are between -4 and -5.5 ppb/K, green are between -2.5 and -4 ppb/K, cyan are between -1 and -2.5 ppb/K, and blue are greater than -1 ppb/K.

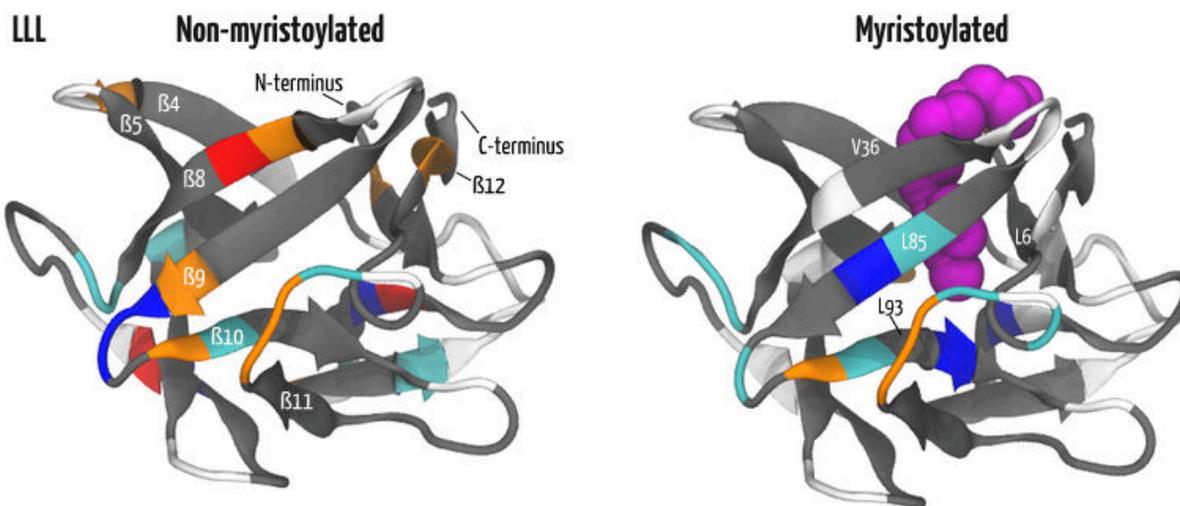


Figure 4.24: Change in temperature coefficients induced by changing the pH from 6.2 to 7.7 in LLL mapped onto hisactophilin's structure for non-myristoylated (left panel) and myristoylated (right panel) protein. The myristoyl group is in magenta. Residues in white are those for which temperature coefficients could not be obtained at one or both pH values for that form. The difference in temperature coefficients were determined by subtracting the temperature coefficients at pH 6.2 from those at pH 7.7. Positive values indicate that a residue has greater structural stability at pH 7.7 than pH 6.2, and negative values indicate that a residue has greater structural stability at pH 6.2 than pH 7.7. Grey residues show a small change in temperature coefficient of between -1 and 1 ppb/K, orange residues have a change of temperature coefficient of between -1 and -3 ppb/K, and red residues have a change in temperature coefficient more negative than -3 ppb/K. The temperature coefficient of cyan residues increases by between 1 and 3 ppb/K, and blue residues increase by more than 3 ppb/K.

Figure 4.18: Bar chart of the LLL (F6L/I85L/I93L) mutant amide proton temperature coefficients, measured at pH 6.2 for the non-myristoylated (top panel) and myristoylated forms (second panel), and at pH 7.7 for non-myristoylated (third panel) and myristoylated (bottom panel). Residues labeled with an asterisk form elements of β -structure. The values correspond to those mapped onto the backbone in Figure 4.19.

4.6.3 Additional I85L temperature coefficients

The range of temperature coefficients observed in I85L recorded for the two pH values and acylation-states span -10.5 ppb/K, with myristoylated at pH 6.2 as the exception at with a wider range of 12.53 ppb/K (complete list of temperature coefficients are in Appendix A). The ranges are generally close to those for wild-type and LLL, if not slightly narrower. Though, I85L has fewer assignments and measurable temperature coefficients than LLL. The distribution of temperature coefficients over the structure follows similar patterns to wild-type and the loose trend of more negative temperature coefficients toward the boundaries of secondary structure, though it is still a long way from a rule. Also, the trend has begun to emerge in the β -strands that a residue i and residue $i + 2$ may share approximately similar temperature coefficients (see Figure 4.25, or Figure 4.26)—represented by similar colours—given that such residues have similar orientations in the protein (e.g. side chains are oriented to the same side of a β -strand), it follows that they can report on how similar environments change with temperature, whether arising from changes to hydrogen-bonding or the relative positions of ring currents.

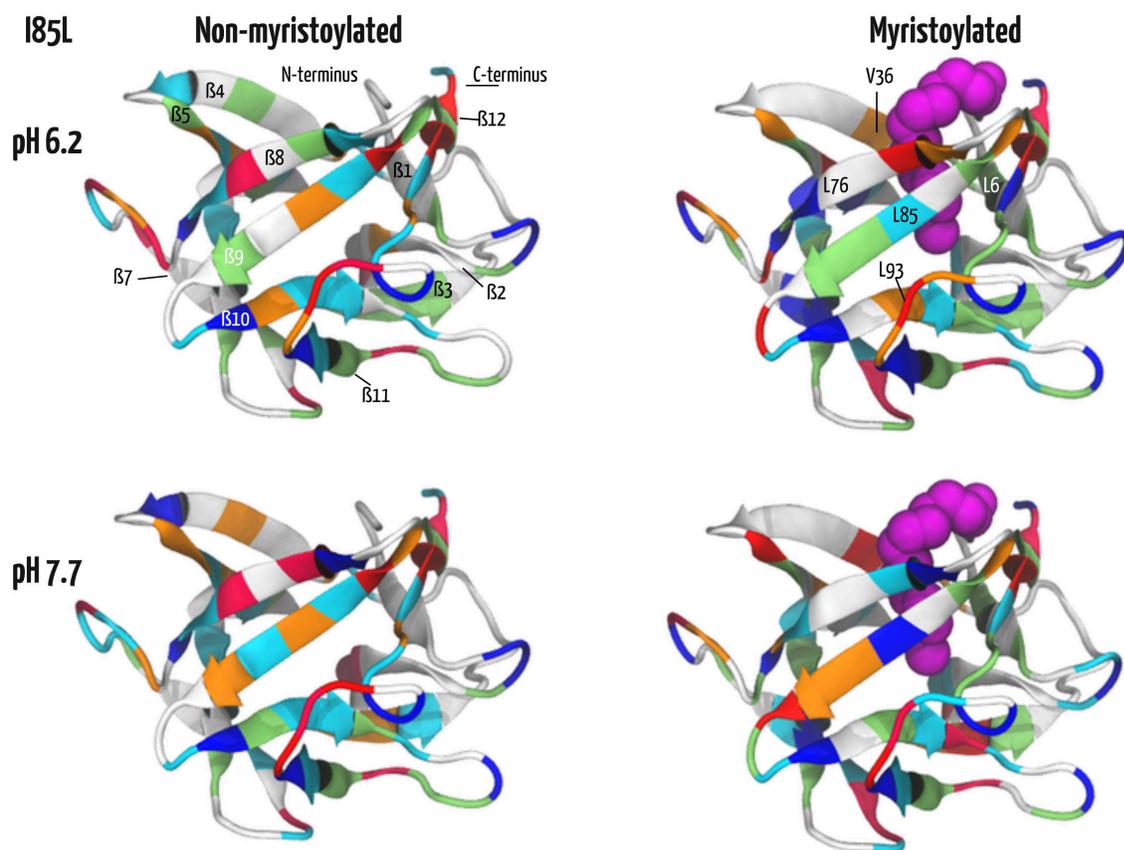


Figure 4.25: Amide proton temperature coefficients for hisactophilin mutant I85L mapped onto the structure. Upper and lower left panels are non-myristoylated at pH 6.2 and pH 7.7, respectively. Upper and lower right panels are myristoylated at pH 6.2 and pH 7.7. The myristoyl group is shown as magenta spheres. Residues in white have no temperature coefficient information from a lack of assignment or reliably tracked peaks. Dark red residues are those with temperature coefficients less than -7 ppb/K, red are between -5.5 and -7 ppb/K, orange are between -4 and -5.5 ppb/K, green are between -2.5 and -4 ppb/K, cyan are between -1 and -2.5 ppb/K, and blue are greater than -1 ppb/K.

185L Temperature Coefficients

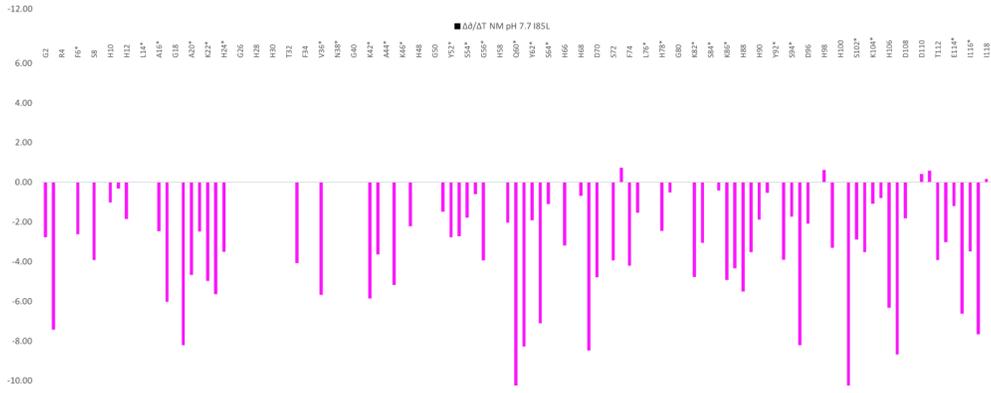
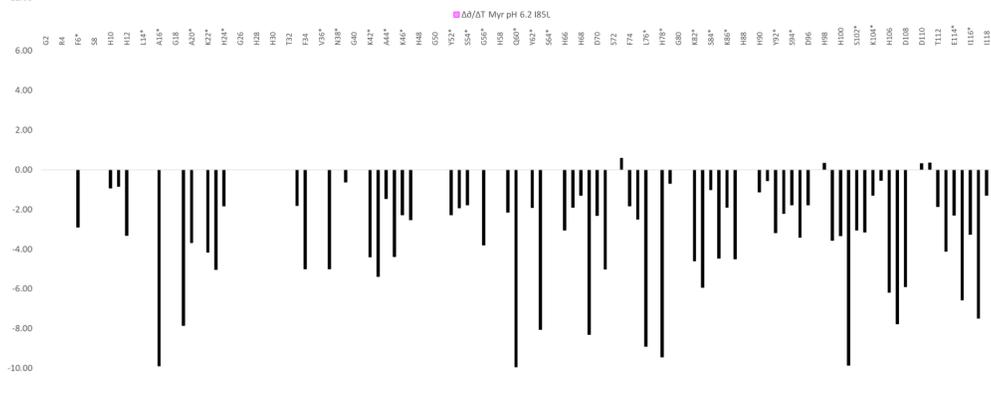
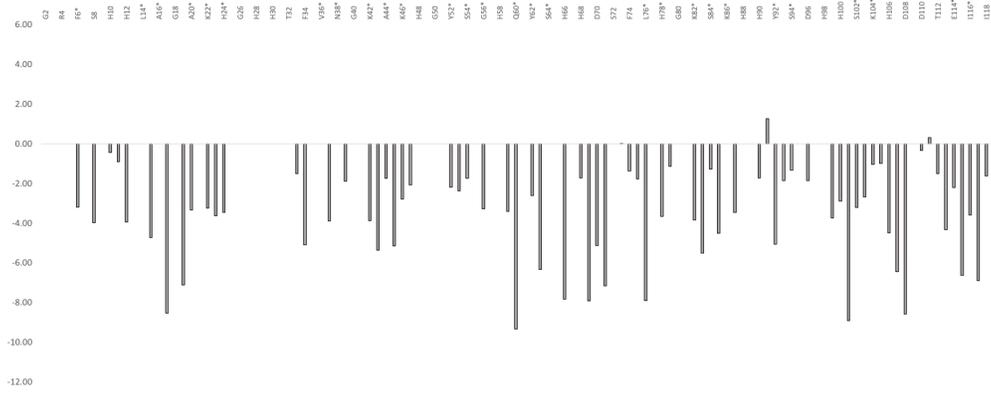


Figure 4.26: Bar chart of the I85L mutant amide proton temperature coefficients, measured at pH 6.2 for the non-myristoylated (top panel) and myristoylated forms (second panel), and at pH 7.7 for non-myristoylated (third panel) and myristoylated (bottom panel). Residues labeled with an asterisk form elements of β -structure. The values correspond to those mapped onto the backbone in Figure 4.25.

Looking at the pH-dependent changes in temperature coefficients for myristoylated and non-myristoylated I85L, the magnitude of the response of temperature coefficients to pH is more muted than in wild-type — local structural stability measured by temperature coefficients is less perturbed by proton binding (note the relative decrease in the number of residues occupying the extremes of perturbation, noted as red and blue in Figure 4.27 compared to Figure 4.12). Myristoylated and non-myristoylated I85L have 54 (of 74) and 52 (of 69) residues, respectively, whose temperature coefficients minimally respond to the pH change (grey residues in Figure 4.27, see also the top two panels of Figure 4.28). That leaves a total of 20 residues for myristoylated I85L that are responsive to pH; 7 decrease in structural stability when the pH was lowered (blue/cyan residues), and the remaining 13 have increased structural stability at pH 6.2 (red/orange residues). Non-myristoylated has 8 residues that are destabilized upon lowering pH (blue/cyan), and 9 that are more structurally stable at pH 6.2 (red/orange). The absolute number of pH-affected residues is similar to wild-type, which suggests the isoleucine to leucine mutation has not broken overall pH-sensitivity. Instead, the mutation may attenuate and decouple local stabilities responsible for driving the equilibrium switch from the myristoyl group (*vide infra*).

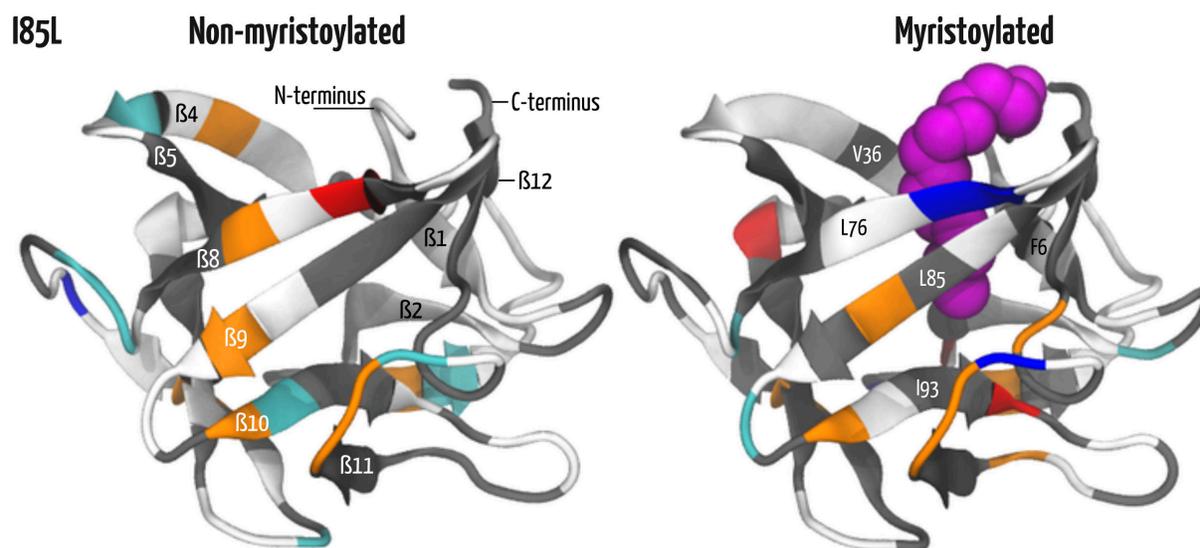


Figure 4.27: Differences in temperature coefficients resulting from changing from pH 7.7 to pH 6.2 in non-myristoylated (left panel) and myristoylated (right panel) hisactophilin mutant I85L. The myristoyl group is in magenta. Residues in white are those for which temperature coefficients could not be obtained at one or both pHs for that form. The difference in temperature coefficients that were mapped were determined by subtracting the temperature coefficients at pH 6.2 from those at pH 7.7. Positive values indicate that a residue has greater structural stability at pH 7.7 than pH 6.2, and negative values indicate that a residue has greater structural stability at pH 6.2 than pH 7.7. Grey residues showed a small change in temperature coefficient of between -1 and 1 ppb/K, orange residues had a change of temperature coefficient of between -1 and -3 ppb/K, and red residues had a change in temperature coefficient more negative than -3 ppb/K. The temperature coefficient of cyan residues increased by between 1 and 3 ppb/K, and blue residues increased by more than 3 ppb/K.

I85L Temperature Coefficients Differences

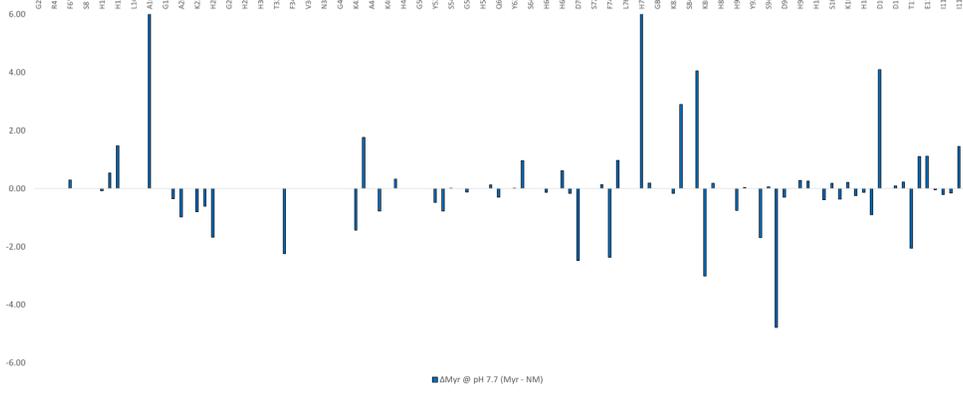
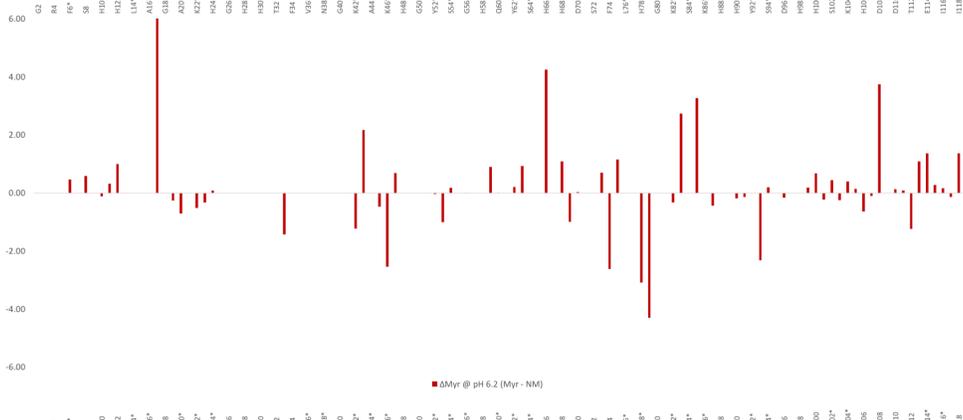
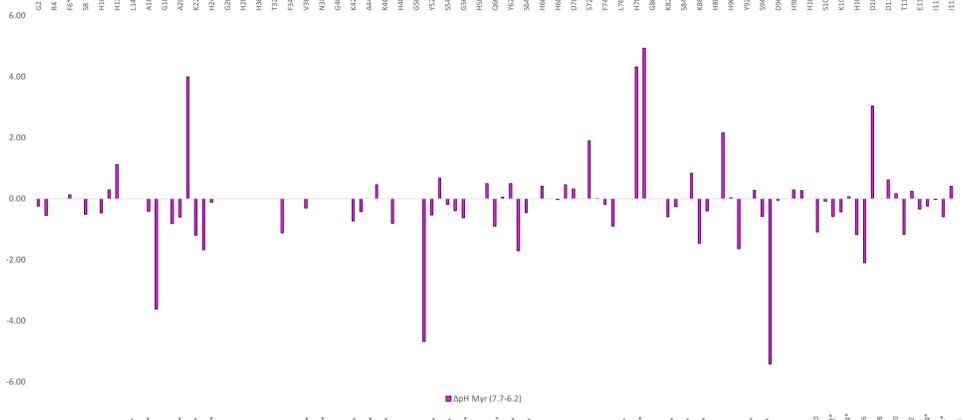
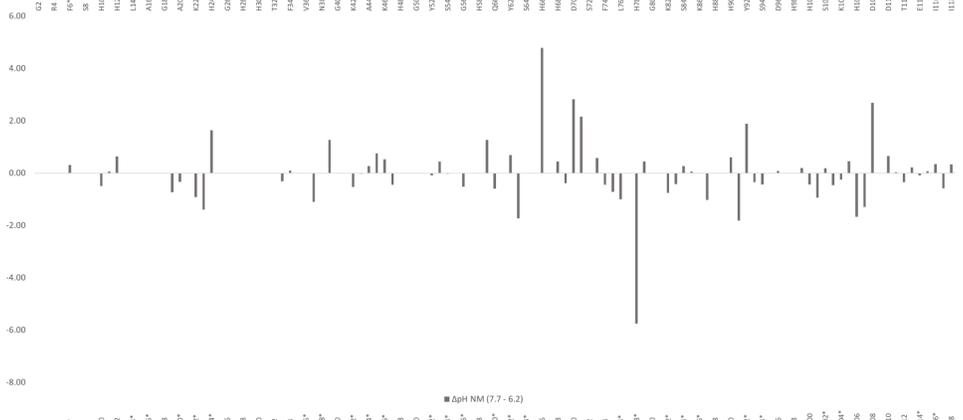


Figure 4.28: Bar chart of the temperature coefficient differences corresponding to each of the edges in the thermodynamic cycle in 3.2 for I85L hisactophilin. (Top & second panels) Difference in temperature coefficients upon pH change for the non-myristoylated (top) and myristoylated forms (second) by subtracting pH 6.2 from 7.7. The values are mapped on hisactophilin's structure in Figure 4.27. (Third & bottom panels) Difference in temperature coefficients upon myristoylation (myr - non-myr) at pH 6.2 (third) and pH 7.7 (bottom), as in Figure 4.22.

Of the residues whose temperature coefficients are affected by myristoylation in I85L, the magnitude of the response is generally attenuated compared to wild-type (see the reduced number of dark reds and blues in Figure 4.22 compared to Figure 4.14. This is true at both pH values. The smaller deflection of the temperature coefficients in response to myristoylation suggests that its contribution to local structural stability also decreased. At pH 7.7, residues stabilized by the myristoyl group are located more in the β -barrel than elsewhere with a slight tendency for higher in the cavity (residues 43, 75, 78, 83, 113, 114, and the c-terminal 118, right panel of Figure 4.22). A source of increased structural stability at pH 7.7 around the top of the barrel could arise from some favourable interactions with the myristoyl group in the accessible state that are not available for the non-myristoylated form. Conversely, those residues whose temperature coefficients are pushed toward values that suggest structural destabilization by the myristoyl group tend to cluster at the bottom of the β -barrel and through some of the hairpins. This reduced structural stability may arise from the isoleucine to leucine mutation creating a less favourable environment (perhaps with less hospitable geometry in the pocket) for the myristoyl group's full insertion, manifested here as the relative structural destabilization of residues on the edge of the pocket, e.g. A95, I93, and K86. Notably, myristoylation affects many residues similarly at pH 6.2 as pH 7.7; though there are some differences, the fact remains that I85L is impacted in fewer places and to a lesser degree by myristoylation than wild-type. Intriguing but logical in the context of the broken thermodynamic switch and the decoupling of the myristoyl group from site(s) of protonation.

4.6.4 Correlating Temperature Coefficients with experiment and simulation

An open question in using and interpreting temperature coefficients is what, physically, they are reporting on. Structural stability, as we have used throughout, is an umbrella term that may encompass various enthalpic interactions or changes in entropy. Earlier work with amide proton chemical shift temperature dependence interpreted them as reporters of hydrogen bonding [136,231–233], and while hydrogen bonding can play a role in determining temperature coefficients, they are not the sole determinants as more recent work attributes temperature-dependent propensity to loss of structure, independent of the presence of intramolecular hydrogen bonding [150,183,230,240]. In an attempt to close in on the physical meaning of what temperature coefficients tell us, and just as importantly what they do not tell us, herein are compared the results (and analyses) of other experiments on hisactophilin to wild-type temperature coefficients. These experiments include the root mean squared fluctuation of the backbone from umbrella sampling simulations, comparing the dis-

tance between each temperature coefficient probe and different parts of the myristoyl moiety, the secondary chemical shift index, and energetics derived from hydrogen/deuterium exchange experiments.

4.6.4.1 Secondary chemical shift

The chemical shift index was developed as a means to identify secondary structure in a protein using only backbone chemical shift data [197]. This secondary structural information is found in the difference between the measured chemical shift and the value for that probe in random coil conformation where, in the simplest sense, chemical shifts that move away from those in random coil likely arise from increasing secondary structure. Thus, if temperature coefficients report on the temperature-dependent loss of folded secondary structure, there should be a relationship between its value and the secondary chemical shift [[183]; [231]]. Figure 4.29 examines this relationship for myristoylated and non-myristoylated hisactophilin at pH 7.7. With a correlation coefficient, r^2 , of 0.10 and 0.07, the secondary chemical shift and temperature coefficients are not significantly correlated. Excluding residues that do not form regular secondary structure does not meaningfully improve this correlation. The Spearman rank-order correlation coefficient to consider whether their relationship can be described by a monotonic function yields $\rho = 0.104$ and 0.093 for myristoylated and non-myristoylated, respectively and provides no evidence of a significant relationship between the two experimental measures (Figure 4.30). A lack of correlation is informative, however; proteins and/or peptides whose temperature coefficients correlate with secondary chemical shift tend to be smaller and undergo extensive conformational averaging that arises from flexibility and/or sampling an unstructured/unfolded state [183,231]. Other reasons for an apparent relationship to temperature coefficients could arise from significant deterioration of the hydrogen bond network that stabilizes the protein, signaling a more global structural perturbation. With an appropriately chosen temperature range, larger, more stably folded (and presumably cooperative) proteins should not undergo global transitions. The observed lack of correlation, then, is also a reassurance that myristoylated and non-myristoylated hisactophilin were well-folded over the temperature range of the experiments and do not appear to measurably sample an unfolded state.

4.6.4.2 Root mean square fluctuation

Then PhD-student, Aron Broom, simulated hisactophilin to gain a deeper understanding of some of our other experimental studies of hisactophilin. One such enterprise used umbrella sampling, a biased simulation that constructs a free energy surface describing a process along some reaction coordinate [293,304,305]. For hisactophilin, the reaction coordinate was the distance from the tip of the myristoyl group to the bottom of the binding pocket and the targeted process was switching, while in an explicit solvent. A harmonic, spring-like force restrained the myristoyl group at a prescribed depth and the potential energy of the system was computed for that approximate depth (plus or minus the harmonic oscillation). Then, the depth of the myristoyl group was adjusted and the calculations repeated. Collating the energetics from each step along the

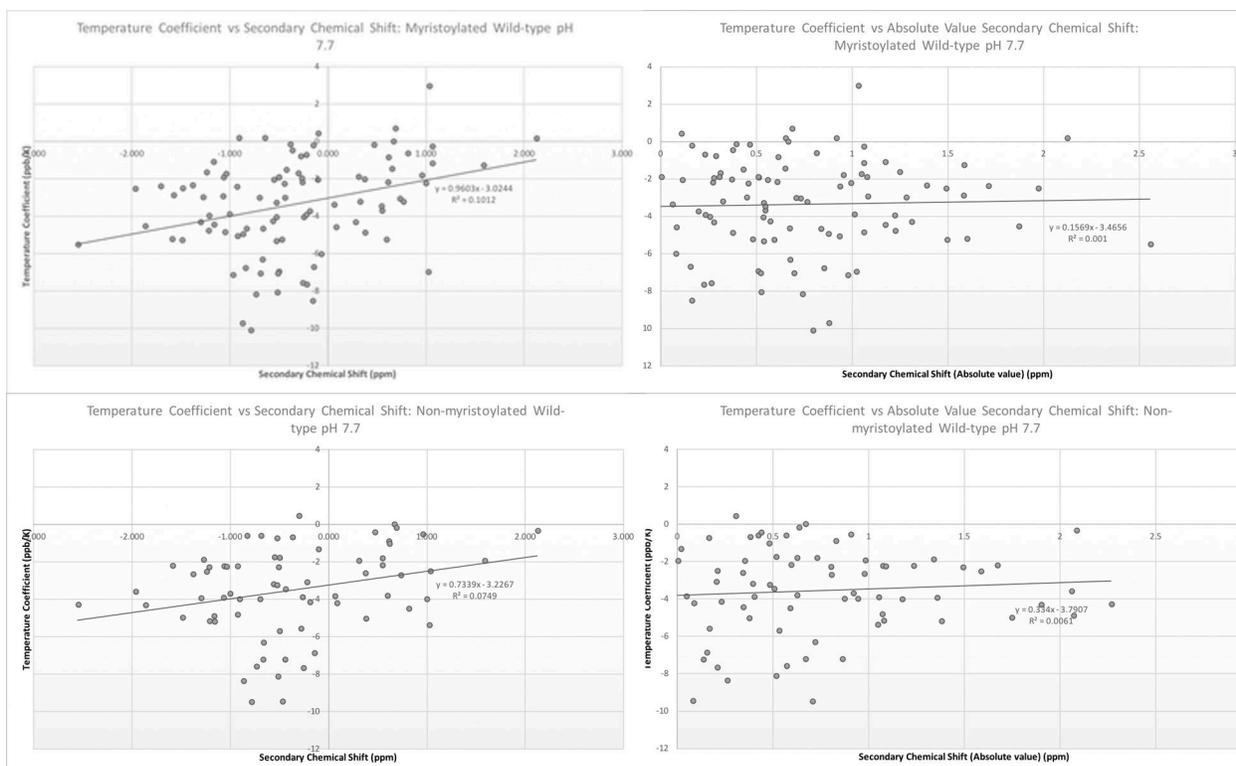


Figure 4.29: Scatterplots of the correlation between temperature coefficients for myristoylated (upper panels) and non-myristoylated (lower panels) hisactophilin and the corresponding secondary chemical shifts. Secondary shifts were calculated using Flemming Poulsen's tools [302,303]. The (left panels) are using the secondary chemical shift as calculated by subtracting the observed chemical shift at 298K from the random coil value calculated at the same temperature. The (right panels) use the absolute value of secondary chemical shifts because the analysis doesn't necessarily ascribe meaning to whether secondary chemical shifts are positive or negative. The correlation coefficients are not suggestive of significant correlation between the measures.

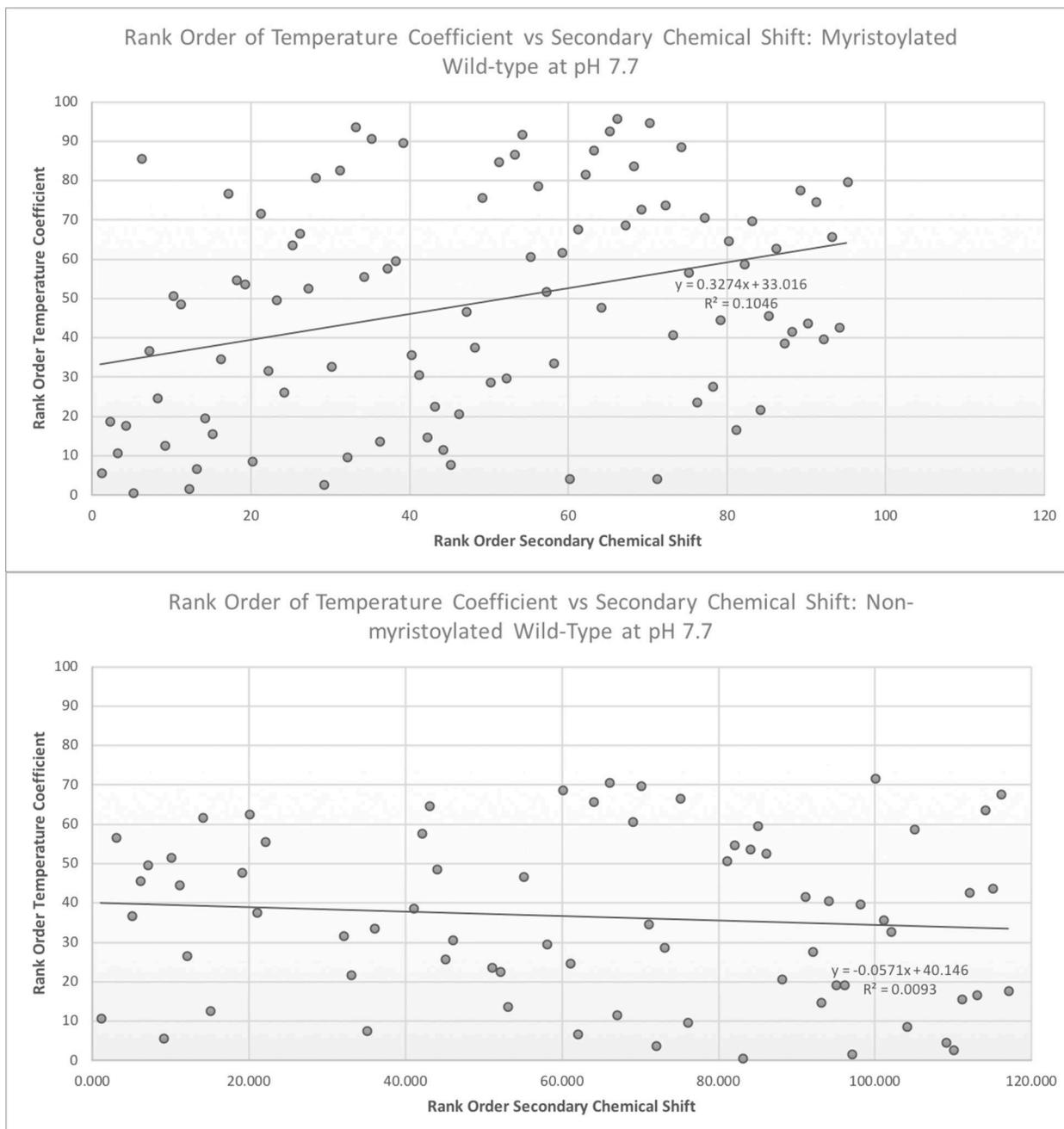


Figure 4.30: Rank order scatterplots showing the lack of correlation between temperature coefficients and secondary chemical shift. Rank order is determined by arranging all temperature coefficients in increasing order and then numbering them from one onward. This is also done for the secondary chemical shifts. Each data point is then plotted according to its rank rather than the value of the variables and then correlated. The correlation coefficient will range from -1 to 1, where -1 is a perfect negative correlation, +1 is a perfect positive correlation, and 0 is no correlation at all. (Upper panel) shows this analysis for myristoylated wild-type at pH 7.7, and the (lower panel) is non-myristoylated wild-type at pH 7.7.

reaction coordinate (i.e. myristoyl depth) constructs the free energy surface for the whole process. In total, the total simulation time was ~ 1200 ns, and each window amounted to 60 ns of simulation time. Simulations offer a rich trove of data that have already served to substantiate aspects of our thermodynamic model of switching (Chapter 2).

The root mean square fluctuation, or RMSF, is a mineable parameter from the umbrella simulations that provides a measure of the amplitude of motion of a given heavy atom. Here, the RMSF reports on the amplitude of motion of the backbone amide nitrogen while the myristoyl moiety is held at a target depth in the binding pocket. Thus, amide nitrogens in loops are expected to (and do) have a greater RMSF than for those that are more constrained by secondary structure. For temperature coefficients, the looming question is how much they too report on local flexibility.

The two measures do not correlate, as shown in Figure 4.31. RMSF values for the backbone nitrogens were extracted for each myristoyl depth, 0 Å (maximally inserted) to 31.5 Å (waving about in solvent). The free energy surface that was constructed for hisactophilin from the umbrella sampling simulations indicated that energy minima at ~ 9 Å and ~ 22 Å likely represent the sequestered and accessible states, respectively [100]. These depths also agree with other simulations (Chapter 2) [73]. RMSF values measured at a myristoyl depth of ~ 9 Å, the sequestered state, ought to correspond to temperature coefficients at pH 7.7 where hisactophilin is thought to be predominantly sequestered. Ideally, experimental measures at pH 6.2 may correspond with the accessible state when the depth of the myristoyl group is ~ 22 Å. It bears stating that pH 6.2 is below hisactophilin's pI and some appreciable fraction of the 31 titratable histidines will be protonated and accounting for these complex electrostatics is beyond the ken of the average simulation. That said, there have been some successes in implementing constant pH molecular dynamics [306,307], but those techniques were not implemented here.

Although it was at first disappointing that there was not a convincing and beautiful correlation to hang my temperature coefficient hat on, it is an explainable thing. One possible reason for the discrepancy is that the umbrella sampling simulations included only heavy atoms, thus we analyzed the amide nitrogen atom which may not necessarily be a perfect proxy for the motion of the amide proton. Furthermore, it is possible that the timescale of the motions reflected in the RMSF values are on different timescales than motions that impact chemical shifts. What this might tell us of temperature coefficients is that they are likely an aggregate measure that includes contributions from flexibility, in addition to the significant contributions from hydrogen bonding, [215,231,233], as well as the probe's immediate environment, like the proximity to charged residues [183] or aromatics [229,232]. A final reason why I believe these RMSF values have missed the mark is because the values were mined from simulations that constrained the myristoyl to a single depth (plus or minus a small oscillation), so expecting a correlation of these RMSFs with temperature coefficients implicitly assumes that the myristoyl group is at a single depth while chemical shifts and temperature coefficients are measured. But

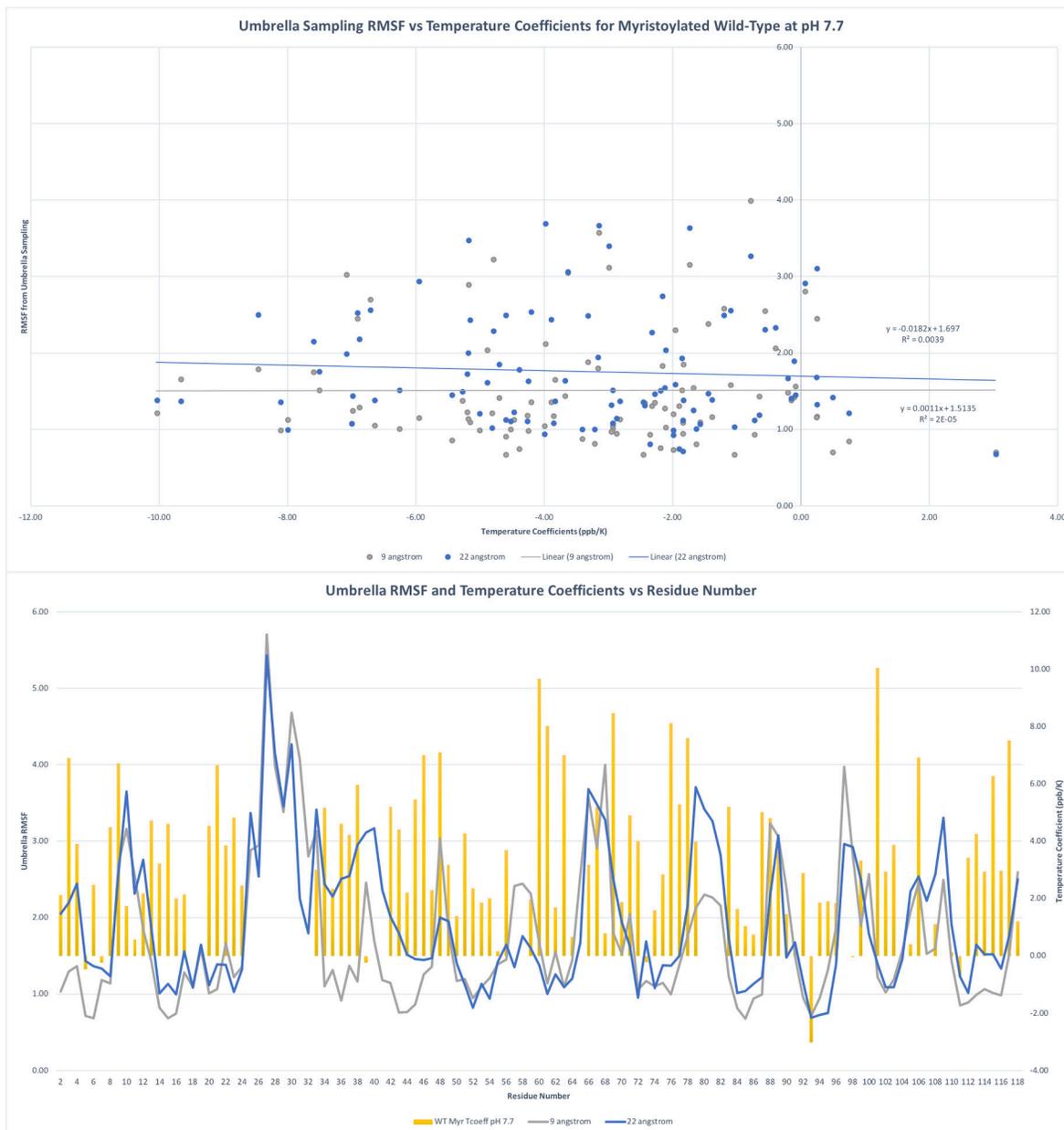


Figure 4.31: Correlations between RMSF values obtained from umbrella sampling and temperature coefficients from myristoylated wild-type hisactophilin at pH 7.7. (Upper panel) Scatterplot of RMSF values with the myristoyl group being held at 9 Å (grey) or 22 Å (blue) from the bottom of the β -barrel, the likely sequestered and accessible states, respectively, vs temperature coefficients. The linear regression for these two RMSF series is shown, and the correlations were both small enough to be considered as having no correlation. (Lower panel) a mixed plot of the RMSF values for 9 Å (grey) and 22 Å (blue), and the temperature coefficients for each residue listed along the x-axis. For ease of comparison, temperature coefficients (yellow bars) were transformed by changing their sign, so here the tallest (most positive) bars represent the least structural stability, which might have matched with areas with the greatest RMSF, though they do not.

the myristoyl group and hisactophilin are dynamic, so these umbrella sampling RMSF values would require some population weighted average of myristoyl depths applied to more accurately reflect temperature coefficients. However, weighting the RMSFs for the sequestered (9 Å) and accessible (22 Å) states based on the expected populations from global energetics [8] did not improve the correlation. There have been some successes with simulating aspects of temperature coefficients with 110 ns of parallel tempering experiments that predicted chemical shifts at varying temperatures [215]. NMR-derived order parameters—measures of very fast dynamics—have been accurately simulated [308] and the performance of various packages for doing so benchmarked, but results diverged from experiment with larger scale motions and variable temperature [309], but simulating longer timescales needed to predict amide proton chemical shifts and other NMR-observables involving structural ensembles is a difficult problem [310]. The current data, however, are a valuable dataset that may help refine simulations of dynamic and switchable systems.

4.6.4.3 Distance to the Myristoyl Group

Interrogating the structural implications of ligand binding or some other protein modification can be examined using chemical shift perturbation analysis [177]. This method extracts information about structural changes that result from the binding/modification event and can be used to highlight residues responsible for directly interacting with the ligand, or to suggest likely allosteric pathways using chemical shift covariance analysis (CHESCA) [186]. Along the same lines, we wonder if, because myristoylation is such a globally stabilizing modification, changes in temperature coefficients may be explained by proximity to the myristoyl group, or if its stabilizing influence is more complex and distributed throughout the protein by some enthalpy/entropy compensation mechanism.

The pairwise distance between each amide proton probe and the myristoyl group in a couple different positions was calculated using a custom Jupyter notebook and the Bio.PDB parser[311]. Distances were calculated between amide protons and the terminal methyl of the myristoyl group and the approximate geometric centre of the fatty acid in energy-minimized structures, one with the myristoyl group 9 Å from the hairpin triplet (sequestered state), and with the myristoyl group in an accessible-like state at 21 Å (left and right panels of Figure 4.32). Pearson and ranked-order correlations (i.e. Spearman) yielded no relationship between the temperature coefficients obtained for myristoylated hisactophilin at pH 7.7 and pH 6.2 (not shown) and distance from the myristoyl group in the three studied conformations. Figure 4.33 rows A-C show poor Pearson correlations for myristoylated wild-type at pH 7.7. Row D, however, does have a weak pattern between the change in temperature coefficient upon myristoylation and the distance to the nearest myristoyl atom, where those residues that line the myristoyl binding pocket (e.g. residues 5-6, 84-86) have more evidence of a relationship than others. This evinces that direct interaction between residues and the myristoyl group accounts for some of the changes in global stability upon myristoylation by increasing structural stability for

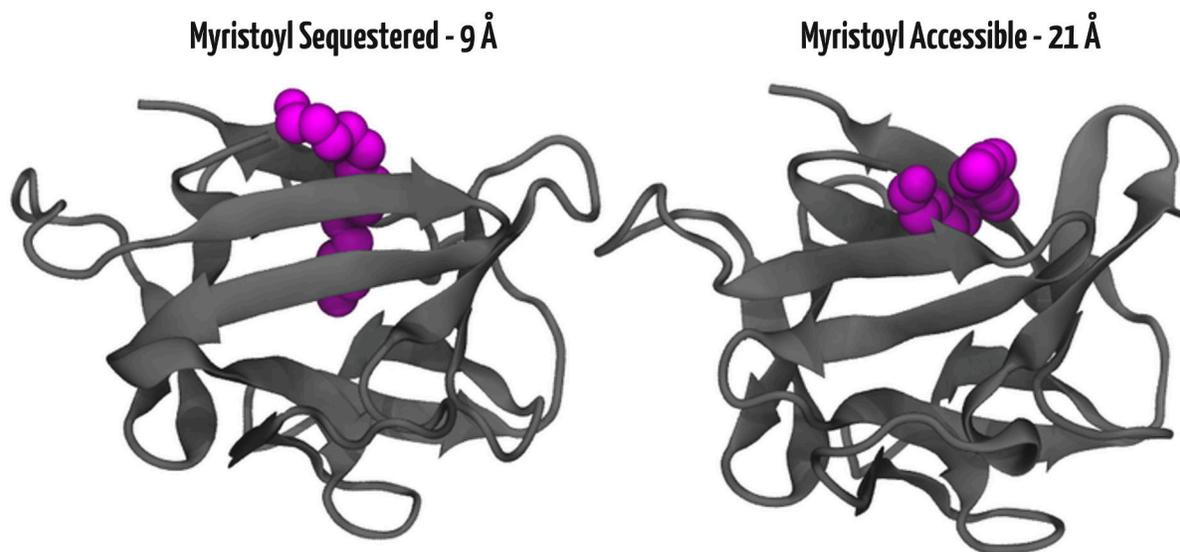


Figure 4.32: Structures of hisactophilin resulting from umbrella sampling. The tip of the myristoyl group is at a distance of $\sim 9\text{\AA}$ (left panel) from the bottom of the hairpin triplet in the sequestered state, and a distance of $\sim 21\text{\AA}$ (right panel) approximates the accessible state. The myristoyl group is shown as magenta spheres.

especially proximal residues, but other stabilizing and destabilizing effects propagate through the structure, perhaps by dynamic/entropic mechanisms, and affecting loops as well as β -strands. Temperature coefficients do appear to be sensitive measures of these subtle thermodynamic changes.

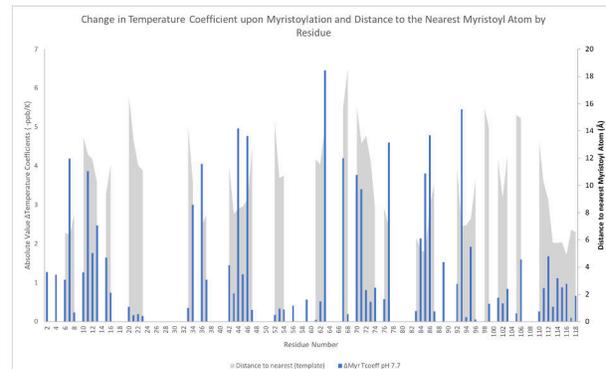
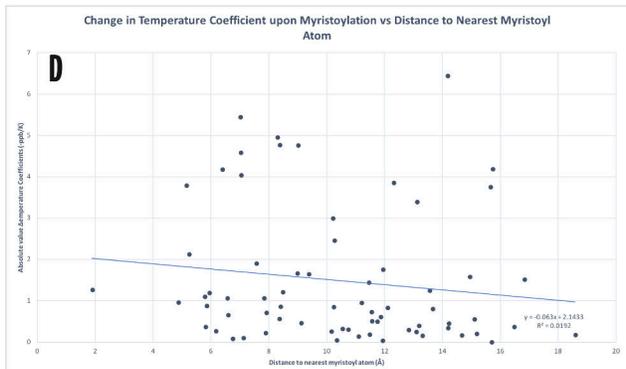
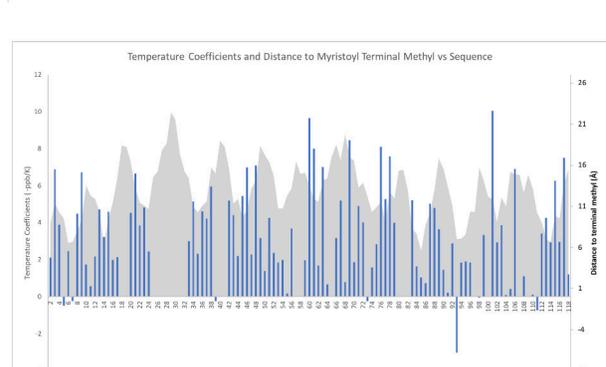
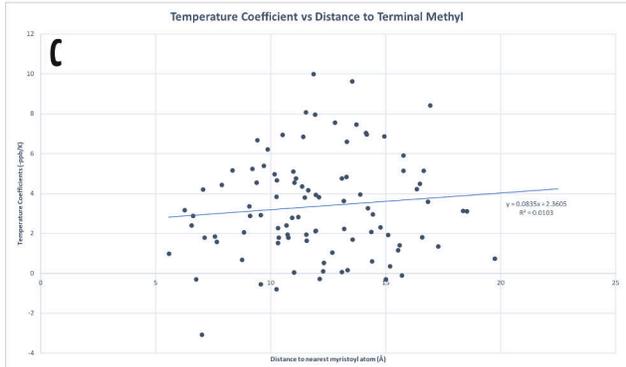
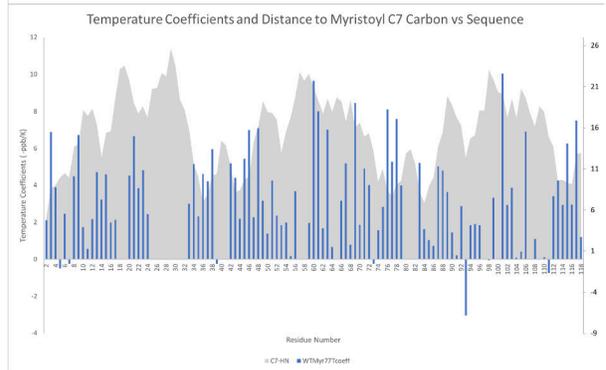
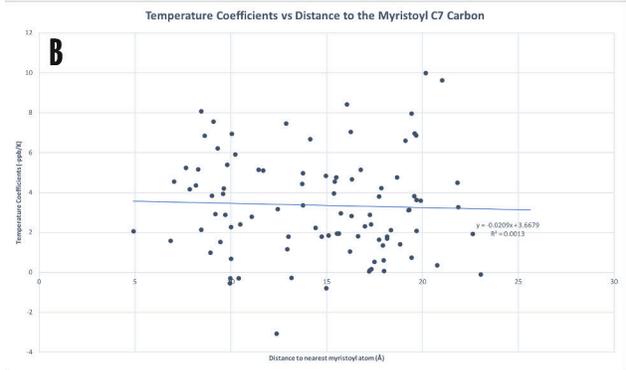
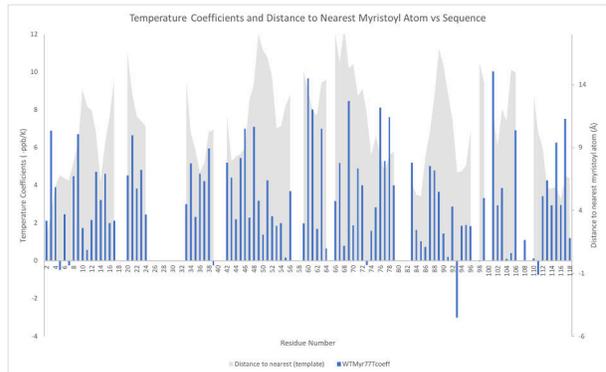
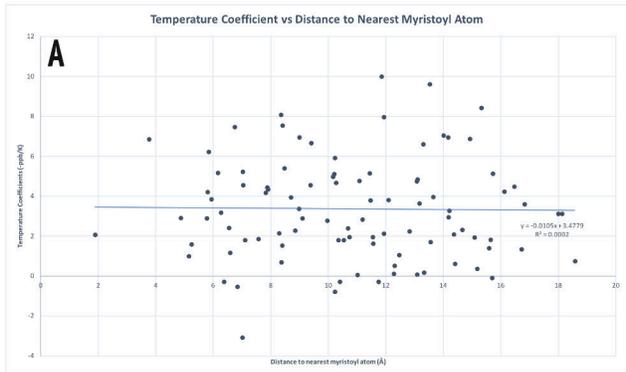


Figure 4.33: (Left column) Scatterplots of the temperature coefficients determined for myristoylated wild-type at pH 7.7 vs the distance to various parts of the myristoyl group; (row A) the nearest atom, (row B) the middle carbon C₇, (row C) the terminal methyl. The model used for calculation was from the umbrella sampling already described, when the myristoyl group was at ~9Å, the appropriate depth for the sequestered state. The same (lack of) correlations to the various depths were evident with the myristoyl group in the accessible state (21 Å) as well, though not shown here. (Row D) looks instead at the whether the change in temperature coefficients upon myristoylation (Δmyr) correlates with the distance to the nearest myristoyl atom. Distance to the C₇ and terminal methyl were also considered, again no trend was observed. (Right column) Combination plots where the blue bars indicate the temperature coefficients for myristoylated wild-type at pH 7.7, and the grey area is the distance between an amide and the given position on the myristoyl group. The scatterplots suggest no meaningful correlation between the two metrics (left column), nor does there seem to be a per-residue pattern of temperature coefficients and the distances (right column). The Δmyr sequence plot does have a few regions that correlate and some of the relative patterns of temperature coefficients are similar to the distance parameters; residues 83-87 for example, which could be logical owing to their especially close relationship with the myristoyl group.

4.6.4.4 Amide Exchange

An established tool in the physical biochemist's toolbox for obtaining site-specific information on local and global (generally slow) structural fluctuations is amide hydrogen-deuterium exchange by NMR, simply called amide exchange going forward. Different rates of exchange of hydrogen for deuterium in a folded protein inform on which areas are protected from exchange by the exclusion of solvent by tertiary structure or by participation of that amide in a hydrogen-bond (e.g. secondary structure). Determining the rates of exchange as a function of temperature allows for some of the thermodynamic parameters associated with any of the relatively large structural fluctuations to be calculated. Exchange rates, protection factors, and the associated thermodynamics were previously painstakingly measured for myristoylated [73,312] and non-myristoylated [88] wild-type hisactophilin.

Amide exchange rates and the calculated protection factors are affected by pH, whether an amide proton is participating in a hydrogen-bond (e.g. secondary structure), or the solvent accessibility of that probe (e.g. tertiary structure) [241,243,313]. Similarly, a major determinant of temperature coefficients is hydrogen-bonding (among the numerous other factors associated with chemical shifts discussed previously), thus it is not unreasonable to wonder if the two measures correlate. As demonstrated in Figure 4.34, temperature coefficients are not correlated with amide exchange rates, protection factors, or the enthalpy, entropy, and ΔG of exchange — correlation coefficients and ranked-order correlations do not exceed $r^2 \sim 0.15$. Residues that exchanged only upon global unfolding [88] (the global exchangers are residues 44, 45, 46, 54, 62, 74, 85, 86, 93, 94, 113, 114, shown in Figure 4.35) make up some of the conserved hydrophobic core of hisactophilin, have temperature coefficients consistent significant local structural stability (Table 4.2 and Appendix A).

The timescales and amplitudes of motion measured by temperature coefficients and amide NH/D exchange are separated by a large chasm, where the former is determined by motions that are relatively fast (up to microsecond to millisecond), and the latter tend to be slower (millisecond to second), a thousand- to million-fold chasm. Amide exchange requires motions that expose labile protons to the solvent and break intramolecular hydrogen bonds;

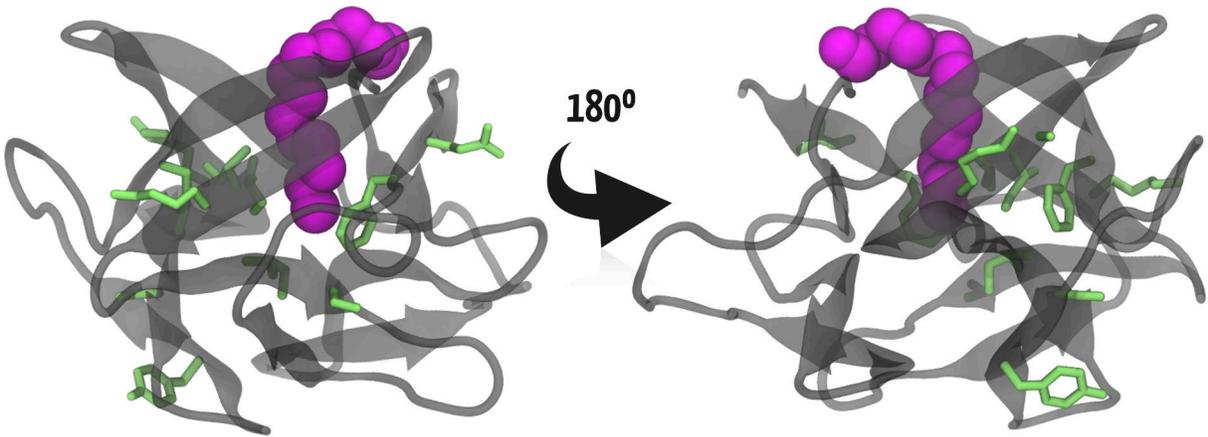


Figure 4.35: Structure of hisactophilin showing the residues that exchange only with global unfolding events, the so-called global exchangers. The right and left panels differ by 180° rotation around the myristoyl group (magenta). The global exchangers are shown as green sticks, they are residues 45, 46, 54, 62, 74, 85, 86, 93, 94, 113, 114.

these can be substantial motions that encompass local fluctuations and global unfolding events. Granted, there is some debate about mechanisms for exchange that involve solvent penetration into the native state, so a grain of sodium chloride and some caution is recommended in interpretation [241,243,313–319]. Measuring temperature coefficients over a range that remains 10–15 K below the protein melting temperature evades significant global-unfolding events [320] and emphasizes thermally induced conformational fluctuations that may involve changes in the vibration/lengthening of hydrogen-bonds or small structural rearrangements and bond angles, which is to say that these motions (and thus temperature coefficients) are not contingent on breaking hydrogen-bonds. Both methods are affected by hydrogen-bonding, however, they may be sensitive to sufficiently different facets of the interaction that they need not necessarily reflect one another. Early studies in temperature coefficients in peptides posited that temperature coefficients also provide information about solvent accessibility [228,229], which is something else temperature coefficients and amide exchange should agree on, but any strong link between solvent accessibility and temperature coefficients has since been downplayed on account of their indifference to pH changes (unless a pH change induces extensive enough conformational/dynamic change) [136,232]. Thus, the lack of correlation with amide exchange data should not be alarming because temperature coefficients may measure a different timescale and involve subtle motions that do not necessitate the disruption of hydrogen bonding or local unfolding of secondary structure.



Figure 4.34: Scatter plots of temperature dependent NMR hydrogen deuterium exchange-derived thermodynamic parameters from fitting to the Eyring model vs the amide proton temperature coefficients for non-myristoylated (upper panels) and myristoylated (lower panels) at pH 6.2 (left panels) and pH 7.7 (right panels). (Red) the change in enthalpy to the transition state of opening [This probably isn't right], (green) change in Gibbs free energy, and (yellow) change in entropy. There is no obvious correlation between the data of the two experiments.

Chapter 5: Conformational Heterogeneity in Myristoyl Switching

5.1 Chapter Abstract

Protein switches are essential in numerous signalling pathways inside the cell, and engineering conformational switching into proteins is a frontier of design because it may involve designing multiple stable, low-energy states with barriers that an environmental change can overcome. Where conformational strain has been identified as a likely driver of myristoyl switching in hisactophilin, the mechanism of how strain is modulated and affects the energy landscape, in addition to other determinants of switching, is an open question. Nonlinear temperature dependence of amide proton chemical shifts is an emerging and straightforward approach for identifying areas of a protein that access functionally relevant low-energy alternate states. We show that there is evidence of a group of residues adjacent to the myristoyl group in the binding pocket that might sample the myristoyl-accessible state at pH 7.7, when the myristoyl-sequestered state is favoured. This supports a population-shift model of switching. High resolution detail of proton binding was obtained by a nuclear magnetic resonance-monitored series of pH titrations to identify which residues' apparent pK_a s are sensitive to switching and may thus modulate conformational strain. We find that a cooperative network of residues throughout the protein couple the myristoyl group and ionization events is dampened in mutants with broken myristoyl switches. These findings highlight the utility of nonlinear temperature dependence of amide proton chemical shifts and their clear application to studying similar switching systems. Furthermore, these results emphasize the importance of robust, high resolution data in understanding and designing advanced systems with switchable conformational states.

5.2 Introduction

There is plenty of evidence that hisactophilin lives a dynamic life and that the myristoyl group is a major effector of its motions [8,73,74], but the contribution of small and rapid movements and vibrations to its role as a pH-sensing switch are still shrouded by the technical difficulty in taking some of those measurements. In Chapter 4, applying a thermodynamic cycle-like analysis to linear temperature coefficients clarified the role of the numerous residues whose local stability is coupled to the myristoyl group and switching. These coupled residues could form a network of communication between sites of proton binding and the myristoyl moiety. The combination of those data with previously published results supports that there is frequent and fast sampling of the sequestered and accessible states, even under conditions where thermodynamics support that one conformation is more favoured than the other (myristoyl sequestered at $pH > 6.95$ and accessible at $pH < 6.95$)[8]. The energy difference between the two states is a paltry ~ 2 kcal/mol, so it is reasonable

that the two states so readily interconvert, however the populations skew toward one of the states. If the two states are accessible and quickly interconverting, i.e. faster than the NMR timescale (say μs -ns), then NMR chemical shifts should represent the population-weighted average of the two and it should be possible to identify residues that are sensitive to or that drive switching between dominant and minor (i.e. excited) conformations.

The strategies to access dynamic data applied herein were both NMR-centric, one leveraging deviations in linear behaviour of the temperature dependence of the chemical shifts acquired in Chapter 4, and pH-titrations similar to those reported previously for wild-type hisactophilin [8] on the two broken switch mutants, LLL and I85L. These techniques reveal the residues that are accessing (presumably) functionally relevant alternate conformations close in energy to the native state and correlate some of them with changes in the apparent pK_a ($\text{pK}_{a,\text{app}}$) of backbone amides. The latter reports on the ionization state of nearby titratable residues and the conformational changes associated with the ionization or switching between sequestered and accessible states, so we gain insight on which ionizations are coupled to the myristoyl group and form the allosteric conduit or, when comparing mutants to wild-type, where coupling is disrupted.

5.2.1 Dynamics and function

To continue the discussion of protein dynamics and their role and relationship to function started in Chapter 4, we'll discuss some of the theoretical considerations behind measuring protein dynamics and a subset of the more popular and widely used methods. When functional aspects of protein behaviour spring from protein dynamics, where higher-energy conformational states are accessed, they can become difficult to measure and characterize because they are, almost by definition, transiently populated by only a small proportion of the protein at any given moment [137–141,143]. The relatively small population of an excited state typically results in the much larger signal from ground state swamping the probe for the excited state. Advances in nuclear magnetic resonance have turned its atomic and amino acid residue-level resolution to cleverly detecting and describing these transiently populated excited states [145]. However, NMR has no single panacean experiment that illuminates all of the fast and slow dynamics and motions of a protein. Functionally relevant dynamics can span the gamut of timescales and a given NMR experiment can generally look only at a subset of timescales at a time (see Figure 5.1), so a diverse toolbox of methods is to the benefit of the biochemist.

The definition of the timescale of protein dynamics is determined by the relative populations and difference in chemical shifts between the ground and alternative states compared to the rate of exchange. If the rate of chemical exchange is much faster than the chemical shift difference between the two states then that motion will fall into the fast exchange regime, manifesting in an NMR spectrum as a single peak at the chemical shift of the population-weighted average of the states (upper part of Figure 5.1). On the other hand, if the rate of exchange is much slower than the chemical shift difference between the two states, then a peak

for each state manifests with intensities that scale with the relative population because a probe resides in a state long enough for the NMR to measure (lower part of Figure 5.1). However, if the rate of exchange between the two states is on the same order as the chemical shift difference, then the exchange rate may enhance the intrinsic transverse relaxation of the signal and broaden the coalesced peak into poorly-defined oblivion (middle part of Figure 5.1). Because the varied timescale regimes each impact the measured spectra differently, the methods for extracting the interesting bits of dynamics are diverse in their approaches and degree of difficulty.

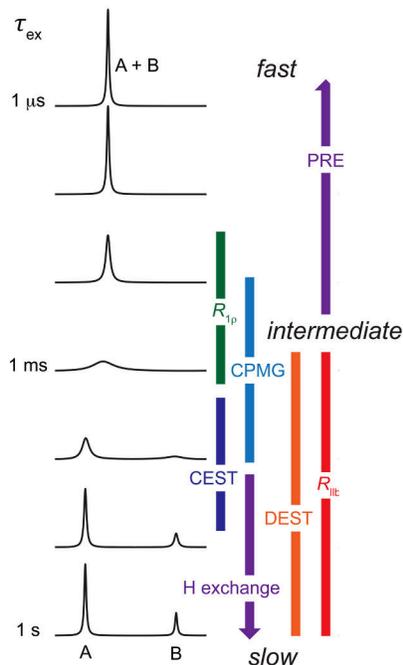


Figure 5.1: Adapted from Anthis & Clore [134]. Shown are seven commonly used techniques for measuring dynamics and exchange with an excited, or dark, state, including Paramagnetic Resonance Enhancement (PRE), Dark-state Exchange Saturation Transfer (DEST), Chemical Exchange Saturation Transfer (CEST), hydrogen/deuterium exchange, Carr-Purcell-Meiboom-Gill (CPMG), rotating frame experiments ($R_{1\rho}$), and lifetime line broadening (R_{1IE}). The left side of the figure shows 1D spectra simulated with the McConnell equations in MATLAB [321] where two states, A and B, are exchanging on timescales of $1\ \mu\text{s}$, $10\ \mu\text{s}$, $100\ \mu\text{s}$, $1\ \text{ms}$, $10\ \text{ms}$, $100\ \text{ms}$, and $1\ \text{s}$ to show how peaks manifest. Here, the chemical shift difference between the two states is $120\ \text{Hz}$, the A state population dominates by a 3:1 ratio, and the intrinsic transverse relaxation rate is $10\ \text{s}^{-1}$. At the bottom of the figure, the slow exchange regime, states give rise to their own peak, however the weaker population might be too small to measure or detect because of the limitations of signal-to-noise. At the top of the figure, where exchange is fast, there is a single peak corresponding to the population-weighted average chemical shift of the two states, and in the intermediate regime, the peak broadens because the rate of chemical exchange is on the same order of transverse relaxation and enhances it.

The spontaneous relaxation of magnetization back to its steady state in the high-field magnet of the NMR spectrometer is a necessary part of the technique and it exerts no small influence on the resulting spectrum in terms of signal intensity and peak width. The transverse relaxation rate is the rate at which the magnetization perpendicular to the external magnet loses coherence and randomizes due to random fluctuations in the local magnetic environment that broaden peaks. These random fluctuations—and inhomogeneities in the magnetic

field—could result from stochastic interactions between tumbling molecules in solution (more for larger molecules), or from non-stochastic interactions of biologically relevant conformational fluctuations occurring on the order of $\sim 200 - 2000 \text{ s}^{-1}$ [134] (which can be extended to $\sim 6000 \text{ s}^{-1}$ with additional chemical shift information [166]). These non-stochastic motions exacerbate transverse relaxation and predictably increase peak width. Carr-Purcell-Meiboom-Gill relaxation dispersion (CPMG RD) experiment [322,323] measures changes in peak-width resulting from conformational fluctuations. The CPMG profile measured for each exchanging/fluctuation nucleus is a complex curve—a function of the rate of exchange, the difference in chemical shifts between the two states, and the repetition rate of the CPMG pulses (the amount of influence exchange is allowed to exert). Analyzing these data is not trivial, but there are thorough and analytical solutions to the underlying functions discussed elsewhere [134,159,324].

Characterizing excited states by CPMD RD requires the state to have a lifetime in the $50 \mu\text{s}$ - 10 ms range but accurately deconvoluting the chemical shift difference of the two states and the relative population terms is impossible. Additional structural information (e.g. the chemical shift) of the excited state or repeating the experiment on a spectrometer with a different field strength (field strength affects the chemical shift difference between the two states) is needed to extract dynamics information from the CPMG RD experiments [325].

Lineshape analysis, also called dynamic NMR, measures motions on the intermediate timescale that occur in the range of $10 \mu\text{s}$ to 100 ms by quantitating changes in peak widths when the system is perturbed. Based on previously published lineshape analysis of hisactophilin, switching is expected to occur on a timescale that ranges somewhere in the intermediate and fast regimes [8]. The order parameter is a long established method for reporting on fast, small amplitude motions of the backbone [326,327] and there has been some success of predicting them by molecular dynamics simulations as well [328,329]. When coupled with paramagnetic relaxation enhancement (PRE), order parameters [330–332] are amenable to characterizing transient excited states [333] by measuring the changes in relaxation properties of major species caused by the strong paramagnetic tag. Where other methods rely on different chemical shifts for the ground and excited states, PRE requires that the two states have different relaxation properties (induced by the positioning of the paramagnetic prosthetic group). The downside is that PRE requires the addition of a paramagnetic centre to the protein, often by covalently modifying a cysteine residue (which may itself have been introduced by mutation) with a spin label. The lifetime of the excited state should be on the order of ~ 250 - $500 \mu\text{s}$ [137]. Interpreting the resulting structural information is also intensive, where measuring the spin-label tag-to-nucleus distances is sometimes required, for example in [334]. Furthermore, high quality structural information of the protein without rapid conformational fluctuations may be necessary, and the excited state must place the spin-label must be closer to the measured probe/nucleus. If these conditions are all met, then PRE is another tool to quantify structural fluctuations on the fast timescale.

Although only a very brief overview of just a few of the plethora of methods for measuring functionally

relevant structural dynamics, selecting the correct experiment may require an existing understanding of the timescales expected to be involved, and the experiments and analysis are not easy (though, nothing worth doing is...). So, there is an appeal to techniques that are fast, measured in tandem with other structural/dynamic experiments, and that do not require extensive existing knowledge or information about the anticipated excited state. Nonlinear amide proton temperature coefficients may just fit the bill...

5.2.2 Nonlinear Temperature Dependence

Linear temperature dependence of amide proton chemical shifts (i.e. temperature coefficients, Chapter 4), offer a glimpse into the local structural stability of each residue in a protein through a relatively straightforward experiment comprising a series of ^1H , ^{15}N HSQCs taken over a range of temperatures terminating well below the melting temperature of a protein. The temperature dependence of amide proton chemical shifts for most amino acid residues follows a straightforward linear relationship (upper panel of Figure 5.2), however there are cases where a straight line is insufficient to describe the relationship between chemical shift and temperature (lower panel of Figure 5.2). A source of the nonlinearity may be conformational heterogeneity from a small population of protein accessing low-lying excited state(s). The conformational richness of the near-native energy landscape contributes to some protein functions, potentially including hisactophilin's myristoyl switch [136,182].

Although requiring only a series of variable temperature HSQC experiments, characterizing low-energy excited states by nonlinear temperature coefficients is an under-utilized approach to studying excited states. Early illustrative examples were to study phosphoglycerate kinase (PGK), lysozyme, and bovine pancreatic trypsin inhibitor (BPTI). Two domains of PGK were thought to move relative to one another for catalysis, however, here they studied the N-terminal domain alone which remained native-like in the absence of the C-terminal domain [335]. Some residues from the N-terminal domain with nonlinear temperature dependences were located in the interface with the missing C-terminal domain, and so might sample conformations relevant to the catalytic movements of the two domains [182]. Other regions with curvature in PGK were localized at the ends of elements of regular secondary structure which might suggest an alternative state with a minor reorganization to that secondary structure—otherwise subtle and difficult-to-detect changes. In lysozyme, some residues with curved temperature dependences formed hydrogen-bonds with side chain acceptors in the native state, thus it is reasonable to figure that the excited state may involve the loss of those hydrogen-bonds. BPTI, a well-characterized system for conformational variability, has a disulphide bond that isomerizes at a rate of $\sim 2,000 \text{ sec}^{-1}$ and significantly disrupts chemical shift of the amide proton of Ala16. Proximal to Ala16, Tyr35 undergoes a 180° aromatic ring-flip correlated with the disulphide isomerization, so a nonlinear temperature dependence of the two residues is likely to result from the two disulphide conformers [182]. In cytochrome C, residues with curving temperature dependences agreed with what is already known of

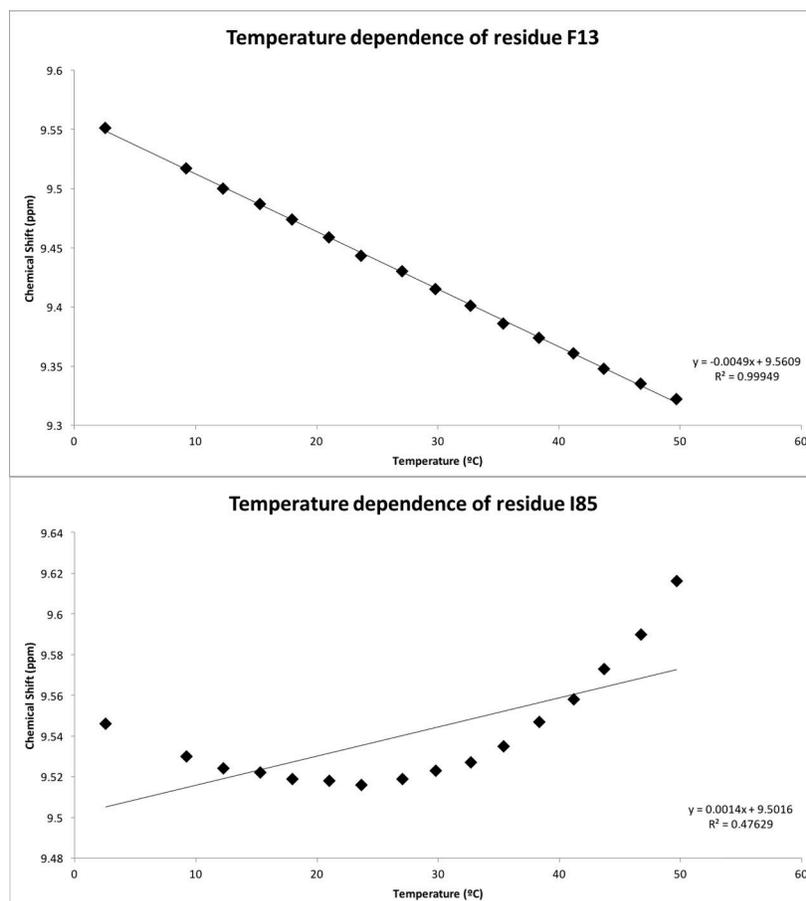


Figure 5.2: Plots illustrating the temperature dependence of the amide proton chemical shifts of two core residues in hisactophilin, F13 and I85. Temperatures were determined from the relationship between the water signal and the internal reference of DSS [259], and chemical shifts were referenced to DSS. Peaks were picked using Bruker TopSpin's peak picking functionality with parabolic interpolation, and plotted in Microsoft Excel. (Upper panel) F13 is demonstrative of a prototypical linear temperature coefficient with a slope of -4.9 ppb/K, corresponding to moderate conformational stability. The lower panel shows the temperature dependence of residue I85, a residue that interacts with the terminal methyl of the myristoyl group. I85 is shown here to demonstrate an especially curved (convex) residue, though there are cases where the deviation from linearity is much less pronounced.

the global unfolding pathway (residues accessing the alternative state were also the first to lose structure upon global unfolding). There was also some curvature in close proximity to Met80, an important haeme ligand that reports folding or changes in oxidation state [336]. In dynein light chain 8, nonlinear temperature dependences were determined at different pH values (NB: germane to the work to follow) and curvature was observed in both monomeric and dimeric forms of the protein. Approximately 25% of the measured residues showed nonlinear behaviour. Some were clustered around the dimer interface which is also a ligand binding site, so it is conceivable that the sensitivity of those residues to environmental changes (like pH) tunes cargo-loading [337,338]. Nonlinear temperature dependences in many of these proteins singled out residues accessing alternative conformations, perhaps involving changed hydrogen bonds or aromatic ring-flips, important for protein function.

Neuronal Calcium Sensor is a myristoylated (like hisactophilin), Ca^{2+} switch protein that was studied in acylated and non-acylated forms [339]. The presence of the myristoyl group increased the number of residues that appeared to access alternative states indicated by nonlinear temperature dependences. Specifically, there were differences in access to excited states around the different Ca^{2+} -binding EF-hands and the conformational heterogeneity was proposed to determine differences in Ca^{2+} affinity. In the studied variant, a stretch of residues whose mutation ablates the Ca^{2+} -myristoyl switch access alternative states, which lends support to the functional importance of the excited state. In the disease-associated protein human SOD1, many residues showed curvature despite the high thermodynamic stability of the fully mature state [150]. And interestingly, the reported conformational heterogeneity was sensitive to mutations in the protein. There was also a perturbation in loop six common among the disease-associated mutants which may operate as a pathway for changes in the protein to propagate into the dimer interface and disrupt protein maturation. These examples further highlight the utility of nonlinear temperature dependences in studying switch proteins as well as that they can report on pathways along which local changes propagate through the protein—important in expanding on the mechanism of myristoyl switching in hisactophilin.

Like the linear temperature coefficients plied in Chapter 4, nonlinear temperature dependences (hereafter referred interchangeably as curvature, possessing curvature, or being curved) rely on chemical shifts as sources of structure and dynamics information. Changes in amide proton chemical shift with temperature may report on changes in hydrogen bonding (e.g. bond length or angle), backbone angles, or relative positioning of aromatic side chains, for example, and those factors can similarly contribute in cases of curvature. Additionally, however, is that residues exploring discrete excited states with populations as small as < 5% can manifest curvature [182]. A ~5% population ensures the alternative state is sufficiently populated to exert influence on the chemical shift, which translates to a free energy difference, $\Delta G_{\text{alternative}}$, between the ground and excited states of 1-5 kcal/mol. For detection, the alternative state must differ in chemical shift from the native state, and it must have a different temperature coefficient with relative populations of the two states

that vary with temperature. Chemical shift differences are fairly trivial because amide proton chemical shifts are exquisitely sensitive to hydrogen bonding, backbone geometry, and the position of nearby aromatics so a small conformational change in an excited state is likely captured by chemical shifts. Mathematically, the chemical shifts of the native and excited states can be represented by equations 5.1 and 5.2,

$$\delta_{native} = \delta_{native}^0 + g_{native}T$$

and

$$\delta_{excited} = \delta_{excited}^0 + g_{excited}T$$

, where δ represents the chemical shift of the native or excited states at temperature T , δ^0 is the reference chemical shift free from the influence of temperature, and g is the temperature coefficient for a given state [336].

If two states are contributing to an observed chemical shift, adding two lines together should result in a third line and not a curve, however, the relative population term has an exponential dependence (vide infra). Equation 5.3 gives the observed chemical shift

$$\delta_{obs} = \delta_{native}f_{native} + \delta_{excited}f_{excited}$$

where δ_{obs} is the observed chemical shift, δ_{native} and $\delta_{excited}$ are the chemical shifts from the previous equation, and f_{native} and $f_{excited}$ are the fraction of molecules in either state. Equations 5.3 & 5.1 highlight that the influence of the excited state on the observed chemical shift is affected both by its population and the chemical shift difference from the native state. The aforementioned exponential term comes from the Gibbs free energy, given in Equation 5.1,

$$\frac{f_{excited}}{f_{native}} = e^{\frac{\Delta G}{RT}}$$

where ΔG is the Gibbs free energy difference between the excited and native states, R is the gas constant, and T is the temperature. Small changes in temperature can lead to an exponentially larger influence on the observed chemical shift of the excited state, resulting in a curved temperature dependence. Upon mutation, curvature of a residue may be attenuated because of a smaller difference in chemical shifts between the native and excited states, e.g. they are now more structurally similar; the temperature dependences of the two states are also more similar, e.g. switching to the alternative state does not impact the local structural stability; new alternative states become accessible whose temperature coefficients are similar to the native state; and lastly, the excited state could increase in energy, and the larger ΔG may also reduce the apparent curvature. Several groups have simulated the curvature for chemical shift, temperature gradient, and energy differences [150,182,336,337,339,340]. The energy difference between the two states, however, is the most influential

determinant of the extent of curvature [339]. In hisactophilin, excited states that curvature reports on could be switching-related because ΔG_{switch} is 2.03 kcal/mol, within optimal 1-5 kcal/mol range.

When chemical shifts are plotted against temperature, in principle curvature can manifest as either concave or convex (lower panel of 2). Convex is more common as it typically results from natives state chemical shifts greater than random coil whose excited state moves toward random coil. Concave may result when the chemical shift of the excited state is downfield of the ground state (more common for helical residues) [182]. Concave residues could also indicate a more structured excited state for native state resonances downfield of random coil. Thus, generally, the direction of curvature can give us additional structural information about the excited state.

Although straightforward and widely applicable, like other NMR measurements, nonlinear temperature dependences are not a panacea for dynamics, they are limited to motions that are faster than chemical shift differences. In hisactophilin, the expected range of timescales for switching is accessible by curvature experiments, and ΔG_{switch} (2.03 kcal/mol) is within the range of 1-5 kcal/mol detectable by curvature. Therefore, residues demonstrating nonlinear temperature dependences are good candidates as being sensitive to and reporters of myristoyl switching. Identifying residues accessing an excited state, perhaps the switched state, provides a window to identify the role of dynamics in hisactophilin's switch and shed light onto allostery.

5.2.3 Titratable residues

As a protein with many ionizable residues (e.g. 31 histidines) and a myristoyl switch tied to small changes in environmental pH, it is of particular interest to examine the $pK_{a,\text{app}}$ of the titratable residues in hisactophilin and monitor their change in response to switching. Altered $pK_{a,\text{app}}$ s can reflect numerous changes in the protein, including changes in local electrostatics [212,341,342], through-bond effects[343,344], conformational change[343,345,346], or the disruption of cooperative/linked equilibria between titratable residues [224,347–349]. Any or all of such changes may indicate changes to the myristoyl group and its preferred environment, sequestered or accessible. Further, changes to $pK_{a,\text{app}}$ may offer insight into the mechanism of the switch; which residues are coupled to the myristoyl group and if disrupting the switch upsets the ionization equilibria of the incognito responsible switch proton(s) binding site. The pH-dependence of chemical shifts has already been determined in the Meiering lab for non-myristoylated [74] and later for myristoylated [8] hisactophilin, so here I follow up by continuing those experiments on myristoylated and non-myristoylated variants of I85L and LLL, mutants with a broken pH-dependent myristoyl switch, to highlight the role of shifting $pK_{a,\text{app}}$ on switching and the correlation to low-energy excited states detected through nonlinear chemical shift temperature dependences.

5.2.3.1 NMR-monitored pH titration

In many experimental contexts, NMR excels in providing residue/site-specific information about a protein, whether that is structural information, temperature dependences, dynamics information, or the $pK_{a,app}$ of ionizable moieties. Because ionization equilibria may impact protein stability, catalysis, and broader function, acquiring higher resolution electrostatic information than what canonical bulk measurements like isothermal titration calorimetry can provide is highly valuable for detailing molecular mechanisms of function, like pH-dependent switching in hisactophilin. However, the labile protons usually involved in acid/base equilibria can pose a problem for NMR, their rapid exchange with water when the moiety is exposed to solvent can suppress their signal [243,349,350] and make them inherently difficult or impossible to measure without resorting to low temperatures, restricted pH ranges [351], and/or specialized pulse sequences that minimize perturbing water magnetization [352]. Detectable residues that do not rapidly exchange with water tend to be buried and shielded from solvent and these specific environments may deflect their pK_a s from those measured for reference compounds [349] in ways that may inform on their function [290].

Labile protons that are not protected from solvent are still of great interest but may require an indirect approach to measuring their macroscopic $pK_{a,app}$ s [74,224]—typically achieved by observing the pH-dependent chemical shift change of nearby, non-labile nuclei like amide protons, amide nitrogens (^{15}N), or any of the available carbons (^{13}C). The changing charge of the titratable group due to adjustments of the solvent pH affects the chemical shifts of nearby residues because of their exquisite sensitivity to small structural alterations that can accompany electrostatic changes [341,353]. The effect of the titration could be transmitted by through-bond induction, through-space electrostatic fields, or by pH-induced conformational changes [224,345]. Many of the surrounding, non-labile nuclei may act as reporters of the titration event, however, with this approach it may be more difficult to assign observed titrations to a single titratable residue with high confidence. Because the effects of a titration can propagate a long way through a protein, reporter nuclei may be party to more than one titration, so extracting microscopic pK_a or $pK_{a,app}$ values needs to be approached with care. Thus, measuring pH-dependent chemical shift changes of backbone amide protons can report on titrating histidines in hisactophilin without measuring them directly.

NMR-monitored pH titration experiments generally involve a series of 1D or 2D experiments (e.g. ^1H , ^{15}N HSQC) over a range of pH values where the protein is functional (or switches...) and stably folded. 2D experiments are advantageous for proteins because of improved signal dispersion that eases tracking resonances as they move with pH and yield arguably better determined chemical shifts [342,354]. A recommendation is that spectra are recorded at least every ~ 0.2 pH units to ensure that the resulting titration curves are accurately and precisely fit while allowing for any multiphasic transitions to also be well-defined [355]. The pH range should be as wide as feasible (bearing in mind protein stability and spectrometer time) so that the baselines and transitions of any titrations are as sampled as possible. Protonation events generally occur at the fast exchange limit so resonance assignments obtained at one pH are readily transposed to spectra at

subsequent pH values. Spectra should also be referenced to some internal standard like DSS to account for any small temperature variation or for a pH-dependent shift in lock signal [355]. After carefully acquiring and processing spectra as described above, $pK_{a,app}$ s are extractable by nonlinear fitting to a suitable equation; in the case of a reporter of a single titration, the chemical shifts are weighted by the Henderson-Hasselbalch equation, Equation 5.4,

$$pH = pK_a + \log \frac{[A^-]}{[HA]}$$

and fit to Equation 5.5

$$\delta_{obs} = \frac{\delta_b + \delta_a \times 10^{pK_{app}-pH}}{1 + 10^{pK_{app}-pH}}$$

, where δ_{obs} is the measured chemical shift for a given nucleus at a given pH, δ_b is the chemical shift at the high pH limit, and δ_a is the chemical shift at the low pH limit, and $pK_{a,app}$ is the apparent pK_a being determined [347]. Should a nucleus observe more than one titration event, expressions involving multiple pK_a s can be used, as in Equation 5.6,

$$\delta_{obs} = \frac{\Delta\delta_a}{1 + 10^{pH-pK_{a1}}} + \frac{\Delta\delta_b}{1 + 10^{pH-pK_{a2}}} + \frac{\Delta\delta_c}{1 + 10^{pH-pK_{a3}}}$$

, where δ_{obs} is the observed chemical shift at a measured pH, $\Delta\delta_a$, $\Delta\delta_b$, and $\Delta\delta_c$ are the chemical shift changes associated with each of the three pK_a values [224]. Unusual or multiphasic titrations can also be fit to a version of the single pK_a equation but with a modified Hill coefficient to account for steeper or shallower than usual titrations, as in Equation 5.7[356],

$$\delta_{obs} = \frac{\Delta\delta_a}{1 + 10^{n(pH-pK_a)}} + \delta_{offset}$$

, where δ_{obs} is the observed chemical shift at a measured pH, $\Delta\delta_a$ is the chemical shift change described by the pK_a , δ_{offset} is the chemical shift at the upper pH limit, and n is the Hill coefficient.

Herein, the $pK_{a,app}$ s for residues in the broken switch mutants LLL and I85L were determined to measure interactions between proton binding and switching compared to wild-type. Titrations curves of the variants of the two hisactophilin mutants were fit, where possible, to a single $pK_{a,app}$ but if that model did not describe the data well, then they were fit to the modified Hill equation to account for the possibility of the nucleus reporting on two overlapped titrations with very similar $pK_{a,app}$ s. In the previously determined $pK_{a,app}$ values for wild-type hisactophilin, burying the myristoyl group causes the $pK_{a,app}$ value for six particular residues (L45, G56, H75, I85, H91, and Y92) to shift by approximately 1 pH unit, from \sim pH 7.0 to \sim 6.0 [8]. So, will two mutants supported by global thermodynamic measurements (Chapters 2 & 3) and amide proton temperature coefficients (Chapter 4) as having ablated pH-dependent switches also have their apparent acid/base equilibria decoupled from myristoyl switching?

5.3 Materials and methods

This section pertains to NMR-centric experiments acquired simultaneously with and the result of variable temperature experiments, or by an arduous series of pH titration experiments. Thus, the purification and sample preparation of ^{15}N -labeled protein is essentially identical to the process outlined in Chapter 4 except where indicated.

5.3.1 Expressing ^{15}N -Labeled Hisactophilin

As described in Chapter 4.

5.3.2 Purifying ^{15}N -labeled hisactophilin

Employed the same purification technique as described in Chapter 4.

5.3.3 Measuring nonlinear temperature dependences

Nonlinear temperature coefficients, for the purposes of inferring conformational heterogeneity, involved an alternate analysis of a subset of the temperature dependence of chemical shifts experiments acquired in Chapter 4. They did not require an independent experiment, so the methods used can be found in Section: 4.3.3.4 Variable Temperature Nuclear Magnetic Resonance Experiments, and a more detailed discussion of the considerations for maximizing the success of the experiment and confidence in defining nonlinearity are given in Chapter 6.

5.3.4 Analyzing nonlinear temperature dependences

Identifying residues that exhibit nonlinear temperature dependence of amide proton chemical shifts first requires tracking chemical shifts as a function of temperature and fit to a straight line, exactly as described in Chapter 4. In principle, the residuals of the linear fit can be visually inspected for those that have a systematic and non-random distribution around 0, usually manifesting as a parabola. Depending on the protein and conditions under investigation, some residuals are on the verge of whether or not there is significant evidence of nonlinearity. While a practiced and experienced eye might reach a reasonable qualitative conclusion, we instead developed a methodology to try to remove subjectivity and determine the statistical significance of whether a linear fit sufficiently describes the data. This is discussed in greater detail in Chapter 6, with key aspects elaborated below. In essence, an R script was written [357] to apply an F-test to the fit of the residuals to linear and polynomial regressions—if the linear fit is substantively worse than the polynomial ($p < 0.01$),

then the data are deemed to not be sufficiently well-described by a linear model. In an attempt to increase the robustness to experimental noise or uncertainty in chemical shifts (arising perhaps from overlapping peaks), model evaluation was bootstrapped (a method for statistical resampling) to reduce the influence of any outliers that might anchor the fit and exert an undue influence on the regression and residuals. What that means is that for each residue, a set of temperature dependence data comprising 14-17 data points is randomly sampled with replacement to construct a new data set that are subjected to the F-test described above. The resampling could result in some points being absent from a given fit while others exist in duplicate, triplicate, or more (rare). The process was repeated another 999 times so that each residue was tested 1000 times, where a subset of those repetitions have had any anchoring points omitted and so their overall effect on determining nonlinearity diminished.

The scores resulting from the goodness-of-fit test can then be filtered by the critical value (α) of your choosing. Our most sensitive and refined workflow yielded good discrimination and agreement with visual inspection at $\alpha = 0.01$. Picking the appropriate critical value (α) requires considering and possibly measuring how likely apparent curvature is to arise from noise in the data, so the better determined the chemical shifts and the more points included in the fit, the more discrimination is possible, but we have found that $\alpha = 0.05$ or $\alpha = 0.01$ are sufficient (for hisactophilin, threefoil, and human super oxide dismutase 1, at least).

After evaluation by the bootstrap script, the results were treated as a binary yes/no answer to the question, "Is there curvature in the residual?". Residues that tested positive were mapped onto the structure using VMD for visualization as described for linear temperature coefficients [52] to identify any spatially relevant clustering. Beyond that, the extent of curvature, or the a-value from a quadratic fit, can report on different aspect of conformational heterogeneity, for example the relative populations/the energy difference between near-degenerate conformations, or the chemical shift difference between the states. So, changes in curvature upon myristoylation, pH change, and mutation were studied for structural/thermodynamic insight into sparsely populated excited states and myristoyl switching.

5.3.5 NMR-monitored pH titrations

NMR samples for the pH titration were conducted by dissolving ~16 mg of one of lyophilized myristoylated or non-myristoylated hisactophilin mutants F6L/I85L/I93L or I85L into 750 μ L of 50 mM potassium phosphate containing 1 mM EDTA, 1 mM DTT, 1 mM DSS, and 10% D₂O, with the pH carefully adjusted to close to 10 (Orion™ Ross™ Sure-Flow™ pH electrode, Fisherbrand™ accumet™ AB15 base) once all components were present and dissolved. The protein in buffer solution was transferred to a 5 mm NMR tube appropriate for the triple resonance TXI probe (Bruker, Billerica, MA)-equipped Bruker Avance 600 MHz spectrometer. Data acquisition and spectra processing were handled by Bruker XWIN-NMR and Bruker Topspin 3.5. Two stocks of hydrochloric acid (Bioshop, Burlington, ON) for adjusting the pH through the

course of the experiment were prepared at 200 mM and 40 mM. Each step of the titration involved thermally equilibrating the sample to 298 K inside the spectrometer, measuring a 1D spectrum with water suppression by excitation sculpting [258] for referencing DSS, measuring an 8-scan ^1H , ^{15}N HSQC [248–251], removing the sample from the NMR tube, reading and adjusting the pH with one of the two HCl stocks (whichever allows for adding small volumes on the order $\sim 1 - 15 \mu\text{L}$), returning the sample to the tube, replacing it in the magnet, and repeating the process. The target for pH increments was $\sim 0.15\text{-}0.20$ pH units from pH 10.0 - 7.80, and slightly smaller increments of 0.10 - 0.15 from pH 7.80 - 5.80 where the various histidines of hisactophilin are most likely to titrate and pH-sensitive spectral changes are most pronounced. To compensate for sample volume loss on the pH probe, additional buffer mixture was periodically added to the LLL samples, however the I85L samples were made in a larger starting volume so there was no need to compensate for volume losses over the course of the experiment.

5.3.6 Analyzing NMR pH Titrations

1D ^1H and 2D ^{15}N , ^1H HSQC spectra were processed in the same way as for linear temperature coefficients: using Bruker TopSpin v 3.5; 90° shifted sine bell window function, quad baseline correction in the direct dimension, linear prediction of the indirect dimension, and re-referenced to DSS. Processed spectra were then imported into CCPN V2.4 (Collaborative Computing Project for NMR)[263] as a data series and resonance assignments were transposed from those determined in Chapter 4 to their corresponding pH values for the myristoylated and non-myristoylated forms of I85L and LLL. Resonance assignments were propagated to subsequent pHs by inspection and chemical shifts were extracted and plotted against the measured pH. In Origin (OriginLab, Northampton, MA), titrating residues were fit to a model for a single titrating residue [74,224,347], Equation 5.5. Where necessary and possible, titrating residues were also fit to a modified Hill equation, Equation 5.7 that is better equipped to cope with irregular titration curves [224,356]. The resulting $\text{pK}_{\text{a,app}}$ s and changes in $\text{pK}_{\text{a,app}}$ ($\Delta\text{pK}_{\text{a}}$ upon myristoylation or mutation) were mapped onto hisactophilin's structure by writing the values into the B-factor field of the PDB file and visualized in VMD [52].

5.4 Results and Discussion

5.4.1 Curvature in Wild-Type

Mapping the residues with nonlinear amide proton temperature coefficients, indicating sampling of low energy excited states, onto hisactophilin's structure may identify clusters of spatially proximate residues that access some functionally relevant excited state. It is possible that a lone residue accesses a single excited state, however, there is more precedent for small stretches of amino acids (e.g. stretches of five) that together

access the alternative state [336]. Hisactophilin is a protein with many functional hats because although it undergoes a pH-dependent myristoyl switch for reversible membrane binding, it is also regulated by phosphorylation[[39]; [358]], binds and bundles actin[358,359], and may also function as an osmoprotectant [66]. Because of its diverse functions, it is not possible to immediately point at a residue (or stretch of residues) that participate in conformational heterogeneity and conclude that it is important in switching because it is possible that some dynamics might be innate to β -trefoils, some for actin-binding, others for binding a kinase, etc. So, using a similar cycle formalism to analyze curvature as we used for temperature coefficients and global energetics (Figure 3.2), we will consider the change in residues with curvature upon myristoylation or pH change. Residues with curvature are listed in Appendix A as well as A-values (extent of curvature) for residues deemed significantly curved.

5.4.1.1 Curvature in Non-myristoylated Hisactophilin

Non-myristoylated hisactophilin at pH 7.7 (lower left panel of Figure 5.3; all curved residues are given in Appendix A) has 27 residues with curved residuals at $p \leq 0.01$ based on the bootstrapped model selection test. These residues are spread out over the whole protein with a mild tendency to localize to the ends of regular secondary structure which, as other have observed [182,339], may indicate an alternative state where the ends of secondary fray or undergo a minor rearrangement in hydrogen-bonding. This behaviour could be consistent with a highly dynamic protein like hisactophilin. Of the total number of residues that were tracked with confidence over the temperature range (72 residues), more than 35% are curved—much more than other proteins, which have proportions closer to 20% [150,182], and closer to the 28% that was seen for another myristoylated protein, neuronal calcium sensor 1 [339].

The direction of curvature, convex or concave, may provide additional structural information on the excited state; convex curvature can arise as a result of an excited state with a chemical shift upfield of the ground state, which for a predominantly β structure would suggest the excited state is closer to random coil and is the more common manifestation of curvature [[150]; [177]]. Concave curvature, on the other hand, can result when the excited state has a chemical shift downfield of the ground state, which in β -structure suggests that the excited state is more structured (or, at least, farther from random coil). In this variant of hisactophilin at pH 7.7, 21 residues are convex and 6 are concave. 5 of the 6 concave residues are concentrated in trefoil one, namely residues R4, S8, F13, S15, and K22. The sidechains of these residues generally point away from the core of hisactophilin and may instead interact with the loop that spans residues 25-33 (unassigned in all spectra). Given the suspected flexibility and dynamic nature of this loop (see the simulated RMSF by umbrella sampling in Figure 4.4), it may be that an excited state of the loop interacts favourably with these N-terminal residues. The other 21 residues, however, manifest the more common convex shape and are likely indicative of the expected more random coil-like excited state (which could simply mean a slightly longer

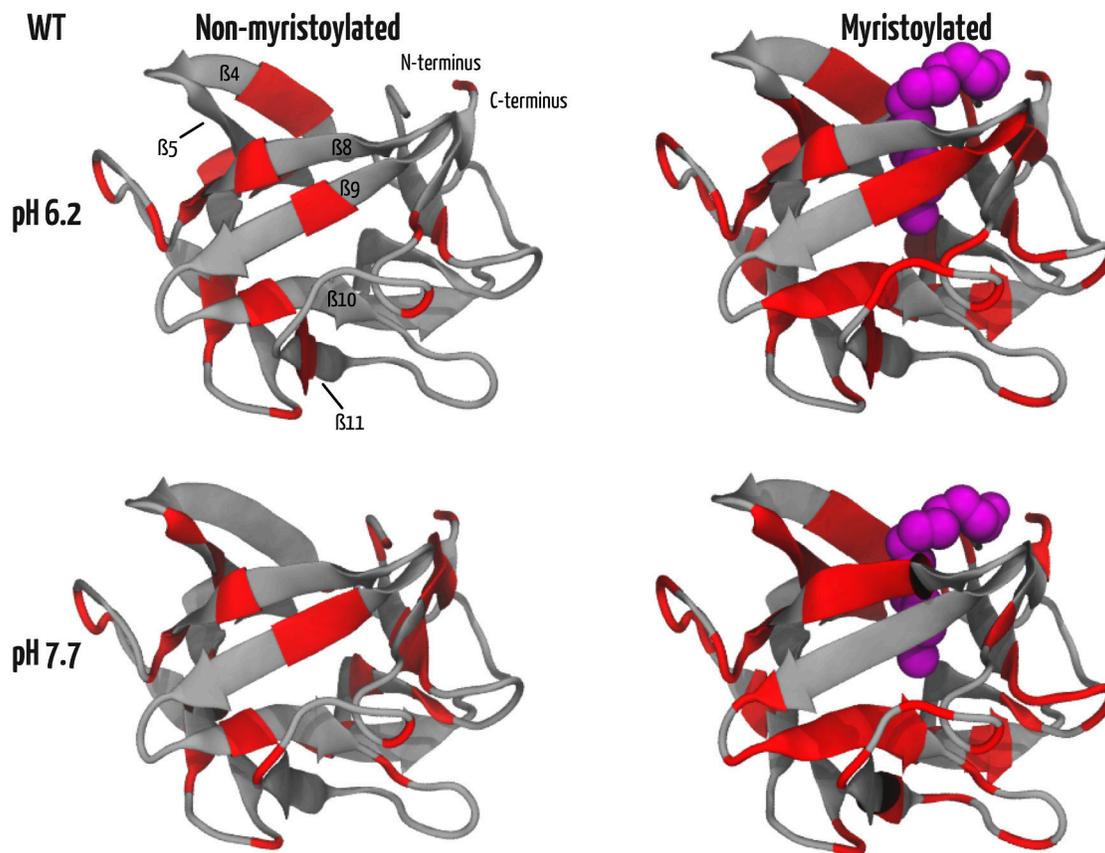


Figure 5.3: Structures of wild-type hisactophilin with and without the myristoyl group (in magenta) illustrating residues that have measurable curvature in the residual resulting from a linear regression at $p \leq 0.01$ (shown in red) for a statistically resampled dataset with replacement. Data from myristoylated hisactophilin are on the right and non-myristoylated on the left. The top row was measured at pH 6.2 and the bottom at pH 7.7. Similar to other proteins, residues with curvature are distributed throughout the protein with a slight enrichment toward the ends of regular secondary structure. Non-myristoylated at pH 6.2 had 23 curved residues and 27 at pH 7.7, while myristoylated at pH 6.2 had 48 curved residues and 46 at pH 7.7.

hydrogen bond or more oblique hydrogen bond angle).

Lowering the pH to 6.2 for the non-myristoylated protein revealed that many of the same residues showed curvature as at pH 7.7 (black residues in Figure 5.5). There was a similar trend of a higher propensity for residues at the boundary of regular secondary structure to show curvature. Particularly interesting changes in curved residues occur in the aforementioned first trefoil (yellow residues in Figure 5.5); at pH 6.2, only S8 remained curved while R4, F13, S15, and K22 appear to have lost their conformational heterogeneity on this timescale/energy level. It bears mentioning, however, that the absence of evidence is not evidence of absence; these residues may continue to be dynamic and access alternate conformations on a slower timescale or have a greater energy gap relative to the ground state and so be invisible by variable temperature NMR experiments. However, previous hydrogen-deuterium amide exchange experiments [88] found that S8 had a ΔG_{ex} on similar to that of residues that exchange only upon global opening/unfolding (10.2 kcal/mol for S8 compared to 10.2 - 11.1 kcal/mol for other global exchangers at pH 7.8), so this does not support a shift in dynamics to a slower regime. A loss of conformational heterogeneity here may suggest that alternative conformations involving trefoil 1 and the long $\beta 3$ - $\beta 4$ loop are unlikely. Residues 9, 10, and 12 are histidines in the $\beta 2$ - $\beta 3$ loop, and although there is no $\text{pK}_{\text{a,app}}$ information available for residue 9, the $\text{pK}_{\text{a,app}}$ for H10 is 7.2 and H12 is 6.47 [8,312], so it is conceivable that their protonation has created some coulombic repulsion with the $\beta 3$ - $\beta 4$ loop or induced some other conformational change that disfavors conformational heterogeneity.

Another interesting change occurs in strand $\beta 4$ at the opening of β -barrel in trefoil 2, residues V36 and E37 become curved at lower pH (violet in Figure 5.5). In isolation, this observation is not particularly interesting, however, V36 is implicated in switching in Chapter 4 by conformational tension, and here it appears as though this might relate to its manifestation of conformational heterogeneity and relieving that tension. The curvature here is concave for V36, which may support hisactophilin forming a locally more structured excited state that encourages interactions with the myristoyl group (that is not present here). The extent of curvature is only slight, so it suggests this excited state is closer the limits of detectable energy differences (either ~ 1 or ~ 5 kcal/mol, as others have found [150,339]). When myristoylated (*vide infra*), V36 is yet more curved (and still concave) which is consistent with a reduction in the energy difference between excited and ground states because of now-present favourable interactions with the myristoyl group. These results suggest that some of the dynamics and conformations required for switching are pre-existing and accessible to non-myristoylated hisactophilin, though the energetics of sampling them are less favourable.

A final consideration for the pH 6.2 curved residues pertains to those in the hairpins that close the bottom of the β -barrel. An additional six residues become curved upon proton binding and this may relate to sampling a conformation that is actin-binding competent. However, without high-resolution actin-binding data, for the time being this suggestion remains speculation.

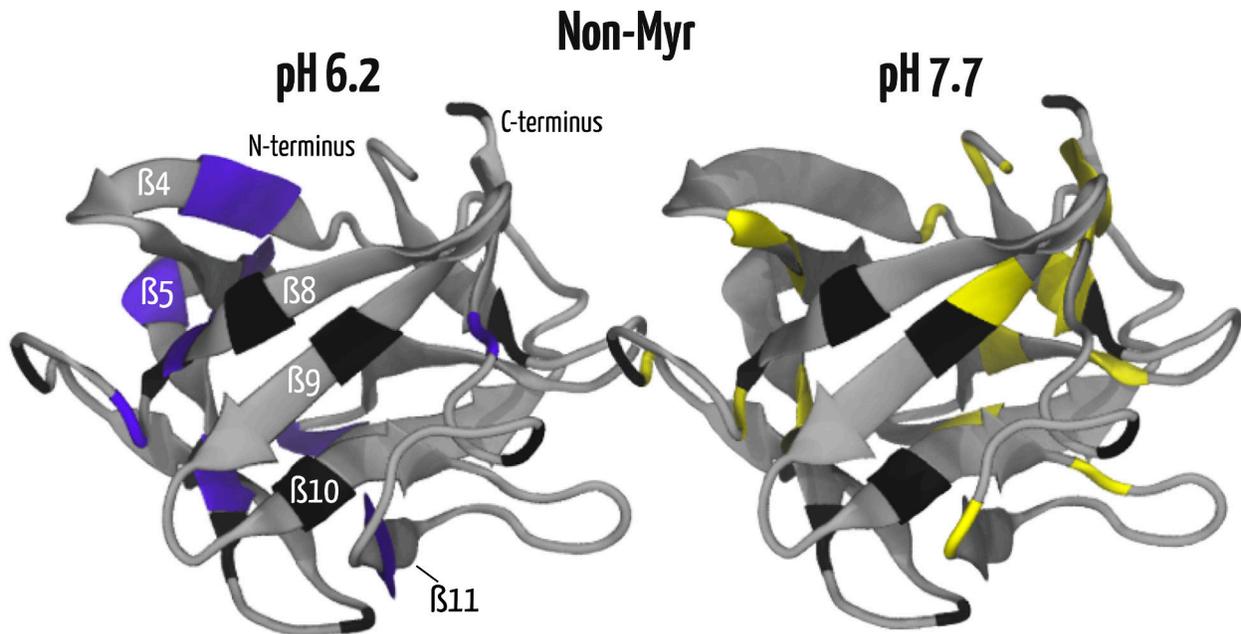


Figure 5.5: Residues that experience curvature in non-myristoylated hisactophilin, coloured to highlight which residues are curved at both pH values (black residues), and where curvature is gained or lost. At pH 6.2 (left panel), residues in violet are those whose curvature is unique to low pH. At pH 7.7 (right panel), residues in yellow are uniquely curved at high pH. There is a concentration of yellow residues in trefoil one (discussed in main text) which disappear at pH 6.2 and there's an increase in curvature in trefoil two at pH 7.7.

5.4.1.2 Curvature in Myristoylated Hisactophilin

Myristoylated hisactophilin at pH 7.7 (lower right panel of Figure 5.3) has 46 residues whose temperature dependences cannot be described by a straight line, significantly more residues than the non-myristoylated form. Interestingly, increased curvature was also observed upon myristoylation for Neuronal Calcium Sensor 1 [339]), and this observation for hisactophilin fits with the finding that myristoylation increases hisactophilin dynamics, owing to the strain of an overpacked core [8].

Residues showing curvature at pH 7.7 for hisactophilin localize to the ends of secondary structure like in the non-myristoylated form, but notably are also located in loops and the middle of secondary structure too. Thus, the dynamic repercussions of the sequestered myristoyl group is propagated throughout the protein and felt by approximately half of the assigned residues. Based on linear temperature coefficient analysis in Chapter 4, the mechanism of allostery and switching was posited to be via a population shift driven by relaxing the conformational tension induced by sequestering the myristoyl group into an overpacked core at high pH; so, a small proportion of hisactophilin sampled the myristoyl-accessible state. The curvature of some residues at bottom of the myristoyl binding pocket (H88, H90, H91, I93, V101) indicates they access a potentially more structured excited state. If the myristoyl group is sampling the accessible state and relaxing tension in the core is consistent with residues that interact transiently with the tip of the myristoyl group becoming

more structured in the absence of the source of tension (the myristoyl group). Residues V36 and E37 and the top of the β -barrel in trefoil 1 appear to access a more structured excited state at pH 7.7, perhaps owing to favourable interactions with the myristoyl group transiently favouring the top of the β -barrel (e.g. right panel of Figure 4.32).

Proton-binding to myristoylated hisactophilin shuffles the identity of some of the curved residues, however, the total number of residues is nearly constant at 48. Many residues are curved at the two pH values (black residues in Figure 5.6), particularly strand β 10, forming part of the deepest part of the myristoyl binding pocket and home to I93 (one of the altered residues in the broken switch mutant F6L/I85L/I93L). Also shared are stretches of strands β 4 (residence of V36 and E37) and β 5 (V43, A44, L45) which may form favourable interactions with the myristoyl group in the accessible state. Unique to pH 6.2 is a stretch from K82 to I85 in β 9, as well as H107 and T112 that pack against β 9 (Figure 5.6). At pH 6.2, the majority of the population is likely in the myristoyl-accessible state, the excited state could constitute adjustments to the core to allow transient sampling of the sequestered state to mitigate the energetic penalty of solvating some of the hydrophobic myristoyl group. Or, because the accessible state likely corresponds to a membrane-binding competent conformation, heterogeneity in the upper sections of the barrel may allow positively charged, basic residues extra freedom to move and interact with the negative charge of a membrane inner leaflet and begin anchoring [358].

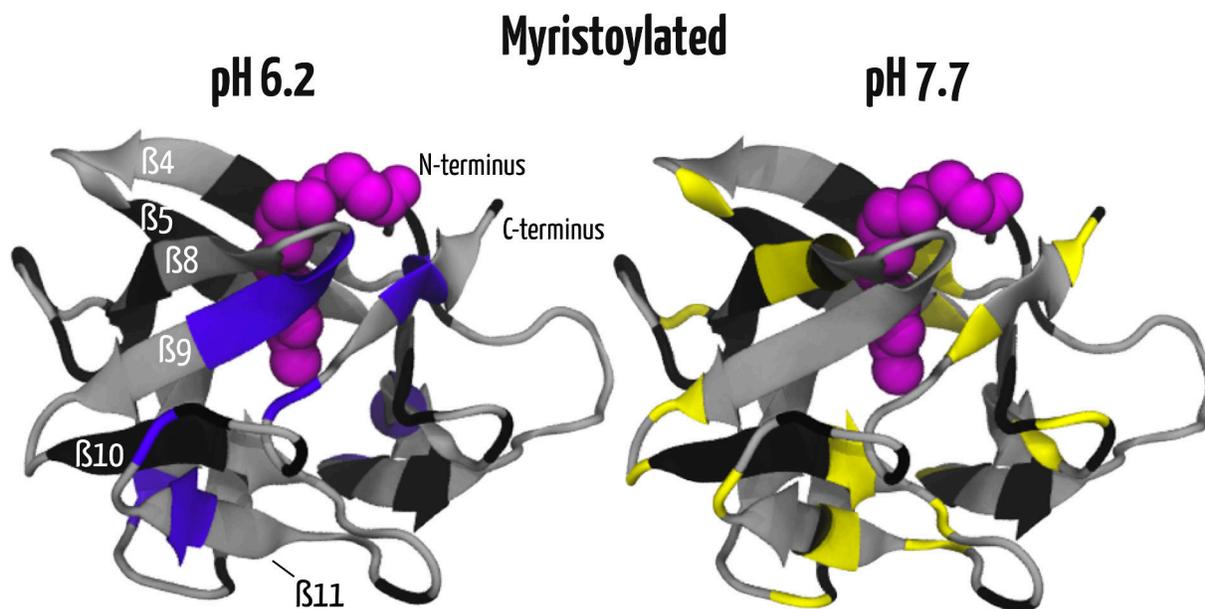


Figure 5.6: Residues that exhibit curvature in myristoylated hisactophilin at pH 6.2 (mainly myristoyl accessible) and pH 7.7 (mainly myristoyl sequestered). Residues that are curved at both pH values are in black, residues unique to pH 6.2 are in violet while those in yellow are unique to pH 7.7. The myristoyl group is shown in magenta. 30 curved residues are shared at both pH values.

5.4.1.3 Myristoylation influences curvature

As already mentioned, myristoylating hisactophilin greatly increases protein dynamics and burying the myristoyl group may increase conformational tension. Directly comparing residues that experience curvature in myristoylated and non-myristoylated forms at the same pH sheds light on where the fatty acid induces conformational heterogeneity. High pH is especially interesting because based on global thermodynamic measurements and NMR data, the myristoyl group is sequestered in the core of hisactophilin [8], however a native state energy landscape compressed by the myristoyl group may allow transient sampling of the accessible state at non-permissive (high) pH and curvature may offer insight into which residues see and effect switching.

The right panel of Figure 5.8 residues that are curved in myristoylated hisactophilin at pH 7.7 (black and orange)—residues in black are shared with the non-myristoylated form. The extensive increase in myristoyl-induced conformational heterogeneity is evident in the orange residues spread throughout the protein. The alternate states around the mouth of the β -barrel might correspond to sampling the myristoyl-accessible state. Strand β 10, home of I93, is also more dynamic, which may arise from core residues needing to adjust packing in relation to movement of the myristoyl moiety.

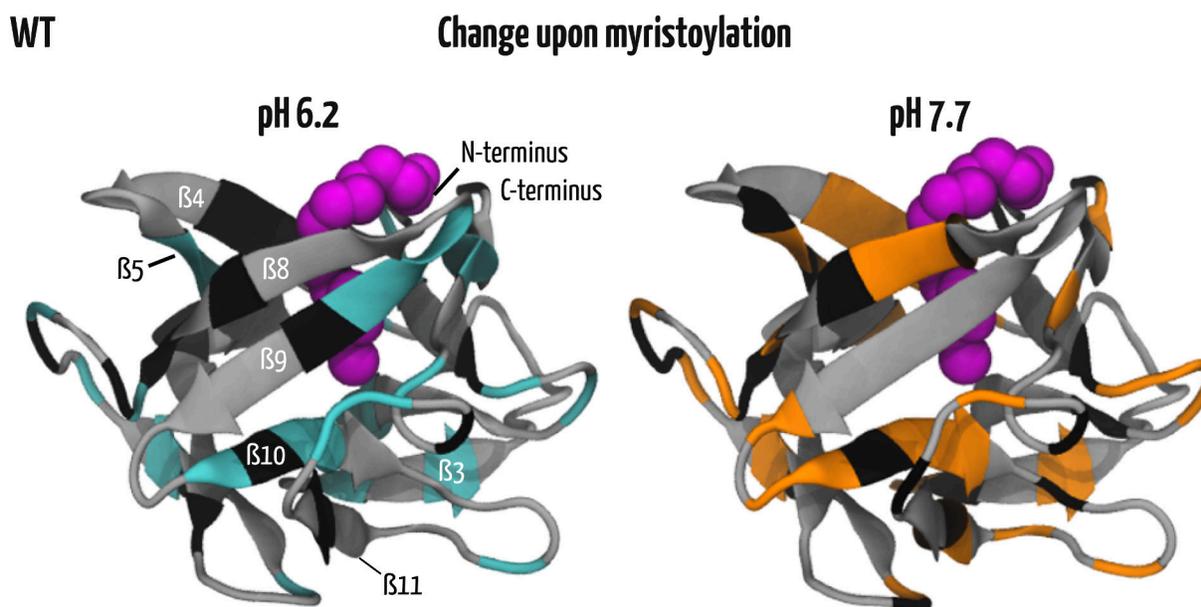


Figure 5.8: The structure of myristoylated hisactophilin at pH 6.2 (left panel) and pH 7.7 (right panel) highlighting the residues with nonlinear temperature coefficients where those in black are shared with the non-myristoylated form and those in cyan or orange are unique to the myristoylated form at each pH. At pH 7.7, myristoylation confers much additional conformational heterogeneity to hisactophilin, with β -strands 4 and 10 especially interesting as temperature coefficients and global thermodynamics implicate some of their constituent residues as important to folding and switching (i.e. V36, I93). All residues with curvature are given in Appendix A.

The left panels of Figure 5.9 show a top-down view of hisactophilin where the side chains of residues

with curvature are displayed. This illustrates that the conformational heterogeneity induced by myristoylation occurs throughout the structure, including into many of the loops. Increased dynamics in the loops might provide a mechanism for hisactophilin to compensate for overpacking the core, working as pressure release valves. Conformational heterogeneity in the hairpin loops at the bottom of the β -barrel could also relate to some of hisactophilin's other functions, such as sampling a state competent for actin binding. The right panels focus on residues within 9 Å of the myristoyl moiety and thus more likely to interact directly with it. On the side of hisactophilin opposite the N- and C-termini (trefoil 3), a cluster of residues that become more dynamic when the myristoyl group is present (orange residues) and a small number that are innately dynamic (black, shared with the non-myristoylated form) form a conduit that runs from top to bottom along the inside of the β -barrel. That the dynamics of these residues are only evident when the myristoyl group is present suggests that the myristoyl group has altered the near-native energy landscape such that they may access their accessible-state conformations and they are likely involved in switching (e.g. V36, V43, L45, H78, and I93). The black residues support that some switching motions of the backbone may be accessible to the non-myristoylated protein, especially at pH 6.2 where residues like V36, L76, and I85 have shared curvature between myristoylated and non-myristoylated forms. Though not included in the core, it is interesting that H88, H90, and H91 are curved because their $pK_{a,app}$ s shift from ~ 7 to ~ 6 when the myristoyl group is buried (vide infra), so their conformational heterogeneity might reflect changes in ionization and associated conformational consequences.

One of the principal determinants of the extent of curvature is the energy difference between the native and excited state [339]. Some residues in the conduit region have similar amounts of curvature to one another which may indicate they access the same excited state (red cluster in the left panel of Figure 5.10). I85, I93, and D108 are grouped by extent of curvature, as well as V83, S84, H106, and F113. The residues with similar extents of curvature (a-value in Appendix A) may indicate concerted access to a more structured excited state that accommodates the myristoyl group and may represent one of the mechanisms for relieving switch-related strain.

To summarize, the nonlinear temperature dependence of amide proton chemical shifts described above offers a window into the diverse near-native energy landscape of hisactophilin with measurable differences in heterogeneity between different pH values, acylation states, and where residues expected to report on switching do so. Going forward, the expectation is that hisactophilin mutants with abrogated switches will have different energy landscapes around the native state that reflect the decoupling between proton-binding and myristoyl switching and impact which residues have curvature.

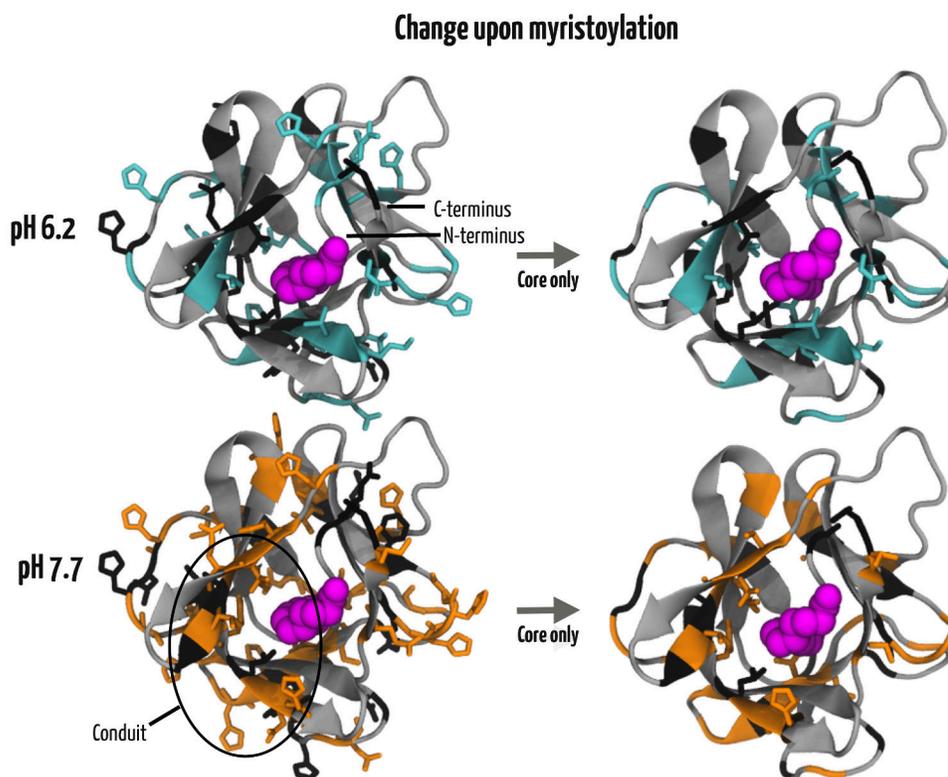


Figure 5.9: Myristoylated hisactophilin at pH 6.2 (top panels) and pH 7.7 (bottom panels) showing residues that are curved in myristoylated hisactophilin. Residues are coloured black if both myristoylated and non-myristoylated are curved at the given pH, and orange/cyan if they are unique to myristoylated. The leftmost panels show stick representations of the sidechains of all curved residues, whereas the rightmost panels show residues for which sidechain atoms are within 9 Å of the myristoyl group to focus mostly on residues that compose hisactophilin's core.

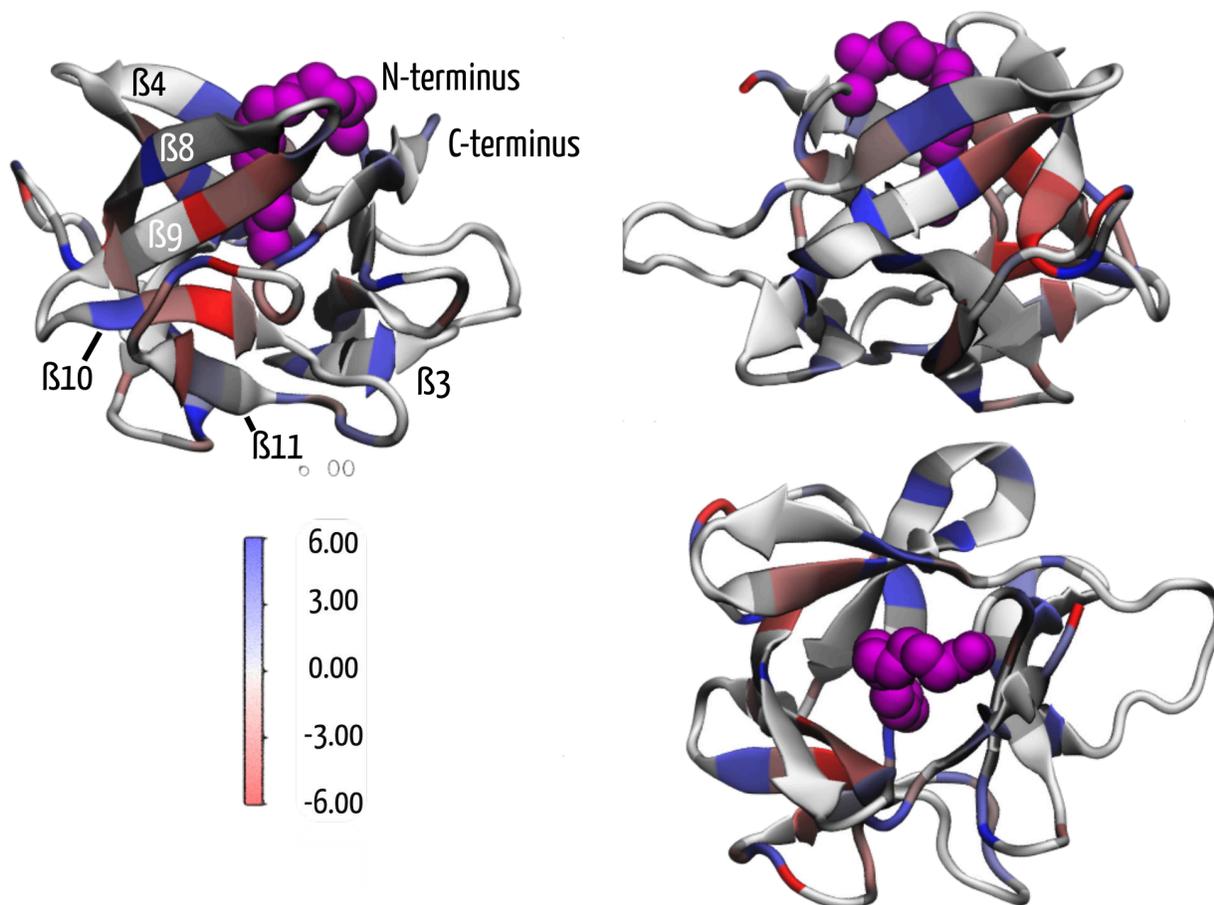


Figure 5.10: Extent of curvature (a-value) for myristoylated wild-type at pH 6.2. The myristoyl group is shown as magenta spheres. Residues in red are concave, residues in blue are convex, and grey/white are where there was no curvature or data. A-values were the quadratic term determined by fitting the residual of the linear regression of the temperature dependence of amide proton chemical shifts to a second order polynomial. Only residues that passed the bootstrapped hypothesis test were assigned a-values. The scale on the figure omits " $\times 10^{-5}$ ". A complete list of a-values is given in Appendix A.

5.4.2 Curvature in F6L/I85L/I93L

F6L/I85L/I93L (LLL) is a broken switch mutant whose global thermodynamics are discussed in Chapter 2 & 3 [5,73] and that was also analyzed by variable temperature NMR to measure local conformational stability by linear temperature coefficients (Chapter 4). These results for LLL show that the interactions between the myristoyl group and proton binding are decoupled, leading to a dearth of pH-dependent global stabilization or perturbation of local energetics by the myristoyl group. Furthermore, the global energetics suggest that LLL favours a conformation where the myristoyl group stabilizes hisactophilin irrespective of pH, similar to the myristoyl-sequestered state of wild-type hisactophilin. Like in wild-type, variable temperature NMR experiments for LLL yielded a number of residues whose chemical shift temperature dependences were nonlinear, but unlike in wild-type, curvature is markedly diminished—notably in residues packed against the myristoyl group—consistent with the lost of pH-dependent switching.

Figure 5.11 shows residues in LLL with nonlinear amide proton temperature coefficients (generally comparable to the figure for wild-type, Figure 5.3). The four variants of LLL have a commensurate number of curved residues: non-myristoylated at pH 7.7 had 34 and pH 6.2 had 32, myristoylated had 37 at pH 7.7 and 35 at pH 6.2, making for ~35% of the assigned residues manifesting curvature (residues listed in Appendix A). That is already a departure from wild-type where there was an appreciable difference in the number of curved residues between myristoylated and non-myristoylated. Our current theory is that the combination of three mutations muddled the geometry of the myristoyl pocket so that proton binding events do not generate enough tension to disfavour the sequestered state. Having nearly the same number of residues reporting on conformational heterogeneity with and without the myristoyl group suggests that burying the myristoyl group is less profoundly altering the near-native state dynamics and that pH minimally tunes them.

Changes in curvature upon myristoylation can provide a clearer image of how adding the myristoyl group alters the energy landscape of hisactophilin (see Supplementary Figures S5.1 and S5.2). Recalling that myristoylation correlated with an increase in the number of residues accessing alternative states the whole protein over in wild-type (27 residues in non-myristoylated and 46 in myristoylated at pH7.7, concentrated especially in the core), at pH 7.7 22 of myristoylated LLL's 37 curved residues are shared with non-myristoylated LLL, far fewer that are new and unique compared to wild-type (see the right panel of Figure 5.14 and the lower panels of Figure 5.15). The reduced impact of the myristoyl group on the energy landscape in LLL is consistent with our model for the broken switch. Even more interesting are the changes in conformational heterogeneity in LLL's core (see the lower right panel in Figure 5.15). Excepting residue L93 which is present in myristoylated and non-myristoylated LLL, the conduit of myristoyl-adjacent dynamic residues opposite the N- and C-termini observed in wild-type (Figure 5.9) is now silent, suggesting a loss of dynamics that might be consistent with the broken pH-dependent switch.

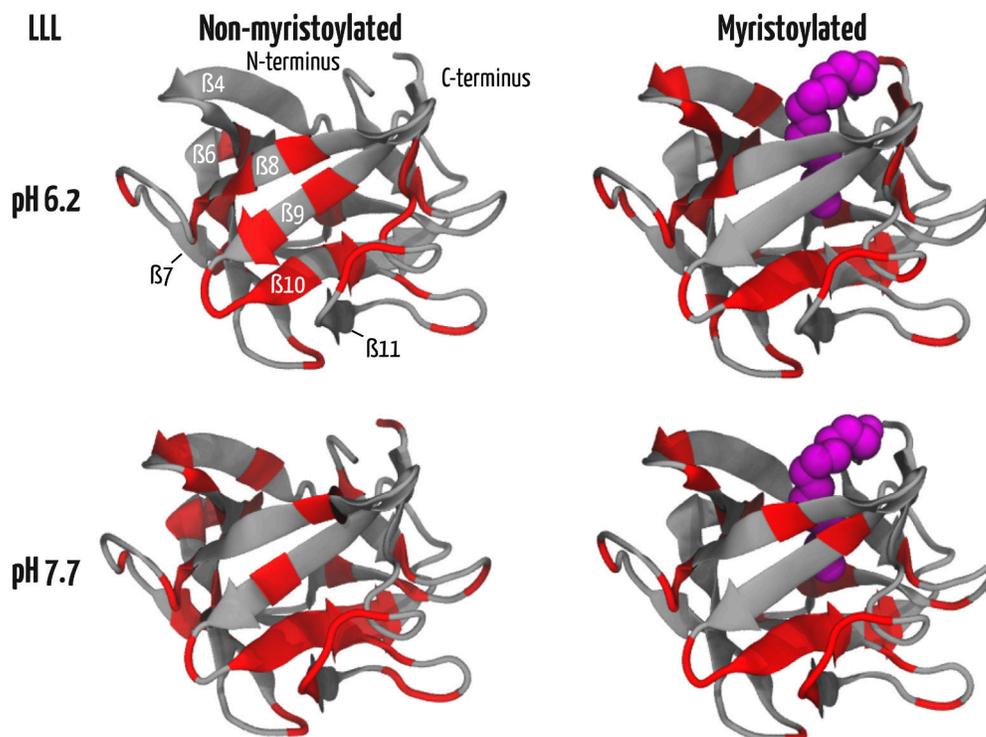


Figure 5.11: Structure of hisactophilin representing the LLL mutant with residues that are statistically significantly curved shown in red. The leftmost panels are for non-myristoylated and the rightmost panels are myristoylated (myristoyl group in magenta) at the pH values as marked. Curvature was determined after fitting the temperature dependence of chemical shifts (range $\sim 10 - 50$ °C) to a straight line by the least squares method, and then subjecting the residual to a bootstrapped (N=1000) F-test comparing the quality of fit of the residual to a line vs a higher order polynomial. Data were good enough to allow discrimination at the 99% confidence level ($p < 0.01$). For non-myristoylated LLL, 34 residues were curved at pH 7.7 and 32 at pH 6.2. Myristoylated LLL had 37 and 35 curved residues at the corresponding pHs. Structures were prepared in VMD.

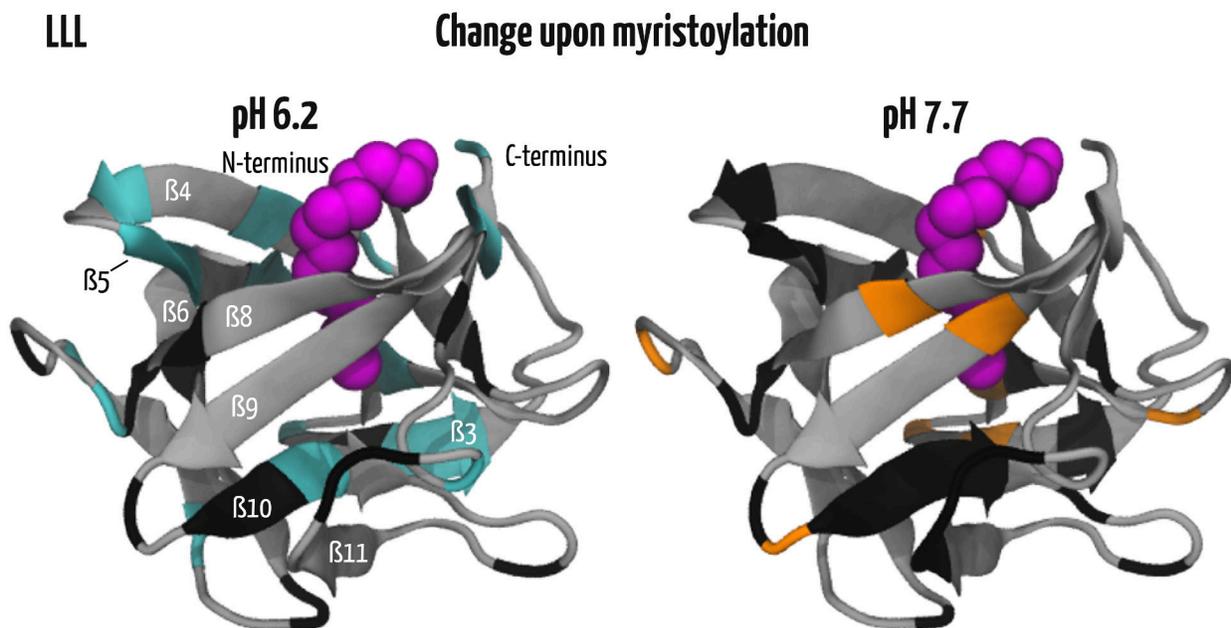


Figure 5.14: These structures are to demonstrate which residues of myristoylated LLL are curved, where those in black are shared with the non-myristoylated form at the given pH (pH 6.2 in the left panel and pH 7.7 in the right panel), and cyan and orange are those residues unique to the myristoylated form at the indicated pH. The myristoyl group is shown in magenta. At pH 7.7, 22 of the myristoylated's curved residues are shared with non-myristoylated, suggesting that the fatty acid moiety's effect on the energy landscape is somewhat less than for wild-type.

The attenuation of apparent conformational heterogeneity in the core and throughout the protein upon myristoylation in LLL could arise from the energy difference between the excited and ground states (i.e. myristoyl-accessible and myristoyl-sequestered states) exceeding ~ 5 kcal/mol in the mutant and thus the population of alternative state is too small to reliably measure with these methods. Or, curvature could disappear if timescales shift to a regime where chemical shift measurements are not good indicators (e.g. into the intermediate regime where peak-width matters). There is also the caveat that if the ground and alternative states have the same temperature coefficients and/or chemical shifts, no curvature in the temperature dependences will manifest. The sensitivity of amide proton chemical shifts to even small conformational changes makes this last point less likely. Comparing residues with curvature in LLL against wild-type offers compelling evidence that one of the consequences of the mutations is an altered energy landscape wherein the residues lining the myristoyl-binding pocket no longer report a change in environment consistent with accessing an excited state.

5.4.3 Curvature in I85L

The other broken switch mutant, I85L, substitutes an isoleucine in strand $\beta 9$ for leucine. When Dr. Martin Smith and Dr. Aron Broom made the model for myristoyl-sequestered hisactophilin, the myristoyl group

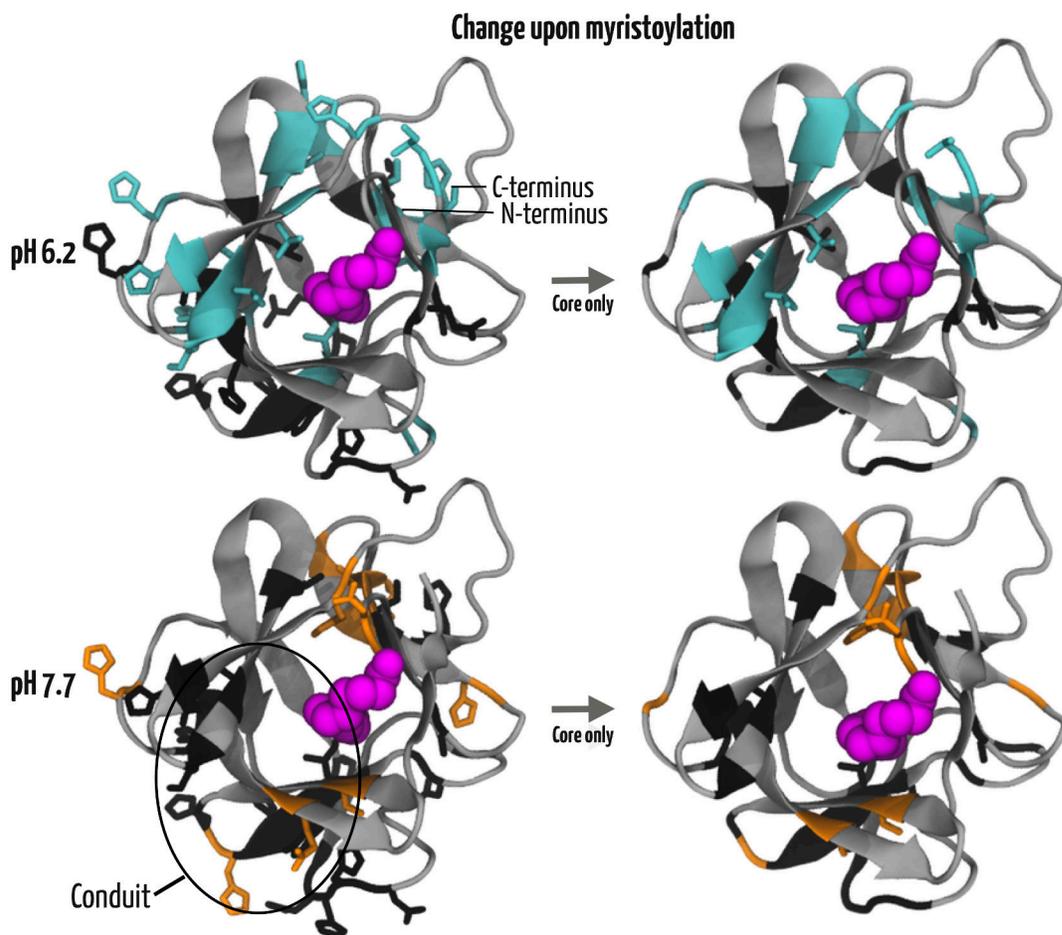


Figure 5.15: Top-down view of the structure of hisactophilin showing myristoylated LLL with curving residues at pH 6.2 (upper panels) and pH 7.7 (lower panels) mapped onto it. Residues with curvature in myristoylated and non-myristoylated are shown in black, and those unique to pH 7.7 are in orange while those unique to pH 6.2 are shown in cyan. In the leftmost panels, the sidechains are shown for all residues with curvature in the appropriate colour. In the rightmost panels, the sidechains are shown only for residues within 9 Å of the myristoyl moiety, intended to highlight residues that point into the core and are more likely to interact with the myristoyl moiety. Note that there are very few for which this is true. The myristoyl group is shown in magenta.

was localized by NMR NOE-derived distance constraints. These distance constraints placed the terminal methyl group of the myristoyl moiety within 5 Å of the backbone amides of F6, I85, I93, and F113. Although not the original motivation for making the I85L mutation (it was more of a symmetry in β -trefoils argument, discussed in Chapter 3 and [105]), its proximity to the myristoyl group impelled its characterization by equilibrium denaturation experiments [73], kinetic folding experiments, and in-depth NMR experiments, including those to assign the ^1H , ^{15}N HSQC spectrum, measure the temperature (Chapter 4) and pH dependence of its chemical shifts (*vide infra*). What is so tantalizing about I85L is that a seemingly straightforward mutation that alters only the position of a methyl group is disruptive enough to ablate the pH-dependent myristoyl switch. Temperature coefficients gave us an initial view of why and how the switch is broken at residue resolution, and with nonlinear temperature dependences we gain additional insight into how its energy landscape is shaped by the myristoyl group and pH changes.

In Figure 5.16, residues with curvature in I85L are mapped in red onto hisactophilin's backbone. Between the four forms of I85L, there are sets of residues whose curvature is shared among three or four forms, and many of these residues are also curved in wild-type. Some of the loop positions are the more evident cases, for example the $\beta 6$ - $\beta 7$ and $\beta 7$ - $\beta 8$ loops; residues K59, Q60, L67, H68, H71, and L73 are curved in nearly all four forms of I85L. G99 and H100 in the hairpin loops are also universally curved, and similarly with the $\beta 11$ - $\beta 12$ loop that packs against β -strands 9 and 10. However, these residues could define a set of curving residues whose dynamics arise from innate structural feature of hisactophilin rather than for explicitly functional reasons. Many of these residues are themselves histidines or immediately adjacent to one, so the conformational heterogeneity here could reflect the local electrostatic effects of ionization and/or associated conformational changes. Based on previous pH titrations on wild-type and the pH work to follow with the broken switch mutants (*vide infra*), the $\text{pK}_{\text{a,app}}$ of most of the listed residues are between 6 and 6.5 (complete list of $\text{pK}_{\text{a,app}}$ s is in Appendix B), with the exception of non-myristoylated wild-type where some are in the pH ~ 7 range (e.g. K59, Q60, H71, G99). Curvature at pH 7.7 of residues with $\text{pK}_{\text{a,app}}$ s closer to 6 is unlikely to arise from an excited state driven by transient protonation [360,361]. Furthermore, in the mutants, these $\text{pK}_{\text{a,app}}$ s are minimally deflected/perturbed by myristoylation. Residues in two different conformations may have a different $\text{pK}_{\text{a,app}}$ for each state, so values that do not change upon myristoylation suggest that they are reporting on similar structural ensembles [290,362]. Heterogeneity that leads to curvature may report on changing electrostatics if the pK_{a} of a nearby residue (which can shift with temperature [360]) is close to the solvent pH, thus measuring $\text{pK}_{\text{a,app}}$ might measure the same process. Alternatively, the $\text{pK}_{\text{a,app}}$ change might drive a conformational shift to which curvature is sensitive, thus curvature and $\text{pK}_{\text{a,app}}$ are more complementary measure of dynamics. Primarily considering differences in curvature that result from myristoylation, proton binding, or mutation skirts some of these potential pitfalls, resulting in a robust measure of changes to access of the excited state.

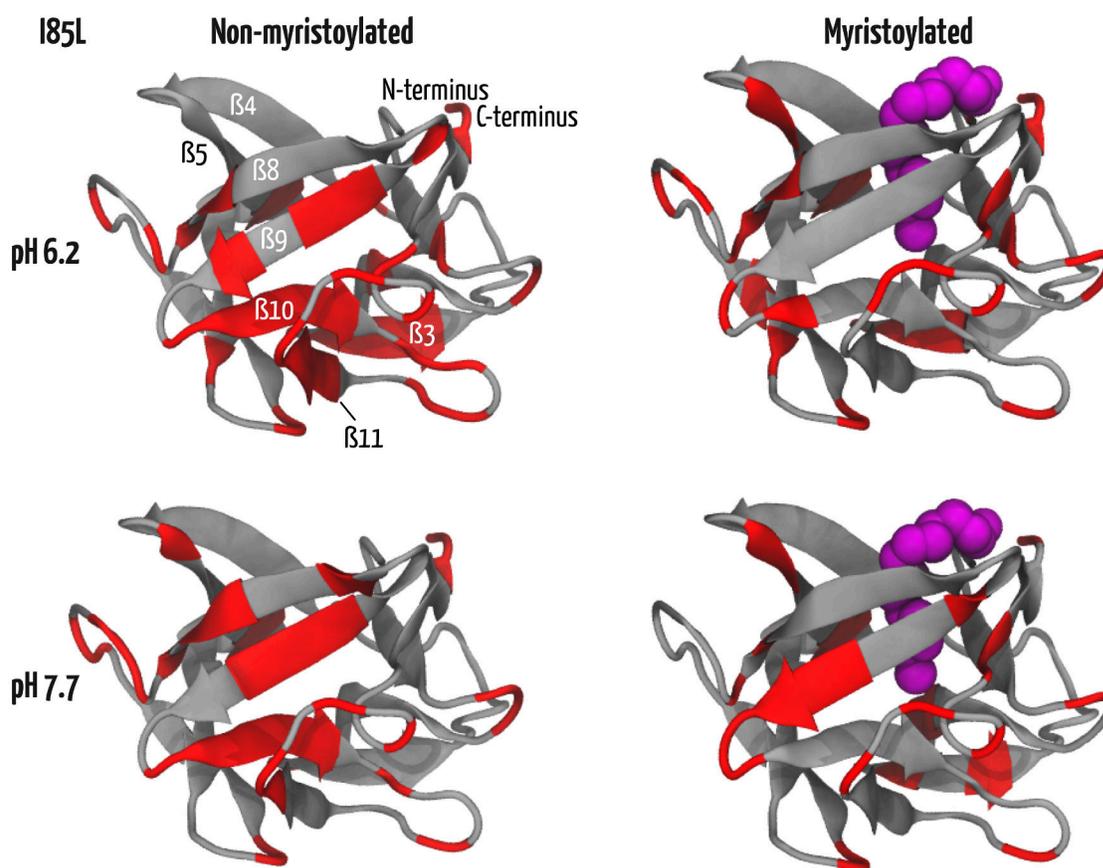


Figure 5.16: Structure of hisactophilin with residues that have statistically significant curvature in I85L shown in red. The myristoyl group is shown as magenta spheres. The panels on the left are for the non-myristoylated form and the panels on the right are myristoylated at the pH 6.2 and 7.7. An F-test of the data was bootstrapped (N=1000) to determine whether the data were best described by a linear or higher order polynomial with a filter at $p < 0.01$. For non-myristoylated hisactophilin, 36 residues were curved at pH 7.7 and 41 at pH 6.2. In myristoylated hisactophilin, 27 residues were curved at pH 7.7 and 31 at pH 6.2 which is the opposite behaviour as wild-type wherein myristoylation increased the number of residues accessing an alternative state.

The number of residues with nonlinear temperature dependences in I85L is unique compared to WT and LLL in that the myristoylated forms at both pH values have fewer curved residues than non-myristoylated. The non-myristoylated form has 41 and 36 such residues at pH 6.2 and 7.7, respectively, and myristoylated has 31 and 27. We note that the number of curved residues in myristoylated I85L is comparable to LLL, so this might be an expected range of residues for a broken switch. The non-myristoylated form of I85L, therefore, is where the largest change has occurred. Equilibrium unfolding experiment place non-myristoylated I85L's global stability on par with wild-type (4.5 kcal/mol vs 4.3 kcal/mol at pH 6.2, Table 4.2), so I85L is unlikely to be sampling more of its unfolding pathway it is not anomalously unstable. In fact, at high pH, non-myristoylated I85L is slightly stable than wild-type (8.0 kcal/mol and 7.6 kcal/mol for I85L and wild-type). The prevailing theory was that I85L breaks the switch by tipping the balance of switching toward the myristoyl-accessible stable, and it may be that the movement of the methyl group by the mutation of the side chain may occlude the myristoyl binding pocket while creating new interactions with proximal residues that stabilize the protein (thus, non-myristoylated I85L is more globally stable than wild-type). The now-occluded pocket may disfavour full insertion of the myristoyl group and block typical residues from favourably interacting with the myristoyl group, consistent with myristoylated I85L less stable than wild-type.

The localization of the curved residues may hold a clue as to how the near-native energy landscape reflects the broken switch. The residues of the oft-discussed strand $\beta 9$ (includes I/L85, adjacent to the tip of the sequestered myristoyl group) are nearly entirely curved along the full strand in myristoylated wild-type at both pH values and minimally curved in the non-myristoylated form (Figure 5.3). Thus, this curvature pattern may be diagnostic of switching. In I85L, $\beta 9$ is curved in the non-myristoylated form and not in the myristoylated form. There are numerous aromatics deep in the myristoyl pocket (e.g. F6, F74, Y92, and F113), so the myristoyl group favouring the accessible might enable some of these aromatics to move into an excited state. If L85 mimics some of the interactions normally made by the myristoyl group, then some aromatic groups might be inhibited from accessing a switch-like excited state and propagating conformational heterogeneity throughout the protein.

Although the I85L mutation appears to increase the measurable conformational heterogeneity in the non-myristoylated form, adding the myristoyl group suppresses or reorganizes some of it. Areas identified in wild-type as being of particular interest to switching (Figure 5.9) have few residues in myristoylated I85L with curvature (Figure 5.18). Gazing down into the core (Figure 5.19, particularly the rightmost panels>), the conduit of curvature opposite the N- and C-termini observed in wild-type (Figure 5.9) is absent in I85L, similar to the other broken-switch mutant, LLL (Figure 5.15, all three variants are illustrated in Figure 5.20). Beyond the area that forms the conduit connecting the incognito proton binding site(s) to the myristoyl group in wild-type (including but not limited to residues I85, I93, and V36), there are other residues with evidence of accessing an excited state that point into the core of I85L, V83, F113, and E114 for example. It is noteworthy

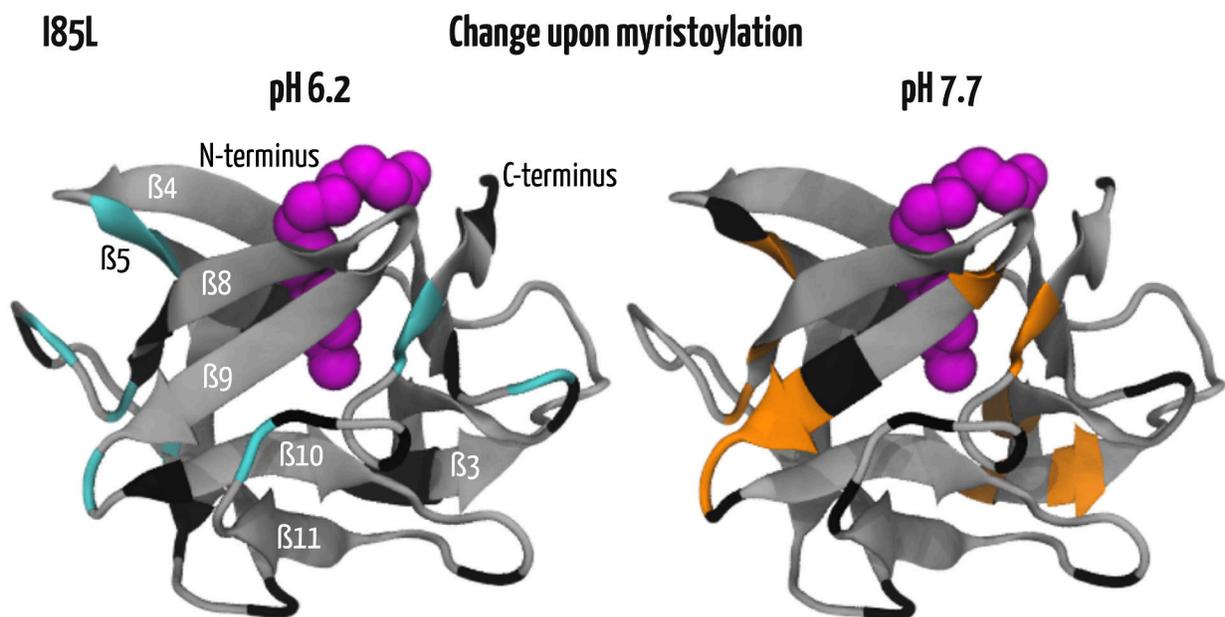


Figure 5.18: Side view of a representative model of I85L hisactophilin illustrating the residues that are curved in the myristoylated form. The myristoyl group is shown as magenta spheres. Residues in black are those that are curved in both the myristoylated and non-myristoylated form, and those in cyan and orange are unique to each pH 6.2 and 7.7, respectively. Few residues in areas thought to be functionally relevant to the switch exhibit conformational heterogeneity.

that that curvature is observed in the upper region of the barrel, where the myristoyl group is thought to be generally situated when in the accessible conformation [73]. Umbrella sampling simulations developed an energy surface for hisactophilin when the myristoyl group was held at different depths in the pocket [293]. In wild-type, two energy minima emerged that likely correspond to the sequestered (the deeper well) and accessible states (black trace in Figure 5.21). The sequestered state well in I85L was relatively shallower (orange trace in Figure 5.21), consistent with the myristoyl group favouring the accessible state. A flatter free energy surface is also consistent less curvature because there isn't a dominant excited state to affect the observed chemical shift. Thus, the myristoyl group in I85L is in an accessible conformation-weighted average position, leading to fewer residues to report nonlinear temperature dependences, but some that do are located toward the top of the barrel where the myristoyl group prefers.

5.4.4 Recapitulating Curvature

Variable temperature NMR experiments have the benefit of simplicity in their execution, amenability to partially automated analysis (Chapters 4 & 6), and they offer insight into local conformational stability as well as the functionally relevant near-native energy landscape from the same experiments. In hisactophilin, many residues have nonlinear temperature dependences pointing to significant conformational heterogeneity, with the extent of curvature more than some other studied proteins [150,336]. The curved residues in hisactophilin

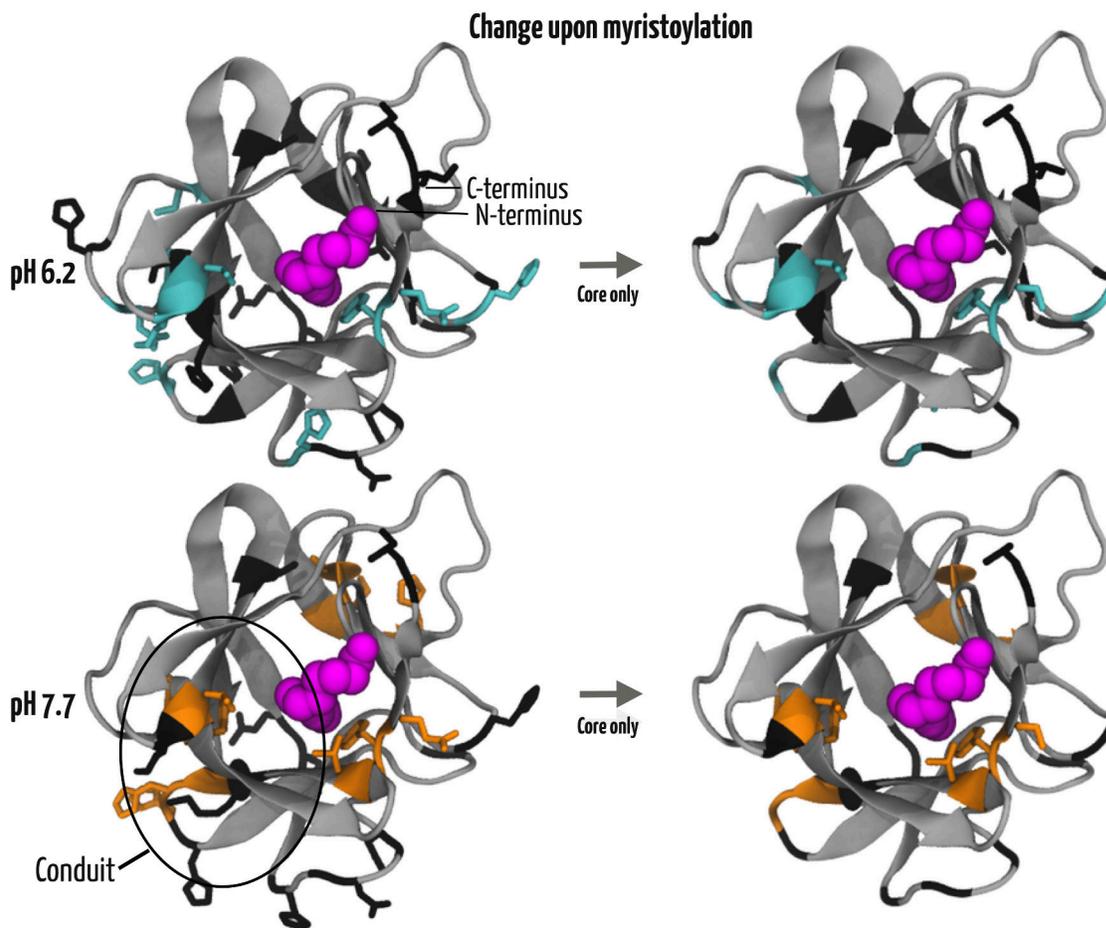


Figure 5.19: Top-down view of myristoylated I85L illustrating the residues that show curvature in the myristoylated form. The myristoyl group is shown as magenta spheres. Residues in black are also curved in the non-myristoylated form, while those in orange and cyan are uniquely curved at pH 7.7 and 6.2 respectively. The leftmost panels also show the sidechains as sticks for residues with curvature in colours that match the backbone colour scheme, while the rightmost panels show sidechain sticks only for those sidechains within 9 Å of the myristoyl group to highlight those most likely to interact directly with the myristoyl group.

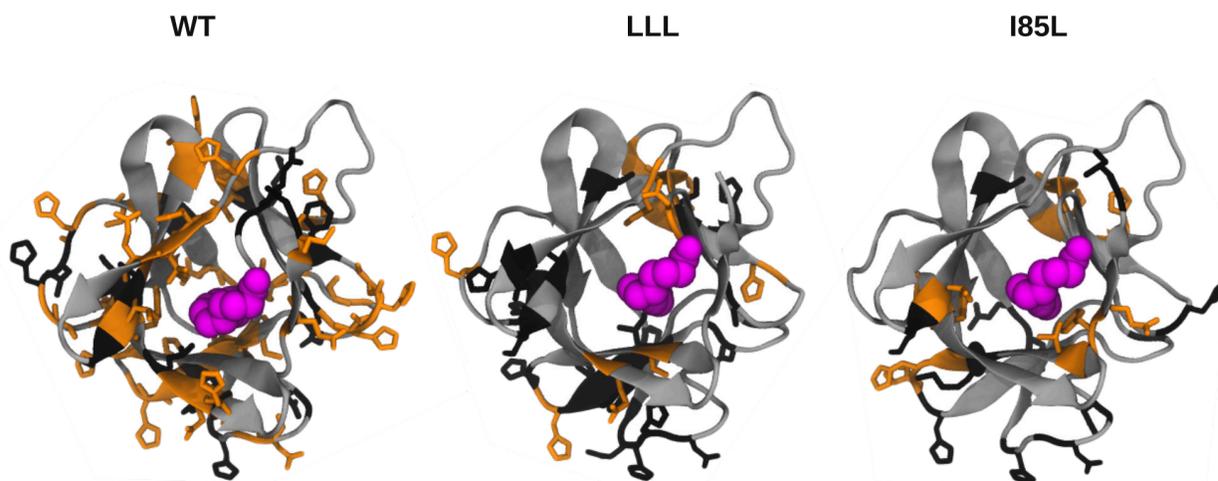


Figure 5.20: Side-by-side comparison of curvature in myristoylated WT, LLL, and I85L at pH 7.7. The myristoyl group is shown as magenta spheres. Black residues are curved in myristoylated and non-myristoylated forms, and orange residues are exclusively curved in the myristoylated form of that variant. The two broken switch mutations have notably fewer residues for which myristoylation induced conformational heterogeneity (orange).

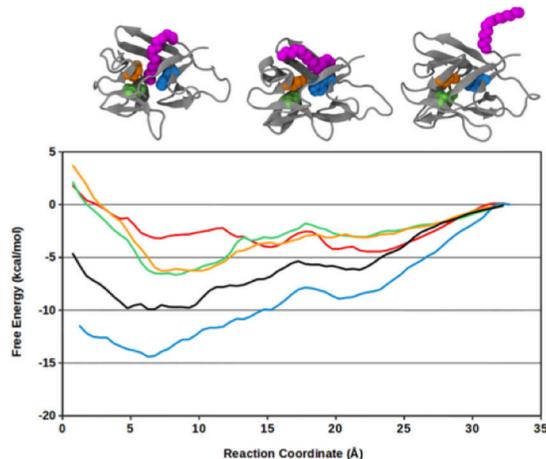


Figure 5.21: Free Energy Surface of myristoylated wild-type hisactophilin and mutants. FES generated from umbrella sampling are shown for the myristoylated form of WT (black trace), F6L (blue), I85L (orange), I93L (green), and LLL (red) mutants. The reaction coordinate is the distance of the tip of the myristoyl group to the bottom of the hairpin triplet. Representative structures from MD of the WT protein along the reaction coordinate are shown with WT residues colored in the same manner as the FES plots. There appears to be an energy minimum fairly deep within the pocket (≈ 5 to 10 Å, first structure on left), termed the “sequestered” state. Additionally there is a local minimum that occurs with very shallow binding (≈ 20 to 25 Å, middle structure), termed the “accessible” state. Finally, an example of the fully “exposed” state of the myristoyl group is shown at the right. FES are aligned such that this “exposed” state all have an energy of 0 kcal/mol, since mutations within the binding pocket should not affect this state. Figure was adapted from [293].

are distributed over the protein but enriched at the boundaries of regular secondary structural elements where an excited state could comprise a local conformational rearrangement in the hydrogen-bonding of the β -sheet elements. In addition, there are residues clustered in the protein core opposite the chain termini, packed against the myristoyl group, that share similar α -values of curvature and therefore appear to access the same alternative state. The ΔG_{switch} was previously determined to be on the order of ~ 2.0 kcal/mol [8], a value within the range of energetics that can give rise to curved temperature dependences [182,336], so it is conceivable that a conduit of residues reports on conformational heterogeneity related to myristoyl switching. When subjected to the same analysis, the broken switch mutants LLL and I85L exhibit markedly attenuated curved temperature dependences in the core, indicating alterations to their function-related near-native energy landscapes that correlate with their broken switches. Furthermore, curvature analysis supports the theory that LLL's switch is tipped toward myristoyl-sequestered while I85L's favours myristoyl-accessible where residues V36 and V83 (among others) and the terminus contribute to stabilizing that particular conformation.

So far, this work has looked primarily at the effects of mutations in the protein core and understanding how they manifest in switching. Proton binding and electrostatics also play a crucial role in the myristoyl switch. Previous work has characterized the importance of conformation-dependent $pK_{a,\text{app}}$ s to switching [8], and in the following, the same approach will be extended to the two mutants to further characterize the mechanism of switching by analyzing ways in which it can be broken.

5.4.5 Macroscopic pK_a s of LLL and I85L

As already mentioned, the pH dependence of amide chemical shifts was measured by previous students in our lab. The $pK_{a,\text{app}}$ s of the non-myristoylated form were published in 1998 [74], and the myristoylated form in 2010 [8]. Together, these two titrations define the $pK_{a,\text{app}}$ s for residues throughout the protein, and notable was that they shifted from pH $\sim 7.0 - 7.5$ to ~ 6.0 for residues L45, G56, I85, Y92, H75, and H91 upon addition of the myristoyl. In the thermodynamic cycle analysis of hisactophilin's switch, the non-myristoylated form is an ersatz representation of the myristoyl out/accessible state [5,8], thus the $pK_{a,\text{app}}$ data imply that switching between the accessible and sequestered state is concomitant with a shift in pK_a of coupled residues from ~ 7 to ~ 6 and binding ~ 1.5 protons. That the pK_a of some residue(s) must shift if a protein conformational change is pH-dependent was theorized by Tanford [363], and simulations have successfully recapitulated observed chemical shifts and the macroscopic $pK_{a,\text{app}}$ s for Nitrophorin4 by accounting for the microscopic pK_a of an aspartate in open and closed conformations [364]. Though, the thermodynamic coupling between ionization and global conformational equilibria/folding is a more seasoned area of study (for example, [365,366]). Regardless, if a hisactophilin mutant hobbles or breaks pH-dependent switching, $pK_{a,\text{app}}$ s should reflect the loss of coupling and are therefore unaffected by the addition of a myristoyl group, specifically for residues whose $pK_{a,\text{app}}$ was sensitive in wild-type.

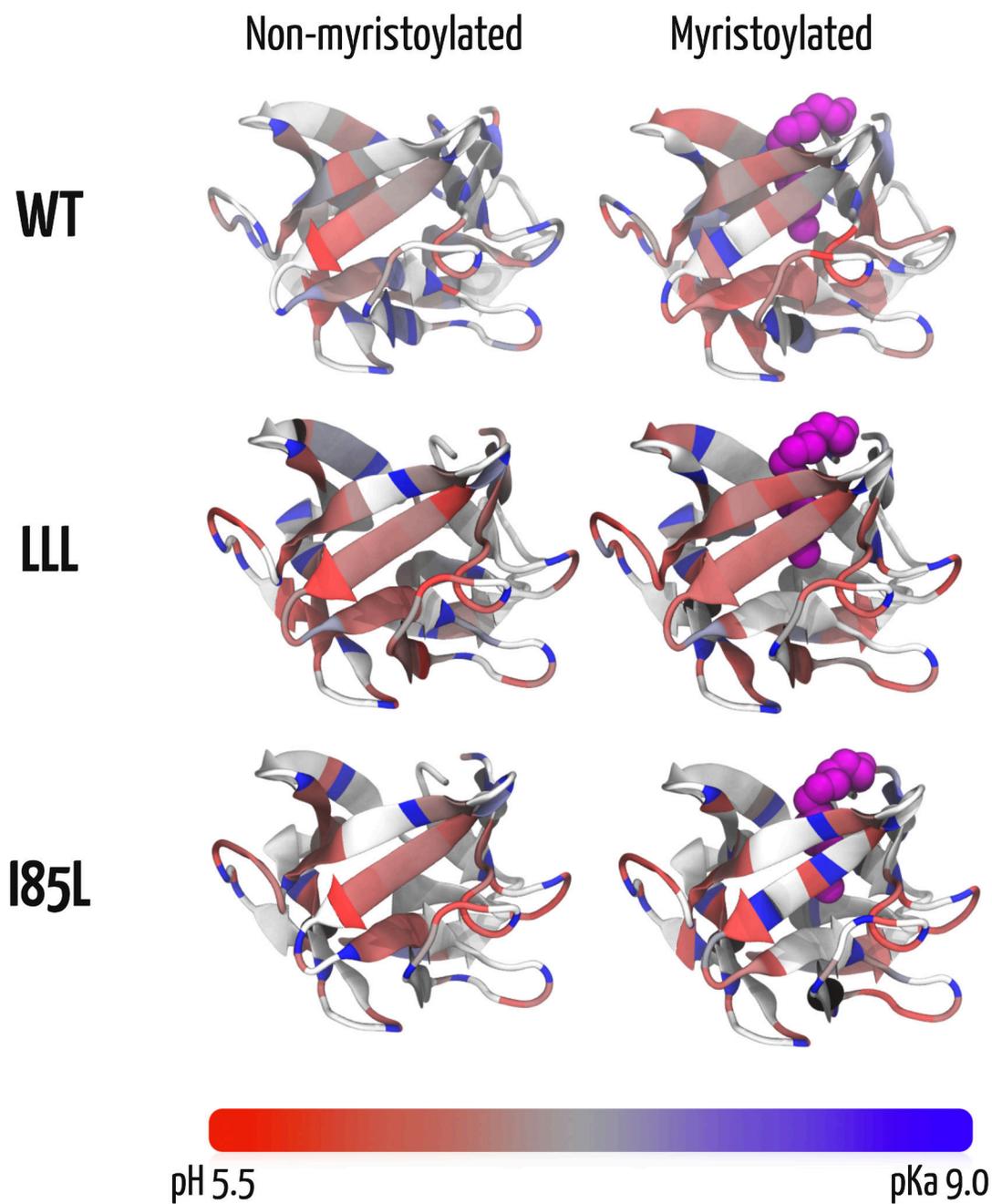


Figure 5.22: Wild-type and mutant hisactophilin apparent $pK_{a,app}$ s. (top row) Wild-type $pK_{a,app}$ s mapped onto the structure for non-myristoylated (left) [74] and myristoylated (right) [8] forms. The myristoyl group is shown as magenta spheres. The middle and bottom rows are the newly measured $pK_{a,app}$ s for broken switch mutants LLL and I85L, as labelled. $pK_{a,app}$ s are coloured by a red-gray-blue spectrum constructed by VMD over the range of 5.5 to 9.0. Grey residues have $pK_{a,app}$ s of 7.25. Residues in white are those for which we are lacking assignments or the pH-dependence could not be fit to a single $pK_{a,app}$.

Figure 5.22 maps the previously measured $pK_{a,app}$ s onto the structure of wild-type as well as the newly determined data for mutants LLL and I85L, focusing on strands $\beta 8$, $\beta 9$, and $\beta 10$ opposite the N and C-termini. A table of $pK_{a,app}$ s can be found in Appendix B. Not unexpectedly, there are many residues reporting $pK_{a,app}$ s below $pH \sim 7.25$ (red colour regime), likely owing to the numerous histidines that titrate in the pH 6-7 range. As already mentioned, shifts in the measured $pK_{a,app}$ s upon myristoylation are more informative for switching, thus Figure 5.23 maps the absolute value of changes in wild-type's $pK_{a,app}$ onto the hisactophilin structure. This representation makes the widespread response of $pK_{a,app}$ s to the presence of the myristoyl group apparent. The $pK_{a,app}$ s of 36 residues are perturbed by more than 0.5 pH units upon addition/burying of the myristoyl group (orange and red residues in Figure 5.23). Applying these same criteria, only 10 residues experience the same magnitude of perturbation in LLL, and a paltry 4 in I85L (caveat: there were fewer probes and assignments in I85L). Nevertheless, in the mutants, fewer residues notice the presence of the myristoyl group.

The muted response of the $pK_{a,app}$ s of the two broken-switch mutants to myristoylation further highlights the chain of interactions necessary for switching and how/that they can break. The NMR-monitored pH titrations of five of the six residues identified in wild-type as candidates for being the switch reporters, L45, G56, H75, I85, H91, and Y92, are shown in Figure 5.24 for wild-type as well as LLL and I85L (Y92 was omitted because there were no data available for the myristoylated forms of the mutants, so no comparison between acyl-states is possible). The wild-type panels illustrate the leftward shift of the titration curve upon myristoylation for these residues. However, the same trend is not evident in LLL and I85L, where although there may be a small vertical offset in some cases, the $pK_{a,app}$ s largely match between the myristoylated and non-myristoylated forms. What these residues indicate, however, is that they no longer report coupling of protonation to the myristoyl group and report on the broken switch.

To show that there was no systemic offset to the data that would account for non-existent $pK_{a,app}$ shifts, Figure 5.25 shows the titrations of two representative residues in LLL and I85L, V101 and H106, that in wild-type did not experience a shift in $pK_{a,app}$ upon myristoylation. The mutant data exhibit comparable behaviour to the wild-type.

Some residues in LLL whose $pK_{a,app}$ s do still undergo some shift upon myristoylation (the orange and red residues in Figure 5.23) are located in a few loops in the hairpins that close the bottom of the β -barrel. The shifted $pK_{a,app}$ s of these hairpin residues are similar in LLL and WT and may indicate independent equilibria with a local conformation that is competent for binding actin, which could involve specific charge-charge interactions with the actin molecule, though this is speculative. I85L lacks shifts in $pK_{a,app}$ s in the same region, however, changes may be unrepresented due to a dearth of data. In WT and LLL, pH may still be coupled to other functions of hisactophilin, but mutations have disrupted communication with the myristoyl group—conformation switching may still occur, but it is not coordinated with proton binding.

Keeping in mind that with a pH -driven conformational change, one or more residues' $pK_{a,app}$ s will shift

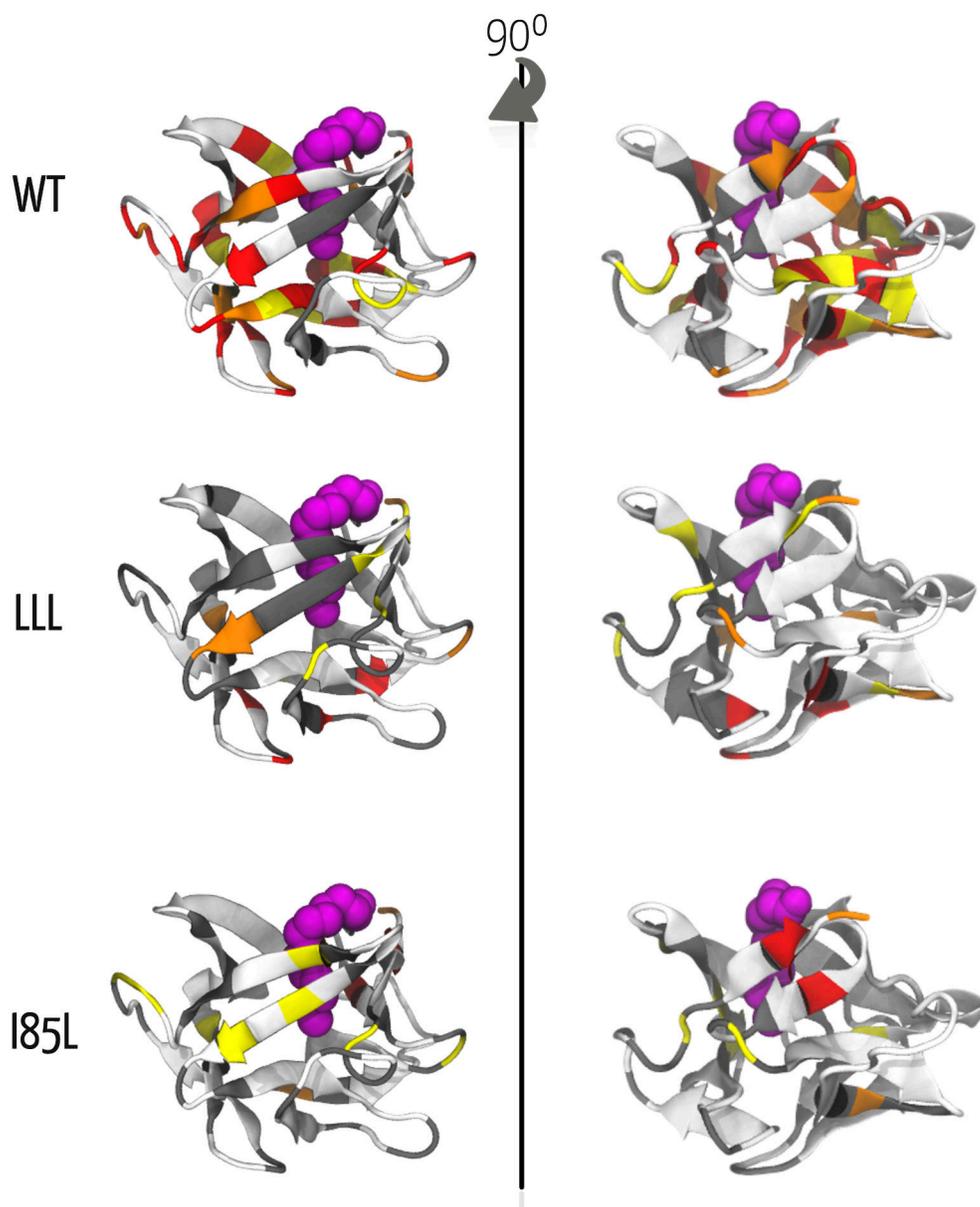


Figure 5.23: The absolute values of changes in $pK_{a,app}$ s upon myristoylation, mapped onto hisactophilin. Wild-type (top row), LLL (middle row), and I85L (bottom row). The myristoyl group is shown as magenta spheres. The left and right sides are a 90° rotation about the vertical axis. Residues in white are those without assignments or whose titration cannot be fit to a single $pK_{a,app}$ the non-myristoylated and/or myristoylated forms. Residues in grey represent those whose $pK_{a,app}$ s shift less than 0.25 pH units, yellow shift by between 0.25 - 0.50, orange by 0.50 - 0.75, and red by greater than 0.75 pH units. Residues in the two broken-switch mutants that are sensitive to myristoylation are generally less perturbed than in wild-type.

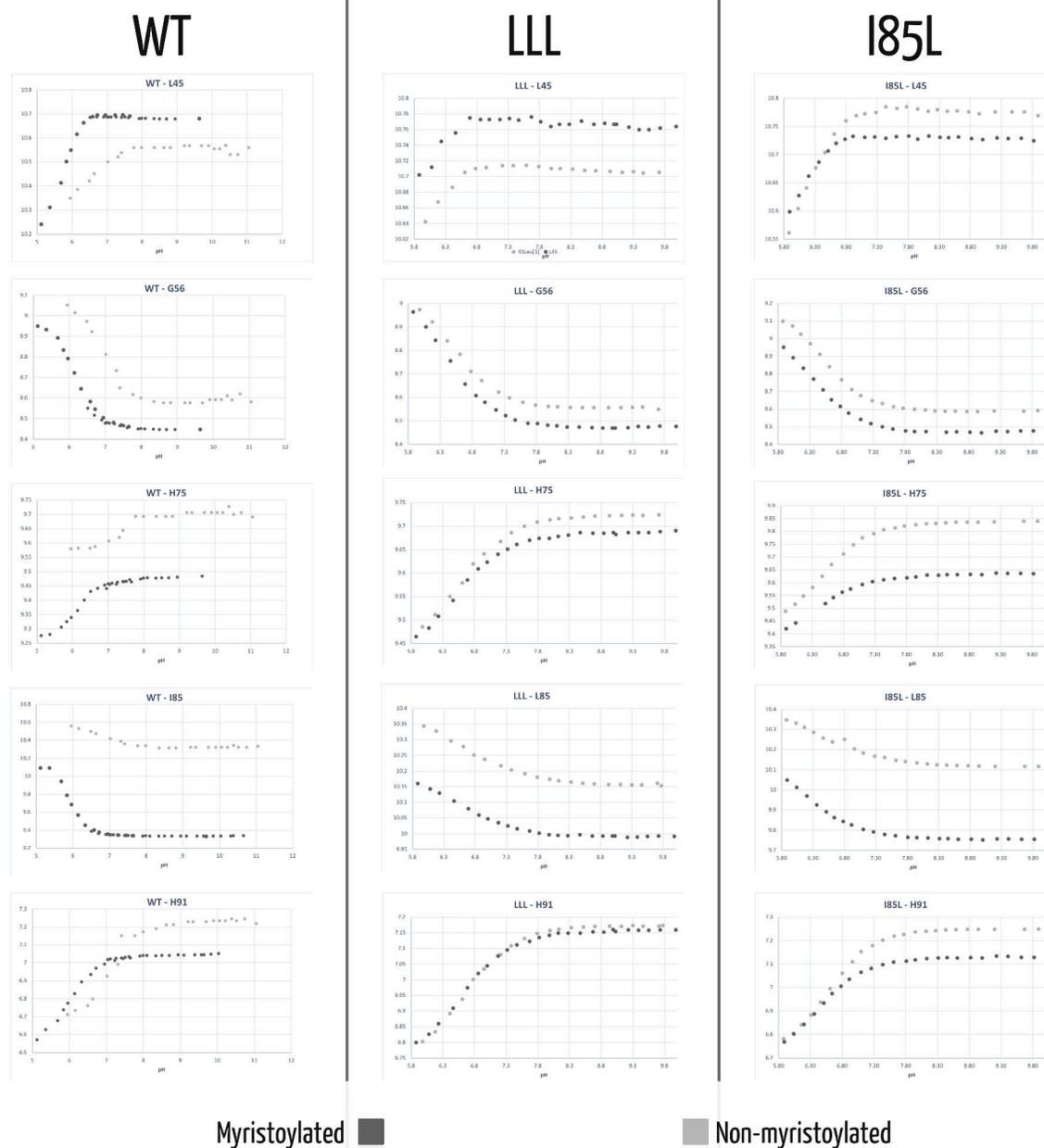


Figure 5.24: Shown are scatter plots of the pH dependence of chemical shifts for wild-type and mutants. Residues are L45, G56, H75, I/L85, and H91 for wild-type (WT, leftmost column), LLL (middle column), and I85L (rightmost column). WT data are from previously published work in our lab [8,74]. The chemical shifts of the myristoylated forms are dark grey while the non-myristoylated forms light grey. In wild-type, the presence of the myristoyl group clearly shifts the sigmoid left, to a lower pH. In the mutants, upon myristoylation, the same residues, though sometimes offset vertically, have nearly identical $pK_{a,app}$ s when fit to a model with a single ionizable group (values are listed in Appendix B).

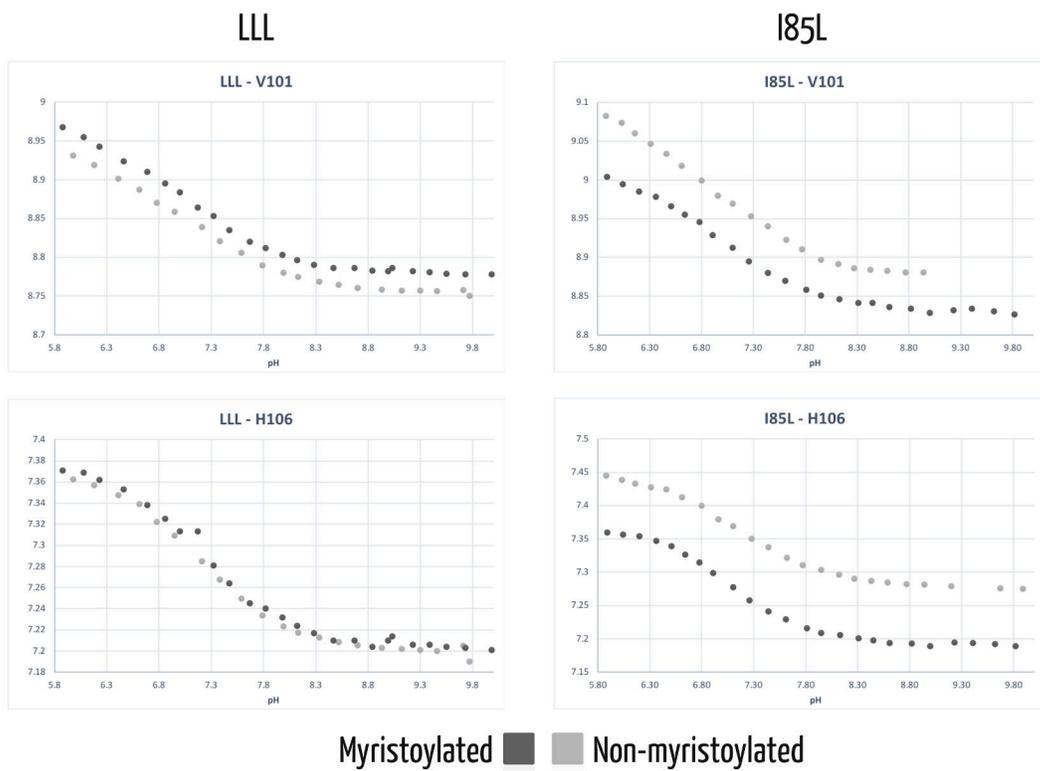


Figure 5.25: The pH dependence of chemical shifts of residues V101 and H106 in mutants LLL and I85L. The data for the myristoylated form are dark grey and the non-myristoylated form are light grey. Data from wild-type inform the expectation that there should not be a shift of $pK_{a,app}$ at these loci upon myristoylation, and it appears to bear out for LLL and I85L.

as a result of the change because each conformation will have its own, unique pK_a values(s). In wild-type hisactophilin, a buried myristoyl group corresponds to pK_a s of the controlling residues of ~ 6 such that at physiologically relevant pHs greater than 7, they will remain largely neutral [8]. However, switching to the accessible state (similar to non-myristoylated hisactophilin because the state has relatively few interactions between the fatty acid and the rest of the protein, though there is residual stabilization of the accessible by the myristoyl group that differentiates it from non-myristoylated [8]) involves pK_a s shifting to greater than 7, which results in the residues becoming largely charged. Functionally, this contributes to hisactophilin's high sensitivity over a narrow pH range—thus small changes to cytosolic pH effect significant cooperative changes in the ionization state of the protein owing in part to shifting pK_a s, which in turn contribute to the conformational changes of the myristoyl switch.

The preceding analysis has been interpreted previously for wild-type as reflecting ionization of one or two residues likely being responsible for switching, though involvement of additional residues could not be ruled out [8]. There is, however, precedent for a cloud of spatially clustered titrating residues being the basis pH-dependent function [367]. For example, Nucleopolyhedrovirus GP64 protein has a loop that was identified as important for membrane fusion, and three histidines in that loop had to be simultaneously mutated to abrogate function. Single mutations at those same positions were not enough, it had to be the combination of all three [368]. For the envelope protein of West Nile virus, another pH-dependent system, mutating single histidines attenuated function but did not altogether break the pH dependence, which the authors claimed meant that one histidine is not responsible for the switch and/or that other features of the protein help "pull the trigger". The Bohr effect in haemoglobin was revisited recently and, not only do multiple histidines make significant contributions to the effect [113], but the alkaline Bohr effect is modulated ($\sim 30\%$) by a valine and twenty-four other surface histidines [369]. There is no single ionizable culprit for Haemoglobin either.

For hisactophilin, a judicious attack with Occam's Razor may suggest the simpler model involving mainly one residue whose ionization is coupled to switching. However, as the wild-type panel in Figure 5.23 shows, residues spread throughout the protein are responsive to the myristoyl group, rather than effects being concentrated to a single area. This suggests that multiple residues make fractional contributions toward the ~ 1.5 protons that drive switching. The two sets of broken-switch mutations suppress many of the myristoyl-dependent $pK_{a,app}$ shifts, but some of the residual pH-dependence may arise because of the small contributions of minor ionizable sites to switching. Though there may not be a single "smoking gun" histidine, it might yet be possible to locate the significant contributors. Some of the residues whose $pK_{a,app}$ s shift upon myristoylation are proximate to one another and may report on similar environmental changes and/or a common ionization event(s). Playing a starring role yet are β -strands $\beta 8$, $\beta 9$, and $\beta 10$. The termini and short loop connecting β -strands 9 and 10 is rich in histidines and $pK_{a,app}$ s are greatly perturbed here as well. It is reasonable, then, to envision proton binding near the $\beta 9$ - $\beta 10$ loop and propagating its effects through the other residues with al-

tered $pK_{a,app}$ s to the myristoyl group (and down to a potential actin-binding surface as well). Having mutated residues that contact the myristoyl group, the broken switchers LLL and I85L still show a small-to-moderate $pK_{a,app}$ perturbation adjacent to the $\beta 9$ - $\beta 10$ loop so a small and local proton binding-induced equilibrium may still exist, but signal propagation to the myristoyl group appears to be disrupted.

For the sake of argument, if the shifted $pK_{a,app}$ s of the residues (with the largest associated chemical shifts) are diagnostic of the myristoyl accessible and sequestered states (pH 7-7.50 and pH ~6, respectively), then we may be able to make a structural argument for the disposition of the myristoyl group in the mutants based on the similarities of their $pK_{a,app}$ s to wild-type's. First, LLL. The $pK_{a,app}$ for L45, H75, and H91 for the myristoylated and non-myristoylated forms are all closer to wild-type's myristoylated form (Table in Appendix B). L85, the site of one of the mutations, is more akin to non-myristoylated wild-type, and G56 is midway between the two $pK_{a,app}$ s (Figure 5.24). The agreement between many of the $pK_{a,app}$ s of the myristoylated form of LLL and wild-type supports that the myristoyl group could be predominantly favouring the sequestered state.

That the $pK_{a,app}$ s of non-myristoylated LLL tend to agree with myristoylated wild-type is at first puzzling, but it might suggest that the mutations favour interactions in the pocket/communication network that mimic the sequestered conformation, even in the absence of the myristoyl moiety. Residue L85 is interesting because, unlike some of the other residues discussed, it's $pK_{a,app}$ s more closely resemble non-myristoylated wild-type. A lack of interaction between the myristoyl group and whatever ionization event L85 reports on may explain this, and this, too, fits with the model for LLL's broken switch — by increasing the space for the myristoyl group in the pocket (and relieving conformational tension), more accessible conformations for the myristoyl group may manifest in the $pK_{a,app}$ s in the bottom of the barrel appearing like they hardly see the myristoyl group at all, while those further up the barrel (e.g. L45 and H75) interact with parts of the myristoyl group that are more geometrically constrained and thus could have $pK_{a,app}$ s more similar to myristoylated wild-type.

In I85L, both forms have $pK_{a,app}$ s that approach those of myristoylated WT for L45, H75, L85, and H91 (G56 is again in between the wild-type limits, like in LLL. Figure 5.24). The non-myristoylated form of I85L is more globally stable than expected (closer in stability to wild-type than LLL, 4.3), potentially on account of the mutation positioning the sidechain methyl to fill/occlude part of the cavity at the bottom of the binding pocket that would otherwise be filled by the myristoyl group. That same argument could also justify the similarity of the $pK_{a,app}$ s in that area matching myristoylated hisactophilin (the sequestered state) because although the myristoyl group is expected to be stuck favouring the accessible state, the I to L mutation may mimic the tip of the myristoyl group and trick the nearby apparent $pK_{a,app}$ s into thinking the myristoyl group is buried. The $pK_{a,app}$ s of a few other residues in I85L still shift upon myristoylation, so parts of the protein may still switch, but changes to I85L disrupted communication to the myristoyl binding pocket.

5.5 Conclusions

The effects of pH on proteins are a flummoxing problem that is the subject of rigorous study because of how significantly ionization and charge can affect catalysis [370–372], ligand binding [292,373–375], and conformational changes [8,290,346,376,377], to name a few. Predicting and simulating pH dependence and microscopic pK_a s are also interesting from a protein engineering perspective [373,375,376,378–380] for designing switches and pH-responsive elements into proteins. For example, introducing a titratable residue into the core of Staphylococcal Nuclease (mutant V23E) resulted in a pH-dependent conformational change that opened the hydrophobic core to hydrate the newly charged moiety. The energetics consequences of trying to bury a charged residue into the core was sufficient to convert what would otherwise have been a transient state into the dominant conformation [290]. The bigger picture interpretations here are that the experimentally measured ionization constants are a mixture of the microscopic pK_a s of the closed and open states and that the change of the $pK_{a,app}$ is governed by a conformational change. Changes in $pK_{a,app}$ resulting from conformational change was studied in peptides [381], systematically studied in Nitrophorin 4, accompanied by $pK_{a,app}$ predictions by constant pH molecular dynamics simulations [364,382], and in staphylococcal nuclease again to correlate contributions of local flexibility to macroscopic $pK_{a,app}$ s by introducing glycine residues [362].

In hisactophilin, previously published pH dependence results also suggest that the conformational changes that enable switching correlate to changes in ionization constant upon myristoylation for the controlling residues [8]. Here, subjecting the two mutants with broken thermodynamic switches (LLL and I85L) to an NMR-monitored pH titration, the myristoyl-induced shift in ionization constants is ablated in many residues throughout the proteins. This supports the decoupling of the myristoyl group and its conformation from proton binding. β -strands 9, 10 and the loop connecting them is a likely locus for some of the switching proton to bind because it is one of the regions where the two sets of mutations most significantly affected the $\Delta pK_{a,app}$ s. Comparisons between residues' apparent ionization constants between mutant and wild-type support the hypothesis based on global stability measurements [5], and modelling [73]—that the myristoyl group prefers the sequestered conformation in LLL at all pH values due to what may be more space in the binding pocket, and prefers the accessible state in I85L because the mutation has blocked access to the bottom of the binding pocket such that it is no longer as energetically favourable to sequester itself. Moreover, the lack of response of many residues' $pK_{a,app}$ in mutants to myristoylation suggests that the contribution of different myristoyl conformations may no longer change upon myristoylation, which may further support the hypothesis that disrupting the allosteric communication pathway decouples pH from myristoyl switching.

5.5.1 Integrating conformational heterogeneity and apparent pK_a

Nonlinear temperature coefficients and $pK_{a,app}$ s both offer a measure of equilibria that exist in a protein; the former reports on conformational heterogeneity within 5 kcal/mol of the native state and conformational shifts are fast on the NMR timescale. $pK_{a,app}$ s, on the other hand, offer a glimpse into the fast equilibria of ionizable residues (and conformational rearrangements that could ensue) that are detected by chemical shift changes of nearby amide protons. Maybe there are similarities enough that nonlinear temperature dependences and shifts in $pK_{a,app}$ s are complementary in revealing some of the motional personality of hisactophilin. In principle, areas that bear the hallmarks of conformational heterogeneity and notice local ionization events may be reporting on the same electrostatic or conformational transitions, which is to say that it is a consistent and compelling story if curvature correlates with areas whose $pK_{a,app}$ s shift upon myristoylation.

In myristoylated wild-type hisactophilin, residues with curvature at either pH 7.7 or pH 6.2 tend also to have $pK_{a,app}$ s less than 7.0. It is not a rule, but it is generally true for the majority of curved residues. Many residues' $pK_{a,app}$ s in myristoylated wild-type are in the 6.5-6.8 range (see Appendix B). To a first approximation, for residues ionizing in this range, at pH 7.7 the population should split as approximately 90% deprotonated and 10% protonated, and 10% is above the lower limit of detection by curvature, so it is possible that ionization changes or the temperature dependence of $pK_{a,app}$ contribute to curvature [360]. Some residues far from an ionizable group still exhibit curvature, as well as residues with sidechain $pK_{a,s} > 7.7$. Curvature in these cases may indicate conformational heterogeneity arising from more than just acid/base equilibria, for example a conformational change associated with protonation, i.e. switching. As previously discussed, a pH-dependent conformational change should be associated with a shift in $pK_{a,app}$ s for the involved residues, so cross-referencing residues with curvature in the myristoylated form against those whose $pK_{a,app}$ shifts could highlight residues that drive the switch because their curvature may result from sampling the myristoyl-accessible and -sequestered states. V36, L45, L76, I85, H90, and H91 form a network of curved residues that could be reporting on myristoyl switching. The $pK_{a,app}$ s of these residues decrease by more than 0.5 ppm upon myristoylation (Appendix B). The correlation of the two independent measures is compelling evidence that they both observe the same pH-driven event and that it is likely to be switching.

In the two broken switch mutants, curvature in residues along the β -barrel communication network is markedly attenuated, nor do those residues experience a myristoylation-dependent shift in $pK_{a,app}$ s which reinforces the point. Some $pK_{a,app}$ s in the mutants still shift upon myristoylation which might suggest that some switching still occurs, but that it is not coordinated with proton(s) binding. The disappearance of curvature may indicate that the energy gap between the ground and excited/switched states is outside the ~1-5 kcal/mol range, that the ground and excited states now have the same amide proton chemical shift temperature dependence, or that the timescale of interconversion has slowed by several orders of magnitude. It is possible

that the ground and excited states have the same temperature coefficients, however, given how sensitive amide proton chemical shifts are to small conformational aberrations, this would itself be diagnostic of a broken switch if true. Lastly, a slower rate of interconversion between states could result in peak broadening or the appearance of a second, weaker set of peaks (which could, granted, be difficult to detect given the signal-to-noise ratio) but neither behaviour was evident in the variable temperature spectra.

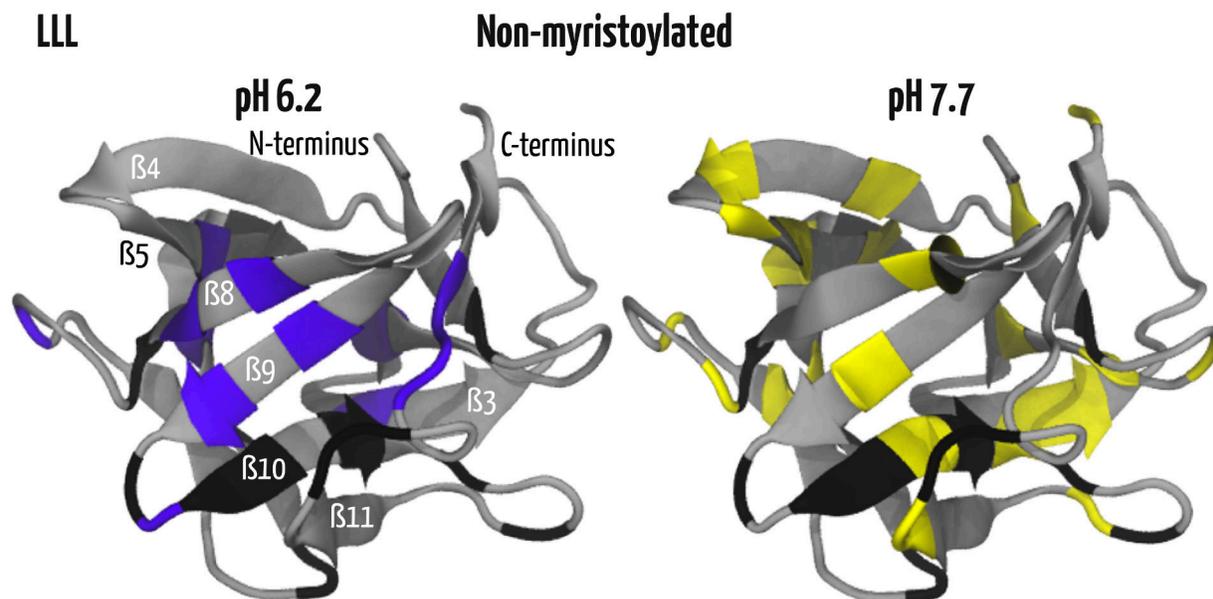
Together, changes in macroscopic $pK_{a,app}$ s and nonlinear temperature dependences of chemical shifts corroborate why the switches are broken and the likely pathway through which the allosteric signal propagates from the site(s) of proton binding to the myristoyl group in a functioning switch. Myristoyl-induced conformational heterogeneity on one side of the β -barrel makes up the likely pathway, and in fact, the residues share similar extents of curvature (α -values) which may mean they access similar, if not the same, switch-related excited state. This speaks to the cooperative nature of switching, the web of interactions necessary to make it happen, and the fine-tuning of the thermodynamics underpinning the communication network. These findings are not likely to be unique to hisactophilin; there are other pH switches small allosteric proteins in nature, as well as a growing interest in designing pH-dependent proteins to which this mechanism or methods might apply. In combination with the linear temperature coefficients, we have a fairly comprehensive picture of the near-native energy landscape and how hisactophilin achieves its finely tuned function. With the interest in the field to design pH-dependence into protein

5.6 Supplementary

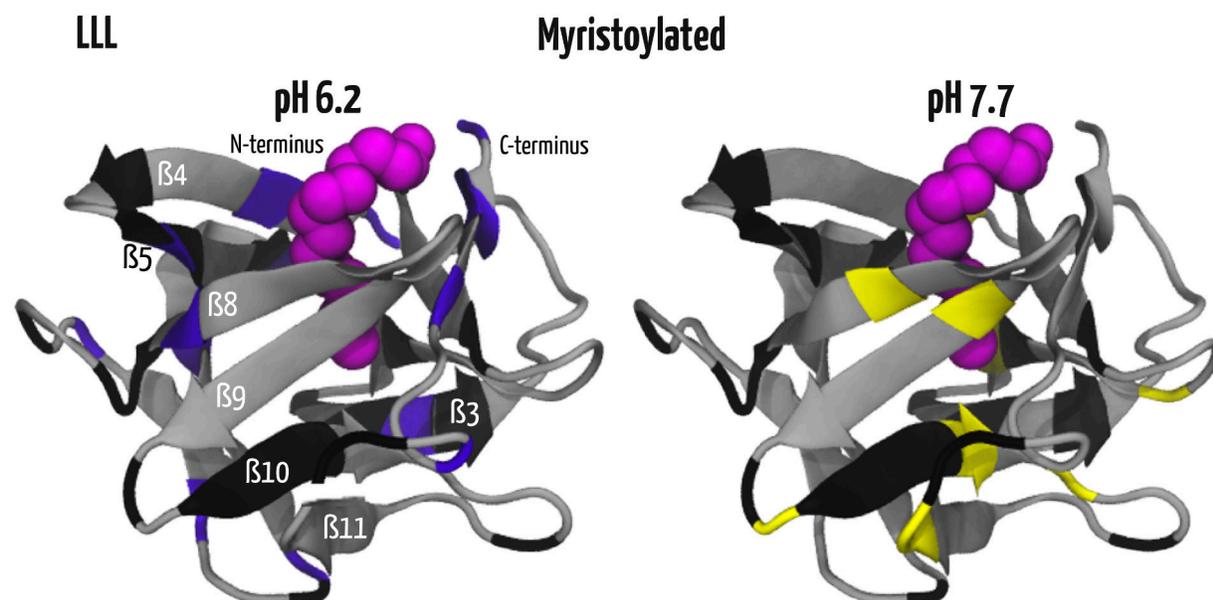
The differences and similarities in curved residues between pHs for non-myristoylated LLL are straightforward: the few similarities (see the black residues in Supplementary Figure S5.1) are enriched in the hairpin triplet end of the protein, especially strand β 10 and a sections of a few loops. The curving residues that are unique to each pH are more diverse, spanning most of the protein and may report on non-switching, innate dynamics. Non-myristoylated LLL is globally destabilized relative to wild-type (by ~ 1.5 kcal/mol; see Figure 4.17), so the increase in curving residues may be due to residues sampling conformations along the unfolding pathway [132,336,339,383].

Myristoylated LLL, on the other hand, has many more curved residues shared between pH 7.7 and pH 6.2 (see the black residues in Supplementary Figure S5.2). Along with β -strand 10, residues with curvature at both pH values also include some of the hairpin triplet at the bottom of the β -barrel, the boundaries of secondary structure at the top of the barrel and in trefoil 1. This supports that the myristoyl group could be a significant modulator of the near-native energy landscape, unifying the dynamics to common outlets rather than the more stochastic distribution seen here in the non-myristoylated form.

pH-dependent changes to curvature in non-myristoylated I85L are relatively few—some additional

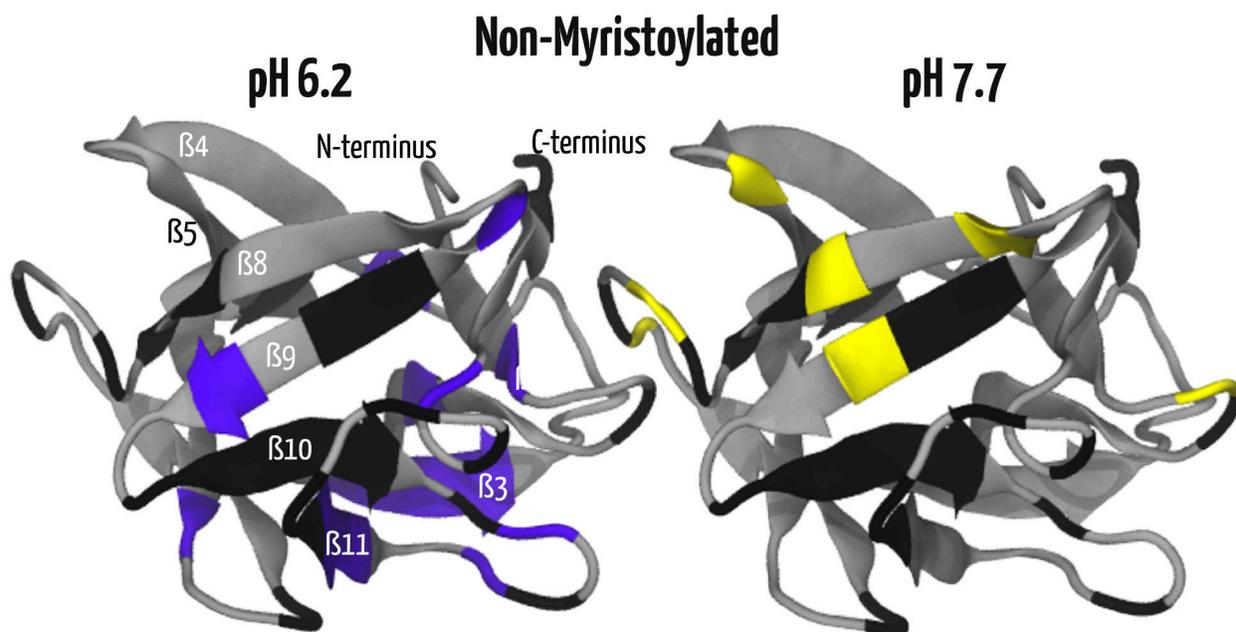


Supplementary Figure S5.1: Structures representing non-myristoylated LLL hisactophilin meant to highlight the pH-dependent differences in which residues are curved. In both structures, residues shown in black are those where curvature is shared at both pH values, residues unique to pH 6.2 are in violet, and those unique to pH 7.7 are in yellow. Shared residues exist mostly in the hairpin triplet region at the bottom of the structure, and otherwise there is not significant agreement between the two; thus, the curved residues may represent some amount of dynamics inherent to non-switching aspects of the protein.

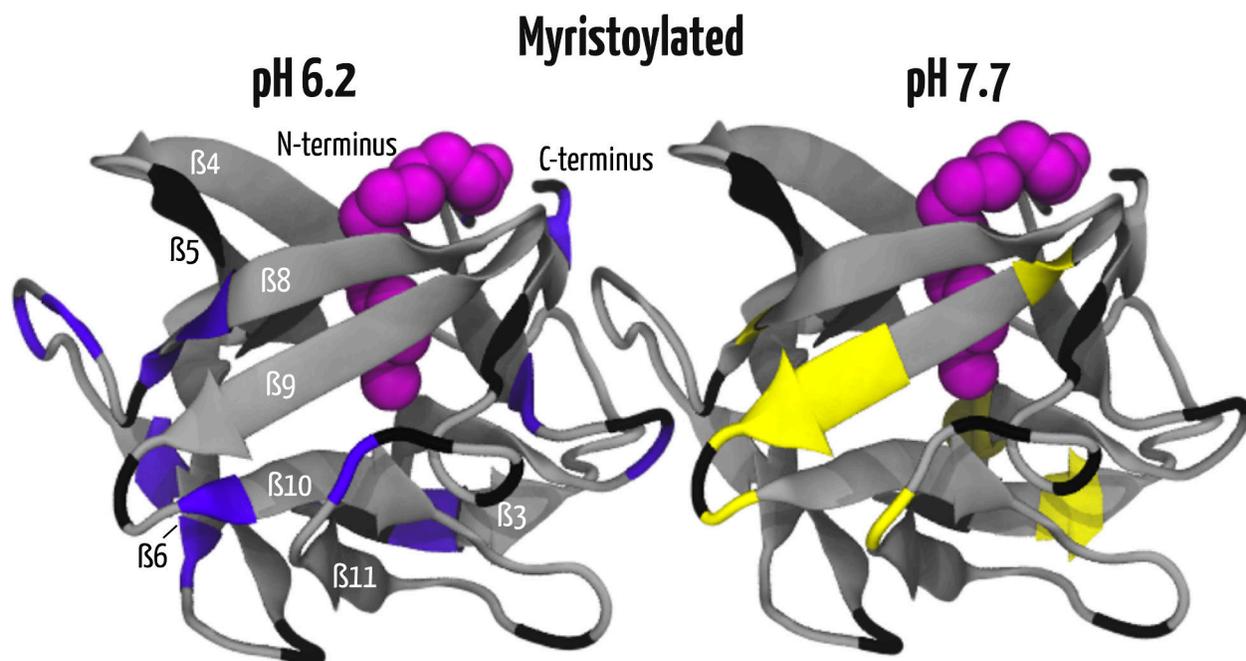


Supplementary Figure S5.2: Structures representing the myristoylated form of LLL hisactophilin to show the pH dependence of which residues are curved. Black residues represent those curved at both pH 6.2 (left) and pH 7.7 (right), violet residues are unique to pH 6.2 and yellow are unique to pH 7.7. The myristoyl group is shown in magenta. Residues that are curved in both are much more widespread than for non-myristoylated, with representatives in many different parts of the structure.

residues become curved at pH 6.2 in the hairpins which, again, could relate to a degree of conformational plasticity required to bind actin. At pH 7.7, additional curvature occurs almost exclusively in the peripheral loops which, incidentally, places those residues adjacent to histidines, or at the boundaries of secondary structure (see Figure S5.3). The myristoylated form (Figure S5.4) tells much the same story, the residues that become curved at either pH are in close proximity to ionizable residues and may be responding to changes in local electrostatics rather than switching as residues adjacent to the myristoyl group show almost no conformational heterogeneity that might report on two distinct states of the fatty acid.



Supplementary Figure S5.3 This shows the pH-dependent changes in the residues with curvature in the I85L mutant of non-myristoylated hisactophilin. Black residues are curved at both pH 7.7 and 6.2, those in violet are uniquely curved at pH 6.2 (left panel) and those in yellow are uniquely curved at pH 7.7 (right panel). Changes with pH in non-myristoylated I85L are relatively modest.



Supplementary Figure S5.4: pH-dependent changes in the residues with curvature in I85L are mapped onto the structure of hisactophilin. The myristoyl group is shown as magenta spheres. Black is curved at both pH 6.2 and pH 7.7. Those in violet are uniquely curved at pH 6.2 (left panel) and those in yellow are curved only at pH 7.7 (right panel). Here, the expected trend of curvature manifesting at the boundaries of secondary structural elements is particularly evident; there are very few incidents of curvature anywhere else. Changes with pH in myristoylated I85L are modest.

Chapter 6: Method Development: automation and bootstraps

6.1 Chapter Abstract

Variable temperature nuclear magnetic resonance is a technique for extracting local conformational stability as well as identifying residues that access to low-energy alternate states in a single series of straightforward two-dimensional heteronuclear single quantum correlation experiments. In an effort to increase the accessibility of the experiments and reproducibility of the subsequent analysis, I developed a workflow and evaluated the fidelity of automating aspects of the analysis. Additionally, we devised a robust hypothesis test for discerning nonlinear amide proton chemical shift temperature dependences using statistical resampling to reduce subjectivity and advantages of experience in residuals. We found that automated tracking of resonances as a function of temperature and applying statistics to detect nonlinearity performed as well as manually tracked data, thus increasing the reproducibility of analysis, enabling high-throughput analysis, and lowering the barrier of entry to the variable temperature NMR experiments.

6.2 Introduction

The uses and boons of variable temperature NMR to measure local stability and characterize the near-native energy landscape have been discussed at some length in Chapters 4 and 5. Though data acquisition and analysis appear straightforward, the latter can be time-consuming, tedious, and determining whether temperature dependences are convincingly nonlinear may rely on expertise and previous experience. This reliance is a hurdle for standardizing the method and interpretation as well as harnessing a graduate student's most robust resource—an undergraduate student. Thus, automating the analysis of temperature coefficients to extract local stability and conformational heterogeneity would speed the process, reduce tedium, and diminish the contributions of individual biases in chemical shift peak tracking and decision-making regarding nonlinearities. Variable temperature NMR data are hearty, but that two people could look at the same data set and draw different conclusions for fringe cases of curvature is distressing. Combined with the implied challenge by Dr. Williamson [336], "...a residual plot, which was examined by eye to determine whether it was curved. This proved a more reliable and robust method than any computer-based method I could devise.", it behooved me to try to improve the workflow for analyzing these data and implement/evaluate automation.

There is no shortage of custom tools for analyzing NMR data, such as TITAN for 2D lineshape analysis [384], Trace in Track (TINT) for tracking peaks during stepwise perturbations of a system [385], NMRFAM-Sparky [386], CARA [225], or CCPN [263] for assigning and general analysis, ENSEMBLE [387] for analyzing the unfolded ensemble of structures and/or studying intrinsically disordered proteins, ISThms [388,389]

for reconstructing non-uniformly sampled data sets, β P [390] for determining the secondary structure propensity, RELAX [391] and CATIA [392] for analyzing faster protein dynamics, and CHESCA [186,268] for identifying allosteric pathways using chemical shift covariance upon perturbation. Though nowhere near an exhaustive list, for every new method or technique there is a new, custom application, script, or workflow for analyzing data, and analyzing the temperature dependence of amide chemical shifts is no different!

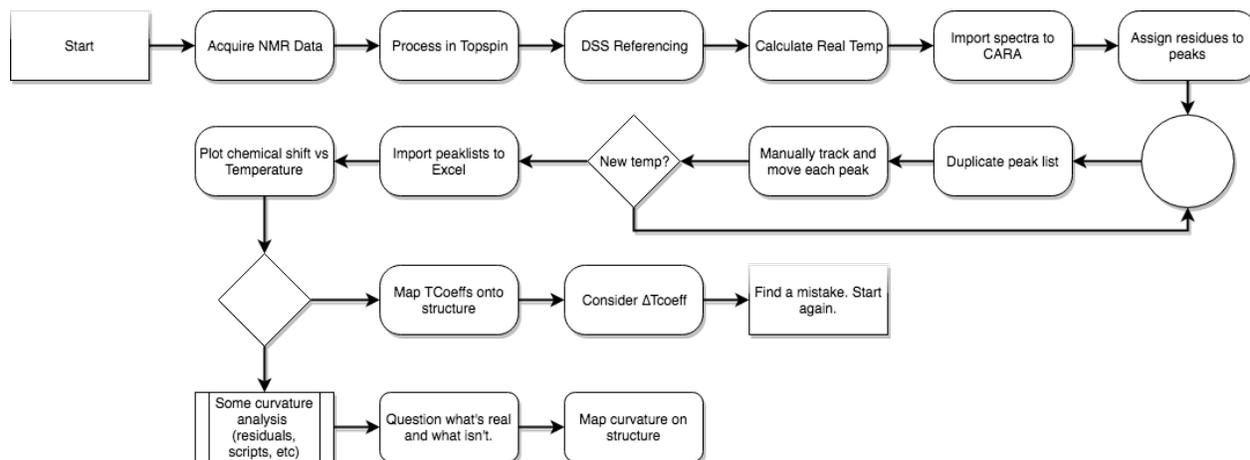


Figure 6.1: Flowchart demonstrating the first iteration and least efficient general workflow for analyzing variable temperature data. Data were acquired by NMR with 12 or more temperature points, ending no less than 15 degrees below the melting temperature of the protein with increments spread approximately evenly over the range. Spectra are processed to balance signal and resolution in the program of choice (e.g. Topspin, NMRPipe [393]) and referenced to DSS. The chemical shift of water is used to calculate the internal temperature [259] to be used in subsequent regressions. Spectra are then imported into CARA [225] for manual peak picking, transferring assignments, and tracking peaks across temperatures. The resulting chemical shift data were exported to Excel for linear regression by least squares and the slopes constituted the temperature coefficients that were mapped onto hisactophilin's backbone. The results of the regression were also used for determining nonlinear temperature dependences by plotting the residuals.

The workflow for analyzing variable temperature data is to acquire and reference the NMR spectra, calculate the sample temperature based on the chemical shift of water [259], transfer existing resonance assignments to the spectrum with the temperature closest to that of the assignments, propagate those assignments to the remaining variable temperature spectra, extract the chemical shifts for each peak at each temperature, plot the data and apply a linear regression, and finally calculate and plot the residuals. This manually intensive workflow is summarized in the flowchart in Figure 6.1. The largest obstacle for rapid data analysis is picking peaks and propagating resonance assignments for each spectrum (~100 assignments for each of 12 - 21 temperatures). The tradeoff for the lack of speed, however, is that each peak/chemical shift checked, audited, and as accurate as possible. Thus, these data are a benchmark against which automation can be compared and evaluated.

6.3 Evaluating Automation

Kyle Trainor, a graduate student in the Meiering Lab, wrote a Python script and GUI (Graphical User Interface) tentatively called Shift-T that I alpha tested with hisactophilin data. Shift-T automates propagating a large percentage of assignments across the set of temperatures/spectra. While manual intervention is still required for 15-25% of residues (overlapping peaks for a subset of spectra presents a difficult problem for automation), the time saved is significant. The input for Shift-T are the actual experimental temperatures (determined from the chemical shift separation difference of water and DSS [259]), the chemical shift of DSS, the resonance assignments, and most importantly, unassigned peaklists for each spectrum. The source of the peaklists is not restrictive; they can derive from any capable program or algorithm. Peaks chosen automatically by Bruker TopSpin 3.5's built-in algorithm were benchmarked against the manually chosen peaks. Figure 6.2 (upper panels) shows two representative residues, V43 and Y62 with the respective linear regressions of the two sets of peaks. The automated and manually produced peak sets yielded very similar temperature coefficients. When evaluated over all residues of hisactophilin, most temperature coefficients varied by only 1-5% with only a few cases at 5-10% between the manually and automatically generated peaks. On one hand, this speaks to the robustness of temperature coefficients against small variations in chemical shift. On the other hand, individual points on the plots from automatically picked peaks are objectively noisier as evinced by lower correlation coefficients.

Beyond the linear temperature dependences that give temperature coefficients, variable temperature experiments are also candidates for analysis for nonlinearity (i.e. curvature). Confidently detecting curvature requires high experimental sensitivity that is enhanced by more temperature data than most other published variable temperature experiments (we recommend 12 or more evenly spaced temperatures) [182,336,394,395]. Noise in determining the chemical shifts of some peaks may obscure the subtle effects of a 1-5% population of excited state on the chemical shifts. The lower panels of Figure 6.2 show the residuals of the data following subtraction of the original data from the linear regression. For residue V43, the dark grey points of the manually picked data demonstrate a non-random distribution around zero (appearing as an inverted parabola), indicating that the residue may access a near-native excited state. However, when the data are processed with the automatic peaks (light grey), the apparent curvature is obfuscated.

Solving the inefficiency and mire of manually tracking temperature-dependent peak movements with Shift-T yielded robust temperature coefficients, but created the problem of processed data that were not sufficiently accurate for robust curvature analysis. The most curved residues were still evident, but subtle curvature that might arise from an excited state whose energy level is closer to the ~ 5 kcal/mol limit were lost. In principle, Shift-T searches the input peak lists for resonances that moved distances within an empirical cutoff and that followed an approximately linear trajectory, so there is nothing implicit to its algorithm that

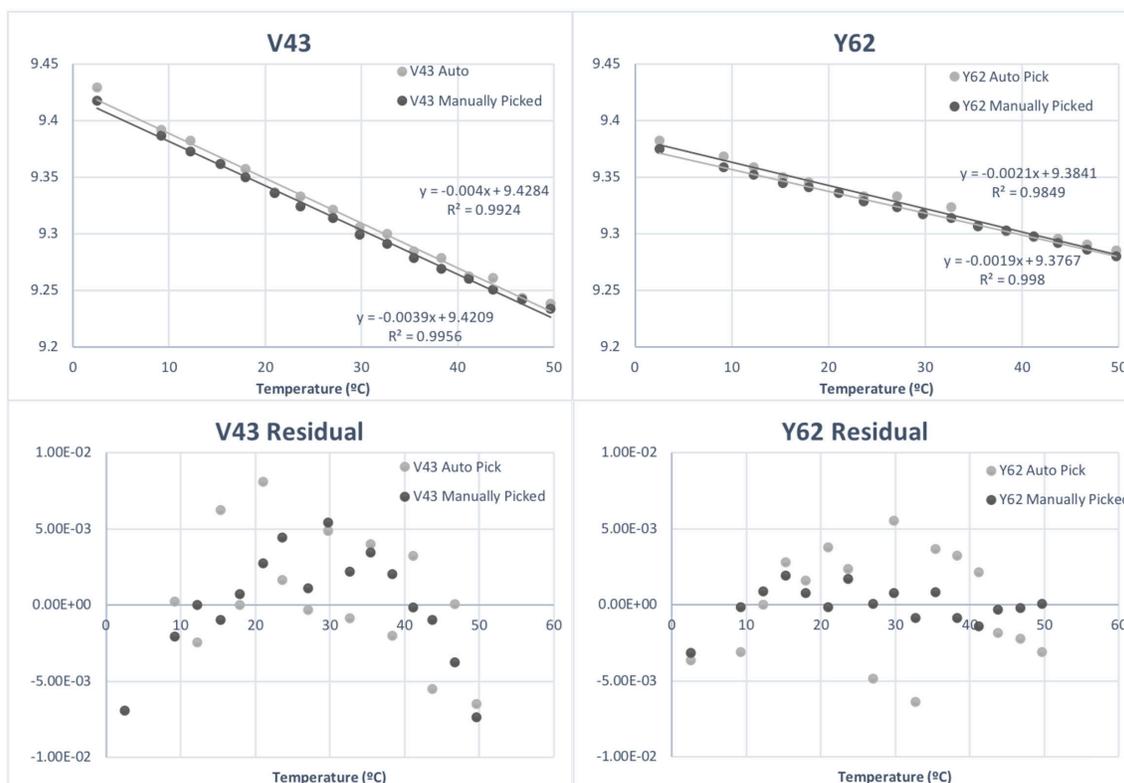


Figure 6.2: Plots of the temperature dependence of amide proton chemical shifts for residues V43 and Y62 (upper panels) for myristoylated wild-type at pH 6.2 from manually picked peaks (black) and those generated by Bruker TopSpin 3.5 (light grey). These two residues represent a typical chemical shift temperature dependence. The slopes of the lines (the temperature coefficients) are very similar which speaks to the robustness of these experiments and analyses, however, the generated peaks are objectively noisier which can be seen in the slightly worse correlation coefficients. The lower panels are the residuals after subtracting the linear regression for the acquired data. The manually picked peaks for V43 show signs of curvature while it is obscured in the automatically picked peaks by the uncertainty in picking.

should muddy the peak positions. In the spirit of the “garbage in, garbage out” principle of computing, the input peaklist derived from the peak picking algorithm was the first suspect.

6.4 Peak Picking Algorithms Impact Performance

An investigation yielded that barreling ahead with analyzing temperature dependence data under time pressure may lead one to overlook a critical option when generating peaklists using TopSpin’s algorithms. A spectrum comprises discrete data points that approximate the shape of absorption peaks. However, it is possible that no discrete point was measured at the inferred maximum of the peak. By default, TopSpin searches for peaks by locating maxima among the discrete points of the spectrum that are greater than a threshold. The result might be peaks picked at the measured point closest to the maximum, which in practice may not be representative of the maximum in the context of an absorption peak shape. Depending on how data points define a peak, the maximum might be chosen differently one experiment to the next and introduce enough noise to render curvature detection very difficult.

More sophisticated peak detection solved the problem. There is the option in TopSpin to pick peaks based on parabolic interpolation where points near the maximum are fit to a quadratic function in each dimension and the maximum is determined regardless of whether a real point was acquired there. All else being equal, enabling interpolation on a well-phased spectrum should provide more accurate peak localization and chemical shifts, and decrease the contribution of noise from peak picking in analyzed data.

In Figure 6.3, the temperature dependences of residues V43 and Y62 are shown for the manually picked dataset (black) and the peaks automatically detected with parabolic interpolation enabled (light grey). There is strong agreement between the manually and automatically picked peaks in this case and the temperature coefficients are, for all intents and purposes, identical. Visually, scatter in the automated peaks is lessened compared to the dataset that was not interpolated (Figure 6.2). The residuals in the lower panels indicate a heartening agreement between the two methods of peak picking. Not only do many residual points in the two sets for V43 overlap, but interpolation conserved the character of the plot—V43 appeared parabolic in V43 and fairly random for Y62. Thus, automating peak picking and tracking across temperatures, when carried out correctly, can faithfully recapture the results from tedious and painstaking manual workflows. Analyzing temperature dependence data is effectively sped up by days (if not weeks) without appreciable degradation of accuracy or precision of the data. The workflow shown in Figure 6.1 was changed to Figure 6.4.

Peak picking algorithms are a feature of many spectral analysis software packages (e.g. CCPN, NMRFAM-Sparky) as well as TopSpin, and alternative software packages that employ similar algorithms may perform comparably. Specifically, peak finding in CCPN has tuneable parameters for sensitivity and tools for finding peaks in crowded areas so it may provide a viable alternative for those opposed to using

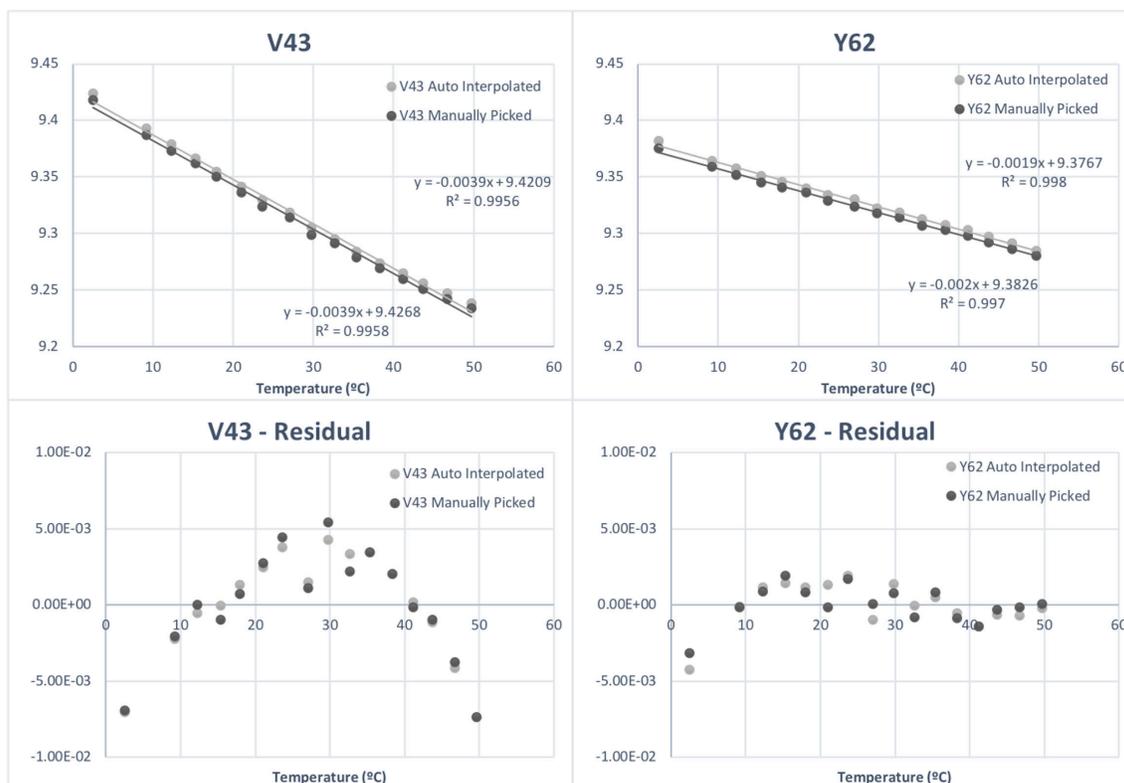


Figure 6.3: Plots of the temperature dependence of amide proton chemical shifts for residues V43 and Y62 (upper panels) for myristoylated wild-type at pH 6.2. Black points represent those peaks that were manually picked, while light grey are data from a peak picking algorithm that used parabolic interpolation. These two residues are fairly typical cases. The slopes of the lines (the temperature coefficients) for the two residues are nearly identical with very good agreement between correlation coefficients as well. The lower panels show the residuals following subtraction of the data from the linear regression. Unlike for the automated data without interpolation, here the residuals match extremely well with the manually picked data and the character of the residual is preserved (curved or more random, for example).

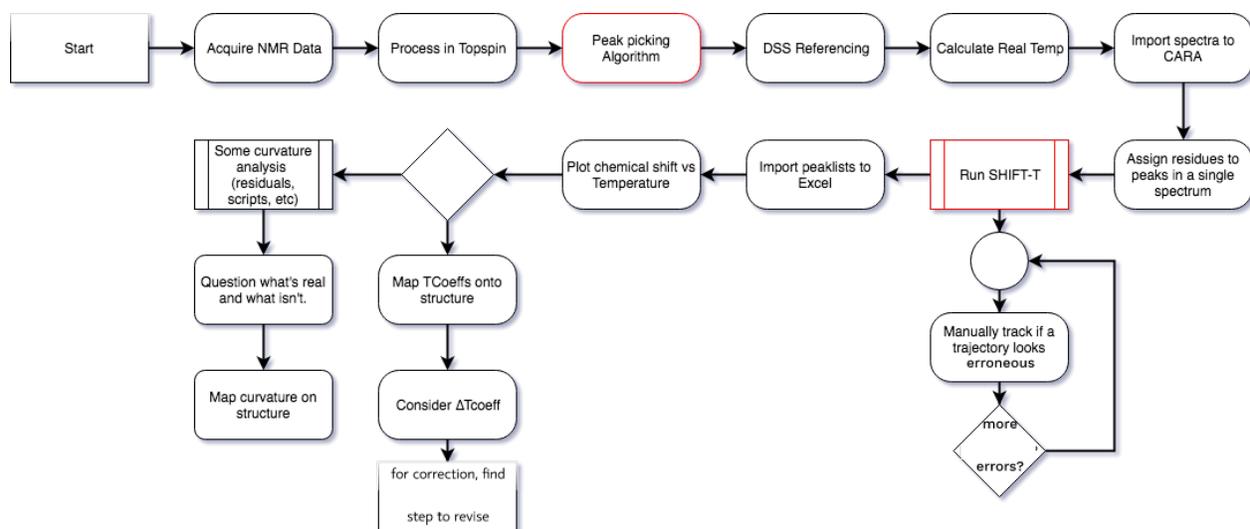


Figure 6.4: Flowchart demonstrating a further iteration of the workflow for semi-automated chemical shift temperature dependence analysis, looking at local stability (linear temperature dependence) and conformational heterogeneity (nonlinear temperature dependence). Contrasting with the first iteration in Figure 6.1, a peak picking algorithm is implemented early in the process, peak tracking is executed by Shift-T with some manual intervention to resolve ambiguities. The bottleneck is now shifted to resolving ambiguities and analyzing the data and their structural interpretation.

TopSpin for analysis, or that prefer to process spectra with NMRPipe [393] before peak finding/picking.

6.5 Excising Experience

Automating the analysis of variable temperature data with Kyle's Shift-T script accelerated work on his-actophilin and simultaneously reduced some of the expertise involved in consistent peak-finding. However, nonlinear amide proton temperature dependences (i.e. curvature analysis) may suffer from the greatest variation due to individual differences in determining whether temperature dependences are curved. As already mentioned, I am interested in reducing individual biases to increase reproducibility and make temperature dependence experiments as accessible as possible.

Defining and determining curvature breaks down into two parts: first is to mathematically and statistically identify curvature; second is to make identification robust and with high confidence. The purpose of the latter is to ensure curvature does not randomly arise from noise nor that it is anchored or unduly influenced by an outlier. The first approach for identifying curvature involved the residual plots that follow linear regression. While there are many informative plots that can help illustrate the goodness-of-fit of a model (e.g. a quantile-quantile, Q-Q, plot), they require a researcher to subjectively determine whether the plot shows signs of nonlinearity, heteroscedasticity, etc. However, the overarching goal was to excise the researcher from as much of the process as possible and develop an analysis that gives numerical output for a straightforward hypothesis test.

In principle, if the data are well-fit by a model, the residual should be approximately normally distributed around zero without a discernible pattern. Conversely, if data are not well described by a model, the residual should show some systematic, non-random deviation from the fit. Testing a residual for normality offers an avenue for determining whether temperature dependences are sufficiently well-described by a linear model or require higher-order considerations.

Some independent research and guidance from the Statistical Consulting and Collaborative Research Unit at the University of Waterloo led us to try three tests that evaluate a residual for characteristics of normality: the Jarque-Bera [396], the Kolmogorov-Smirnov [397,398], and the Shapiro-Wilk [399]. The tests were applied in R [357] from the in-built `{stats}` and the `{tseries}` packages [400]. In brief, the Jarque-Bera test endeavors to test for normality by matching the kurtosis and skewness of the linear regression residuals to those expected of a normal distribution, which is to say that the shape of the distribution is close to symmetric and that tails are not over-large or -small. The Kolmogorov-Smirnov goodness-of-fit test takes a similar tack and evaluates whether two data sets have the same distribution, and if one of them is known to be normal the other is implicitly tested for normality. Lastly, the Shapiro-Wilk test looks at whether the population from which a random sample is drawn is itself normally distributed.

The trial data set included regressions that were highly linear and others whose residuals were highly curved to determine how discriminating the three tests were. Disappointingly, with temperature dependence data, all tests indicated normality, even in cases where there was visible curvature. Further research discovered many possible reasons for the discrepancy. The Jarque-Bera test, is primarily deployed for large data sets because other normality tests can fail when $n > 2,000$ - the implication being that our temperature series of $n \sim 15$ might not be sufficiently large for the test to be effective. The Kolmogorov-Smirnov requires some knowledge of the shape parameters of the distribution and estimating them significantly weakens the test. Lastly, the Shapiro-Wilk is considered a high-power test appropriate for use with small sample sizes, but it does still suffer from a sample size bias wherein more data points are better. A physical biochemist's idea of a large sample size is not the same as a statistician's, so in this context and for this purpose, these techniques are not suited for so (arguably) few points.

We next adopted a new strategy for defining curvature where we evaluate whether an alternative to a linear regression is a significantly better descriptor of the data. The method involved a few steps, first of which was determining the linear regression residual (the model we are testing). If the data are well-described by a linear relationship, then the residual will random, as it approximates the error in the relationship between temperature and chemical shift. If a linear model is insufficient, then a non-random structure should appear in the residual and a polynomial should yield a better fit of the data. In practical terms, we applied an analysis of variance (ANOVA) to compare the linear and polynomial regressions of the residual to test the hypothesis that the polynomial (i.e. quadratic) is a significantly better fit.

After developing an R script to evaluate the two models, we discovered that there is a standard R package, `{car}` (R companion to applied regression) [401] with our test built-in to the `residualPlots` function under the name of Tukey’s test for non-additivity [402]. Tukey’s test for non-additivity (note that there are a few different tests named for Tukey) assesses whether there is an additive relationship between the factor variable and the expected response variable. In other words, are the residuals of a regression better fit to a higher order polynomial. Feeding a test set of temperature dependences into the script yielded positive results for curvature at the 99% confidence level for the same residues identified by an experienced eye. To ensure that the effectiveness of this test was not limited to hisactophilin, we applied it to two other proteins (*vide infra*) and reached the same conclusion.

6.6 Anomalies, Anchors, and Application

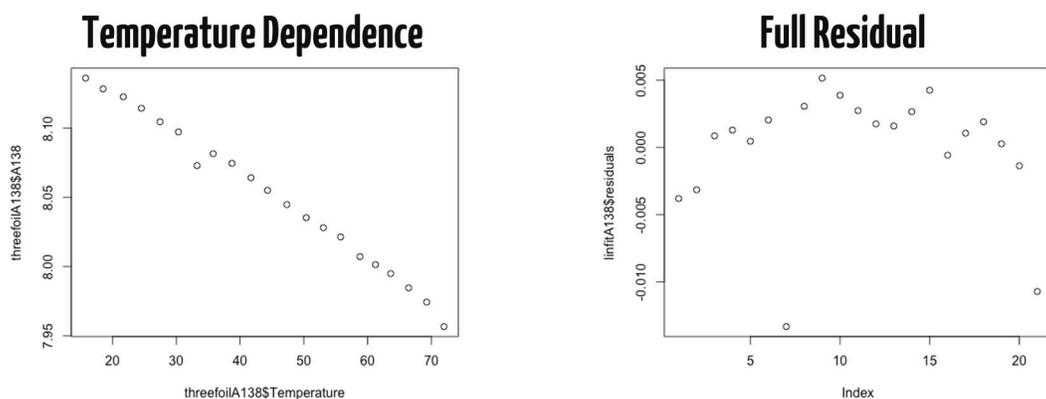
Having addressed the bottleneck of peak finding and tracking as well as reducing the amount of individual bias in determining curvature, one of the boons of using curved temperature dependences is also its bane — sensitivity. Small systematic deflections from linearity provide a wealth of information on the near-native energy landscape, however, experimental noise or uncertainty can obscure these cases (as in Figure 6.2). Therein lies the next target for optimization—confidence. We sought a method to test that observed curvature did not arise stochastically from noise and uncertainty or from the disproportionate weight of points that might be outliers without resorting to extensive replication studies. We decided to deploy the statistical resampling method called bootstrapping [403]. To briefly explain how it works, imagine a data set of 20 temperatures is evaluated for curvature. Next, a new bootstrap sample of size 20 is randomly selected from the original data with replacement, meaning that some points might be omitted while others might be chosen more than once. The statistics are evaluated and recorded for the bootstrap sample, and then repeated by randomly drawing another sample of size 20. This procedure is repeated hundreds or thousands of times and each replication produces a new test statistic. This method relies on the principle that bootstrapped samples approximate the distribution of the test statistic in the same way that a robust, random sample approximates the distribution of test statistic in a population [404]. Random sampling with replacement from an experimental sample is almost like random sampling from the true population. With a sufficiently large number of bootstrap samples, the bootstrap statistic and a Monte Carlo approximation are nearly the same (discussed in a great deal more depth and with greater eloquence in [405]). Furthermore, from a statistics standpoint, bootstrapping is advantageous over other methods of statistical inference because it makes no assumption about the distribution of the data set (e.g. normality) and doesn’t require the distribution to first be measured [406].

To bootstrap the temperature dependence data sets, the `boot` package and function in R [407,408] were integrated into the aforementioned curvature analysis script (now called `Boot-T`) with instructions to create

1000 bootstrap resamples. 1000 resamples is the basic recommendation for bootstrapping because it balances the computation time of the demanding calculations with the statistical power and still large enough for confidence interval estimation [403,405]. The lower section of Figure 6.5 demonstrates eight bootstrap residuals from the set of 1000 resamples. The full range of temperatures comprises 21 values spanning approximately 10 - 70 °C for the engineered protein Threefoil [100] acquired with Dalia Naser, who kindly provided one of her hard-fought-for ¹⁵N-labelled samples for the experiment. Some of the bootstrap residuals shown had as few as seven unique points (meaning the same points were resampled many times) or as many as thirteen, however it was also possible that one resampling could have one occurrence of each of the 21 temperatures. The full temperature dependence has one clearly aberrant point between 30 and 40 °C (NB: it was due to a referencing error, but that is not germane here). As expected, that point is obvious in many of the residuals and, in the third residual, exerts enough influence to reverse the direction of curvature away from many of the other points. Some of the usefulness of bootstrapping is evident in that the problematic point is included in the fitting of only a subset of the bootstrap residuals, so over the 1000 different resamples, the influence of that one point on curvature is significantly decreased. The Boot-T analysis argues that the residual of this amide proton is not best described by a parabola at the 95% confidence level. This is especially interesting because the traditional analysis by a t-test including all data points would consider the residue curved at the same confidence level.

The discrepancy between the bootstrapped statistic and a more traditionally applied t-test (or determining by eye) further highlights why the test is beneficial. Fringe cases where, by eye, slight curvature could as readily be noise as real is where a systematic treatment of the data is most valuable. Threefoil is rife with such cases. Figure 6.6 demonstrates select fringe cases. Plots labelled amides 1, 2, and 3 were all considered candidates for having curvature. By eye they appeared like they could be curved and the non-bootstrapped hypothesis test (t-test) agreed. However, Boot-T disagreed, indicating that the data were insufficient to confidently invoke curvature in the residual. By eye, amides 4, 5, and 6 were not curved by eye, but this time the t-test supported curvature for all three. Boot-T can resolve the disagreement — it assesses amides 5 and 6 as not curved and residue 4 as curved. So, the bootstrap analysis is sufficiently sensitive and statistically justified to discern cases where temperature dependences are on boundary between curved and not and where a less-experienced (even an experienced) eye may struggle for consistency. When the less robust t-test was applied to the whole set of (as yet unassigned) Threefoil temperature dependences, 32 (of 52 amides) met the criteria for curvature. Boot-T reduced that number to 19, which was in line with the 18 determined by a trained eye. As discussed elsewhere, nonlinear temperature coefficients may indicate conformational heterogeneity, extensive signs of which are not necessarily expected for Threefoil because of its incredibly high kinetic stability and resistance to protease digestion [99,409].

Copper, zinc human superoxide dismutase 1 (SOD1) is a metalloprotein protein associated with familial



Bootstrap Residuals

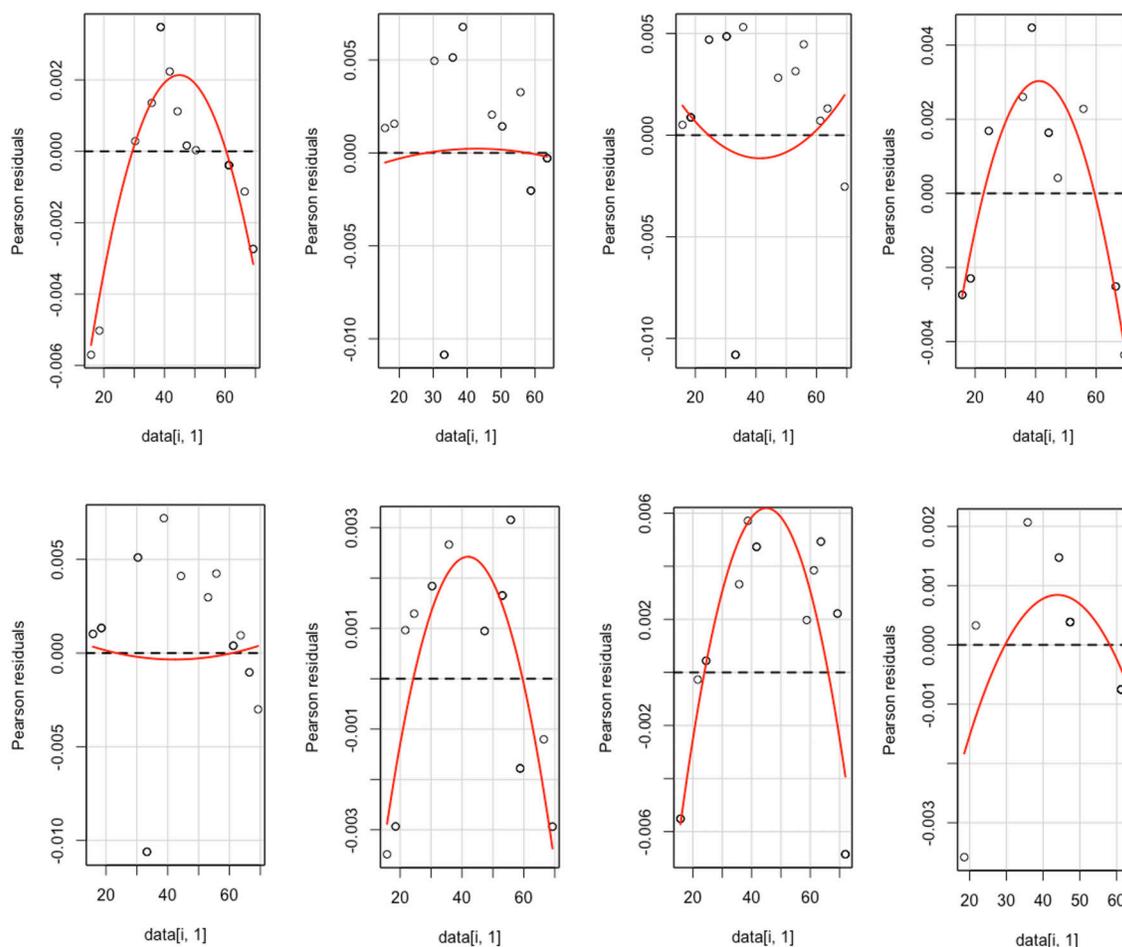


Figure 6.5: The upper section of this figure shows the full 21 point temperature dependence of an unassigned amide proton from the engineered protein Threefoil. The protein was purified with great effort by Dalia Naser with some consultation, and we collaborated on the temperature dependence experiment. Data were acquired at 2.5 °C intervals with real temperatures that ranged from ~15 - 72 °C. They were referenced to DSS following experiments on a Bruker Avance 600 MHz spectrometer. Data were acquired with software Xwin NMR, processed in TopSpin 3.5 and peak trajectories were traced with Shift-T. The plot in the top right of the figure is the residual following linear regression for all points. Below the break are 8 example bootstrap residuals derived from using the boot function in R. The red line represents an attempt to fit the residual to a quadratic function.

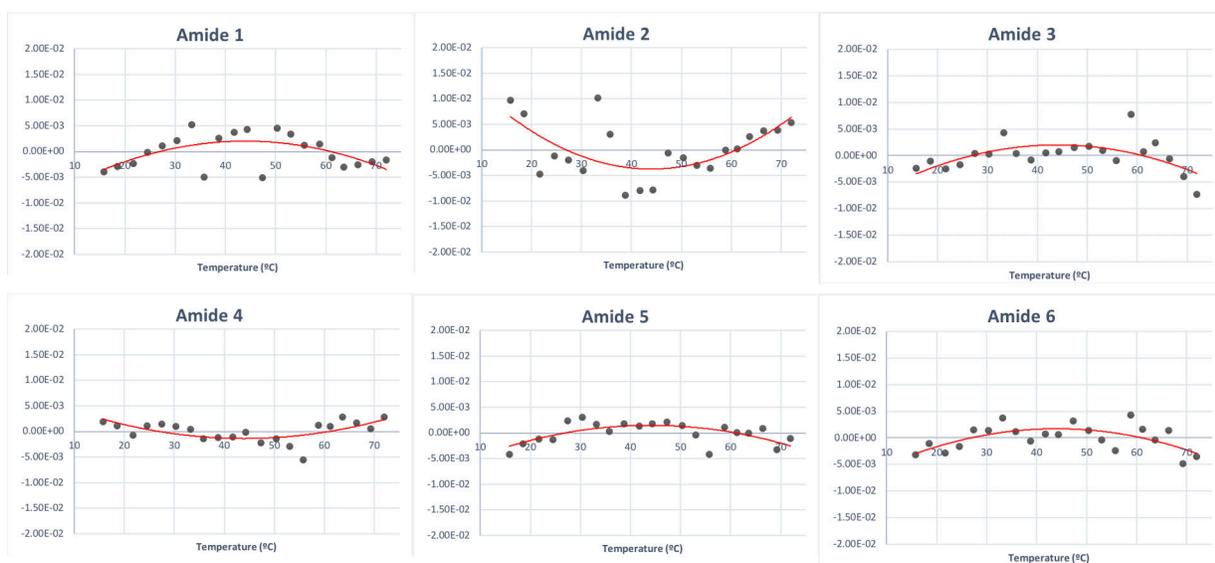


Figure 6.6: Amide proton temperature dependences with ambiguous curvature, which by eye and a more canonical t-test analysis disagreed with one another or with the bootstrap analysis. By eye, amides 1, 2, and 3 were thought to curve and a t-test concurred. Bootstrap analysis, however, deemed the data insufficient for confidence in curvature. Residues 4, 5, and 6 were flagged as curved by the t-test but that disagreed with the eye. For residues 5 and 6, the bootstrap sided with not curved and residue 4 as curved. Here, the bootstrap functions as the robust arbiter of curvature, as free as possible from researcher bias and expectation.

amyotrophic lateral sclerosis and a subject of intense study in the Meiering lab. SOD1 matures from an apo monomer to binding zinc and copper, with disulphide bond formation (via help of a chaperone), and dimerization, though the order of steps is not yet fully clear [410,411]. It has been proposed that SOD1's role in disease involves immature forms of the protein and how disease-associated mutations impact those forms [412–415]. Temperature dependence data have been already reported for the fully mature holo SOD1, as well as the least mature reduced apo form [150]. Some of the technical challenges of studying SOD1 at intermediate stages of maturation (i.e. with one zinc bound per monomer) were overcome by another graduate student in the lab, Harmeen Deol [416]. With her and an undergraduate student, Michael Tarasca, we measured the temperature dependence of amide proton chemical shifts for the disulphide-reduced pseudo wild-type SOD1 with one zinc bound from 10 °C to 50 °C at 2.5 °C increments and a final point at 55 °C. Deploying the aforementioned methods, scripts, and strategies, analyzing the data for the 96 assigned resonances in conjunction with the undergraduate Jeffrey Palumbo was relatively straightforward and painless. Boot-T enabled Jeffrey, with minimal training or experience in studying curvature, to confidently localize regions that accessed low-lying excited states.

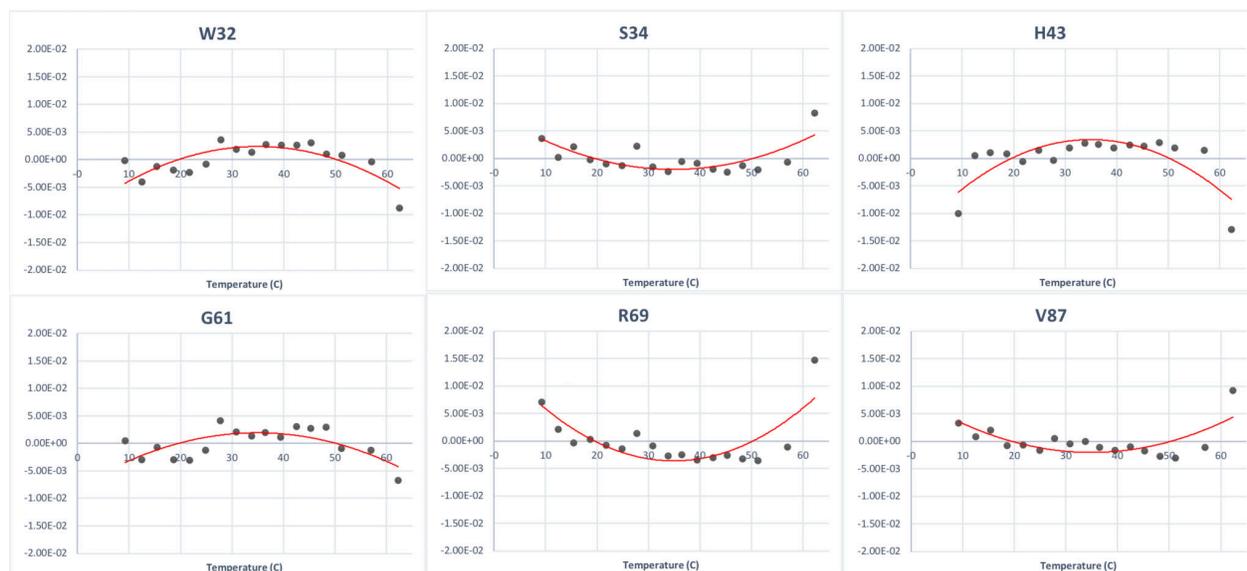


Figure 6.7: Residuals resulting from temperature dependence experiments on reduced pseudo wild-type SOD1 with approximately one stoichiometric equivalent of ZnSO_4 added in HEPES buffer with TCEP to maintain the reducing environment. When tested with traditional t-test described elsewhere [150], the temperature dependence for each of the six residues (W32, S34, H43, G61, R69, and V87) exhibited nonlinear behaviour. When tested with the Boot-T script (1000 bootstrap resamples), these residues were reclassified as having insufficient evidence for nonlinearity. Second order polynomial regression is shown as a red line. The temperature range for the experiment was 9 - 62 °C.

Using data from the one zinc-bound SOD, compared to the traditional t-test, Boot-T analysis had a significant impact on the number of residues labelled with curvature. Note, however, that for the t-test, there

is no preliminary evaluation by eye, all residues are assigned a p-value regardless of whether or not there is reason to believe they are curved (cutoff of $p < 0.01$). In principle, the t-test could be used to corroborate the trained eye's assessment. Over the entire dataset, the t-test nominated a staggering 75 residues as curved. However, the residuals of some of those residues are visibly ambiguous or the curvature attributable to just a few anchoring points. Figure 6.7 highlights six residues that the t-test called curved but that subsequent analysis with Boot-T deemed were not meaningfully curved. Visual inspection makes it understandable why the traditional t-test would support curvature; in H43, for example, the deviation of the first and last points from the fit strongly impact the character of the residual, seeming to force a parabolic shape when the intervening points are more random and scattered. The cases of individual points affecting the residual is also apparent in V87, R69, S34, and W32. G61, on the other hand appears more uniformly curved, which may in fact be true but the data are too noisy to make that assertion with confidence.

Boot-T was applied to all the measured temperature dependences of zinc-bound reduced SOD1, the number of curved residues decreased from 75 to 53 at a 95% confidence level. As an interesting functional aside, the large proportion of curved residues are likely to reflect significant conformational changes: this form of SOD1 has a dimer dissociation constant of $51 \mu\text{M}$ (ΔG of 5.8 kcal/mol, at the limit of detection by curvature) at 25 °C [135], and the dimer interface is extensively coupled with the rest of the protein structure. Conformational heterogeneity here may attest to the plasticity required to facilitate interactions with varied binding partners, either another SOD1 monomer for dimerization (which is dominant under these conditions [416]) or the copper chaperone for SOD1 (CCS)—thought to be responsible for delivering copper to immature SOD1 [411]. CCS binds to the SOD1 monomer with much overlap with the native dimer interface due to structural similarities between the CCS and SOD1 binding surfaces, so it is logical that copper-deficient SOD1 might have an excited state competent for or reflective of CCS binding. Regardless, simply and confidently detecting nonlinear temperature dependences of chemical shifts is a tantalizing tool for characterizing the near-native energy landscape as it relates to biologically relevant conformational excursions. Boot-T has robustified the analysis for curvature, and while many of the 22 residues in SOD1 that lost status would likely have been caught by a trained eye, Boot-T was discriminating enough that a trained eye was not necessary. In Threefoil, the ambiguity and inherent difficulty of classifying many residuals was resolved and made consistent with Boot-T.

6.7 Concluding remarks

Temperature dependence of amide proton chemical shifts, as discussed ad nauseum in Chapters 4 and 5, offers a window into local structural stability and conformational heterogeneity by plying the exquisite sensitivity of amide proton chemical shifts to identify structural perturbations. Coupled with the goal to apply it to

many variants of many different proteins as a component of a standard toolbox, it behooved me (and others) to remove as much of the tedium from analysis as possible and lower the extent of onboarding before being able to provide confident, meaningful, and reproducible analysis methods for new researchers. Thus, by analyzing full data sets determined manually as a benchmark for subsequent efforts at automation, it should not be necessary for any other student to be subjected to that extent of manual processing going forward. I showed that automatic peak picking algorithms, when properly deployed, lead to robust temperature coefficients with little enough noise to accurately detect conformational heterogeneity by curvature analysis. I also sought to rise to the challenge issued by the Williamson group [[336]} and recruit computational methods that were better than (or at least comparable to) the eye. Granted, fifteen years' worth of technological advancement might have helped. In any case, computational analysis of curvature has reduced the necessity for a highly trained eye to detect nonlinear temperature dependences and thus increased the accessibility of the method. Lastly, adding statistical resampling with replacement (ie bootstrapping) to the curvature analysis workflow has robustified the highly sensitive findings against undue influence from experimental noise and uncertainties in determining chemical shifts. Deploying the new workflows in other proteins like SOD1 and Threefoil speak to the generalizable application of the methods, and preliminary analysis of a fresh variable temperature dataset may now reveal the first glimpses of a protein's secrets in a busy afternoon rather than (an optimistic) one to (a realistic) two weeks.

Chapter 7: Proof of concept and future works

7.1 Chapter Abstract

Regulated protein interactions with other proteins or a membrane is an important aspect of myristoyl switching in signalling activity. High resolution structural detail of membrane binding has been elusive in some myristoyl switching proteins owing to the difficulties of crystallizing highly dynamic and/or membrane-associated proteins, or to the size of the complex in solution-state nuclear magnetic resonance, greatly limiting our overall understanding of this class of proteins. Here we present a preliminary nuclear magnetic resonance study of hisactophilin encapsulated in a reverse micelle to obtain novel insights into its membrane-bound structure. Encapsulating hisactophilin revealed that the structure of membrane-bound myristoylated hisactophilin extensively overlaps with the non-myristoylated form, consistent with limited or no interactions between the myristoyl group and the rest of the protein. Results support the predictions and modelling that extrusion of the myristoyl group does not significantly impact the overall structure of hisactophilin. Furthermore, amide proton chemical shift changes upon titration of cardiolipin into the reverse micelle to mimic the negative charge of some biological membranes suggests that cardiolipin may interact with charged residues around the top of the β -barrel. The success of the proof-of-principle experiments drastically increases the scope of investigations on hisactophilin to include further structural and dynamic studies while encapsulated in a reverse micelle, and could additionally allow hisactophilin's other binding interactions, for example with actin, to be investigated.

7.2 Introduction

To date, studies of hisactophilin have focused predominantly on its cytosolic form in the accessible or sequestered states (dashed box in Figure 1.4). When the pH is below 6.9 and natively expressed hisactophilin would bind a membrane, however, none of the stability measurements or NMR studies presented thus far have included lipids to which hisactophilin can anchor. Interactions at and with the membrane interface are one of the next available frontiers for discovery in hisactophilin. There are tools available for studying membrane binding events that have successfully been deployed in other systems, including but not limited to: the myristoylated ADP ribosylation factor in bicelles [286]; the binding activity of the Sec-Y interacting protein, Syd, using nanodiscs [417,418]; numerous studies with large and small unilamellar vesicles, including hisactophilin [11], cathelicidin antimicrobial peptides [419], intrinsically disordered α -synuclein [420]; and the folding of DsbB in micelles [421,422]. Each membrane mimetic has its boons and banes that make it amenable or worth avoiding for a given protein. These tools and more are discussed in the following 2011

review [423], particularly how they apply in NMR.

Exploratory attempts were made previously to observe binding between membrane and hisactophilin using DMPG/DHPC/DMPC-containing bicelles [312]. At that time, though circular dichroism showed evidence of reversible binding, the poor quality and high-unreadable NMR spectra precluded any further structural studies with that bicelle formulation. Recently, however, I had the opportunity to collaborate with Dr. A. Joshua Wand at the University of Pennsylvania. One of Dr. Wand's specialties is encapsulating proteins, large and small, detergent-bound and free, in reverse micelles [424–427]. Hisactophilin, a peripheral membrane protein, made an alluring candidate for encapsulation and study by NMR. Reverse micelles might succeed where bicelles failed, capturing high resolution structural information of a myristoylated pH switch protein bound to a membrane and pushing the envelope of understanding hisactophilin's mechanism of switching one step farther.

7.2.1 Reverse micelles

Standard micelles arise when the concentration of surfactant is greater than its critical micelle concentration and the lipids coalesce in an aqueous solvent. The hydrophobic effect with the entropic penalty of solvating the long hydrophobic tails of surfactant molecules drive micelle formation [428]. A reverse micelle inverts this principle by changing the primary solvent to an alkane. Now, the hydrophobic tails of the surfactant are favourably solvated by the bulk solvent while the hydrophilic moiety is driven to the inside of the (reverse) micelle where it would interact with water and/or other hydrophilic molecules. Figure 7.1 depicts a schematic overview of a reverse micelle with a guest protein (left panel), and the structure of a reverse micelle based on molecular dynamics parameterized by small angle X-ray and neutron scattering (right panel). Overall, a reverse micelle typically contains 100-400 primary surfactant molecules, 1000-5000 water molecules that form a native-like hydration shell [429], and a single protein molecule.

The options for surfactants and solvents that produce reverse micelles are numerous, without any one being the panacea for: studying proteins that are too large for traditional solution NMR [432,433], high throughput screening of fragment libraries for binding in the nanoscale reverse micelle environment, or characterizing membrane anchored proteins [427]. Often, screening different mixtures and ratios of surfactants is necessary to determine which works for a target protein [424]. Earlier work that looked at an encapsulated polypeptide by NMR used the detergent AOT (bis(2-ethylhexyl)sulfosuccinate) [434,435]; however, its use with proteins is limited to ubiquitin as it was the only protein found that is hearty enough to withstand the strong denaturing effect of AOT [425,436]. This encouraged the development of gentler reverse micelle systems that are more likely to preserve a protein's native structure.

One of the more recently developed reverse micelle systems is 10MAG/LDAO (decylmonoacylglycerol/lauryldimethylamine N-oxide). It has a low charge density at the surfactant/water interface and it is

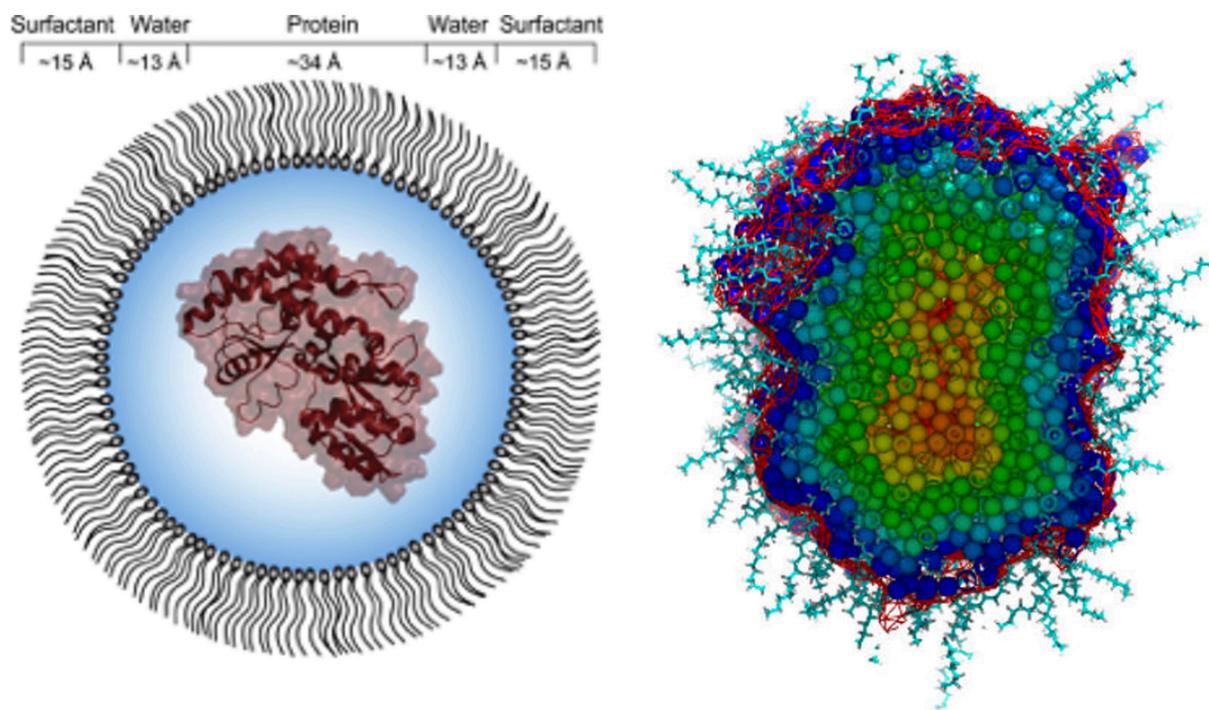


Figure 7.1: Panels shown were adapted from some of A. Joshua Wand's publications [425,430]. The left panel is an illustration of a prototypical reverse micelle with maltose binding protein (MBP) (PDB ID 2H25) as the guest, drawn to scale. The size of the protein was measured in molecular visualization software, the interior of the reverse micelle was measured from a series of experiments of MBP encapsulated in a CTAB/hexanol reverse micelle and solvents of varying viscosity [429], and the surfactant layer thickness was based on estimates of a protein-free reverse micelle [431]. The right panel shows the model of a CTAB/hexanol reverse micelle developed using molecular dynamics that were parameterized with experimental data from small angle X-ray and neutron scattering. Reverse micelles with a water loading ratio of 20 are approximately spherical and comprise 240 hexanol and ~150 CTAB molecules. The core of reverse micelle (spheres coloured according to distance from the surfactant interface) contains on the order of 3000 water molecules with a gradient of translational diffusion that reaches that of bulk water in the centre. At the interface of the reverse micelle and the polar molecules of the core is a ~5 Å region of Br⁻, hexanol hydroxyl group-, and CTAB headgroup- enriched layer of water that accounts for about a third of the total water.

amenable to encapsulating proteins up to 80 kDa with pIs that range between 4 and 11 (its applicability extends to nucleic acids as well [425]). To demonstrate that native structure is likely preserved in the reverse micelle, the aqueous and reverse-micelle-encapsulated spectra of arginine kinase (40 kDa) overlapped with high fidelity—chemical shifts correlated with an R^2 of 0.999. Additionally, the cosurfactant hexanol is required for stable encapsulation, which allows favourable curvature in the reverse micelle [431]. The titratable headgroup of LDAO is a convenient mechanism for controlling the pH inside the nanoscale water environment of the reverse micelle. The pH of the solubilized reverse micelle surfactant solution prior to drying and resuspension determines the pH of the reverse micelle interior [425,437]. It is as important to control pH in reverse micelles as it is in bulk solution, so pH-dependent conformation or activity is conserved.

Another commonly used formulation is a combination of CTAB (Cetyltrimethylammonium bromide) and hexanol (as depicted in the right panel of Figure 7.1). A CTAB reverse micelle is approximately spherical, in which the polar molecules closest to the hydrophobic boundary are enriched in bromide ions and the most constrained in their motion while the gradient of translational diffusion reaches that of bulk water in the centre. With a water loading ratio of 20 (a measure of the water to lipid ratio), the reverse micelle comprises 240 hexanol, 150 CTAB, and ~ 3000 water molecules. This formulation encapsulated membrane-associated myristoylated HIV matrix protein with phosphatidylinositol-4,5-bisphosphate, and separately, recoverin for structural characterization [427]. Given that the CTAB/hexanol surfactant cocktail does not have titratable headgroups, the pH of the interior of the reverse micelle is determined largely by the pH of the injected protein solution, even though a reverse micelle may contain only a few buffer molecules.

Beyond providing an environment for studying membrane interactions by NMR, reverse micelles offer an avenue to raise the ceiling on the size limit of solution NMR beyond 25 - 30 kDa [438,439]. Larger proteins tend to have broader peaks and inefficient magnetization transfer, both of which yield poor spectra and limit the usefulness of multidimensional experiments [433]. Broad peaks relate to the rate of transverse relaxation which itself is a function of molecular tumbling in solution. Tumbling can be well-described by the Stokes-Einstein equation, which relates tumbling (correlation time, τ_m) to protein hydrodynamic radius (r), solution viscosity (η), and Temperature (T) (along with the Boltzmann constant, K_B , and π), as shown in equation 7.1,

$$\tau_m = \frac{4\pi\eta r^3}{3K_B T}$$

While increasing temperature may increase τ_m , the relatively limited range over which proteins are stable and/or functional reduces its effectiveness as a solution to the relaxation problem. Instead, suspending reverse micelles in low viscosity solvents like ethane, propane, and pentane forms part of the strategy to extend solution-state NMR to proteins up to 100 kDa [432,433]. In liquid ethane, a 100 kDa protein may behave as though it's as small as 10 kDa and the aforementioned surfactant combinations of CTAB/hexanol and LDAO/10MAG are as amenable to solvation by ethane as they are by n-pentane [440,441]. The increased

upper size limit for NMR may enable characterizing binding and bundling F-actin dimers by hisactophilin [442,443]. Here, we show the proof-of-concept that the cutting-edge technology of reverse micelles are amenable to elucidating unstudied membrane- and actin-binding events in hisactophilin with atomic detail.

7.3 Materials and Methods

7.3.1 Materials

Proteins used for reverse micelle experiments were ^{15}N -labelled myristoylated and non-myristoylated hisactophilin. The protein was purified and lyophilized as described in Section 4.3.3.1 Sample Preparation.

All solutions and reagents used for the reverse micelle experiments were from the Wand lab at University of Pennsylvania. Their source was Sigma Aldrich (St. Louis, Missouri), Coretecnet (Brooklyn, New York), or Avanti Polar Lipids (Alabaster, Alabama) unless otherwise specified.

7.3.2 Reverse micelle encapsulation

All the experimental work with reverse micelles was completed with extensive help from Dr. Brian Fuglestad, a postdoctoral fellow in Dr. Josh Wand's lab, and Nikki Kerstetter, a graduate student.

7.3.2.1 Encapsulation in 10MAG/LDAO

The first attempts at encapsulation were made in the Wand lab's newer reverse micelle cocktails; a three- or four-component system comprising 10MAG (decylmonoacylglycerol), LDAO (lauryldimethylamine N-oxide), sometimes DTAB (dodecyltrimethylammonium bromide), hexanol, and 90% n-pentane: 10% d_{12} -pentane as the bulk solvent. The detergents of this reverse micelle formulation have titratable groups that buffer the aqueous phase at a target pH, which requires that the lipids are first dissolved in a buffer at the target pH and lyophilized overnight before use in experiments. 5 mg of lyophilized protein was dissolved in 185 μL of 20 mM TRIS, 50 mM NaCl, and 2 mM DTT. Protein was mixed with different preparations/ratios of surfactant and titrated with hexanol starting from 20 mM at 2.5 mM increments to optimize the proportions of surfactants and cosurfactants for hisactophilin. Surfactant ratios tested included 60:40 10MAG:LDAO, 60:20:20 10MAG:LDAO:DTAB, 60:40 10MAG:DTAB, and a target water loading ratio of 15 - 30 [427]. Final protein concentrations in the initial tests were 40 and 80 μM , however the goal in a final formulation is close to 200 μM .

1D ^1H NMR spectra (using bruker pulse sequence *zgpr* with solvent presaturation) were measured to check for successful encapsulation of folded hisactophilin on their Bruker 500 MHz Avance III spectrometer with 5mM TXI cryoprobe and TopSpin version 3.0. The 10MAG/LDAO system was generally unsuccessful

with varying amounts of DTAB, protein, hexanol, and lipids that had been adjusted to pH 7 and 8. The 1D spectra of these attempts had minimal or no signal in the amide region which indicated ineffective encapsulation (*vide infra*).

7.3.2.2 Encapsulation in CTAB/Hexanol

Encapsulation in the two-component (CTAB:Hexanol) reverse micelle cocktail was a success. Proof of concept experiments involved 50 μ M myristoylated hisactophilin at pH 8.1, 75 mM CTAB (Cetyltrimethylammonium bromide) (Avanti Polar Lipids, Alabaster, Alabama), 425 mM hexanol, and a final water loading ratio of 15. For improved signal-to-noise, the CTAB was scaled to 300 mM, 725 mM hexanol, and 200 μ M myristoylated hisactophilin at pH 8.1. Encapsulation at \sim pH 6.1 was very fortuitous in that the same amounts of protein and CTAB were effective while the hexanol was scaled to 750 mM. In a bout of continuing luck, non-myristoylated hisactophilin was encapsulated at pH 8.1 under conditions that matched the myristoylated form except they required 775 mM hexanol, and 800 mM hexanol at pH 6.1 before they formed a stable reverse micelle.

Diagnostic 1D spectra were carried out on the 500 MHz spectrometer in an identical fashion to the 10MAG/LDAO samples. 2D ^1H , ^{15}N HSQC (bruker pulse program: `hsqcetpf3gpsi2`) spectra were measured on the same spectrometer in 5 mm screw-cap NMR tubes with an approximate volume of 500 μ L. The spectrometer was locked to the methyls of d_{12} -pentane and used a flip-back pulse for solvent suppression. The number of data points were 2048 in the direct dimension, 128 or 256 in the indirect dimension, and 32 scans. Spectra were processed in TopSpin 3.5 and/or NMRpipe [393], and then visualized in SPARKY [444] or CARRA [225].

3D ^{15}N -edited NOESY - HSQC experiments for assignment were obtained by a pulse sequence based on Bruker `noesyhsqcetpf3gpsi3d`, but modified for Poisson-gap non-uniform sampling schemes to greatly reduce acquisition times [388] and subsequently reconstructed using iterative soft thresholding methods as implemented in NMRpipe [388,389,393,445,446]. The sample comprised the aforementioned optimized CTAB cocktail for myristoylated hisactophilin at \sim pH 8, except that the surfactant, pentane, and hexanol were fully deuterated to maximize the signal-to-noise ratio. The mixing time was 100 ms and, with the non-uniform sampling, the acquisition time was 17 hours. CARRA NMR was used for backbone resonance assignment with the assistance of undergraduate Christopher Leo.

7.3.2.3 Titration with cardiolipin

To explore the effects of an additional charged component to the reverse micelle membrane, cardiolipin was titrated into a stable reverse micelle encapsulated sample. The encapsulated guest was 200 μ M myristoylated hisactophilin at pH 6.2, 300 mM CTAB, and 750 mM hexanol. A cardiolipin stock (150 mM) was titrated into the sample for a 4:1 ratio of cardiolipin to hisactophilin, a 16:1 ratio, and then finally a 64:1 ratio.

2D ^1H , ^{15}N HSQC spectra were acquired identically as in Section 7.3.2.2 Encapsulation in CTAB/Hexanol, processed in TopSpin 3, and visualized in CARRA NMR.

7.4 Results and Discussion

It has long been a goal to obtain structural information of hisactophilin bound to a membrane, and while others have had success characterizing binding events [11,447], high resolution details and molecular mechanism of membrane binding remained elusive [312]. Collaborating with Dr. Joshua Wand from University of Pennsylvania to encapsulate and anchor myristoylated hisactophilin in reverse micelles offered a unique opportunity to obtain the first tantalizing glimpse into membrane binding.

Reverse micelles, a different approach to biomolecular solution NMR, can take considerable optimization to successfully encapsulate a protein, but in exchange, reverse micelles offer some conveniences. First, the dielectric constant of the typical alkane solvents is much lower than water, so, gains in non-lossiness result in a 2 to 3 fold signal-to-noise enhancement with use of a cryogenically cooled NMR probe [425,448]. This, in part, compensates for the small protein concentrations that are encapsulated (hundreds of micromolar compared to millimolar in aqueous experiments) [425]. Furthermore, the small number of water molecules encapsulated (<2.5 M) with the guest protein do not require special efforts to suppress in simple experiments. In multi-dimensional homonuclear experiments, suppression is effected by a flipback pulse [449], and in heteronuclear experiments, water is inherently handled by the X-nucleus editing.

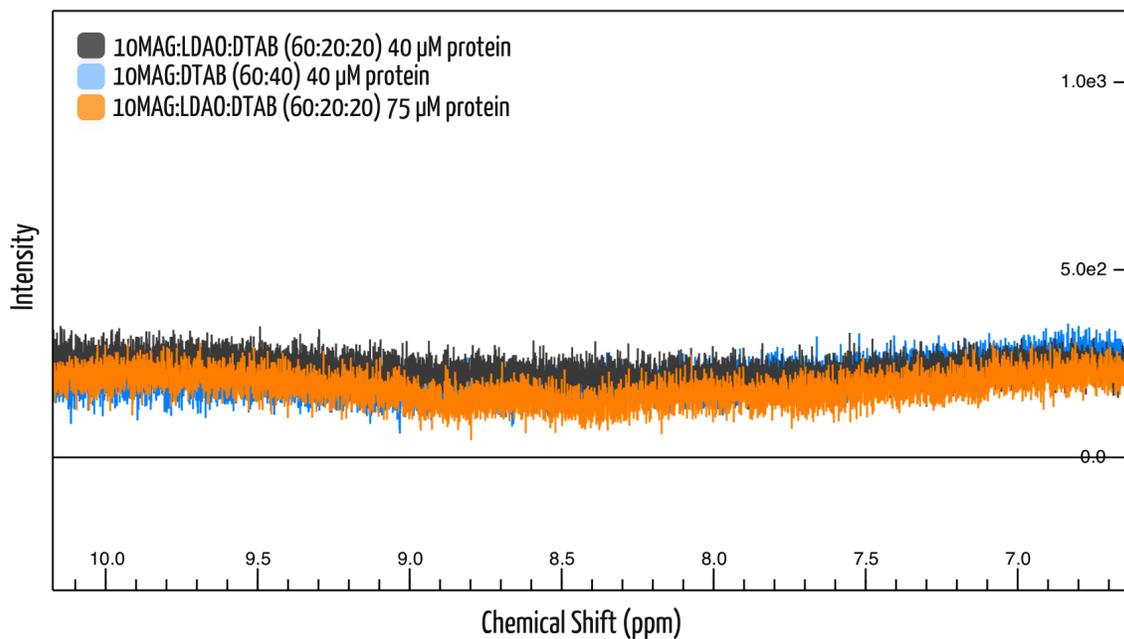
The stability of reverse micelles is dependent on the amount of water encapsulated (water loading), temperature, pressure, concentration, and the ratio of surfactants. The optimal formulations may vary from protein to protein [431,450]. Water loading [424], surfactant identity and ratios depend on the encapsulated protein [451], which illustrates that there is no omnipotent reverse micelle formulation that will encapsulate every protein. Thus, it is an optimization problem, one we successfully navigated for myristoylated and non-myristoylated hisactophilin at high and low pH on a narrow timeline while visiting Dr. Joshua Wand's lab. A post-doctoral fellow, Dr. Brian Fuglestad, and his significant expertise was indispensable in encapsulating hisactophilin where each of the four forms could, in principle, have had (very) different permissive conditions.

Encapsulation was first attempted in the Wand group's latest three- and four-component reverse micelle systems composed of 10MAG (decylmonoacylglycerol), the cosurfactant hexanol, and combinations of LDAO (lauryldimethylamine N-oxide) and DTAB (dodecyltrimethylammonium bromide). Because hisactophilin is so pH sensitive, the titratable headgroup of LDAO provides means for robust control of the pH of the nanoscale water inside the reverse micelle [424,425,437]. Overlaid in Figure 7.2 (upper panel) are 1D ^1H spectra of some attempted encapsulations of myristoylated hisactophilin in cocktails with 1:1 LDAO and

DTAB, or DTAB with 10MAG and hexanol. The lack of peaks in the amide region, even when the protein concentration was increased from 40 to 75 μM indicates encapsulation failed. Note: upfield of the water resonance is not shown because it is dominated by the surfactants and solvents.

As the purpose of the experiment was to determine feasibility rather than optimize the 10MAG system (which might eventually have worked), we switched to the two-component reverse micelle system—CTAB and hexanol [440]. The change in tack was likely to be fruitful because HIV-1 matrix protein and recoverin, two myristoylated proteins, were cheerfully encapsulated in the same surfactant [427]. The lower panel of Figure 7.2 shows the 1D plane of a ^1H , ^{15}N HSQC to compare to the upper panel. Peaks in the amide region are more evident in the CTAB system, while the grey trace is from an attempt in 10MAG and there is much less protein signal. Shown in purple is an early attempt with CTAB - the signal appears weak, but it was acquired with only 16 scans compared to the 64 scans for the green and red spectra, so, observing peaks with relatively few scans was promising and prompted attempts with increased protein and surfactant concentrations (green and red were the spectra with the greatest signal).

Overlay of hisactophilin 1D spectra in 3 component reverse micelle formulations



Overlay of hisactophilin 1D planes of HSQC in 2 and 3 component formulations

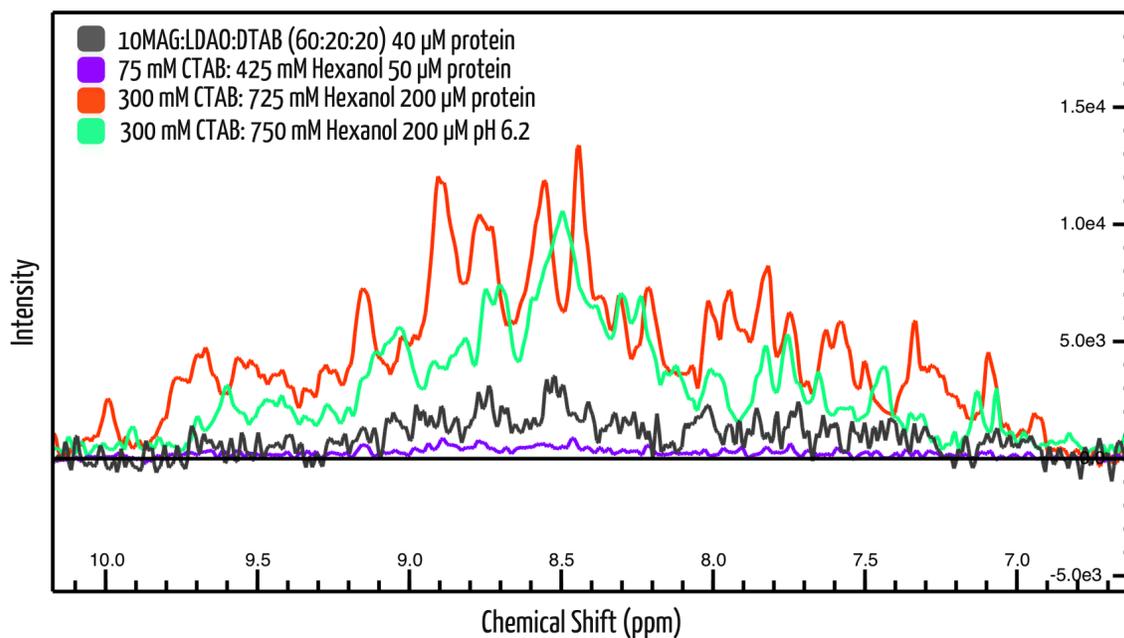


Figure 7.2: Representative NMR spectra of reverse micelles. (upper panel) 1D spectrum or 1D spectrum corresponding to the first increment of the ^1H , ^{15}N HSQC (lower panel) reverse micelle formulations while attempting to encapsulate myristoylated hisactophilin at $\sim\text{pH}$ 8 unless specified otherwise. (upper panel) 1D ^1H spectrum of myristoylated wild-type at $\text{pH} \sim 8$ in different reverse micelle systems for the amide region of the spectrum. An encapsulated, well-folded and tumbling protein in a reverse micelle presents an envelope of overlapping peaks in the amide region. A lack of measurable signal or broad peaks suggest that further optimization is necessary. The protein from a failed encapsulation is generally irrecoverable. Spectra were acquired on a 500 MHz Bruker spectrometer with a 5 mm cryoprobe, processed using TopSpin 3.0, visualized in NMRFAM and CCPN. The pulse sequence used was zgpr with 4-16 scans while locked to the methyl groups of d_{12} -pentane. Different encapsulation conditions are shown in grey, blue, and orange as indicated on the legend. Mixtures with 10MAG were not tractable for hisactophilin given the lack of amide peaks. (lower panel) overlays of the 1D spectrum corresponding to the first increment of the ^1H , ^{15}N HSQC. Pulse sequence was hsqcetfpf3gpsi2 with 16 or 64 scans, locked to deuterated pentane. When hisactophilin is encapsulated, standard 1D spectra are difficult to phase and observe peaks, so for illustrative purposes 1D spectrum of the first increment of the 2D experiments are shown. In grey, a 10MAG sample with little distinguishable peaks relative to the noise. The remaining spectra (as indicated on the legend) are the result of encapsulation in CTAB and hexanol. The red and green traces were the best cases that produced excellent 2D spectra and were the basis for preparing a deuterated sample for a 3D experiment.

With indications in the 1D spectra that encapsulation was successful, a 2D ^1H , ^{15}N HSQC was acquired to compare with aqueous spectra to ensure that the encapsulated protein is correctly folded. The top left panel of Figure 7.3 shows the spectrum for encapsulated protein. Based on the success with myristoylated hisactophilin at $\sim\text{pH}$ 8, the same formulation but with an additional 25 mM hexanol (total concentration was 750 mM) at $\sim\text{pH}$ 6.1 was also tested, as illustrated in Figure 7.3. Remarkably, the spectrum for encapsulated myristoylated hisactophilin is extremely similar to that of the non-myristoylated form in aqueous solution at the same pH. Non-myristoylated hisactophilin was successfully encapsulated in CTAB with 775 mM and 800 mM hexanol for pH 8 and pH 6.1, respectively (Figure 7.3).

Reverse Micelle Encapsulated Hisactophilin in CTAB/Hexanol

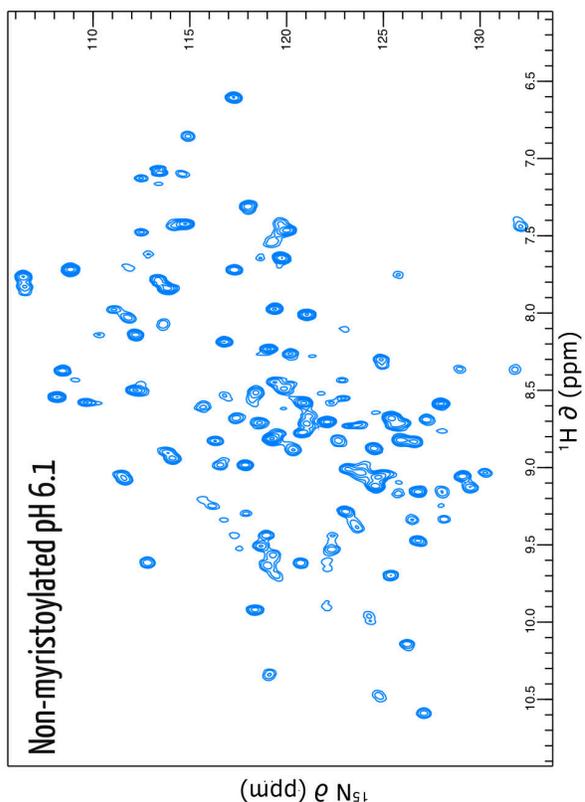
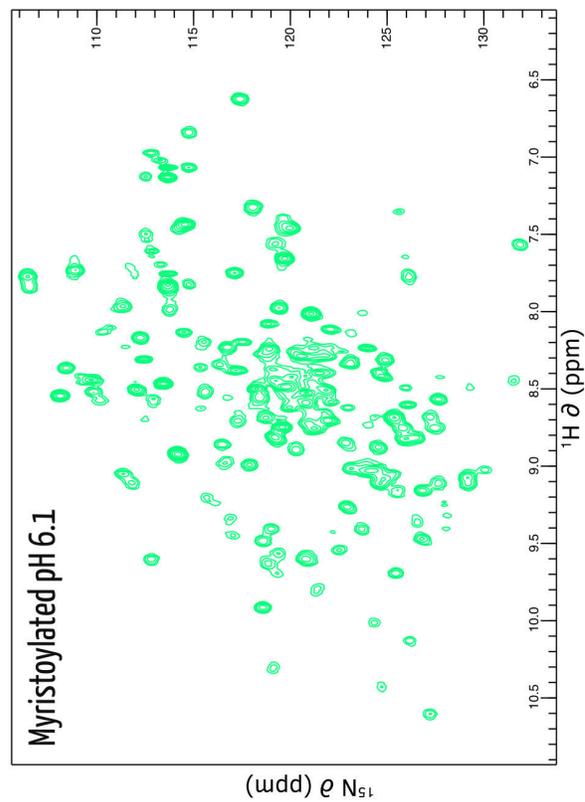
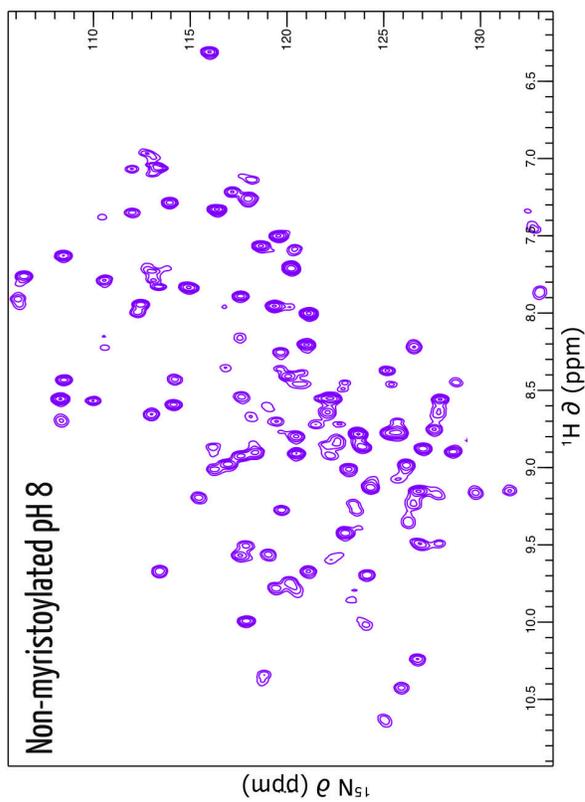
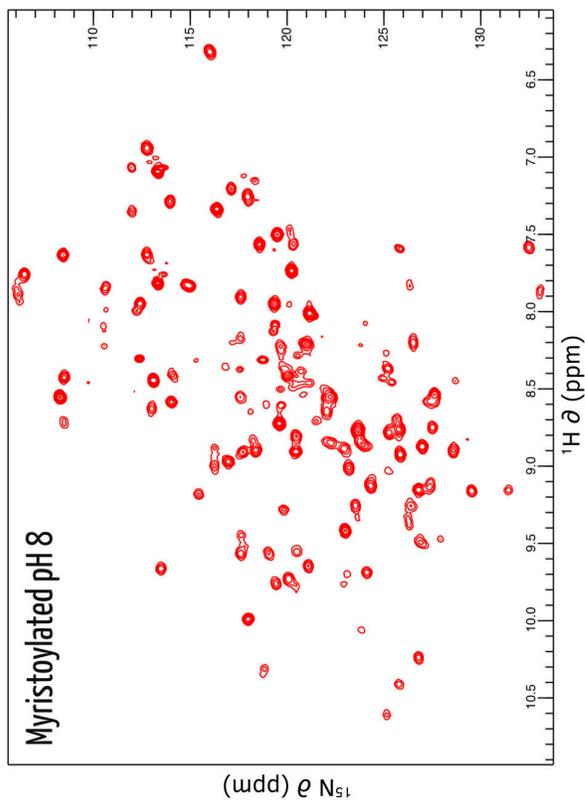


Figure 7.3: ^1H , ^{15}N HSQC of encapsulated myristoylated hisactophilin at $\sim\text{pH}$ 8.1 (upper left), $\sim\text{pH}$ 6.1 (lower left), and non-myristoylated hisactophilin at $\sim\text{pH}$ 8.1 (upper right), and $\sim\text{pH}$ 6.1 (lower right). The samples contained 200 μM of each respective protein in 20 mM TRIS, 50 mM NaCl, 2 mM DTT, 300 mM CTAB, and 725, 750, 775, or 800 mM hexanol, and n-pentane (10% d_{12} -pentane for lock) with a final water loading ranging from 15-22. The pH values are expected to match those of the aqueous protein solution preceding encapsulation. Spectra were acquired on a Bruker 500 MHz spectrometer with a cryoprobe.

The structural information from the 2D spectra gave the first details of the disposition of encapsulated hisactophilin. Figure 7.4 compares encapsulated myristoylated hisactophilin with aqueous myristoylated and non-myristoylated forms, as well as encapsulated non-myristoylated protein. The overlay in the top panel of encapsulated (red) and aqueous (black) spectra of myristoylated hisactophilin shows significant and widespread chemical shift differences. Interestingly, there was extensive agreement between chemical shifts of encapsulated myristoylated hisactophilin with the aqueous non-myristoylated form (middle panel). Chemical shifts deviate most significantly around the terminal residues G2 and I118, but the resonances of the two spectra are very similar. Lastly, the lower panel compares the encapsulated myristoylated form (red) with the encapsulated non-myristoylated form (purple). Like with the aqueous non-myristoylated spectrum, the agreement with the myristoylated form is very extensive and the difference at the C-terminal I118 is no longer observed. This may suggest an interaction between the C-terminus and the reverse micelle, but that other residues around the opening of the barrel are mostly unaffected suggests that the interactions are not extensive and might therefore arise from altered structure or dynamics in the terminus.

Together, these results give structural insight into what is occurring inside the reverse micelle. First, encapsulation in a CTAB reverse micelle has very little effect on the chemical shifts in the ^1H , ^{15}N HSQC spectra and, given the sensitivity of amide proton chemical shifts to minor changes in structure (be it bond angles, hydrogen bond donor/acceptor angles, hydrogen bond length, etc.), hisactophilin appears to maintain its native fold. Secondly, widespread differences between spectra for myristoylated aqueous and encapsulated hisactophilin (upper panel) show changes in key structures/interactions in hisactophilin, however, overlap between myristoylated encapsulated protein and non-myristoylated (aqueous or encapsulated) clarifies the source of the change. Inside the reverse micelle, the myristoyl group appears to no longer interact with the binding pocket, so, the myristoyl group is likely anchoring the protein to the reverse micelle lipid phase. Strikingly, very similar results are obtained at pH 6.1 as at pH 8.1, i.e. the myristoylated hisactophilin spectra in reverse micelles look very similar to those for non-myristoylated protein in aqueous buffer at the same pH. This observation may be unexpected in the framework of our understanding of hisactophilin's switch because association with a membrane should occur only at $\text{pH} < 6.9$ (and not at pH 8.1, as shown in Figure 7.4) [8]. When myristoylated recoverin—a Ca^{2+} -dependent switch—was encapsulated in a CTAB/hexanol reverse micelle while in the sequestered state (no Ca^{2+} present), its NMR spectrum showed evidence of mixed conformations, including those with the myristoyl group extruded. Similarly, the equilibrium of HIV

matrix protein was pushed toward the solvent-accessible state by the reverse micelle in the absence of its ligand [427]. The makeup of the lipid phase and the interior environment of the reverse micelle is a very different environment than the cell—the nanoscale water environment, lack of crowding, and reported extreme hydrophobicity of the lipid-phase [427,430] are probably sufficient to overcome the thermodynamics of hisactophilin and other proteins' switch and induce the accessible/membrane-bound state.

2D HSQC of Aqueous Hisactophilin at pH 7.7

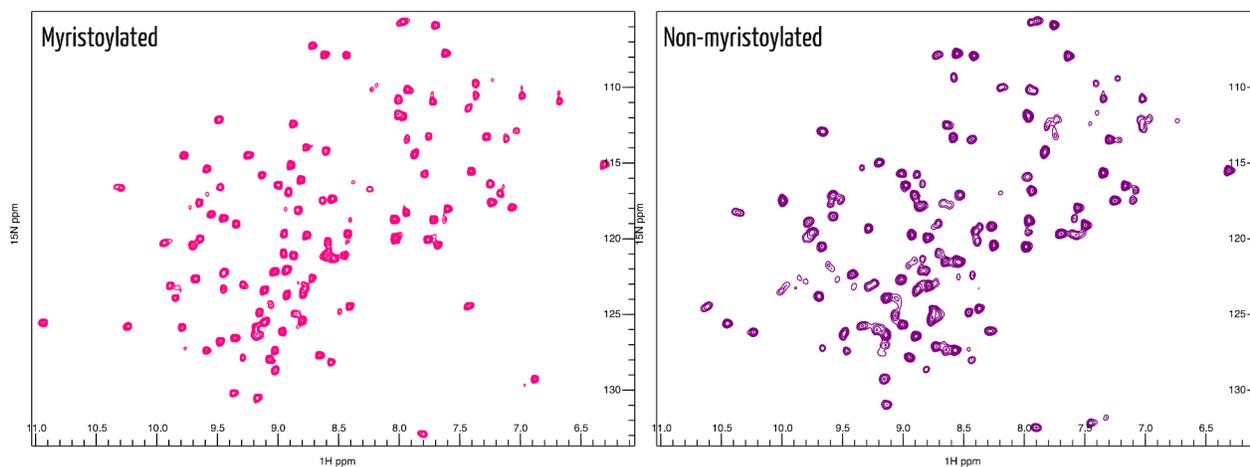


Figure 7.5: ^1H , ^{15}N HSQC spectra of myristoylated (left) and non-myristoylated (right) hisactophilin at pH 7.7, acquired on a Bruker 600 MHz spectrometer with a 5 mm TXI probe at 298 K (25 °C). These spectra were acquired as a part of the wild-type temperature coefficient data in Chapter 4. They are illustrative of the spectra of aqueous, folded hisactophilin.

That the spectrum of encapsulated myristoylated protein matches that of non-myristoylated also validates our reported model of calculating switch energies using a double mutant cycles [5,8]. Because at high pH it is not possible to gather stability data exclusively of the myristoyl accessible conformation when hisactophilin is predominantly in the sequestered state, the non-myristoylated variant is a facsimile for the protein that lacks interactions with the myristoyl group. The reverse micelle confirms it is a well-founded assumption because, when anchored to a membrane by the myristoyl group, the rest of the protein appears as essentially structurally identical to the non-myristoylated form.

Comparisons with Encapsulated Myristoylated Hisactophilin

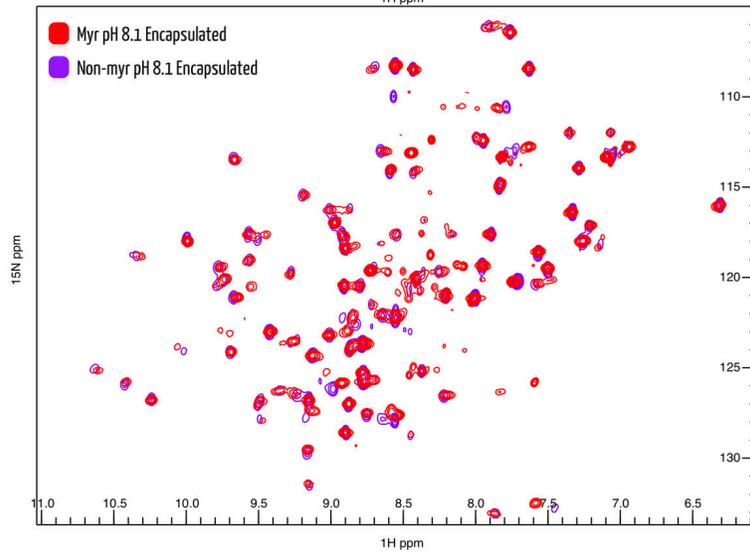
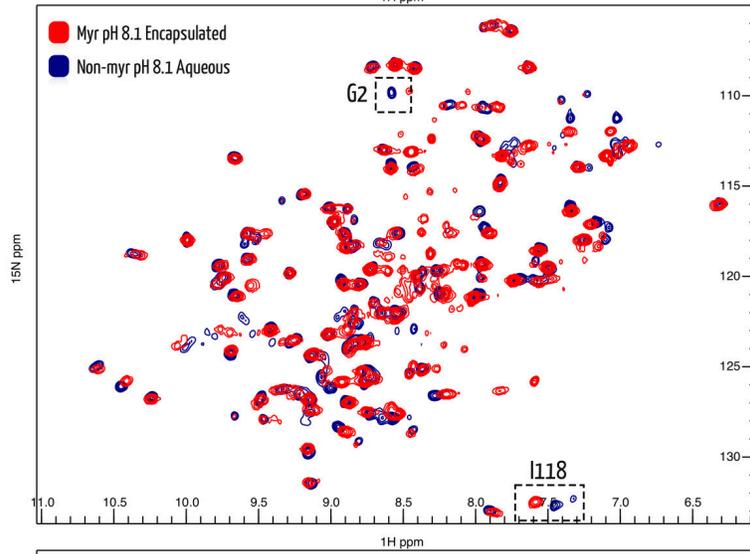
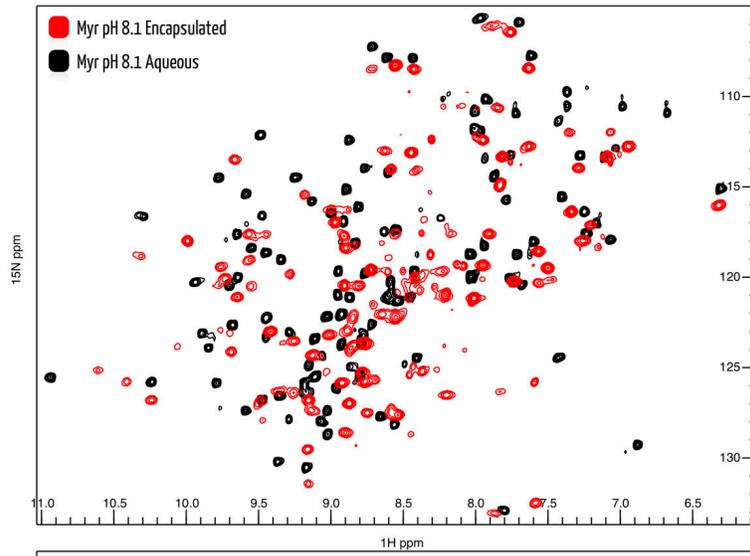


Figure 7.4: Overlay of ^1H , ^{15}N HSQC spectra of encapsulated myristoylated hisactophilin with aqueous or non-myristoylated forms of the protein. Encapsulated myristoylated hisactophilin at $\sim\text{pH}$ 8 appears in each panel. Spectra were acquired on a 500 MHz Bruker spectrometer with a cryoprobe. Encapsulated hisactophilin was in CTAB, 725 mM hexanol (myristoylated) or 775 mM (non-myristoylated), n-pentane (10% d_{12} -pentane), and 200 μM protein. Samples of aqueous hisactophilin were prepared with 2 mM protein, 50 mM potassium phosphate, and 10% D_2O at 298 K (25 $^\circ\text{C}$), and spectra were acquired on a Bruker 600 MHz spectrometer. Aqueous myristoylated hisactophilin at pH 7.7 is in black in the upper panel, aqueous non-myristoylated hisactophilin at pH 7.7 is in blue in the middle panel, and encapsulated non-myristoylated hisactophilin at $\sim\text{pH}$ 8 is in purple in the lower panel.

As for assigning the resonances in the reverse micelle, it was assumed that the assignments should transfer directly from those previously measured for non-myristoylated hisactophilin. This assumption was verified by measuring a 3D ^{15}N -edited NOESY-HSQC of reverse-micelle encapsulated sample. Data were acquired for seventeen hours with Poisson-gap non-uniform sampling implemented to maximize data acquisition within the time constraints. While the spectra were not as well-resolved or with as good a signal-to-noise ratio as the traditionally sampled 3D experiments in Chapter 4, they were sufficient to assign ~ 70 residues which all matched the corresponding residues in the non-myristoylated protein. Additional spectrometer time would, I expect, produce spectra at least comparable to a traditionally sampled spectrum of an aqueous sample of hisactophilin.

The final exploratory experiment in reverse-micelles was focused on replicating the polybasic interactions that, in addition to those of the myristoyl group, secure hisactophilin to the membrane [11], the so-called myristoyl-electrostatic switch [37,39,40]. Interactions between basic residues around the opening of the hisactophilin β -barrel and the negatively charged inner leaflet components of the membrane contribute to anchoring to the membrane [11]. In *Dictyostelium discoideum*, that component is likely to be phosphatidylserine, however, phosphatidylglycerol is a commonly used alternative in a laboratory setting. Neither of those surfactants, however, were part of the reverse micelle formulation. Previously, cardiolipin, a negatively charged lipid, was successfully incorporated into reverse micelles to measure interactions with cytochrome c during the apoptotic trigger [452], and phosphatidylinositol-4,5-bisphosphate was included for binding HIV-1 matrix protein [427]. Accordingly, we added cardiolipin to hisactophilin-containing reverse micelles in an attempt to replicate some of the electrostatic interactions that might occur during membrane-binding.

Figure 7.6 shows encapsulated myristoylated hisactophilin with (pink) and without (blue) cardiolipin (64:1 cardiolipin to hisactophilin ratio). The general similarity of chemical shifts and relative peak intensities of the two spectra supports that incorporating cardiolipin does not compromise the integrity of the reverse micelle or disrupt hisactophilin's fold. That some of the protein peaks shift suggests it is likely the cardiolipin incorporated into reverse micelles that host hisactophilin. The chemical shifts of residues A44, K59, S64, and H68 were some of those perturbed by cardiolipin. They are spatially clustered around the $\beta 4$ - $\beta 5$ turn at the top of the barrel where interactions with the membrane are expected to occur. There are also chemical shift perturbations in the hairpins at the bottom of the protein and, although purely speculative, these may suggest

some allosteric messaging from the membrane binding surface to the supposed actin-binding surface. Given that membrane- and actin-binding events occur at approximately the same pH, coupling may function to strengthen the interactions between actin and hisactophilin, and so promote locomotion of *Dictyostelium discoideum*. Or, perhaps negative coupling so that the energetics of forming lamellipodia are easily reversed.

Titrating Cardiolipin into Reverse Micelles

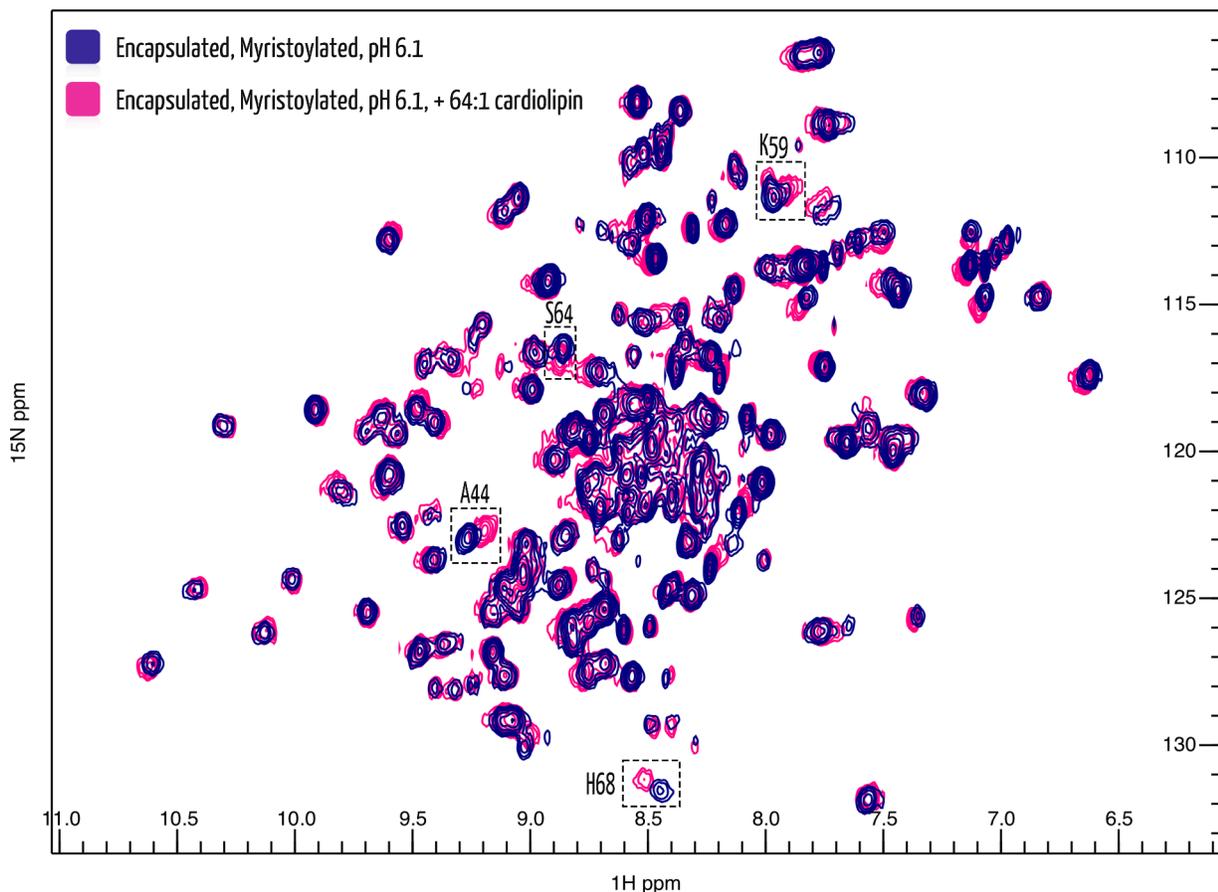


Figure 7.6: Spectra of encapsulated myristoylated hisactophilin at pH 6.1 (dark blue) overlaid by encapsulated hisactophilin at pH 6.1 with 64:1 cardiolipin to hisactophilin (pink). The overlap between the two spectra suggests that the cardiolipin has not noticeably or measurably had a negative impact on encapsulation, nor has the folded spectrum been significantly perturbed. That said, there are differences between the two spectra that suggest cardiolipin has some effect, ranging from small changes in chemical shift or changes in signal intensity/peak width. Examples of residues that experienced chemical shift changes were A44, K59, S64, and H68, residues clustered near the β 4- β 5 turn which is at the opening of the β -barrel, so the chemical shift changes may respond to interactions with nearby charged residues and lipids/cardiophilin. Some residues experienced a change in peak intensity in the presence of cardiolipin, for example, H68 increased in intensity, which could indicate changes in dynamics.

While reverse micelles most likely are not perfect facsimile of biological membranes, they offer a path to never-before-seen high-resolution structural data of membrane-bound hisactophilin with an extruded myristoyl moiety. Agreement between existing structural data and reverse micelle-encapsulated hisactophilin is

conducive to further experiments with reverse micelles. Surfactant concentrations at or beyond ~300 mM may partially align reverse micelles [425,453], therefore it is possible to measure structural and motional information with residual dipolar couplings (RDCs) [454]. For example, RDCs might aid to solve a high resolution structure of encapsulated hisactophilin (in conjunction with other standard NMR experiments), as well as measure the relative angle of the myristoyl insertion into the membrane. Furthermore, the size of actin has limited its study by NMR, however, examining the mechanism of binding to hisactophilin may be possible with reverse micelles. Hisactophilin bound to actin presents layers of difficulty: first, hisactophilin bundles actin so the complex would likely grow too large for traditional NMR. This leads to the second difficulty, rapid relaxation of large molecules in the NMR. Reverse micelles can be made in low-viscosity solvents (e.g. ethane) that increase the rate of tumbling, in which spectra of complexes up to 100 kDa are measurable [425,429,433,441]. Combining the reverse micelle technology with polymerization-deficient F-actin dimers [442,443] could enable structural characterization of binding, starting with chemical shift perturbations, followed by measuring how actin- and/or membrane-binding modulate the structure, dynamics, and entropy [455] of hisactophilin.

7.5 Additional future studies on hisactophilin

Other work presented in this thesis characterized the effects of proton binding and/or myristoylation on the local conformational stability of hisactophilin, the functional energy landscape, the global energetics, the folding landscape, and in some cases repurposed (new) tools to do so. In the complex behavioural framework of hisactophilin's pH-dependent myristoyl switch, the protein reversibly binds membranes, is modulated by phosphorylation, reversibly binds and bundles actin, and may function as an osmoprotectant. Phosphorylation, osmoprotection, and membrane- and actin-binding remain relatively uncharacterized. However, measuring the temperature dependence of chemical shifts as a function of salt is a straightforward way to broach osmoprotection and electrostatics. Membrane- and actin-binding are amenable to further investigation further using reverse micelles in the NMR and calorimetry (e.g. isothermal titration calorimetry) using other membrane mimetics. Similar experiments as those presented in this thesis could be applied to the phosphorylated protein, the study of which may be simplified by a phospho-mimetic mutation to avoid having to identify and/or isolate the kinase. The additional post-translational modification could allow for us to ascertain the effects of adjusting surface charge on hisactophilin's numerous binding and switching events.

The NMR-monitored pH titrations provided a wealth of apparent pK_a s that may be of interest to the burgeoning field of constant pH molecular dynamics simulations [379]. The data in Chapter 5 may prove useful as benchmarks for models of dynamic systems while offering, in return, structural bases for the observed switch-induced pK_a shifts (e.g. by continued collaboration with Dr. Charles L. Brooks III). Continuing with

dynamics, nonlinear temperature dependences of amide proton chemical shifts are under-utilized and not fully understood. In particular, expanding our understanding of the extent and sign of curvature (the α -value) and how it reports on the structure of the excited state and its relative energy level would further empower nonlinear temperature dependences as a tool for characterizing excited states. Work presented in this thesis expanded our understanding of the thermodynamic and structural determinants of switching, so applying these findings by grafting the myristoyl binding pocket and NMT recognition sequence into another protein or designing a myristoyl switch de novo would present a challenge and further test the state of our understanding of switches. Hisactophilin, as a model system for studying complex behaviours and allostery in a small, relatively well-behaved system, is not exhausted yet. Many roads are left to travel.

Bibliography

1. Fersht A: *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*. 1999.
2. Resh MD: **Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins.** *Biochimica et biophysica acta* 1999, **1451**:1–16.
3. Dill KA: **Dominant forces in protein folding.** *Biochemistry* 1990, **29**:7133–7155 <https://doi.org/10.1021/bi00483a001>.
4. Nicholls A, Sharp KA, Honig B: **Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons.** *Proteins: Structure, Function, and Genetics* 1991, **11**:281–296 <https://doi.org/10.1002/prot.340110407>.
5. Smith MTJ, Mackenzie DWS, Meiering EM: **Dissecting the molecular determinants of ligand-binding-induced macromolecular switching using thermodynamic cycles.** *Protein Engineering, Design and Selection* 2011, **24**:213–217 <https://doi.org/10.1093/protein/gzq099>.
6. Horovitz A, Fersht AR: **Strategy for analysing the co-operativity of intramolecular interactions in peptides and proteins.** *Journal of Molecular Biology* 1990, **214**:613–617 [https://doi.org/10.1016/0022-2836\(90\)90275-Q](https://doi.org/10.1016/0022-2836(90)90275-Q).
7. Süel GM, Lockless SW, Wall MA, Ranganathan R: **Evolutionarily conserved networks of residues mediate allosteric communication in proteins.** *Nature Structural Biology* 2003, **10**:59–69 <https://doi.org/10.1038/nsb881>.
8. Smith MTJ, Meissner J, Esmonde S, Wong HJ, Meiering EM: **Energetics and mechanisms of folding and flipping the myristoyl switch in the β -trefoil protein, hisactophilin.** *Proceedings of the National Academy of Sciences* 2010, **107**:20952–20957 <https://doi.org/10.1073/pnas.1008026107>.
9. Nelson ML, Kang H-S, Lee GM, Blaszcak AG, Lau DKW, McIntosh LP, Graves BJ: **Ras signaling requires dynamic properties of Ets1 for phosphorylation-enhanced binding to coactivator CBP.** *Proceedings of the National Academy of Sciences* 2010, **107**:10026–10031 <https://doi.org/10.1073/pnas.0915137107>.
10. Shental-Bechor D, Levy Y: **Folding of glycoproteins: toward understanding the biophysics of the glycosylation code.** *Current Opinion in Structural Biology* 2009, **19**:524–533 <https://doi.org/10.1016/j.sbi.2009.07.002>.
11. Hanakam F, Gerisch G, Lotz S, Alt T, Seelig A: **Binding of hisactophilin I and II to lipid membranes is controlled by a pH-dependent myristoyl-histidine switch.** *Biochemistry* 1996, **35**:11036–11044 <https://doi.org/10.1021/bi960789j>.
12. Ames JB, Porumb T, Tanaka T, Ikura M, Stryer L: **Amino-terminal Myristoylation Induces Cooperative Calcium Binding to Recoverin.** *Journal of Biological Chemistry* 1995, **270**:4526–4533 <https://doi.org/10.1074/jbc.270.9.4526>.
13. Resh MD: **Trafficking and signaling by fatty-acylated and prenylated proteins.** *Nature chemical biology* 2006, **2**:584–90 <https://doi.org/10.1038/nchembio834>.
14. Farazi TA, Waksman G, Gordon JI: **The Biology and Enzymology of Protein N-Myristoylation.** *Journal of Biological Chemistry* 2001, **276**:39501–39504 <https://doi.org/10.1074/jbc.R100042200>.
15. Duronio RJ, Towler D a, Heuckeroth RO, Gordon JI: **Disruption of the yeast N-myristoyl transferase gene causes recessive lethality.** *Science* 1989, **243**:796–800 <https://doi.org/10.1126/science.2644694>.
16. Ntwasa M, Aapies S, Schiffmann D a, Gay NJ: **Drosophila embryos lacking N-myristoyltransferase have multiple developmental defects.** *Experimental cell research* 2001, **262**:134–44 <https://doi.org/10.1006/excr.2000.5086>.
17. Ducker CE: **Two N-Myristoyltransferase Isozymes Play Unique Roles in Protein Myristoylation, Proliferation, and Apoptosis.** *Molecular Cancer Research* 2005, **3**:463–476 <https://doi.org/10.1158/1541-7786.MCR-05-0037>.
18. Yang SH, Shrivastav A, Kosinski C, Sharma RK, Chen M-H, Berthiaume LG, Peters LL, Chuang P-T, Young SG, Bergo MO: **N-Myristoyltransferase 1 Is Essential in Early Mouse Development.** *Journal of Biological Chemistry* 2005, **280**:18990–18995 <https://doi.org/10.1074/jbc.M412917200>.
19. Meissner J: **Effects of Myristoylation on the Structure and Stability of Hisactophilin.** 2007, [no volume].
20. Wong HJ: **The role of temperature, osmolytes and myristoylation on the stability and folding of hisactophilin.** 2002, [no volume].
21. Færgeman NJ, Black PN, Zhao XD, Knudsen J, DiRusso CC: **The Acyl-CoA Synthetases Encoded within FAA1 and FAA4 in Saccharomyces cerevisiae Function as Components of the Fatty Acid Transport System Linking Import, Activation, and Intracellular Utilization.** *Journal of Biological Chemistry* 2001, **276**:37051–37059 <https://doi.org/10.1074/jbc.M100884200>.
22. Potvin-Fournier K, Lefèvre T, Picard-Lafond A, Valois-Paillard G, Cantin L, Salesse C, Auger M: **The Thermal Stability of Recoverin Depends on Calcium Binding and Its Myristoyl Moiety As Revealed by Infrared Spectroscopy.** *Biochemistry* 2014, **53**:48–56 <https://doi.org/10.1021/bi401336g>.
23. Schröder T, Lilie H, Lange C: **The myristoylation of guanylate cyclase-activating protein-2 causes an increase in thermodynamic stability in the presence but not in the absence of Ca²⁺.** *Protein Science* 2011, **20**:1155–1165 <https://doi.org/10.1002/pro.643>.
24. Dell'Orco D, Behnen P, Linse S, Koch K-W: **Calcium binding, structural stability and guanylate cyclase activation in GCAP1 variants associated with human cone dystrophy.** *Cellular and Molecular Life Sciences* 2010, **67**:973–984 <https://doi.org/10.1007/s00018-009-0243-8>.

25. Orban T, Bereta G, Miyagi M, Wang B, Chance MR, Sousa MC, Palczewski K: **Conformational Changes in Guanylate Cyclase-Activating Protein 1 Induced by Ca²⁺ and N-Terminal Fatty Acid Acylation.** *Structure (London, England : 1993)* 2010, **18**:116–126 <https://doi.org/10.1016/j.str.2009.11.008>.
26. Kennedy MT, Brockman H, Rusnak F: **Contributions of Myristoylation to Calcineurin Structure/Function.** *Journal of Biological Chemistry* 1996, **271**:26517–26521 <https://doi.org/10.1074/jbc.271.43.26517>.
27. Wu Z, Alexandratos J, Ericksen B, Lubkowski J, Gallo RC, Lu W: **Total chemical synthesis of N-myristoylated HIV-1 matrix protein p17: Structural and mechanistic implications of p17 myristoylation.** *Proceedings of the National Academy of Sciences* 2004, **101**:11587–11592 <https://doi.org/10.1073/pnas.0404649101>.
28. Kuhlman B, Baker D: **Exploring folding free energy landscapes using computational protein design.** *Current Opinion in Structural Biology* 2004, **14**:89–95 <https://doi.org/10.1016/j.sbi.2004.01.002>.
29. Strecker T, Maisa A, Daffis S, Eichler R, Lenz O, Garten W: **The role of myristoylation in the membrane association of the Lassa virus matrix protein Z.** *Virology Journal* 2006, **3**:1–10 <https://doi.org/10.1186/1743-422X-3-93>.
30. Du Y, Zuckermann FA, Yoo D: **Myristoylation of the small envelope protein of porcine reproductive and respiratory syndrome virus is non-essential for virus infectivity but promotes its growth.** *Virus Research* 2010, **147**:294–299 <https://doi.org/10.1016/j.virusres.2009.11.016>.
31. Thaa B, Kabatek A, Zevenhoven-Dobbe JC, Snijder EJ, Herrmann A, Veit M: **Myristoylation of the arterivirus E protein: the fatty acid modification is not essential for membrane association but contributes significantly to virus infectivity.** *Journal of General Virology* 2009, **90**:2704–2712 <https://doi.org/10.1099/vir.0.011957-0>.
32. Ames JB, Ishima R, Tanaka T, Gordon JI, Stryer L, Ikura M: **Molecular mechanics of calcium-myristoyl switches.** *Nature* 1997, **389**:198–202 <https://doi.org/10.1038/38310>.
33. Griffith JP, Kim JL, Kim EE, Sintchak MD, Thomson JA, Fitzgibbon MJ, Fleming MA, Caron PR, Hsiao K, Navia MA: **X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex.** *Cell* 1995, **82**:507–522 [https://doi.org/10.1016/0092-8674\(95\)90439-5](https://doi.org/10.1016/0092-8674(95)90439-5).
34. Stephen R, Bereta G, Golczak M, Palczewski K, Sousa MC, Protein GC-a, Stephen R, Bereta G, Golczak M, Palczewski K, et al.: **Stabilizing function for myristoyl group revealed by the crystal structure of a neuronal calcium sensor, guanylate cyclase-activating protein 1.** *Structure (London, England : 1993)* 2007, **15**:1392–1402 <https://doi.org/10.1016/j.str.2007.09.013>.
35. Tang C, Loeliger E, Luncsford P, Kinde I, Beckett D, Summers MF: **Entropic switch regulates myristate exposure in the HIV-1 matrix protein.** *Proceedings of the National Academy of Sciences* 2004, **101**:517–522 <https://doi.org/10.1073/pnas.0305665101>.
36. Resh MD: **Covalent lipid modifications of proteins.** *Current Biology* 2013, **23**:R431–R435 <https://doi.org/10.1016/j.cub.2013.04.024>.
37. Zhou W, Parent LJ, Wills JW, Resh MD: **Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids.** *J. Virol.* 1994, **68**:2556–2569.
38. Peitzsch RM, McLaughlin S: **Binding of acylated peptides and fatty acids to phospholipid vesicles: Pertinence to myristoylated proteins.** *Biochemistry* 1993, **32**:10436–10443 <https://doi.org/10.1021/bi00090a020>.
39. McLaughlin S, Aderem A: **The myristoyl-electrostatic switch: a modulator of reversible protein-membrane interactions.** *Trends in Biochemical Sciences* 1995, **20**:272–276 [https://doi.org/10.1016/S0968-0004\(00\)89042-8](https://doi.org/10.1016/S0968-0004(00)89042-8).
40. Sigal CT, Zhou W, Buser C a, McLaughlin S, Resh MD: **Amino-terminal basic residues of Src mediate membrane binding through electrostatic interaction with acidic phospholipids.** *Proceedings of the National Academy of Sciences of the United States of America* 1994, **91**:12253–7.
41. Goldberg J: **Structural Basis for Activation of ARF GTPase.** *Cell* 1998, **95**:237–248 [https://doi.org/10.1016/S0092-8674\(00\)81754-7](https://doi.org/10.1016/S0092-8674(00)81754-7).
42. Habazettl J, Gondol D, Wiltschek R, Otlewski J, Schleicher M, Holak T a: **Structure of hisactophilin is similar to interleukin-1 beta and fibroblast growth factor.** *Nature* 1992, **359**:855–858 <https://doi.org/10.1038/359855a0>.
43. Van Haastert PJM, De Vries MJ, Penning LC, Roovers E, Van der Kaay J, Erneux C, Van Lookeren Campagne MM: **Chemoattractant and guanosine 5'-[γ -thio]triphosphate induce the accumulation of inositol 1,4,5-trisphosphate in Dictyostelium cells that are labelled with [³H]inositol by electroporation.** *Biochemical Journal* 1989, **258**:577–586 <https://doi.org/10.1042/bj2580577>.
44. Hanakam F, Albrecht R, Eckerskorn C, Matzner M, Gerisch G: **Myristoylated and non-myristoylated forms of the pH sensor protein hisactophilin II: intracellular shuttling to plasma membrane and nucleus monitored in real time by a fusion with green fluorescent protein.** *The EMBO journal* 1996, **15**:2935–43.
45. Kumita JR, Smart OS, Woolley GA: **Photo-control of helix content in a short peptide.** *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**:3803–3808 <https://doi.org/10.1073/pnas.97.8.3803>.
46. Beharry AA, Woolley GA: **Azobenzene photoswitches for biomolecules.** *Chemical Society Reviews* 2011, **40**:4422 <https://doi.org/10.1039/c1cs15023e>.
47. Reis JM, Burns DC, Woolley GA: **Optical Control of Protein–Protein Interactions via Blue Light-Induced Domain Swapping.** *Biochemistry* 2014, **53**:5008–5016 <https://doi.org/10.1021/bi500622x>.
48. Morgan S-A, Al-Abdul-Wahid S, Woolley GA: **Structure-based design of a photocontrolled DNA binding protein.** *Journal of molecular biology* 2010, **399**:94–112 <https://doi.org/10.1016/j.jmb.2010.03.053>.
49. Fan HY, Morgan SA, Brechun KE, Chen YY, Jaikaran ASI, Woolley GA: **Improving a designed photocontrolled DNA-binding protein.** *Biochemistry* 2011, **50**:1226–1237 <https://doi.org/10.1021/bi101432p>.
50. Samanta S, Beharry AA, Sadvoski O, McCormick TM, Babalhavaeji A, Tropepe V, Woolley GA: **Photoswitching Azo Compounds in Vivo with Red Light.** *Journal of the American Chemical Society* 2013, **135**:9777–9784 <https://doi.org/10.1021/ja402220t>.
51. Tanaka T, Amest JB, Harvey TS, Stryer L, Ikura M: **Sequestration of the membrane-targeting myristoyl group of recoverin in the calcium-free state.** *Nature* 1995, **376**:444–447 <https://doi.org/10.1038/376444a0>.
52. Humphrey W, Dalke A, Schulten K: **VMD: Visual molecular dynamics.** *Journal of Molecular Graphics* 1996, **14**:33–38 [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
53. Bryngelson JD, Onuchic JN, Succi ND, Wolynes PG: **Funnels, pathways, and the energy landscape of protein folding: a synthesis.** *Proteins* 1995, **21**:167–95 <https://doi.org/10.1002/prot.340210302>.

54. Levinthal C: **Are there pathways for protein folding?** *Journal de Chimie Physique* 1968, <https://doi.org/10.1051/jcp/1968650044>.
55. Fersht AR, Matouschek A, Sancho J, Serrano L, Vuilleumier S: **Pathway of protein folding.** *Faraday Discussions* 1992, **93**:183 <https://doi.org/10.1039/fd9929300183>.
56. Bowler BE: **Residual structure in unfolded proteins.** *Current Opinion in Structural Biology* 2012, **22**:4–13 <https://doi.org/10.1016/j.sbi.2011.09.002>.
57. Daggett V, Li A, Itzhaki LS, Otzen DE, Fersht AR: **Structure of the Transition State for Folding of a Protein Derived from Experiment and Simulation.** *Journal of Molecular Biology* 1996, **257**:430–440 <https://doi.org/10.1006/jmbi.1996.0173>.
58. Yoo TY, Adhikari A, Xia Z, Huynh T, Freed KF, Zhou R, Sosnick TR: **The Folding Transition State of Protein L Is Extensive with Nonnative Interactions (and Not Small and Polarized).** *Journal of Molecular Biology* 2012, **420**:220–234 <https://doi.org/10.1016/j.jmb.2012.04.013>.
59. Capaldi AP, Kleanthous C, Radford SE: **Im7 folding mechanism: Misfolding on a path to the native state.** *Nature Structural Biology* 2002, **9**:209–216 <https://doi.org/10.1038/nsb757>.
60. Di Nardo A a, Korzhnev DM, Stogios PJ, Zarrine-Afsar A, Kay LE, Davidson AR: **Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**:7954–9 <https://doi.org/10.1073/pnas.0400550101>.
61. Zarrine-Afsar A, Wallin S, Neculai a M, Neudecker P, Howell PL, Davidson AR, Chan HS: **Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding.** *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**:9999–10004 <https://doi.org/10.1073/pnas.0801874105>.
62. Ferreira DU, Hegler JA, Komives EA, Wolynes PG: **On the role of frustration in the energy landscapes of allosteric proteins.** *Proceedings of the National Academy of Sciences* 2011, **108**:3499–3503 <https://doi.org/10.1073/pnas.1018980108>.
63. Shental-Bechor D, Levy Y: **Effect of glycosylation on protein folding: a close look at thermodynamic stabilization.** *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**:8256–61 <https://doi.org/10.1073/pnas.0801340105>.
64. Li L, Mirny LA, Shakhnovich EI: **Kinetics, thermodynamics and evolution of non-native interactions in a protein folding nucleus.** *Nature Structural Biology* 2000, **7**:336–342 https://doi.org/10.1038/nsb0400_336.
65. Zhang Z, Chan HS: **Competition between native topology and nonnative interactions in simple and complex folding kinetics of natural and designed proteins.** *Proceedings of the National Academy of Sciences* 2010, **107**:2920–2925 <https://doi.org/10.1073/pnas.0911844107>.
66. Pintsch T, Zischka H, Schuster SC: **Hisactophilin is involved in osmoprotection in Dictyostelium.** *BMC biochemistry* 2002, **3**:10 <https://doi.org/10.1186/1471-2091-3-10>.
67. Murzin AG, Lesk AM, Chothia C: **β -Trefoil fold: Patterns of Structure and Sequence in the Kunitz Inhibitors, Interleukins-1 β and 1 α , and Fibroblast Growth Factors.** *Journal of Molecular Biology* 1992, **223**:531–543 [https://doi.org/10.1016/0022-2836\(92\)90668-A](https://doi.org/10.1016/0022-2836(92)90668-A).
68. Lingel A, Weiss TM, Niebuhr M, Pan B, Appleton BA, Wiesmann C, Bazan JF, Fairbrother WJ: **Structure of IL-33 and Its Interaction with the ST2 and IL-1RACp Receptors-Insight into Heterotrimeric IL-1 Signaling Complexes.** *Structure* 2009, **17**:1398–1410 <https://doi.org/10.1016/j.str.2009.08.009>.
69. Chavez LL, Gosavi S, Jennings P a, Onuchic JN: **Multiple routes lead to the native state in the energy landscape of the beta-trefoil family.** *Proceedings of the National Academy of Sciences* 2006, **103**:10254–10258 <https://doi.org/10.1073/pnas.0510110103>.
70. Ponting CP, Russell RB: **Identification of distant homologues of fibroblast growth factors suggests a common ancestor for all beta-trefoil proteins.** *Journal of molecular biology* 2000, **302**:1041–7 <https://doi.org/10.1006/jmbi.2000.4087>.
71. Liu C, Gaspar JA, Wong HJ, Meiering EM: **Conserved and nonconserved features of the folding pathway of hisactophilin, a beta-trefoil protein.** *Protein Science* 2002, **11**:669–79 <https://doi.org/10.1110/ps.31702>.
72. Wong HJ, Stathopoulos PB, Bonner JM, Sawyer M, Meiering EM: **Non-linear effects of temperature and urea on the thermodynamics and kinetics of folding and unfolding of hisactophilin.** *Journal of molecular biology* 2004, **344**:1089–107 <https://doi.org/10.1016/j.jmb.2004.09.091>.
73. Shental-Bechor D, Smith MTJ, MacKenzie D, Broom A, Marcovitz A, Ghashut F, Go C, Bralha F, Meiering EM, Levy Y: **Nonnative interactions regulate folding and switching of myristoylated protein.** *Proceedings of the National Academy of Sciences* 2012, **109**:17839–17844 <https://doi.org/10.1073/pnas.1201803109>.
74. Hammond MS, Houlston RS, Meiering EM: **Two-dimensional 1H and 15 N NMR titration studies of hisactophilin.** *Biochemistry and Cell Biology* 1998, **76**:294–301 <https://doi.org/10.1139/o98-032>.
75. Onuchic JN, Wolynes PG: **Theory of protein folding.** *Current Opinion in Structural Biology* 2004, **14**:70–75 <https://doi.org/10.1016/j.sbi.2004.01.009>.
76. Baker D: **A surprising simplicity to protein folding.** *Nature* 2000, **405**:39–42 <https://doi.org/10.1038/35011000>.
77. Levy Y, Cho SS, Onuchic JN, Wolynes PG: **A Survey of Flexible Protein Binding Mechanisms and their Transition States Using Native Topology Based Energy Landscapes.** *Journal of Molecular Biology* 2005, **346**:1121–1145 <https://doi.org/10.1016/j.jmb.2004.12.021>.
78. Hegler J a, Ferreira DU, Wolynes PG, Sutto L, La J, Lätzer J, Hegler J a, Ferreira DU, Wolynes PG: **Consequences of localized frustration for the folding mechanism of the IM7 protein.** *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**:19825–30 <https://doi.org/10.1073/pnas.0709922104>.
79. Shea J-E, Onuchic JN, Brooks CL: **Exploring the origins of topological frustration: Design of a minimally frustrated model of fragment B of protein A.** *Proceedings of the National Academy of Sciences* 1999, **96**:12512–12517 <https://doi.org/10.1073/pnas.96.22.12512>.
80. Clementi C, Plotkin SS: **The effects of nonnative interactions on protein folding rates: Theory and simulation.** *Protein Science* 2004, **13**:1750–1766 <https://doi.org/10.1110/ps.03580104>.
81. Gosavi S, Chavez LL, Jennings PA, Onuchic JN: **Topological frustration and the folding of interleukin-1 β .** *Journal of Molecular Biology* 2006, **357**:986–996 <https://doi.org/10.1016/j.jmb.2005.11.074>.

82. Capraro DT, Roy M, Onuchic JN, Gosavi S, Jennings PA: **β -Bulge triggers route-switching on the functional landscape of interleukin-1 β .** *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**:1490–3 <https://doi.org/10.1073/pnas.1114430109>.
83. Cho J-H, Sato S, Horng J-C, Anil B, Raleigh DP: **Electrostatic interactions in the denatured state ensemble: Their effect upon protein folding and protein stability.** *Archives of Biochemistry and Biophysics* 2008, **469**:20–28 <https://doi.org/10.1016/j.abb.2007.08.004>.
84. Weinkam P, Pletneva EV, Gray HB, Winkler JR, Wolynes PG: **Electrostatic effects on funneled landscapes and structural diversity in denatured protein ensembles.** *Proceedings of the National Academy of Sciences* 2009, **106**:1796–1801 <https://doi.org/10.1073/pnas.0813120106>.
85. Azia A, Levy Y: **Nonnative electrostatic interactions can modulate protein folding: molecular dynamics with a grain of salt.** *Journal of molecular biology* 2009, **393**:527–42 <https://doi.org/10.1016/j.jmb.2009.08.010>.
86. Thirumalai D, Hyeon C: **RNA and protein folding: Common themes and variations.** *Biochemistry* 2005, **44**:4957–4970 <https://doi.org/10.1021/bi047314+>.
87. Liu C, Chu D, Wideman RD, Houliston RS, Wong HJ, Meiering EM: **Thermodynamics of denaturation of hisactophilin, a beta-trefoil protein.** *Biochemistry* 2001, **40**:3817–27.
88. Houliston RS, Liu C, Singh LMR, Meiering EM: **pH and Urea Dependence of Amide Hydrogen–Deuterium Exchange Rates in the β -Trefoil Protein Hisactophilin †.** *Biochemistry* 2002, **41**:1182–1194 <https://doi.org/10.1021/bi0115838>.
89. Clementi C, Nymeyer H, Onuchic JN: **Topological and energetic factors: what determines the structural details of the transition state ensemble and “en-route” intermediates for protein folding? an investigation for small globular proteins.** *Journal of Molecular Biology* 2000, **298**:937–953 <https://doi.org/10.1006/jmbi.2000.3693>.
90. Case DA, Pearlman DA, Caldwell JW, Cheatham III TE, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, et al.: **AMBER 6.** 1999, [no volume].
91. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, et al.: **A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations.** *Journal of computational chemistry* 2003, **24**:1999–2012 <https://doi.org/10.1002/jcc.10349>.
92. Biyun S, Cho SS, Thirumalai D: **Folding of human telomerase RNA pseudoknot using ion-jump and temperature-quench simulations.** *Journal of the American Chemical Society* 2011, **133**:20634–43 <https://doi.org/10.1021/ja2092823>.
93. Hagai T, Levy Y: **Ubiquitin not only serves as a tag but also assists degradation by inducing protein unfolding.** *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**:2001–6 <https://doi.org/10.1073/pnas.0912335107>.
94. Namanja AT, Wang XJ, Xu B, Mercedes-camacho AY, Wilson K a, Eitzkorn F a, Peng JW: **Stereospecific gating of functional motions in Pin1.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**:12289–94 <https://doi.org/10.1073/pnas.1019382108>.
95. Marlow MS, Dogan J, Frederick KK, Valentine KG, Wand AJ: **The role of conformational entropy in molecular recognition by calmodulin.** *Nature chemical biology* 2010, **6**:352–358 <https://doi.org/10.1038/nchembio.347>.
96. Tzeng S-R, Kalodimos CG: **Protein dynamics and allostery: an NMR view.** *Current opinion in structural biology* 2011, **21**:62–7 <https://doi.org/10.1016/j.sbi.2010.10.007>.
97. Wensley BG, Batey S, Bone F a C, Chan ZM, Tumelty NR, Steward A, Kwa LG, Borgia A, Clarke J: **Experimental evidence for a frustrated energy landscape in a three-helix-bundle protein family.** *Nature* 2010, **463**:685–8 <https://doi.org/10.1038/nature08743>.
98. Friel CT, Smith DA, Vendruscolo M, Gsponer J, Radford SE: **The mechanism of folding of Im7 reveals competition between functional and kinetic evolutionary constraints.** *Nature structural & molecular biology* 2009, **16**:318–24 <https://doi.org/10.1038/nsmb.1562>.
99. Broom A, Ma SM, Xia K, Rafalia H, Trainor K, Colón W, Gosavi S, Meiering EM: **Designed protein reveals structural determinants of extreme kinetic stability.** *Proceedings of the National Academy of Sciences of the United States of America* 2015, **112**:14605–10 <https://doi.org/10.1073/pnas.1510748112>.
100. Broom A, Doxey AC, Lobsanov YD, Berthin LG, Rose DR, Howell PL, McConkey BJ, Meiering EM: **Modular evolution and the origins of symmetry: Reconstruction of a three-fold symmetric globular protein.** *Structure* 2012, **20**:161–171 <https://doi.org/10.1016/j.str.2011.10.021>.
101. Lee J, Blaber M: **Experimental support for the evolution of symmetric protein architecture from a simple peptide motif.** *Proceedings of the National Academy of Sciences* 2011, **108**:126–130 <https://doi.org/10.1073/pnas.1015032108>.
102. Borgia MB, Borgia A, Best RB, Steward A, Nettels D, Wunderlich B, Schuler B, Clarke J: **Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins.** *Nature* 2011, **474**:662–665 <https://doi.org/10.1038/nature10099>.
103. Zheng W, Schafer NP, Wolynes PG: **Frustration in the energy landscapes of multidomain protein misfolding.** *Proceedings of the National Academy of Sciences* 2013, **110**:1680–1685 <https://doi.org/10.1073/pnas.1222130110>.
104. Wright CF, Teichmann SA, Clarke J, Dobson CM: **The importance of sequence diversity in the aggregation and evolution Wright, C. F., Teichmann, S. A., Clarke, J., & Dobson, C. M. (2005). The importance of sequence diversity in the aggregation and evolution of proteins. Nature, 438(7069), 878–81. http://. Nature 2005, 438:878–81 https://doi.org/10.1038/nature04195.**
105. Ghashut F: **Role of Symmetry in Hisactophilin Folding.** 2009, [no volume].
106. Liu C: **Thermodynamic and kinetic folding studies of hisactophilin, a beta-trefoil protein.** 1999, [no volume].
107. Fersht AR: *Structure and Mechanism in Protein Science. A guide to Enzyme Catalysis and Protein Folding.* Macmillan; 1999.
108. Lunzer M, Golding GB, Dean AM: **Pervasive Cryptic Epistasis in Molecular Evolution.** *PLoS Genetics* 2010, **6**:e1001162 <https://doi.org/10.1371/journal.pgen.1001162>.
109. Breen MS, Kemena C, Vlasov PK, Notredame C, Kondrashov FA: **Epistasis as the primary factor in molecular evolution.** *Nature* 2012, **490**:535–538 <https://doi.org/10.1038/nature11510>.
110. Kim J, Brych SR, Lee J, Logan TM, Blaber M: **Identification of a key structural element for protein folding within β -hairpin turns.** *Journal of Molecular Biology* 2003, **328**:951–961 [https://doi.org/10.1016/S0022-2836\(03\)00321-8](https://doi.org/10.1016/S0022-2836(03)00321-8).

111. Smith CK, Regan L: **Construction and Design of β -Sheets**. *Accounts of Chemical Research* 1997, **30**:153–161 <https://doi.org/10.1021/ar9601048>.
112. Brych SR, Dubey VK, Bienkiewicz E, Lee J, Logan TM, Blaber M: **Symmetric primary and tertiary structure mutations within a symmetric superfold: A solution, not a constraint, to achieve a foldable polypeptide**. *Journal of Molecular Biology* 2004, **344**:769–780 <https://doi.org/10.1016/j.jmb.2004.09.060>.
113. Zheng G, Schaefer M, Karplus M: **Hemoglobin Bohr Effects: Atomic Origin of the Histidine Residue Contributions**. *Biochemistry* 2013, **52**:8539–8555 <https://doi.org/10.1021/bi401126z>.
114. Nelson S, Poddar S, Lin T-Y, Pierson TC: **Protonation of Individual Histidine Residues Is Not Required for the pH-Dependent Entry of West Nile Virus: Evaluation of the "Histidine Switch" Hypothesis**. *Journal of Virology* 2009, **83**:12631–12635 <https://doi.org/10.1128/JVI.01072-09>.
115. Anil B, Sato S, Cho J-h, Raleigh DP: **Fine structure analysis of a protein folding transition state; distinguishing between hydrophobic stabilization and specific packing**. *Journal of molecular biology* 2005, **354**:693–705 <https://doi.org/10.1016/j.jmb.2005.08.054>.
116. Zwanzig R: **Two-state models of protein folding kinetics**. *Proceedings of the National Academy of Sciences* 1997, **94**:148–150 <https://doi.org/10.1073/pnas.94.1.148>.
117. Creighton TE: **Toward a better understanding of protein folding pathways**. *Proceedings of the National Academy of Sciences* 1988, **85**:5082–5086 <https://doi.org/10.1073/pnas.85.14.5082>.
118. Zarrine-afzar A, Zhang Z, Schweiker KL, Makhatadze GI, Davidson AR, Chan HS: **Kinetic consequences of native state optimization of surface-exposed electrostatic interactions in the Fyn SH3 domain**. *Proteins* 2012, **80**:858–70 <https://doi.org/10.1002/prot.23243>.
119. Cho J-H, Raleigh DP: **Electrostatic interactions in the denatured state and in the transition state for protein folding: effects of denatured state interactions on the analysis of transition state structure**. *Journal of molecular biology* 2006, **359**:1437–1446 <https://doi.org/10.1016/j.jmb.2006.04.038>.
120. Tripathi S, Garcia AE, Makhatadze GI: **Alterations of Nonconserved Residues Affect Protein Stability and Folding Dynamics through Charge-Charge Interactions**. *Journal of Physical Chemistry B* 2015, **119**:13103–13112 <https://doi.org/10.1021/acs.jpcc.5b08527>.
121. Tzul FO, Schweiker KL, Makhatadze GI: **Modulation of folding energy landscape by charge–charge interactions: Linking experiments with computational modeling**. *Proceedings of the National Academy of Sciences* 2015, **2**:201410424 <https://doi.org/10.1073/pnas.1410424112>.
122. Fisher KM, Haglund E, Noel JK, Hailey KL, Onuchic JN, Jennings PA: **Geometrical Frustration in Interleukin-33 Decouples the Dynamics of the Functional Element from the Folding Transition State Ensemble**. *PLOS ONE* 2015, **10**:e0144067 <https://doi.org/10.1371/journal.pone.0144067>.
123. Giri Rao VVH, Gosavi S: **Structural Perturbations Present in the Folding Cores of Interleukin-33 and Interleukin-1 β Correlate to Differences in Their Function**. *The Journal of Physical Chemistry B* 2015, **119**:11203–11214 <https://doi.org/10.1021/acs.jpcc.5b03111>.
124. Kiefhaber T, Quaas R, Hahn U, Schmid FX: **Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization**. *Biochemistry* 1990, **29**:3053–61.
125. Kiefhaber T: **Kinetic traps in lysozyme folding**. *Proceedings of the National Academy of Sciences* 1995, **92**:9029–9033 <https://doi.org/10.1073/pnas.92.20.9029>.
126. Buchner J, Kiefhaber T: *Protein Folding Handbook*. Wiley-VCH Verlag GmbH; 2005.
127. Starr TN, Thornton JW: **Epistasis in protein evolution**. *Protein Science* 2016, **25**:1204–1218 <https://doi.org/10.1002/pro.2897>.
128. Phillips PC: **Epistasis — the essential role of gene interactions in the structure and evolution of genetic systems**. *Nature Reviews Genetics* 2008, **9**:855–867 <https://doi.org/10.1038/nrg2452>.
129. Olson CA, Wu NC, Sun R: **A Comprehensive Biophysical Description of Pairwise Epistasis throughout an Entire Protein Domain**. *Current Biology* 2014, **24**:2643–2651 <https://doi.org/10.1016/j.cub.2014.09.072>.
130. Podgornaia AI, Laub MT: **Pervasive degeneracy and epistasis in a protein-protein interface**. *Science* 2015, **347**:673–677 <https://doi.org/10.1126/science.1257360>.
131. Bridgham JT, Keay J, Ortlund EA, Thornton JW: **Vestigialization of an Allosteric Switch: Genetic and Structural Mechanisms for the Evolution of Constitutive Activity in a Steroid Hormone Receptor**. *PLoS Genetics* 2014, **10** <https://doi.org/10.1371/journal.pgen.1004058>.
132. Mohan Poluri MK: **NMR Investigations on Ruggedness of Native State Energy Landscape in Folded Proteins**. In *Protein-protein interactions - computational and experimental tools*. InTech; 2012.
133. Mulder FAA, Mittermaier A, Hon B, Dahlquist FW, Kay LE: **Studying excited states of proteins by NMR spectroscopy**. *Nature Struct. Biol.* 2001, **8**:932–935 <https://doi.org/10.1038/nsb1101-932>.
134. Anthis NJ, Clore GM: **Visualizing transient dark states by NMR spectroscopy**. *Q Rev Biophys.* 2015, **1**:35–116 <https://doi.org/10.1017/S0033583514000122>.
135. Culik RM, Sekhar A, Nagesh J, Deol H, Rumpfolt JAO, Meiering EM, Kay LE: **Effects of maturation on the conformational free-energy landscape of SOD1**. *Proceedings of the National Academy of Sciences* 2018, **115**:E2546–E2555 <https://doi.org/10.1073/pnas.1721022115>.
136. Baxter NJ, Williamson MP: **Temperature dependence of ^1H chemical shifts in proteins**. *Journal of Biomolecular NMR* 1997, **9**:359–369 <https://doi.org/10.1023/A:1018334207887>.
137. Clore GM: **Exploring sparsely populated states of macromolecules by diamagnetic and paramagnetic NMR relaxation**. *Protein Science* 2011, **20**:229–246 <https://doi.org/10.1002/pro.576>.
138. Stafford KA, Trbovic N, Butterwick JA, Abel R, Friesner RA, Palmer AG: **Conformational preferences underlying reduced activity of a thermophilic ribonuclease H**. *Journal of Molecular Biology* 2015, **427**:853–866 <https://doi.org/10.1016/j.jmb.2014.11.023>.
139. Kerns SJ, Agafonov RV, Cho Y-J, Pontiggia F, Otten R, Pachov DV, Kutter S, Phung LA, Murphy PN, Thai V, et al.: **The energy landscape of adenylate kinase during catalysis**. *Nature structural & molecular biology* 2015, **22**:124–31 <https://doi.org/10.1038/nsmb.2941>.
140. Brünschweiler S, Konrat R, Tollinger M: **Allosteric communication in the KIX domain proceeds through dynamic repacking of the hydrophobic core**. *ACS Chemical Biology* 2013, **8**:1600–1610 <https://doi.org/10.1021/cb4002188>.

141. Tzeng S-R, Kalodimos CG: **Allosteric inhibition through suppression of transient conformational states.** *Nature chemical biology* 2013, **9**:462–5 <https://doi.org/10.1038/nchembio.1250>.
142. Tzeng S-R, Kalodimos CG: **Dynamic activation of an allosteric regulatory protein.** *Nature* 2009, **462**:368–72 <https://doi.org/10.1038/nature08560>.
143. Akimoto M, Selvaratnam R, McNicholl ET, Verma G, Taylor SS, Melacini G: **Signaling through dynamic linkers as revealed by PKA.** *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**:14231–6 <https://doi.org/10.1073/pnas.1312644110>.
144. Teilum K, Olsen JG, Kragelund BB: **Protein stability, flexibility and function.** *Biochimica et Biophysica Acta - Proteins and Proteomics* 2011, **1814**:969–976 <https://doi.org/10.1016/j.bbapap.2010.11.005>.
145. Sekhar A, Kay LE: **NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers.** *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**:12867–74 <https://doi.org/10.1073/pnas.1305688110>.
146. McDonald LR, Whitley MJ, Boyer JA, Lee AL: **Colocalization of fast and slow timescale dynamics in the allosteric signaling protein CheY.** *Journal of Molecular Biology* 2013, **425**:2372–2381 <https://doi.org/10.1016/j.jmb.2013.04.029>.
147. Sabo TM, Trent JO, Lee D: **Population shuffling between ground and high energy excited states.** *Protein Science* 2015, **24**:1714–1719 <https://doi.org/10.1002/pro.2797>.
148. Haupt C, Weininger U, Kovermann M, Balbach J: **Local and coupled thermodynamic stability of the two-domain and bifunctional enzyme SlyD from Escherichia coli.** *Biochemistry* 2011, **50**:7321–7329 <https://doi.org/10.1021/bi2000627>.
149. Geitner AJ, Weininger U, Paulsen H, Balbach J, Kovermann M: **Structure-Based Insights into the Dynamics and Function of Two-Domain SlpA from Escherichia coli.** *Biochemistry* 2017, **56**:6533–6543 <https://doi.org/10.1021/acs.biochem.7b00786>.
150. Doyle CM, Rumpfheldt JA, Broom HR, Sekhar A, Kay LE, Meiering EM: **Concurrent Increases and Decreases in Local Stability and Conformational Heterogeneity in Cu, Zn Superoxide Dismutase Variants Revealed by Temperature-Dependence of Amide Chemical Shifts.** *Biochemistry* 2016, **55**:1346–1361 <https://doi.org/10.1021/acs.biochem.5b01133>.
151. Sekhar A, Rumpfheldt JAO, Broom HR, Doyle CM, Bouvignies G, Meiering EM, Kay LE: **Thermal fluctuations of immature SOD1 lead to separate folding and misfolding pathways.** *eLife* 2015, **4**:e07296 <https://doi.org/10.7554/eLife.07296>.
152. Jia Z, Vandonselaar M, Quail JW, Delbaere LTJ: **Active-centre torsion-angle strain revealed in 1.6 Å-resolution structure of histidine-containing phosphocarrier protein.** *Nature* 1993, **361**:94–97 <https://doi.org/10.1038/361094a0>.
153. Ventura S, Cristina Vega M, Lacroix E, Angrand I, Spagnolo L, Serrano L: **Conformational strain in the hydrophobic core and its implications for protein folding and design.** *Nature Structural Biology* 2002, **9**:485–493 <https://doi.org/10.1038/nsb799>.
154. Vega MC, Serrano L, Martínez JC: **Thermodynamic and structural characterization of Asn and Ala residues in the disallowed II' region of the Ramachandran plot.** *Protein Science* 2000, **9**:2322–2328 <https://doi.org/10.1110/ps.9.12.2322>.
155. Schuler B, Fürst F, Osterroth F, Steinbacher S, Huber R, Seckler R: **Plasticity and steric strain in a parallel β -helix: Rational mutations in the P22 tailspike protein.** *Proteins: Structure, Function and Genetics* 2000, **39**:89–101 [https://doi.org/10.1002/\(SICI\)1097-0134\(20000401\)39:1%3C89::AID-PROT10%3E3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1097-0134(20000401)39:1%3C89::AID-PROT10%3E3.0.CO;2-Q).
156. Broom A, Jacobi Z, Trainor K, Meiering EM: **Computational tools help improve protein stability but with a solubility tradeoff.** *Journal of Biological Chemistry* 2017, **292**:14349–14361 <https://doi.org/10.1074/jbc.M117.784165>.
157. Whitford PC, Miyashita O, Levy Y, Onuchic JN: **Conformational Transitions of Adenylate Kinase: Switching by Cracking.** *Journal of Molecular Biology* 2007, **366**:1661–1671 <https://doi.org/10.1016/j.jmb.2006.11.085>.
158. Stojanoski V, Adamski CJ, Hu L, Mehta SC, Sankaran B, Zwart P, Prasad BVV, Palzkill T: **Removal of the Side Chain at the Active-Site Serine by a Glycine Substitution Increases the Stability of a Wide Range of Serine β -Lactamases by Relieving Steric Strain.** *Biochemistry* 2016, **55**:2479–2490 <https://doi.org/10.1021/acs.biochem.6b00056>.
159. Kleckner IR, Foster MP: **An introduction to NMR-based approaches for measuring protein dynamics.** *Biochimica et Biophysica Acta - Proteins and Proteomics* 2011, **1814**:942–968 <https://doi.org/10.1016/j.bbapap.2010.10.012>.
160. Wolynes PG: **Energy landscapes and solved protein-folding problems.** *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 2005, **363**:453–464; discussion 464–467 <https://doi.org/10.1098/rsta.2004.1502>.
161. Henzler-Wildman K, Kern D: **Dynamic personalities of proteins.** *Nature* 2007, **450**:964–972 <https://doi.org/10.1038/nature06522>.
162. Wüthrich K: *NMR of Proteins and Nucleic Acids.* 1986.
163. Kim PS, Baldwin RL: **Structural Intermediates Trapped During the Folding of Ribonuclease A by Amide Proton Exchange.** *Biochemistry* 1980, **19**:6124–6129 <https://doi.org/10.1021/bi00567a027>.
164. Roder H, Elöve GA, Englander SW: **Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR.** *Nature* 1988, **335**:700–704 <https://doi.org/10.1038/335700a0>.
165. Schmid FX, Baldwin RL: **Detection of an early intermediate in the folding of ribonuclease A by protection of amide protons against exchange.** *Journal of Molecular Biology* 1979, **135**:199–215 [https://doi.org/10.1016/0022-2836\(79\)90347-4](https://doi.org/10.1016/0022-2836(79)90347-4).
166. Vallurupalli P, Bouvignies G, Kay LE: **Increasing the exchange time-scale that can be probed by CPMG relaxation dispersion NMR.** *Journal of Physical Chemistry B* 2011, **115**:14891–14900 <https://doi.org/10.1021/jp209610v>.
167. Bouvignies G, Vallurupalli P, Kay LE: **Visualizing side chains of invisible protein conformers by solution NMR.** *Journal of Molecular Biology* 2014, **426**:763–774 <https://doi.org/10.1016/j.jmb.2013.10.041>.
168. Palmer AG, Kroenke CD, Patrick Loria J: **Nuclear Magnetic Resonance Methods for Quantifying Microsecond-to-Millisecond Motions in Biological Macromolecules.** In *Methods in enzymology.* Elsevier Masson SAS; 2001:204–238.
169. Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, Kern D: **Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair.** *Nature Structural and Molecular Biology* 2004, **11**:945–949 <https://doi.org/10.1038/nsmb821>.

170. Kay LE, Torchia DA, Bax A: **Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease.** *Biochemistry* 1989, **28**:8972–8979 <https://doi.org/10.1021/bi00449a003>.
171. Wang T, Frederick KK, Igumenova TI, Wand AJ, Zuiderweg ERP: **Changes in Calmodulin Main-Chain Dynamics upon Ligand Binding Revealed by Cross-Correlated NMR Relaxation Measurements.** *Journal of the American Chemical Society* 2005, **127**:828–829 <https://doi.org/10.1021/ja045743p>.
172. Lundström P, Mulder FAA, Akke M: **Correlated dynamics of consecutive residues reveal transient and cooperative unfolding of secondary structure in proteins.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**:16984–16989 <https://doi.org/10.1073/pnas.0504361102>.
173. Pelupessy P, Ravindranathan S, Bodenhausen G: **Correlated motions of successive amide N-H bonds in proteins.** *Journal of Biomolecular NMR* 2003, **25**:265–280 <https://doi.org/10.1023/A:1023076212536>.
174. Evans JNS: *Biomolecular NMR Spectroscopy.* Oxford University Press; 1995.
175. Rule GS, Hitchens TK: *Fundamentals of protein nmr spectroscopy.* Springer Science & Business Media; 2006.
176. Hafsa NE, Arndt D, Wishart DS: **Accessible surface area from NMR chemical shifts.** *Journal of Biomolecular NMR* 2015, **62**:387–401 <https://doi.org/10.1007/s10858-015-9957-0>.
177. Williamson MP: **Using chemical shift perturbation to characterise ligand binding.** 2013, **73**:1–16 <https://doi.org/10.1016/j.pnmrs.2013.02.001>.
178. Jeener J, Meier BH, Bachmann P, Ernst RR: **Investigation of exchange processes by two-dimensional NMR spectroscopy.** *The Journal of Chemical Physics* 1979, **71**:4546–4553 <https://doi.org/10.1063/1.438208>.
179. Wishart DS: **Interpreting protein chemical shift data.** *Progress in Nuclear Magnetic Resonance Spectroscopy* 2011, **58**:62–87 <https://doi.org/10.1016/j.pnmrs.2010.07.004>.
180. Han B, Liu Y, Ginzinger SW, Wishart DS: **SHIFTX2: Significantly improved protein chemical shift prediction.** *Journal of Biomolecular NMR* 2011, **50**:43–57 <https://doi.org/10.1007/s10858-011-9478-4>.
181. Hafsa NE, Berjanskii MV, Arndt D, Wishart DS: **Rapid and reliable protein structure determination via chemical shift threading.** *Journal of Biomolecular NMR* 2018, **70**:33–51 <https://doi.org/10.1007/s10858-017-0154-1>.
182. Baxter NJ, Hosszu LLP, Waltho JP, Williamson MP: **Characterisation of low free-energy excited states of folded proteins.** *Journal of Molecular Biology* 1998, **284**:1625–1639 <https://doi.org/10.1006/jmbi.1998.2265>.
183. Tomlinson JH, Williamson MP: **Amide temperature coefficients in the protein G B1 domain.** *Journal of Biomolecular NMR* 2012, **52**:57–64 <https://doi.org/10.1007/s10858-011-9583-4>.
184. Berjanskii M, Tang P, Liang J, Cruz J a, Zhou J, Zhou Y, Bassett E, MacDonell C, Lu P, Lin G, et al.: **GeNMR: a web server for rapid NMR-based protein structure determination.** *Nucleic acids research* 2009, **37**:W670–7 <https://doi.org/10.1093/nar/gkp280>.
185. Cui DS, Beaumont V, Ginther PS, Lipchock JM, Loria JP: **Leveraging Reciprocity to Identify and Characterize Unknown Allosteric Sites in Protein Tyrosine Phosphatases.** *Journal of Molecular Biology* 2017, **429**:2360–2372 <https://doi.org/10.1016/j.jmb.2017.06.009>.
186. Selvaratnam R, Chowdhury S, VanSchouwen B, Melacini G: **Mapping allostery through the covariance analysis of NMR chemical shifts.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**:6133–8 <https://doi.org/10.1073/pnas.1017311108>.
187. Sternlicht H, Wilson D: **Magnetic Resonance Studies of Macromolecules. I. Aromatic-Methyl Interactions and Helical Structure Effects in Lysozyme *.** *Biochemistry* 1967, **6**:2881–2892 <https://doi.org/10.1021/bi00861a032>.
188. Hafsa NE, Arndt D, Wishart DS: **CSI 3.0: a web server for identifying secondary and super-secondary structure in proteins using NMR chemical shifts.** *Nucleic Acids Research* 2015, **43**:W370–W377 <https://doi.org/10.1093/nar/gkv494>.
189. Berjanskii MV, Wishart DS: **Unraveling the meaning of chemical shifts in protein NMR.** *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 2017, **1865**:1564–1576 <https://doi.org/10.1016/j.bbapap.2017.07.005>.
190. Neal S, Nip AM, Zhang H, Wishart DS: **Rapid and accurate calculation of protein 1H, 13C and 15N chemical shifts.** *Journal of Biomolecular NMR* 2003, **26**:215–240 <https://doi.org/10.1023/A:1023812930288>.
191. Meiler J: **PROSHIFT: Protein chemical shift prediction using artificial neural networks.** *Journal of Biomolecular NMR* 2003, **26**:25–37 <https://doi.org/10.1023/A:1023060720156>.
192. Cavalli A, Salvatella X, Dobson CM, Vendruscolo M: **Protein structure determination from NMR chemical shifts.** *Proceedings of the National Academy of Sciences* 2007, **104**:9615–9620 <https://doi.org/10.1073/pnas.0610313104>.
193. Martin OA, Vila JA, Scheraga HA: **CheShift-2: Graphic validation of protein structures.** *Bioinformatics* 2012, **28**:1538–1539 <https://doi.org/10.1093/bioinformatics/bts179>.
194. Iwadate M, Asakura T, Williamson MP: **C(α) and C(β) carbon-13 chemical shifts in proteins from an empirical database.** *Journal of Biomolecular NMR* 1999, **13**:199–211 <https://doi.org/10.1023/A:1008376710086>.
195. Shen Y, Bax A: **SPARTA+: A modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network.** *Journal of Biomolecular NMR* 2010, **48**:13–22 <https://doi.org/10.1007/s10858-010-9433-9>.
196. Wishart DS, Watson MS, Boyko RF, Sykes BD: **Automated H-1 and C-13 chemical shift prediction using the BioMagResBank.** *Journal of Biomolecular NMR* 1997, **10**:329–336.
197. Wishart DS, Sykes BD, Richards FM: **The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy.** *Biochemistry* 1992, **31**:1647–1651 <https://doi.org/10.1021/bi00121a010>.
198. Eghbalian HR, Wang L, Bahrami A, Assadi A, Markley JL: **Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements.** *Journal of Biomolecular NMR* 2005, **32**:71–81 <https://doi.org/10.1007/s10858-005-5705-1>.
199. Marsh JA, Singh VK, Jia Z, Forman-Kay JD: **Sensitivity of secondary structure propensities to sequence differences between α - and γ -synuclein: Implications for fibrillation.** *Protein Science* 2006, **15**:2795–2804 <https://doi.org/10.1110/ps.062465306>.
200. Shen Y, Bax A: **Identification of helix capping and β -turn motifs from NMR chemical shifts.** *Journal of Biomolecular NMR* 2012, **52**:211–232 <https://doi.org/10.1007/s10858-012-9602-0>.
201. Camilloni C, De Simone A, Vranken WF, Vendruscolo M: **Determination of Secondary Structure Populations in Disordered States of Proteins Using Nuclear Magnetic Resonance Chemical Shifts.** *Biochemistry* 2012, **51**:2224–2231 <https://doi.org/10.1021/bi3001825>.

202. Berjanskii MV, Wishart DS: **Application of the random coil index to studying protein flexibility.** *Journal of Biomolecular NMR* 2008, **40**:31–48 <https://doi.org/10.1007/s10858-007-9208-0>.
203. Berjanskii MV, Wishart DS: **A simple method to measure protein side-chain mobility using NMR chemical shifts.** *Journal of the American Chemical Society* 2013, **135**:14536–14539 <https://doi.org/10.1021/ja407509z>.
204. Bratholm LA, Christensen AS, Hamelryck T, Jensen JH: **Bayesian inference of protein structure from chemical shift data.** *PeerJ* 2015, **3**:e861 <https://doi.org/10.7717/peerj.861>.
205. Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, et al.: **Consistent blind protein structure generation from NMR chemical shift data.** *Proceedings of the National Academy of Sciences* 2008, **105**:4685–4690 <https://doi.org/10.1073/pnas.0800256105>.
206. Shen Y, Bax A: **Homology modeling of larger proteins guided by chemical shifts.** *Nature Methods* 2015, **12**:747–750 <https://doi.org/10.1038/nmeth.3437>.
207. Berjanskii MV, Neal S, Wishart DS: **PREDITOR: a web server for predicting protein torsion angle restraints.** *Nucleic Acids Research* 2006, **34**:W63–W69 <https://doi.org/10.1093/nar/gkl341>.
208. Shen Y, Delaglio F, Cornilescu G, Bax A: **TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts.** *Journal of Biomolecular NMR* 2009, **44**:213–223 <https://doi.org/10.1007/s10858-009-9333-z>.
209. Shen Y, Bax A: **Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks.** *Journal of Biomolecular NMR* 2013, **56**:227–241 <https://doi.org/10.1007/s10858-013-9741-y>.
210. Christensen AS, Linnet TE, Borg M, Boomsma W, Lindorff-Larsen K, Hamelryck T, Jensen JH: **Protein Structure Validation and Refinement Using Amide Proton Chemical Shifts Derived from Quantum Mechanics.** *PLoS ONE* 2013, **8**:e84123 <https://doi.org/10.1371/journal.pone.0084123>.
211. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD: **¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.** *Journal of Biomolecular NMR* 1995, **5**:67–81 <https://doi.org/10.1007/BF00227471>.
212. Buckingham AD: **Chemical Shifts in the Nuclear Magnetic Resonance Spectra of Molecules Containing Polar Groups.** *Canadian Journal of Chemistry* 1960, **38**:300–307 <https://doi.org/10.1139/v60-040>.
213. Ösapay K, Case DA: **A New Analysis of Proton Chemical Shifts in Proteins.** *Journal of the American Chemical Society* 1991, <https://doi.org/10.1021/ja00025a002>.
214. Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD, Wishart DS: **VADAR: A web server for quantitative evaluation of protein structure quality.** *Nucleic Acids Research* 2003, <https://doi.org/10.1093/nar/gkg565>.
215. Hong J, Jing Q, Yao L: **The protein amide ¹HN chemical shift temperature coefficient reflects thermal expansion of the N–H=O hydrogen bond.** *Journal of Biomolecular NMR* 2013, **55**:71–78 <https://doi.org/10.1007/s10858-012-9689-3>.
216. Wishart DS, Case DA: **Use of Chemical Shifts in Macromolecular Structure Determination.** In *Methods in enzymology*. Elsevier Masson SAS; 2002:3–34.
217. Parker LL, Houk AR, Jensen JH: **Cooperative hydrogen bonding effects are key determinants of backbone amide proton chemical shifts in proteins.** *Journal of the American Chemical Society* 2006, **128**:9863–9872 <https://doi.org/10.1021/ja0617901>.
218. Case DA: **Chemical shifts in biomolecules.** *Current Opinion in Structural Biology* 2013, **23**:172–176 <https://doi.org/10.1016/j.sbi.2013.01.007>.
219. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW: **Discovering High-Affinity Ligands for Proteins: SAR by NMR.** *Science* 1996, **274**:1531–1534 <https://doi.org/10.1126/science.274.5292.1531>.
220. McCoy MA, Wyss DF: **Spatial Localization of Ligand Binding Sites from Electron Current Density Surfaces Calculated from NMR Chemical Shift Perturbations.** *Journal of the American Chemical Society* 2002, **124**:11758–11763 <https://doi.org/10.1021/ja026166c>.
221. Piazza M, Taiakina V, Dieckmann T, Guillemette JG: **Structural Consequences of Calmodulin EF Hand Mutations.** *Biochemistry* 2017, **56**:944–956 <https://doi.org/10.1021/acs.biochem.6b01296>.
222. Piazza M, Guillemette JG, Dieckmann T: **Chemical shift perturbations induced by residue specific mutations of CaM interacting with NOS peptides.** *Biomolecular NMR Assignments* 2015, **9**:299–302 <https://doi.org/10.1007/s12104-015-9596-0>.
223. Kucic P, Farrell D, McIntosh LP, García-Moreno E. B, Jensen KS, Toleikis Z, Teilum K, Nielsen JE: **Protein Dielectric Constants Determined from NMR Chemical Shift Perturbations.** *Journal of the American Chemical Society* 2013, **135**:16968–16976 <https://doi.org/10.1021/ja406995j>.
224. Webb H, Tynan-Connolly BM, Lee GM, Farrell D, O'Meara F, Søndergaard CR, Teilum K, Hewage C, McIntosh LP, Nielsen JE: **Remeasuring HEWL pKa values by NMR spectroscopy: Methods, analysis, accuracy, and implications for theoretical pKa calculations.** *Proteins: Structure, Function, and Bioinformatics* 2011, **79**:685–702 <https://doi.org/10.1002/prot.22886>.
225. Keller RLJ: **Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment.** 2004, <https://doi.org/10.3929/ethz-a-005068942>.
226. Ohnishi M, Urry DW: **Temperature dependence of amide proton chemical shifts: The secondary structures of gramicidin S and valinomycin.** *Biochemical and Biophysical Research Communications* 1969, **36**:194–202 [https://doi.org/10.1016/0006-291X\(69\)90314-3](https://doi.org/10.1016/0006-291X(69)90314-3).
227. Koppke KD, Ohnishi M, Go A: **Conformations of cyclic peptides. IV. Nuclear magnetic resonance studies of cyclo-pentaglycyl-L-leucyl and cyclo-diglycyl-L-histidylglycyl-L-tyrosyl.** *Biochemistry* 1969, **8**:4087–4095 <https://doi.org/10.1021/bi00838a028>.
228. Dyson H, Rance M, Houghten RA, Lerner RA, Wright PE: **Folding of immunogenic peptide fragments of proteins in water solution.** *Journal of Molecular Biology* 1988, **201**:161–200 [https://doi.org/10.1016/0022-2836\(88\)90446-9](https://doi.org/10.1016/0022-2836(88)90446-9).
229. Merutka G, Jane Dyson H, Wright PE: **'Random coil' ¹H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG.** *Journal of Biomolecular NMR* 1995, **5**:14–24 <https://doi.org/10.1007/BF00227466>.

230. Rothmund S, Weißhoff H, Beyermann M, Krause E, Bienert M, Mügge C, Sykes BD, Sönnichsen FD: **Temperature coefficients of amide proton NMR resonance frequencies in trifluoroethanol: A monitor of intramolecular hydrogen bonds in helical peptides?** *Journal of Biomolecular NMR* 1996, **8**:93–97 <https://doi.org/10.1007/BF00198143>.
231. Andersen NH, Neidigh JW, Harris SM, Lee GM, Liu Z, Tong H: **Extracting Information from the Temperature Gradients of Polypeptide NH Chemical Shifts. 1. The Importance of Conformational Averaging.** *Journal of the American Chemical Society* 1997, **119**:8547–8561 <https://doi.org/10.1021/ja963250h>.
232. Cierpicki T, Otlewski J: **Amide proton temperature coefficients as hydrogen bond indicators in proteins.** *Journal of Biomolecular NMR* 2001, **21**:249–261 <https://doi.org/10.1023/A:1012911329730>.
233. Cierpicki T, Zhukov I, Byrd RA, Otlewski J: **Hydrogen Bonds in Human Ubiquitin Reflected in Temperature Coefficients of Amide Protons.** *Journal of Magnetic Resonance* 2002, **157**:178–180 <https://doi.org/10.1006/jmre.2002.2597>.
234. Cordier F, Grzesiek S: **Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR.** *Journal of Molecular Biology* 2002, **317**:739–752 <https://doi.org/10.1006/jmbi.2002.5446>.
235. Miletti T, Di Trani J, Levros LC, Mittermaier A: **Conformational plasticity surrounding the active site of NADH oxidase from *Thermus thermophilus*.** *Protein Science* 2015, **24**:1114–1128 <https://doi.org/10.1002/pro.2693>.
236. Qiao X, Liu Y, Luo L, Chen L, Zhao C, Ai X: **Effects of naturally occurring charged mutations on the structure, stability, and binding of the Pin1 WW domain.** *Biochemical and Biophysical Research Communications* 2017, **487**:470–476 <https://doi.org/10.1016/j.bbrc.2017.04.093>.
237. Huang K-Y, Kingsley CN, Sheil R, Cheng C-Y, Bierma JC, Roskamp KW, Khago D, Martin RW, Han S: **Stability of Protein-Specific Hydration Shell on Crowding.** *Journal of the American Chemical Society* 2016, **138**:5392–5402 <https://doi.org/10.1021/jacs.6b01989>.
238. Fu I, Case DA, Baum J: **Dynamic Water-Mediated Hydrogen Bonding in a Collagen Model Peptide.** *Biochemistry* 2015, **54**:6029–6037 <https://doi.org/10.1021/acs.biochem.5b00622>.
239. Dolenský B, Konvalinka R, Jakubek M, Král V: **Identification of intramolecular hydrogen bonds as the origin of malfunctioning of multitopic receptors.** *Journal of Molecular Structure* 2013, **1035**:124–128 <https://doi.org/10.1016/j.molstruc.2012.09.040>.
240. Bouvignies G, Vallurupalli P, Cordes MHJ, Hansen DF, Kay LE: **Measuring 1HN temperature coefficients in invisible protein states by relaxation dispersion NMR spectroscopy.** *Journal of Biomolecular NMR* 2011, **50**:13–18 <https://doi.org/10.1007/s10858-011-9498-0>.
241. Hvidt A, Nielsen SO: **Hydrogen Exchange in Proteins.** In *Advances in protein chemistry*.. Academic Press; 1966:287–386.
242. Englander SW, Downer NW, Teitelbaum H: **Hydrogen Exchange.** *Annual Review of Biochemistry* 1972, **41**:903–924 <https://doi.org/10.1146/annurev.bi.41.070172.004351>.
243. Englander SW, Kallenbach NR: **Hydrogen exchange and structural dynamics of proteins and nucleic acids.** *Quarterly Reviews of Biophysics* 1983, **16**:521 <https://doi.org/10.1017/S0033583500005217>.
244. Sharma M, Kumar D, Poluri KM: **Unraveling the differential structural stability and dynamics features of T7 endolysin partially folded conformations.** *Biochimica et Biophysica Acta (BBA) - General Subjects* 2018, **1862**:924–935 <https://doi.org/10.1016/j.bbagen.2018.01.004>.
245. Cornilescu G, Hu J-S, Bax A: **Identification of the Hydrogen Bonding Network in a Protein by Scalar Couplings.** *Journal of the American Chemical Society* 1999, **121**:2949–2950 <https://doi.org/10.1021/ja9902221>.
246. Cornilescu G, Ramirez BE, Frank MK, Clore GM, Gronenborn AM, Bax A: **Correlation between 3h J NC ' and Hydrogen Bond Length in Proteins.** *Journal of the American Chemical Society* 1999, **121**:6275–6279 <https://doi.org/10.1021/ja9909024>.
247. Miller JF: **Bacterial transformation by electroporation.** *Methods in Enzymology* 1994, **235**:375–385 [https://doi.org/10.1016/0076-6879\(94\)35156-2](https://doi.org/10.1016/0076-6879(94)35156-2).
248. Grzesiek S, Bax A: **The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements.** *Journal of the American Chemical Society* 1993, **115**:12593–12594 <https://doi.org/10.1021/ja00079a052>.
249. Schleucher J, Schwendinger M, Sattler M, Schmidt P, Schedletsky O, Glaser SJ, Sørensen OW, Griesinger C: **A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients.** *Journal of Biomolecular NMR* 1994, **4**:301–306 <https://doi.org/10.1007/BF00175254>.
250. Kay LE, Keifer P, Saarinen T: **Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity.** *Journal of the American Chemical Society* 1992, **114**:10663–10665 <https://doi.org/10.1021/ja00052a088>.
251. Palmer AG, Cavanagh J, Wright PE, Rance M: **Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy.** *Journal of Magnetic Resonance (1969)* 1991, **93**:151–170 [https://doi.org/10.1016/0022-2364\(91\)90036-S](https://doi.org/10.1016/0022-2364(91)90036-S).
252. Zhang O, Kay LE, Olivier JP, Forman-Kay JD: **Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques.** *Journal of Biomolecular NMR* 1994, **4**:845–858 <https://doi.org/10.1007/BF00398413>.
253. Morris GA, Freeman R: **Enhancement of Nuclear Magnetic Resonance Signals by Polarization Transfer.** 1979, **101**:760–762 <https://doi.org/10.1021/ja00497a058>.
254. Shaka AJ, Barker PB, Freeman R: **Computer-optimized decoupling scheme for wideband applications and low-level operation.** *Journal of Magnetic Resonance (1969)* 1985, **64**:547–552 [https://doi.org/10.1016/0022-2364\(85\)90122-2](https://doi.org/10.1016/0022-2364(85)90122-2).
255. Stone J, Gullingsrud J, Grayson P, Schulten K: **A System for Interactive Molecular Dynamics Simulation.** In *2001 acm symposium on interactive 3D graphics*. Edited by Hughes JF, Séquin CH. ACM SIGGRAPH; 2001:191–194.
256. Stone J: **An Efficient Library for Parallel Ray Tracing and Animation.** 1998, [no volume].
257. Frishman D, Argos P: **Knowledge-based protein secondary structure assignment.** *Proteins: Structure, Function, and Genetics* 1995, **23**:566–579 <https://doi.org/10.1002/prot.340230412>.
258. Hwang TL, Shaka AJ: **Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients.** *Journal of Magnetic Resonance, Series A* 1995, **112**:275–279 <https://doi.org/10.1006/jmra.1995.1047>.

259. Hoffman RE: **Standardization of chemical shifts of TMS and solvent signals in NMR solvents.** *Magnetic Resonance in Chemistry* 2006, **44**:606–616 <https://doi.org/10.1002/mrc.1801>.
260. McKinney W: **Data Structures for Statistical Computing in Python.** *Proceedings of the 9th Python in Science Conference* 2010, **1697900**:51–56.
261. Van Der Walt S, Colbert SC, Varoquaux G: **The NumPy array: A structure for efficient numerical computation.** *Computing in Science and Engineering* 2011, **13**:22–30 <https://doi.org/10.1109/MCSE.2011.37>.
262. Jones E, Oliphant T, Peterson P: **{SciPy}: Open source scientific tools for {Python}.** [date unknown], [no volume].
263. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED: **The CCPN data model for NMR spectroscopy: Development of a software pipeline.** *Proteins: Structure, Function and Genetics* 2005, **59**:687–696 <https://doi.org/10.1002/prot.20449>.
264. Hardman CH, Broadhurst RW, Raine ARC, Grasser KD, Thomas JO, Laue ED: **Structure of the A-Domain of HMG1 and Its Interaction with DNA as Studied by Heteronuclear Three- and Four-Dimensional NMR Spectroscopy.** *Biochemistry* 1995, **34**:16596–16607 <https://doi.org/10.1021/bi00051a007>.
265. Allain FH-T, Yen Y-m, Masse JE, Schultze P, Dieckmann T, Johnson RC, Feigon J: **Solution structure of the HMG protein NHP6A and its interaction with DNA reveals the structural determinants for non-sequence-specific binding.** *The EMBO Journal* 1999, **18**:2563–2579 <https://doi.org/10.1093/emboj/18.9.2563>.
266. Williamson RA, Carr MD, Frenkiel TA, Feeney J, Freedman RB: **Mapping the Binding Site for Matrix Metalloproteinase on the N-Terminal Domain of the Tissue Inhibitor of Metalloproteinases-2 by NMR Chemical Shift Perturbation †.** *Biochemistry* 1997, **36**:13882–13889 <https://doi.org/10.1021/bi9712091>.
267. Coggins BE, Li X, McClerren AL, Hindsgaul O, Raetz CRH, Zhou P: **Structure of the LpxC deacetylase with a bound substrate-analog inhibitor.** *Nature Structural & Molecular Biology* 2003, **10**:645–651 <https://doi.org/10.1038/nsb948>.
268. Boulton S, Akimoto M, Selvaratnam R, Bashiri A, Melacini G: **A tool set to map allosteric networks through the NMR chemical shift covariance analysis.** *Scientific reports* 2014, **4**:7306 <https://doi.org/10.1038/srep07306>.
269. Boulton S, Melacini G: **Advances in NMR Methods To Map Allosteric Sites: From Models to Translation.** *Chemical Reviews* 2016, **116**:6267–6304 <https://doi.org/10.1021/acs.chemrev.5b00718>.
270. Axe JM, Boehr DD: **Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling.** *Journal of Molecular Biology* 2013, **425**:1527–1545 <https://doi.org/10.1016/j.jmb.2013.01.030>.
271. Axe JM, O'Rourke KF, Kerstetter NE, Yezdimer EM, Chan YM, Chasin A, Boehr DD: **Severing of a hydrogen bond disrupts amino acid networks in the catalytically active state of the alpha subunit of tryptophan synthase.** *Protein Science* 2015, **24**:484–494 <https://doi.org/10.1002/pro.2598>.
272. Kim J, Masterson LR, Cembran A, Verardi R, Shi L, Gao J, Taylor SS, Veglia G: **Dysfunctional conformational dynamics of protein kinase A induced by a lethal mutant of phospholamban hinder phosphorylation.** *Proceedings of the National Academy of Sciences* 2015, <https://doi.org/10.1073/pnas.1502299112>.
273. Stollar EJ, Lin H, Davidson AR, Forman-Kay JD: **Differential Dynamic Engagement within 24 SH3 Domain: Peptide Complexes Revealed by Co-Linear Chemical Shift Perturbation Analysis.** *PLoS ONE* 2012, **7**:e51282 <https://doi.org/10.1371/journal.pone.0051282>.
274. Barfield M: **Structural Dependencies of Interresidue Scalar Coupling $h\ 3\ J\ NC'$ and Donor $1\ H$ Chemical Shifts in the Hydrogen Bonding Regions of Proteins.** *Journal of the American Chemical Society* 2002, **124**:4158–4168 <https://doi.org/10.1021/ja012674v>.
275. Sitkoff D, Case DA: **Density functional calculations of proton chemical shifts in model peptides.** *Journal of the American Chemical Society* 1997, **119**:12262–12273 <https://doi.org/10.1021/ja9721430>.
276. Kanaya S, Oobatake M, Liu Y: **Thermal Stability of Escherichia coli Ribonuclease HI and Its Active Site Mutants in the Presence and Absence of the Mg 2+ Ion.** *Journal of Biological Chemistry* 1996, **271**:32729–32736 <https://doi.org/10.1074/jbc.271.51.32729>.
277. Shoichet BK, Baase WA, Kuroki R, Matthews BW: **A relationship between protein stability and protein function.** *Proceedings of the National Academy of Sciences* 1995, **92**:452–456 <https://doi.org/10.1073/pnas.92.2.452>.
278. Warshel A: **Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites.** *Journal of Biological Chemistry* 1998, **273**:27035–27038 <https://doi.org/10.1074/jbc.273.42.27035>.
279. Okazaki K-i, Takada S: **Dynamic energy landscape view of coupled binding and protein conformational change: Induced-fit versus population-shift mechanisms.** *Proceedings of the National Academy of Sciences* 2008, **105**:11182–11187 <https://doi.org/10.1073/pnas.0802524105>.
280. Arora K, Brooks CL: **Large-scale allosteric conformational transitions of adenylate kinase appear to involve a population-shift mechanism.** *Proceedings of the National Academy of Sciences* 2007, **104**:18496–18501 <https://doi.org/10.1073/pnas.0706443104>.
281. Blacklock KM, Yachnin BJ, Woolley GA, Khare SD: **Computational Design of a Photocontrolled Cytosine Deaminase.** *Journal of the American Chemical Society* 2018, **140**:14–17 <https://doi.org/10.1021/jacs.7b08709>.
282. Anthis NJ, Doucleff M, Marius Clore GM: **Transient, sparsely populated compact states of apo and calcium-loaded calmodulin probed by paramagnetic relaxation enhancement: Interplay of conformational selection and induced fit.** *Journal of the American Chemical Society* 2011, **133**:18966–18974 <https://doi.org/10.1021/ja2082813>.
283. Ames JB, Tanaka T, Stryer L, Ikura M: **Secondary structure of myristoylated recoverin determined by three-dimensional heteronuclear NMR: implications for the calcium-myristoyl switch.** *Biochemistry* 1994, **33**:10743–53.
284. Solà M, López-Hernández E, Cronet P, Lacroix E, Serrano L, Coll M, Parraga A: **Towards understanding a molecular switch mechanism: Thermodynamic and crystallographic studies of the signal transduction protein CheY.** *Journal of Molecular Biology* 2000, **303**:213–225 <https://doi.org/10.1006/jmbi.2000.4507>.
285. Liu Y, Kahn RA, Prestegard JH: **Structure and Membrane Interaction of Myristoylated ARF1.** *Structure* 2009, **17**:79–87 <https://doi.org/10.1016/j.str.2008.10.020>.
286. Liu Y, Kahn RA, Prestegard JH: **Dynamic structure of membrane-anchored Arf*GTP.** *Nature structural & molecular biology* 2010, **17**:876–81 <https://doi.org/10.1038/nsmb.1853>.

287. Gambin Y, Schug A, Lemke EA, Lavinder JJ, Ferreone ACM, Magliery TJ, Onuchic JN, Deniz AA: **Direct single-molecule observation of a protein living in two opposed native structures.** *Proceedings of the National Academy of Sciences* 2009, **106**:10153–10158 <https://doi.org/10.1073/pnas.0904461106>.
288. Lockless SW, Ranganathan R: **Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families.** *Science* 1999, **286**:295–299 <https://doi.org/10.1126/science.286.5438.295>.
289. Fuentes EJ, Der CJ, Lee AL: **Ligand-dependent Dynamics and Intramolecular Signaling in a PDZ Domain.** *Journal of Molecular Biology* 2004, **335**:1105–1115 <https://doi.org/10.1016/j.jmb.2003.11.010>.
290. Robinson AC, Majumdar A, Schlessman JL, Garcia-Moreno E B: **Charges in Hydrophobic Environments: A Strategy for Identifying Alternative States in Proteins.** *Biochemistry* 2017, **56**:212–218 <https://doi.org/10.1021/acs.biochem.6b00843>.
291. Richman DE, Majumdar A, Bertrand Garcia-Moreno E: **Conformational Reorganization Coupled to the Ionization of Internal Lys Residues in Proteins.** *Biochemistry* 2015, **54**:5888–5897 <https://doi.org/10.1021/acs.biochem.5b00522>.
292. Strauch E-M, Fleishman SJ, Baker D: **Computational design of a pH-sensitive IgG binding protein.** *Proceedings of the National Academy of Sciences* 2014, **111**:675–680 <https://doi.org/10.1073/pnas.1313605111>.
293. Broom RA: **Computational Design of Protein Structure and Prediction of Ligand Binding.** 2016, [no volume].
294. Bryngelson JD, Wolynes PG: **Spin glasses and the statistical mechanics of protein folding.** *Proceedings of the National Academy of Sciences* 1987, **84**:7524–7528 <https://doi.org/10.1073/pnas.84.21.7524>.
295. Zhao L, Lu HP, Wang J: **Exploration of Multistate Conformational Dynamics upon Ligand Binding of a Monomeric Enzyme Involved in Pyrophosphoryl Transfer.** *The Journal of Physical Chemistry B* 2018, **122**:1885–1897 <https://doi.org/10.1021/acs.jpcc.7b12562>.
296. Choi JH, Laurent AH, Hilser VJ, Ostermeier M: **Design of protein switches based on an ensemble model of allostery.** *Nature communications* 2015, **6**:6968 <https://doi.org/10.1038/ncomms7968>.
297. Kumar S, Ma B, Tsai CJ, Sinha N, Nussinov R: **Folding and binding cascades: dynamic landscapes and population shifts.** *Protein science : a publication of the Protein Society* 2000, **9**:10–9 <https://doi.org/10.1110/ps.9.1.10>.
298. Murtaugh ML, Fanning SW, Sharma TM, Terry AM, Horn JR: **A combinatorial histidine scanning library approach to engineer highly pH-dependent protein switches.** *Protein Science* 2011, **20**:1619–1631 <https://doi.org/10.1002/pro.696>.
299. Xu X, Ishima R, Ames JB: **Conformational dynamics of recoverin's Ca²⁺-myristoyl switch probed by 15N NMR relaxation dispersion and chemical shift analysis.** *Proteins* 2011, **79**:1910–22 <https://doi.org/10.1002/prot.23014>.
300. Preininger AM, Kaya AI, Gilbert JA, Busenlehner LS, Armstrong RN, Hamm HE: **Myristoylation Exerts Direct and Allosteric Effects on G α Conformation and Dynamics in Solution.** *Biochemistry* 2012, **51**:1911–1924 <https://doi.org/10.1021/bi201472c>.
301. Clarkson MW, Lee AL: **Long-Range Dynamic Effects of Point Mutations Propagate through Side Chains in the Serine Protease Inhibitor Eglin c †.** *Biochemistry* 2004, **43**:12448–12458 <https://doi.org/10.1021/bi0494424>.
302. Kjaergaard M, Brander S, Poulsen FM: **Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH.** *Journal of Biomolecular NMR* 2011, **49**:139–149 <https://doi.org/10.1007/s10858-011-9472-x>.
303. Kjaergaard M, Poulsen FM: **Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution.** *Journal of Biomolecular NMR* 2011, **50**:157–165 <https://doi.org/10.1007/s10858-011-9508-2>.
304. Torrie GM, Valleau JP: **Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling.** *Journal of Computational Physics* 1977, **23**:187–199 [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
305. Beutler TC, Gunsteren WF van: **The computation of a potential of mean force: Choice of the biasing potential in the umbrella sampling technique.** *The Journal of Chemical Physics* 1994, **100**:1492–1497 <https://doi.org/10.1063/1.466628>.
306. Arthur EJ, Brooks CL: **Efficient implementation of constant pH molecular dynamics on modern graphics processors.** *Journal of Computational Chemistry* 2016, **107233**:2171–2180 <https://doi.org/10.1002/jcc.24435>.
307. Goh GB, Laricheva EN, Brooks CL: **Uncovering pH-dependent transient states of proteins with buried ionizable residues.** *Journal of the American Chemical Society* 2014, **136**:8496–8499 <https://doi.org/10.1021/ja5012564>.
308. Chan SWS, Yau J, Ing C, Liu K, Farber P, Won A, Bhandari V, Kara-Yacoubian N, Seraphim TV, Chakrabarti N, et al.: **Mechanism of Amyloidogenesis of a Bacterial AAA+ Chaperone.** *Structure* 2016, **24**:1095–1109 <https://doi.org/10.1016/j.str.2016.05.002>.
309. Childers MC, Daggett V: **Validating Molecular Dynamics Simulations against Experimental Observables in Light of Underlying Conformational Ensembles.** *Journal of Physical Chemistry B* 2018, **122**:6673–6689 <https://doi.org/10.1021/acs.jpcc.8b02144>.
310. Nerli S, McShan AC, Sgourakis NG: **Chemical shift-based methods in NMR structure determination.** *Progress in Nuclear Magnetic Resonance Spectroscopy* 2018, **106-107**:1–25 <https://doi.org/10.1016/j.pnmrs.2018.03.002>.
311. Hamelryck T, Manderick B: **PDB file parser and structure class implemented in Python.** *Bioinformatics* 2003, **19**:2308–2310 <https://doi.org/10.1093/bioinformatics/btg299>.
312. Smith M: **Determining the Native and Nonnative Effects of Myristoylation on the Folding and Switching of Hisactophilin (Thesis).** 2012, [no volume].
313. Molday RS, Englander SW, Kallen RG: **Primary structure effects on peptide group hydrogen exchange.** *Biochemistry* 1972, **11**:150–158 <https://doi.org/10.1021/bi00752a003>.
314. Woodward CK, Hilton BD: **Hydrogen Exchange Kinetics and Internal Motions in Proteins and Nucleic Acids.** *Annual Review of Biophysics and Bioengineering* 1979, **8**:99–127 <https://doi.org/10.1146/annurev.bb.08.060179.000531>.
315. Li R, Woodward C: **The hydrogen exchange core and protein folding.** *Protein Science* 1999, **8**:1571–1590 <https://doi.org/10.1110/ps.8.8.1571>.
316. Skinner JJ, Lim WK, Bédard S, Black BE, Englander SW: **Protein dynamics viewed by hydrogen exchange.** *Protein Science* 2012, **21**:996–1005 <https://doi.org/10.1002/pro.2081>.
317. Skinner JJ, Lim WK, Bédard S, Black BE, Englander SW: **Protein hydrogen exchange: Testing current models.** *Protein Science* 2012, **21**:987–995 <https://doi.org/10.1002/pro.2082>.
318. Ferraro DM, Lazo ND, Robertson AD: **EX1 Hydrogen Exchange and Protein Folding †.** *Biochemistry* 2004, **43**:587–594 <https://doi.org/10.1021/bi035943y>.

319. Persson F, Halle B: **How amide hydrogens exchange in native proteins.** *Proceedings of the National Academy of Sciences* 2015, **112**:10383–10388 <https://doi.org/10.1073/pnas.1506079112>.
320. Farber P, Darmawan H, Sprules T, Mittermaier A: **Analyzing Protein Folding Cooperativity by Differential Scanning Calorimetry and NMR Spectroscopy.** *Journal of the American Chemical Society* 2010, **132**:6214–6222 <https://doi.org/10.1021/ja100815a>.
321. McConnell HM: **Reaction Rates by Nuclear Magnetic Resonance.** *The Journal of Chemical Physics* 1958, **28**:430–431 <https://doi.org/10.1063/1.1744152>.
322. Loria JP, Berlow RB, Watt ED: **Characterization of enzyme motions by solution NMR relaxation dispersion.** *Acc Chem Res* 2008, **41**:214–221 <https://doi.org/10.1021/ar700132n>.
323. Loria JP, Rance M, Palmer AG: **A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy [13].** *Journal of the American Chemical Society* 1999, **121**:2331–2332 <https://doi.org/10.1021/ja983961a>.
324. Carver JP, Richards RE: **A general two-site solution for the chemical exchange produced dependence of T2 upon the carr-Purcell pulse separation.** *Journal of Magnetic Resonance (1969)* 1972, **6**:89–105 [https://doi.org/10.1016/0022-2364\(72\)90090-X](https://doi.org/10.1016/0022-2364(72)90090-X).
325. Ili AGP, Palmer AG: **Chemical exchange in biomacromolecules: Past, present, and future.** *Journal of Magnetic Resonance* 2014, **241**:3–17 <https://doi.org/10.1016/j.jmr.2014.01.008>.
326. Akke M, Brueschweiler R, Palmer III AG: **NMR order parameters and free energy: an analytical approach and its application to cooperative Ca2+ binding by Calbindin D.** *Journal of the American Chemical Society* 1993, **115**:9832–9833 <https://doi.org/10.1021/ja00074a073>.
327. Palmer III AG: **Dynamic properties of proteins from NMR spectroscopy.** *Current opinion in biotechnology* 1993, **4**:385–91 [https://doi.org/10.1016/0958-1669\(93\)90002-E](https://doi.org/10.1016/0958-1669(93)90002-E).
328. Zeiske T, Stafford KA, Friesner RA, Palmer AG: **Starting-structure dependence of nanosecond timescale inter-substate transitions and reproducibility of MD-derived order parameters.** *Proteins: Structure, Function and Bioinformatics* 2013, **81**:499–509 <https://doi.org/10.1002/prot.24209>.
329. Lu M, Hou G, Zhang H, Suiter CL, Ahn J, Byeon I-JL, Perilla JR, Langmead CJ, Hung I, Gor'kov PL, et al.: **Dynamic allostery governs cyclophilin A–HIV capsid interplay.** *Proceedings of the National Academy of Sciences* 2015, **112**:14617–14622 <https://doi.org/10.1073/pnas.1516920112>.
330. Lee L, Sykes BD: **Use of lanthanide-induced nuclear magnetic resonance shifts for determination of protein structure in solution: EF calcium binding site of carp parvalbumin.** *Biochemistry* 1983, **22**:4366–4373 <https://doi.org/10.1021/bi00288a004>.
331. Lee L, Sykes BD: **Strategies for the uses of lanthanide NMR shift probes in the determination of protein structure in solution. Application to the EF calcium binding site of carp parvalbumin.** *Biophysical journal* 1980, **32**:193–210 [https://doi.org/10.1016/S0006-3495\(80\)84933-2](https://doi.org/10.1016/S0006-3495(80)84933-2).
332. Solomon I: **Relaxation processes in a system of two spins.** *Physical Review* 1955, **99**:559–565 <https://doi.org/10.1103/PhysRev.99.559>.
333. Keizers PHJ, Ubink M: **Paramagnetic tagging for protein structure and dynamics analysis.** *Progress in Nuclear Magnetic Resonance Spectroscopy* 2011, **58**:88–96 <https://doi.org/10.1016/j.pnmrs.2010.08.001>.
334. Gottstein D, Reckel S, Dötsch V, Güntert P: **Requirements on paramagnetic relaxation enhancement data for membrane protein structure determination by NMR.** *Structure* 2012, **20**:1019–1027 <https://doi.org/10.1016/j.str.2012.03.010>.
335. Hosszu LLP, Craven CJ, Spencer J, Parker MJ, Clarke AR, Kelly M, Waltho JP: **Is the structure of the N-domain of phosphoglycerate kinase affected by isolation from the intact molecule?** *Biochemistry* 1997, **36**:333–340 <https://doi.org/10.1021/bi961784p>.
336. Williamson MP: **Many Residues in Cytochrome c Populate Alternative States Under Equilibrium Conditions.** *Proteins: Structure, Function and Genetics* 2003, **53**:731–739 <https://doi.org/10.1002/prot.10464>.
337. Krishna Mohan PM, Barve M, Chatterjee A, Ghosh-Roy A, Hosur RV: **NMR comparison of the native energy landscapes of DLC8 dimer and monomer.** *Biophysical Chemistry* 2008, **134**:10–19 <https://doi.org/10.1016/j.bpc.2007.12.010>.
338. Liang J, Jaffrey SR, Guo W, Snyder SH, Clardy J: **Structure of the PIN/LC8 dimer with a bound peptide.** *Nature Structural Biology* 1999, **6**:735–740 <https://doi.org/10.1038/11501>.
339. Chandra K, Sharma Y, Chary KVR: **Characterization of low-energy excited states in the native state ensemble of non-myristoylated and myristoylated neuronal calcium sensor-1.** *Biochimica et Biophysica Acta - Proteins and Proteomics* 2011, **1814**:334–344 <https://doi.org/10.1016/j.bbapap.2010.10.007>.
340. Srivastava AK, Chary KVR: **Conformational heterogeneity and dynamics in a $\beta\gamma$ -crystallin from *Hahella chejuensis*.** *Biophysical Chemistry* 2011, **157**:7–15 <https://doi.org/10.1016/j.bpc.2011.04.001>.
341. Hass M a S, Jensen MR, Led JJ: **Probing electric fields in proteins in solution by NMR spectroscopy.** *Proteins* 2008, **72**:333–43 <https://doi.org/10.1002/prot.21929>.
342. Nielsen JE: *Analyzing Enzymatic pH Activity Profiles and Protein Titration Curves Using Structure-Based pKa Calculations and Titration Curve Fitting.* Elsevier Inc. 2009.
343. Tomlinson JH, Green VL, Baker PJ, Williamson MP: **Structural origins of pH-dependent chemical shifts in the B1 domain of protein G.** *Proteins: Structure, Function and Bioinformatics* 2010, **78**:3000–3016 <https://doi.org/10.1002/prot.22825>.
344. Haruyama H, Qian YQ, Wüthrich K: **Static and Transient Hydrogen-Bonding Interactions in Recombinant Desulfatohirudin Studied by ^1H Nuclear Magnetic Resonance Measurements of Amide Proton Exchange Rates and pH-Dependent Chemical Shifts.** *Biochemistry* 1989, **28**:4312–4317 <https://doi.org/10.1021/bi00436a028>.
345. Sakurai K, Goto Y: **Principal component analysis of the pH-dependent conformational transitions of bovine beta-lactoglobulin monitored by heteronuclear NMR.** *Proceedings of the National Academy of Sciences* 2007, **104**:15346–15351 <https://doi.org/10.1073/pnas.0702112104>.
346. Tomlinson JH, Jeremy Craven C, Williamson MP, Pandya MJ: **Dimerization of protein G B1 domain at low pH. A conformational switch caused by loss or a single hydrogen bond.** *Proteins: Structure, Function and Bioinformatics* 2010, **78**:1652–1661 <https://doi.org/10.1002/prot.22683>.

347. Shrager RI, Cohen JS, Heller SR, Sachs DH, Schechter AN: **Mathematical Models for Interacting Groups in Nuclear Magnetic Resonance Titration Curves.** *Biochemistry* 1972, **11**:541–547 <https://doi.org/10.1021/bi00754a010>.
348. Joshi MD, Hedberg A, McIntosh LP: **Complete measurement of the pKa values of the carboxyl and imidazole groups in Bacillus circulans xylanase.** *Protein Science* 2008, **6**:2667–2670 <https://doi.org/10.1002/pro.5560061224>.
349. Platzer G, Okon M, McIntosh LP: **pH-dependent random coil 1H, 13C, and 15N chemical shifts of the ionizable amino acids: a guide for protein pKa measurements.** *Journal of Biomolecular NMR* 2014, **60**:109–129 <https://doi.org/10.1007/s10858-014-9862-y>.
350. Wüthrich K, Wagner G: **Nuclear magnetic resonance of labile protons in the basic pancreatic trypsin inhibitor.** *Journal of Molecular Biology* 1979, **130**:1–18 [https://doi.org/10.1016/0022-2836\(79\)90548-5](https://doi.org/10.1016/0022-2836(79)90548-5).
351. Liepinsh E, Otting G: **Proton exchange rates from amino acid side chains—implications for image contrast.** *Magnetic resonance in medicine* 1996, **35**:30–42.
352. Zheng G, Price WS: **Solvent signal suppression in NMR.** *Progress in Nuclear Magnetic Resonance Spectroscopy* 2010, **56**:267–288 <https://doi.org/10.1016/j.pnmrs.2010.01.001>.
353. LeMaster DM: **Structural determinants of the catalytic reactivity of the buried cysteine of Escherichia coli thioredoxin.** *Biochemistry* 1996, **35**:14876–14881 <https://doi.org/10.1021/bi961607o>.
354. Nielsen JE: **Analyzing Protein NMR pH-Titration Curves.** In *Annual reports in computational chemistry.* Elsevier B.V. 2008:89–106.
355. Farrell D, Miranda ES, Webb H, Georgi N, Crowley PB, McIntosh LP, Nielsen JE: **Titration_DB: Storage and analysis of NMR-monitored protein pH titration curves.** *Proteins: Structure, Function, and Bioinformatics* 2010, **78**:843–857 <https://doi.org/10.1002/prot.22611>.
356. Markley JL: **Observation of histidine residues in proteins by nuclear magnetic resonance spectroscopy.** *Accounts of Chemical Research* 1975, **8**:70–80 <https://doi.org/10.1021/ar50086a004>.
357. R Core Team: *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing; 2018.
358. Hanakam F, Eckerskorn C, Lottspeich F, Muller-Taubenberger A, Schafer W, Gerisch G: **The pH-sensitive actin-binding protein hisactophilin of Dictyostelium exists in two isoforms which both are myristoylated and distributed between plasma membrane and cytoplasm.** 1995, **270**:596–602 <https://doi.org/10.1074/jbc.270.2.596>.
359. Scheel J, Ziegelbauer K, Kupke T, Humbel BM, Noegel AA, Gerisch G, Schleicher M: **Hisactophilin, a histidine-rich actin-binding protein from Dictyostelium discoideum.** *Journal of Biological Chemistry* 1989, **264**:2832–2839.
360. Bhattacharya S, Lecomte JT: **Temperature dependence of histidine ionization constants in myoglobin.** *Biophysical Journal* 1997, **73**:3241–3256 [https://doi.org/10.1016/S0006-3495\(97\)78349-8](https://doi.org/10.1016/S0006-3495(97)78349-8).
361. Di Russo NV, Martí MA, Roitberg AE: **Underlying Thermodynamics of pH-Dependent Allostery.** *Journal of Physical Chemistry B* 2014, **118**:12818–12826 <https://doi.org/10.1021/jp507971v>.
362. Peck MT, Ortega G, De Luca-Johnson JN, Schlessman JL, Robinson AC, García-Moreno E B: **Local Backbone Flexibility as a Determinant of the Apparent pKa Values of Buried Ionizable Groups in Proteins.** *Biochemistry* 2017, **56**:5338–5346 <https://doi.org/10.1021/acs.biochem.7b00678>.
363. Tanford C: **Ionization-linked Changes in Protein Conformation. I. Theory.** *Journal of the American Chemical Society* 1961, **83**:1628–1634 <https://doi.org/10.1021/ja01468a021>.
364. Di Russo NV, Estrin DA, Martí MA, Roitberg AE: **pH-Dependent Conformational Changes in Proteins and Their Effect on Experimental pKas: The Case of Nitrophorin 4.** *PLoS Computational Biology* 2012, **8** <https://doi.org/10.1371/journal.pcbi.1002761>.
365. Whitten ST, Garcia-Moreno E. B, Hilser VJ: **Local conformational fluctuations can modulate the coupling between proton binding and global structural transitions in proteins.** *Proceedings of the National Academy of Sciences* 2005, **102**:4282–4287 <https://doi.org/10.1073/pnas.0407499102>.
366. Karp DA, Gittis AG, Stahley MR, Fitch CA, Stites WE, García-Moreno E. B: **High Apparent Dielectric Constant Inside a Protein Reflects Structural Reorganization Coupled to the Ionization of an Internal Asp.** *Biophysical Journal* 2007, **92**:2041–2053 <https://doi.org/10.1529/biophysj.106.090266>.
367. González W, Riedelsberger J, Morales-Navarro SE, Caballero J, Alzate-Morales JH, González-Nilo FD, Dreyer I: **The pH sensor of the plant K⁺-uptake channel KAT1 is built from a sensory cloud rather than from single key amino acids.** *The Biochemical journal* 2012, **442**:57–63 <https://doi.org/10.1042/BJ20111498>.
368. Li Z, Blissard GW: **Multiple Nucleopolyhedrovirus GP64 Protein: Roles of Histidine Residues in Triggering Membrane Fusion and Fusion Pore Expansion.** *Journal of Virology* 2011, **85**:12492–12504 <https://doi.org/10.1128/JVI.05153-11>.
369. Yuan Y, Tam MF, Simplaceanu V, Ho C: **New Look at Hemoglobin Allostery.** *Chemical Reviews* 2015, **115**:1702–1724 <https://doi.org/10.1021/cr500495x>.
370. Søndergaard CR, McIntosh LP, Pollastri G, Nielsen JE: **Determination of Electrostatic Interaction Energies and Protonation State Populations in Enzyme Active Sites.** *Journal of Molecular Biology* 2008, **376**:269–287 <https://doi.org/10.1016/j.jmb.2007.09.070>.
371. Fersht AR, Leatherbarrow RJ, Wells TNC: **Quantitative analysis of structure-activity relationships in engineered proteins by linear free-energy relationships.** *Nature* 1986, **322**:284–286 <https://doi.org/10.1038/322284a0>.
372. McIntosh LP, Naito D, Baturin SJ, Okon M, Joshi MD, Nielsen JE: **Dissecting electrostatic interactions in Bacillus circulans xylanase through NMR-monitored pH titrations.** *Journal of Biomolecular NMR* 2011, **51**:5–19 <https://doi.org/10.1007/s10858-011-9537-x>.
373. Yang W, Jones LM, Isley L, Ye Y, Lee HW, Wilkins A, Liu Z ren, Hellinga HW, Malchow R, Ghazi M, et al.: **Rational design of a calcium-binding protein.** *Journal of the American Chemical Society* 2003, **125**:6165–6171 <https://doi.org/10.1021/ja034724x>.
374. Kuramitsu S, Ikeda K, Hamaguchi K: **Participation of the catalytic carboxyls, Asp 52 and Glu 35, and Asp 101 in the binding of substrate analogues to hen lysozyme.** *Journal of Biochemistry* 1975, **77**:291–301 <https://doi.org/10.1093/oxfordjournals.jbchem.a130726>.
375. Murtaugh ML, Fanning SW, Sharma TM, Terry AM, Horn JR: **A combinatorial histidine scanning library approach to engineer highly pH-dependent protein switches.** *Protein science: a publication of the Protein Society* 2011, **20**:1619–31 <https://doi.org/10.1002/pro.696>.

376. Watanabe H, Matsumaru H, Ooishi A, Feng YW, Odahara T, Suto K, Honda S: **Optimizing pH response of affinity between protein G and IgG Fc. How electrostatic modulations affect protein-protein interactions.** *Journal of Biological Chemistry* 2009, **284**:12373–12383 <https://doi.org/10.1074/jbc.M809236200>.
377. Zhuang T, Chisholm C, Chen M, Tamm LK: **NMR-based conformational ensembles explain pH-gated opening and closing of OmpG channel.** *Journal of the American Chemical Society* 2013, **135**:15101–15113 <https://doi.org/10.1021/ja408206e>.
378. Jiang L, Althoff EA, Clemente FR, Doyle L, Röthlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF, et al.: **De novo computational design of retro-aldol enzymes.** *Science* 2008, **319**:1387–1391 <https://doi.org/10.1126/science.1152692>.
379. Khandogin J, Brooks CL: **Constant pH molecular dynamics with proton tautomerism.** *Biophysical Journal* 2005, **89**:141–157 <https://doi.org/10.1529/biophysj.105.061341>.
380. Mongan J, Case DA, McCammon JA: **Constant pH molecular dynamics in generalized Born implicit solvent.** *Journal of Computational Chemistry* 2004, **25**:2038–2048 <https://doi.org/10.1002/jcc.20139>.
381. Bundi A, Wüthrich K: **Use of amide1H-nmr titration shifts for studies of polypeptide conformation.** *Biopolymers* 1979, **18**:299–311 <https://doi.org/10.1002/bip.1979.360180207>.
382. Martí MA, Estrin DA, Roitberg AE: **Molecular basis for the pH dependent structural transition of nitrophenol 4.** *Journal of Physical Chemistry B* 2009, **113**:2135–2142 <https://doi.org/10.1021/jp808055e>.
383. Tunnicliffe RB, Waby JL, Williams RJ, Williamson MP: **An experimental investigation of conformational fluctuations in proteins G and L.** *Structure* 2005, **13**:1677–1684 <https://doi.org/10.1016/j.str.2005.08.006>.
384. Waudby CA, Ramos A, Cabrita LD, Christodoulou J: **Two-Dimensional NMR Lineshape Analysis.** *Scientific Reports* 2016, **6**:24826 <https://doi.org/10.1038/srep24826>.
385. Banelli T, Vuano M, Fogolari F, Fusiello A, Esposito G, Corazza A: **Automation of peak-tracking analysis of stepwise perturbed NMR spectra.** *Journal of Biomolecular NMR* 2017, **67**:121–134 <https://doi.org/10.1007/s10858-017-0088-7>.
386. Lee W, Tonelli M, Markley JL: **NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy.** *Bioinformatics* 2015, **31**:1325–1327 <https://doi.org/10.1093/bioinformatics/btu830>.
387. Marsh JA, Forman-Kay JD: **Structure and Disorder in an Unfolded State under Nondenaturing Conditions from Ensemble Models Consistent with a Large Number of Experimental Restraints.** *Journal of Molecular Biology* 2009, **391**:359–374 <https://doi.org/10.1016/j.jmb.2009.06.001>.
388. Hyberts SG, Takeuchi K, Wagner G: **Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data.** *Journal of the American Chemical Society* 2010, **132**:2145–2147 <https://doi.org/10.1021/ja908004w>.
389. Hyberts SG, Milbradt AG, Wagner AB, Arthanari H, Wagner G: **Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling.** *Journal of Biomolecular NMR* 2012, **52**:315–327 <https://doi.org/10.1007/s10858-012-9611-z>.
390. Marsh JA, Singh VK, Jia Z, Forman-Kay JD: **Sensitivity of secondary structure propensities to sequence differences between α - and γ -synuclein: Implications for fibrillation.** *Protein Science* 2006, **15**:2795–2804 <https://doi.org/10.1110/ps.062465306>.
391. Bieri M, D'Auvergne EJ, Gooley PR: **relaxGUI: a new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and μ s motion of proteins.** *Journal of Biomolecular NMR* 2011, **50**:147–155 <https://doi.org/10.1007/s10858-011-9509-1>.
392. Hansen DF, Vallurupalli P, Lundström P, Neudecker P, Kay LE: **Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: How well can we do?** *Journal of the American Chemical Society* 2008, **130**:2667–2675 <https://doi.org/10.1021/ja078337p>.
393. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A: **NMRPipe: A multidimensional spectral processing system based on UNIX pipes.** *Journal of Biomolecular NMR* 1995, **6**:277–293 <https://doi.org/10.1007/BF00197809>.
394. Khairutdinov BI, Ermakova EA, Yushpovych YM, Bessolicina EK, Tarasova NB, Toporkova YY, Kovaleva V, Zuev YF, Nesmelova IV: **NMR structure, conformational dynamics, and biological activity of Ps Def1 defensin from Pinus sylvestris.** *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 2017, **1865**:1085–1094 <https://doi.org/10.1016/j.bbapap.2017.05.012>.
395. Kumar A, Srivastava S, Hosur RV: **NMR Characterization of the Energy Landscape of SUMO-1 in the Native-state Ensemble.** *Journal of Molecular Biology* 2007, **367**:1480–1493 <https://doi.org/10.1016/j.jmb.2007.01.035>.
396. Jarque CM, Bera AK: **Efficient tests for normality, homoscedasticity and serial independence of regression residuals.** *Economics Letters* 1980, **6**:255–259 [https://doi.org/10.1016/0165-1765\(80\)90024-5](https://doi.org/10.1016/0165-1765(80)90024-5).
397. Massey FJ: **The Kolmogorov-Smirnov Test for Goodness of Fit.** *Journal of the American Statistical Association* 1951, **46**:68–78 <https://doi.org/10.1080/01621459.1951.10500769>.
398. Lilliefors HW: **On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown.** *Journal of the American Statistical Association* 1967, **62**:399–402 <https://doi.org/10.1080/01621459.1967.10482916>.
399. Shapiro SS, Wilk MB: **An Analysis of Variance Test for Normality (Complete Samples).** *Biometrika* 1965, **52**:591 <https://doi.org/10.2307/2333709>.
400. Trapletti A, Hornik K: **tseries: Time Series Analysis and Computational Finance** No Title. 2018, [no volume].
401. Fox J, Weisberg S: *An R Companion to Applied Regression.* Sage; 2011.
402. Tukey JW: **One Degree of Freedom for Non-Additivity.** *Biometrics* 1949, **5**:232 <https://doi.org/10.2307/3001938>.
403. Efron B, Tibshirani RJ: *An Introduction to the Bootstrap.* Chapman & Hall; 1993.
404. Singh K, Xie M: **Bootstrap: A Statistical Method.** *International Encyclopedia of Education* 2010, <https://doi.org/www.stat.rutgers.edu/~mxie/rcpapers/bootstrap.pdf>.
405. Chernick MR: *Bootstrap Methods: A Guide for Practitioners and Researchers.* 2007.
406. Fox J, Weisberg S: **Appendix: Bootstrapping Regression Models in R.** *An R Companion to Applied Regression* 2017, [no volume].
407. Davison, A. C. & Hinkley DV: *Bootstrap Methods and Their Applications.* Cambridge University Press; 1997.
408. Canty A, Ripley B: **boot: Bootstrap R (S-Plus) Functions.** 2017, [no volume].

409. Broom A, Gosavi S, Meiering EM: **Protein unfolding rates correlate as strongly as folding rates with native structure.** *Protein science : a publication of the Protein Society* 2014, **24**:1–35 <https://doi.org/10.1002/pro.2606>.
410. Banci L, Barbieri L, Bertini I, Cantini F, Luchinat E: **In-cell NMR in E. coli to monitor maturation steps of hSOD1.** *PLoS ONE* 2011, **6** <https://doi.org/10.1371/journal.pone.0023561>.
411. Banci L, Bertini I, Cantini F, Kozyreva T, Massagni C, Palumaa P, Rubino JT, Zovo K: **Human superoxide dismutase 1 (hSOD1) maturation through interaction with human copper chaperone for SOD1 (hCCS).** *Proceedings of the National Academy of Sciences* 2012, **109**:13555–13560 <https://doi.org/10.1073/pnas.1207493109>.
412. Zetterström P, Graffmo KS, Andersen PM, Brännström T, Marklund SL: **Composition of Soluble Misfolded Superoxide Dismutase-1 in Murine Models of Amyotrophic Lateral Sclerosis.** *NeuroMolecular Medicine* 2013, **15**:147–158 <https://doi.org/10.1007/s12017-012-8204-z>.
413. Sheng Y, Chattopadhyay M, Whitelegge J, Selverstone Valentine J: **SOD1 Aggregation and ALS: Role of Metal-lation States and Disulfide Status.** *Current Topics in Medicinal Chemistry* 2012, **12**:2560–2572 <https://doi.org/10.2174/1568026611212220010>.
414. Furukawa Y, O'Halloran TV: **Amyotrophic Lateral Sclerosis Mutations Have the Greatest Destabilizing Effect on the Apo- and Reduced Form of SOD1, Leading to Unfolding and Oxidative Aggregation.** *Journal of Biological Chemistry* 2005, **280**:17266–17274 <https://doi.org/10.1074/jbc.M500482200>.
415. Broom HR, Rumpfheldt JAO, Meiering EM: **Many roads lead to Rome? Multiple modes of Cu,Zn superoxide dismutase destabilization, misfolding and aggregation in amyotrophic lateral sclerosis.** *Essays In Biochemistry* 2014, **56**:149–165 <https://doi.org/10.1042/bse0560149>.
416. Deol H: **Characterizing the Dimer Interface and Thermal Unfolding of Reduced, Zinc Bound Superoxide Dismutase-1.** 2018, [no volume].
417. Borch J, Hamann T: **The nanodisc: a novel tool for membrane protein studies.** *Biological chemistry* 2009, **390**:805–14 <https://doi.org/10.1515/BC.2009.091>.
418. Dalal K, Nguyen N, Alami M, Tan J, Moraes TF, Lee WC, Maurus R, Sligar SS, Brayer GD, Duong F: **Structure, binding, and activity of Syd, a SecY-interacting protein.** *The Journal of biological chemistry* 2009, **284**:7897–902 <https://doi.org/10.1074/jbc.M808305200>.
419. Andrushchenko VV, Aarabi MH, Nguyen LT, Prenner EJ, Vogel HJ: **Thermodynamics of the interactions of tryptophan-rich cathelicidin antimicrobial peptides with model and natural membranes.** *Biochimica et biophysica acta* 2008, **1778**:1004–14 <https://doi.org/10.1016/j.bbame.2007.12.022>.
420. Bartels T, Ahlstrom LS, Leftin A, Kamp F, Haass C, Brown MF, Beyer K: **The N-terminus of the intrinsically disordered protein alpha-synuclein triggers membrane binding and helix folding.** *Biophysical Journal* 2010, **99**:2116–2124 <https://doi.org/10.1016/j.bpj.2010.06.035>.
421. Otzen DE: **Folding of DsbB in mixed micelles: a kinetic analysis of the stability of a bacterial membrane protein.** *Journal of molecular biology* 2003, **330**:641–649 [https://doi.org/10.1016/S0022-2836\(03\)00624-7](https://doi.org/10.1016/S0022-2836(03)00624-7).
422. Otzen DE: **Mapping the folding pathway of the transmembrane protein DsbB by protein engineering.** *Protein engineering, design & selection : PEDS* 2011, **24**:139–49 <https://doi.org/10.1093/protein/gzq079>.
423. Warschawski DE, Arnold AA, Beaugrand M, Gravel A, Chartrand É, Marcotte I: **Choosing membrane mimetics for NMR structural studies of transmembrane proteins.** *Biochimica et Biophysica Acta - Biomembranes* 2011, **1808**:1957–1974 <https://doi.org/10.1016/j.bbame.2011.03.016>.
424. Dodevski I, Nucci NV, Valentine KG, Sidhu GK, O'Brien ES, Pardi A, Wand AJ: **Optimized Reverse Micelle Surfactant System for High-Resolution NMR Spectroscopy of Encapsulated Proteins and Nucleic Acids Dissolved in Low Viscosity Fluids.** *Journal of the American Chemical Society* 2014, **136**:3465–3474 <https://doi.org/10.1021/ja410716w>.
425. Nucci NV, Valentine KG, Wand AJ: **High-resolution NMR spectroscopy of encapsulated proteins dissolved in low-viscosity fluids.** *Journal of Magnetic Resonance* 2014, **241**:137–147 <https://doi.org/10.1016/j.jmr.2013.10.006>.
426. Nucci NV, Pometun MS, Wand AJ: **Site-resolved measurement of water-protein interactions by solution NMR.** *Nature structural & molecular biology* 2011, **18**:245–9 <https://doi.org/10.1038/nsmb.1955>.
427. Valentine KG, Peterson RW, Saad JS, Summers MF, Xu X, Ames JB, Wand AJ: **Reverse Micelle Encapsulation of Membrane-Anchored Proteins for Solution NMR Studies.** *Structure* 2010, **18**:9–16 <https://doi.org/10.1016/j.str.2009.11.010>.
428. Maibaum L, Dinner AR, Chandler D: **Micelle Formation and the Hydrophobic Effect †.** *The Journal of Physical Chemistry B* 2004, **108**:6778–6781 <https://doi.org/10.1021/jp037487t>.
429. Nucci NV, Marques BS, Bédard S, Dogan J, Gledhill JM, Moorman VR, Peterson RW, Valentine KG, Wand AL, Wand AJ: **Optimization of NMR spectroscopy of encapsulated proteins dissolved in low viscosity fluids.** *Journal of Biomolecular NMR* 2011, **50**:421–430 <https://doi.org/10.1007/s10858-011-9528-y>.
430. Fuglestad B, Gupta K, Wand AJ, Sharp KA: **Characterization of Cetyltrimethylammonium Bromide/Hexanol Reverse Micelles by Experimentally Benchmarked Molecular Dynamics Simulations.** *Langmuir* 2016, **32**:1674–1684 <https://doi.org/10.1021/acs.langmuir.5b03981>.
431. Palazzo G, Carbone L, Colafemmina G, Angelico R, Ceglie A, Giustini M: **The role of the cosurfactant in the CTAB/water/n-pentanol/n-hexane system: Pentanol effect on the phase equilibria and mesophase structure.** *Phys. Chem. Chem. Phys.* 2004, **6**:1423–1429 <https://doi.org/10.1039/B314509N>.
432. Wand AJ, Ehrhardt MR, Flynn PF: **High-resolution NMR of encapsulated proteins dissolved in low-viscosity fluids.** *Proceedings of the National Academy of Sciences of the United States of America* 1998, **95**:15299–15302. <https://doi.org/10.1073/pnas.95.26.15299>.
433. Ehrhardt MR, Flynn PF, Wand AJ: **Preparation of encapsulated proteins dissolved in low viscosity fluids.** *Journal of Biomolecular NMR* 1999, **14**:75–78 <https://doi.org/10.1023/A:1008354507250>.
434. Thompson KF, Gierasch LM: **Conformation of a peptide solubilize in a reversed micelle water pool.** *Journal of the American Chemical Society* 1984, **106**:3648–3652 <https://doi.org/10.1021/ja00324a040>.
435. De Marco A, Zetta L, Menegatti E, Luisi PL: **1H-NMR of reverse micelles. II: Conformational studies of peptides and proteins in the AOT/water/isooctane system.** *Journal of Biochemical and Biophysical Methods* 1986, **12**:335–347 [https://doi.org/10.1016/0165-022X\(86\)90071-0](https://doi.org/10.1016/0165-022X(86)90071-0).
436. Peterson RW, Pometun MS, Shi Z, Wand AJ: **Novel surfactant mixtures for NMR spectroscopy of encapsulated proteins dissolved in low-viscosity fluids.** *Protein Science* 2005, **14**:2919–2921 <https://doi.org/10.1110/ps.051535405>.

437. Marques BS, Nucci NV, Dodevski I, Wang KWC, Athanasoula EA, Jorge C, Wand AJ: **Measurement and Control of pH in the Aqueous Interior of Reverse Micelles.** *The Journal of Physical Chemistry B* 2014, **118**:2020–2031 <https://doi.org/10.1021/jp4103349>.
438. McIntosh LP, Dahlquist FW: **Biosynthetic Incorporation of 15N and 13C for Assignment and Interpretation of Nuclear Magnetic Resonance Spectra of Proteins.** *Quarterly Reviews of Biophysics* 1990, **23**:1 <https://doi.org/10.1017/S0033583500005400>.
439. Cavanagh J, Fairbrother W, Palmer III A, Rance M, Skelton N: *Protein NMR Principles And Practice.* Elsevier; 2007.
440. Lefebvre BG, Liu W, Peterson RW, Valentine KG, Wand AJ: **NMR spectroscopy of proteins encapsulated in a positively charged surfactant.** *Journal of Magnetic Resonance* 2005, **175**:158–162 <https://doi.org/10.1016/j.jmr.2005.03.008>.
441. Peterson RW, Nucci NV, Wand AJ: **Modification of encapsulation pressure of reverse micelles in liquid ethane.** *Journal of magnetic resonance (San Diego, Calif. : 1997)* 2011, **212**:229–33 <https://doi.org/10.1016/j.jmr.2011.06.010>.
442. Morrison SS, Loncar A, Dawson JF: **Non-polymerizing long-pitch actin dimers that interact with myosin.** *Archives of Biochemistry and Biophysics* 2010, **501**:188–194 <https://doi.org/10.1016/j.abb.2010.06.003>.
443. Perieteanu AA, Visschedyk DD, Merrill AR, Dawson JF: **ADP-Ribosylation of Cross-Linked Actin Generates Barbed-End Polymerization-Deficient F-Actin Oligomers.** *Biochemistry* 2010, **49**:8944–8954 <https://doi.org/10.1021/bi1008062>.
444. Goddard TD, Kneller DG: **SPARKY 3.** [date unknown], [no volume].
445. Drori I: **Fast Minimization by Iterative Thresholding for Multidimensional NMR Spectroscopy.** *EURASIP Journal on Advances in Signal Processing* 2007, **2007**:020248 <https://doi.org/10.1155/2007/20248>.
446. Donoho DL: **De-noising by soft-thresholding.** *IEEE Transactions on Information Theory* 1995, **41**:613–627 <https://doi.org/10.1109/18.382009>.
447. Behrisch a, Dietrich C, Noegel a a, Schleicher M, Sackmann E: **The actin-binding protein hisactophilin binds in vitro to partially charged membranes and mediates actin coupling to membranes.** *Biochemistry* 1995, **34**:15182–90.
448. Flynn PF, Mattiello DL, Hill HDW, Wand AJ: **Optimal Use of Cryogenic Probe Technology in NMR Studies of Proteins.** *Journal of the American Chemical Society* 2000, **122**:4823–4824 <https://doi.org/10.1021/ja993743x>.
449. Grzesiek S, Bax A: **The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements.** *Journal of the American Chemical Society* 1993, **115**:12593–12594 <https://doi.org/10.1021/ja00079a052>.
450. Tingey JM, Fulton JL, Matson DW, Smith RD: **Micellar and bicontinuous microemulsions formed in both near-critical and supercritical propane with didodecyldimethylammonium bromide and water.** *Journal of Physical Chemistry* 1991, **95**:1445–1448 <https://doi.org/10.1021/j100156a075>.
451. Senske M, Smith AE, Pielak GJ: **Protein Stability in Reverse Micelles.** *Angewandte Chemie International Edition* 2016, **55**:3586–3589 <https://doi.org/10.1002/anie.201508981>.
452. O'Brien ES, Nucci NV, Fuglestad B, Tommos C, Wand AJ: **Defining the Apoptotic Trigger.** *Journal of Biological Chemistry* 2015, **290**:30879–30887 <https://doi.org/10.1074/jbc.M115.689406>.
453. Valentine KG, Pometun MS, Kielec JM, Baigelman RE, Staub JK, Owens KL, Wand AJ: **Magnetic Susceptibility-Induced Alignment of Proteins in Reverse Micelles.** *Journal of the American Chemical Society* 2006, **128**:15930–15931 <https://doi.org/10.1021/ja061438n>.
454. Bouvignies G, Markwick PRL, Blackledge M: **Simultaneous definition of high resolution protein structure and backbone conformational dynamics using NMR residual dipolar couplings.** *ChemPhysChem* 2007, **8**:1901–1909 <https://doi.org/10.1002/cphc.200700353>.
455. Frederick KK, Marlow MS, Valentine KG, Wand AJ: **Conformational entropy in molecular recognition by proteins.** *Nature* 2007, **448**:325–329 <https://doi.org/10.1038/nature05959>.

Appendices

Appendix A - Temperature Dependence and Curvature

Appendix A Table 1: Table collecting all the measured linear temperature dependences (i.e temperature coefficients, $\Delta\delta/\Delta T$) in ppb/K and extent of curvature for residues in wild-type (WT), LLL, and I85L—each myristoylated (myr) and non-myristoylated (NM) at pH 6.2 and pH 7.7. Temperature coefficients were determined from a least-squares linear regression of chemical shift vs temperature (over a range of 10 °C - 50 °C) and presented as ppb/K. Curvature A-values are given only for those residues that passed the bootstrapped model selection test with a critical value of 0.01. The residuals following linear regression were fit to a second order polynomial, and the a-value is the coefficient ahead of the squared term.

L45	4.55E-05	-6.62	1.93E-05	-5.45		-4.73		-4.23		-5.77		-5.52	1.47E-05	-4.90		-4.59		-5.61		-5.16		-5.14		-4.38	
K46		0.38		-7.00		-3.08		-2.23		1.90E-05	-0.26	-1.13		-2.84		-2.98		-5.34				-2.80		-2.27	
T47	4.59E-05	-2.94		-2.29	-1.40E-05	-1.77		-2.60		1.36E-05	-1.89	4.18E-05	-1.57		-2.37	1.98E-05	-1.72	2.09E-05	-1.39	1.73E-05	-2.20	1.77E-05	-2.08	2.22E-05	-2.53
H48			-8.73E-06	-7.10							-4.41	-5.73													
C49	2.22E-05	-2.85		-3.18									2.41E-05	-5.95	1.01E-05	-3.19									
G50			-1.40											-1.99		-2.00									
K51	4.47E-05	-5.92		-4.27	1.92E-05	2.32								-4.12	7.06E-06	-2.40	3.21	4.88E-06	-1.47						
Y52		-1.65		-2.36		-2.48		-2.19		7.38E-06	-2.37	-2.93		-2.50		-2.98		-2.22		-2.75		-2.19		-2.28	
L53		-2.69		-1.85		-2.97		-2.19		-2.88		-2.38	8.46E-06	-3.10		-2.51		-3.37		-2.71		-2.37		-1.93	
S54		-1.77		-2.00		-2.00	-1.36E-05	-1.69				-2.88	-2.65		-2.41	7.93E-06	-2.07		-1.56		-1.76		-1.74		-1.77
I55		0.93		-0.17	5.93E-05	-2.12			2.43E-05	-0.23	0.35		-0.22		-0.24		-0.19		-0.59						
G56	-1.78E-05	-2.02		-3.69	1.71E-05	-2.73	-3.25E-05	-4.10		-3.44		-3.38		-3.19		-3.64	-2.83E-05	-3.28		-3.91	-1.82E-05	-3.27		-3.79	
D57																									
H58										2.17E-05															
K59	-1.14E-05	-1.81	-2.46E-05	-1.98	-6.65E-05	-5.03	-3.78E-05	-2.54	4.49E-05	-2.42	1.34E-05	-1.97	1.54E-05	-2.39	2.12E-05	-2.26	-2.85E-05	-2.51	-1.99E-05	-2.02	-4.65E-05	-3.41	-2.02E-05	-2.15	
Q60	7.20E-05	-8.63	6.89E-05	-9.67		-1.67				-9.54		-10.31	4.66E-05	-9.40	5.15E-05	-9.80	5.13E-05	-9.32	4.28E-05	-10.22	5.92E-05	-9.33	3.38E-05	-9.93	
V61	9.74E-06	-8.17		-8.01		-8.16		-8.06		-7.73		-8.15		-8.20		-8.04		-8.30		-8.26					
Y62		-1.97		-1.69		-6.91		-2.21		-2.55		-1.92		-2.94		-1.97		-2.40		-1.90		-2.60		-1.92	
L63	4.40E-05	-6.93	2.08E-05	-7.01	-2.36E-05	-0.42	-1.97E-05	-0.55					4.33E-05	-4.56	1.43E-05	-0.82		-5.39		-7.09		-6.32		-8.06	
S64	-1.92E-05	-0.29	-1.12E-05	-0.66		-1.96					-1.27			-1.35	1.13E-04	-7.23	-1.88E-05	-0.63		-1.09					
H65								-7.18	8.24E-06																
H66	7.98E-06	-3.22	-4.56E-06	-3.17						-3.20		-2.95		-3.23		-3.13		-3.58	1.66E-07	-3.17		-7.83		-3.04	
L67		-7.31	-6.07E-05	-5.19		####	-1.21E-04	-9.39	3.44E-05	-6.80		-6.62		-7.06	1.51E-05	-6.85							2.34E-05	-1.89	
G68	-5.19E-05	-0.22	-2.01E-05	-0.80	-4.40E-05	-1.06	-2.91E-05	-0.99		-2.61	1.37E-05	-2.34	3.44E-05	-1.63		-1.61	-4.31E-05	-0.64	-1.12E-05	-0.67	-4.38E-05	-1.73	-1.44E-05	-1.29	
G69	3.05E-05	-8.93	4.05E-05	-8.47		-1.53				-7.02		-7.65		-8.31		-8.17		-8.91		-8.46		-7.91		-8.30	
D70		-0.81		-1.87		3.28		-5.63	2.10E-05	-4.28		-1.89		-4.92		-4.56	1.24E-05	-5.09		-4.77		-5.13	-1.13E-05	-2.30	
H71	5.97E-05	-7.56	3.26E-05	-4.90	-5.64E-05	-8.63		-8.30	4.34E-05	-6.50	2.79E-05	-4.55		-7.57	3.74E-05	-5.18					4.35E-05	-7.16	2.79E-05	-5.01	
S72	-6.70E-05	####	4.29E-05	-4.01		-3.03	-1.33E-05	-3.19	2.77E-05	-4.32	3.08E-05	-3.66	4.29E-05	-5.76	2.67E-05	-4.08		-5.81	3.56E-05	-3.91					
L73	-2.74E-05	0.76	-1.72E-05	0.23	-3.21E-05	-1.04	-4.20E-05	-0.28		-0.12	2.29E-05	0.20	2.30E-05	0.28	2.41E-05	0.37	-3.64E-05	0.71		0.72	-3.17E-05	0.01	-2.00E-05	0.58	
F74	-3.63E-05	-1.15	-1.76E-05	-1.59	3.92E-05	-1.80		-2.46	2.14E-05	-2.18		-2.68		-2.73		-2.57		-3.99	-3.23E-05	-4.19		-1.38		-1.82	
H75		-2.23		-2.83						-1.91		-2.81	2.26E-05	-2.13		-2.73	1.84E-05	-0.62		-1.52	1.43E-05	-1.77	1.83E-05	-2.49	
L76	5.33E-05	-8.93	-5.46E-05	-8.12	3.37E-05	-6.76	-2.15E-05	-7.55		-6.78				-7.83		-8.58						-7.90		-8.90	
E77		-6.81	2.94E-05	-5.28		-1.13		-0.68		-3.11	2.02E-05	-3.22	3.28E-05	-2.16		-7.61									
H78		-6.25	-4.40E-05	-7.60						-7.27		-7.90		-7.22	5.02E-06	-8.48		-6.74		-2.43		-3.67		-9.43	
H79		-1.98		-3.99		-8.80				-0.79		-0.99		-0.99		-0.67		-5.43		-0.50		-1.14	1.23E-05	-0.70	
G80																									
G81																									
K82	1.49E-05	-3.90				-3.71		-4.39		-3.91				-4.03		-4.57		-4.16		-4.76	6.54E-06	-3.84		-4.60	
V83	-2.51E-05	-4.56		-5.21		-4.57		-4.94		-5.72		-6.19		-5.61		-5.87		-2.77	-2.08E-05	-3.04		-5.51		-5.93	
S84	-1.60E-05	-1.80		-1.64		-0.01	3.84E-05	0.50		-1.19	2.90E-05	-1.55		-1.84		-1.66					7.03E-06	-1.27	2.09E-05	-1.00	
I85	-1.12E-04	1.43		-1.05	-1.65E-05	-5.26	-2.72E-05	-4.85		-3.33	-2.19	1.39E-05	-4.37		-3.90		-1.24		-0.40		-1.93E-05	-4.51	2.05E-05	-4.46	
K86		-0.82		-0.74		-0.73		-5.52		-8.97		-4.75		-0.98	2.68E-05	-1.11		-3.44	-8.22E-05	-4.90			2.34E-05	-1.89	
G87	-1.31E-05	-4.38		-5.02		-1.06		-4.75		-4.27		-4.34	2.21E-05	-2.82		-4.05		-3.90	-2.34E-05	-4.31	1.40E-05	-3.46		-4.49	
H88			2.73E-05	-4.81		-0.14			8.22E-05	-4.35		-4.67		-6.78		-0.78					4.19E-05	-5.50			
H89		-7.77		-3.65		-8.48		-2.12		-5.33	6.97E-05	-2.79	3.36E-05	-7.96	5.68E-05	-3.72	-9.41E-05	-5.65	-5.11E-05	-3.50					
H90		-2.37	2.15E-05	-1.46		-5.79			2.43E-05	-1.75	8.67E-06	-0.81	1.94E-05	-1.86		-1.19		-1.90	1.82E-05	-1.87	1.70E-05	-1.72	1.03E-05	-1.12	
H91	3.60E-05	0.05	2.92E-05	-0.22		-5.64	1.15		3.05E-05	1.43	1.98E-05	-0.25	3.19E-05	0.82	1.62E-05	-0.36	3.40E-05	1.12		-0.52	3.91E-05	1.26	2.39E-05	-0.56	
Y92	-1.65E-05	-3.65	-7.96E-06	-2.89	-3.80E-05	-5.65	-4.80E-05	-3.85	1.92E-05	-3.81	1.79E-05	-2.23	3.04E-05	-4.45	2.00E-05	-3.08					-4.41E-05	-5.05	-1.46E-05	-3.17	
I93	-1.54E-04	6.00	-7.29E-05	3.02		-5.29		-2.44		-1.58	2.29E-05	-1.82		-1.42	2.11E-05	-1.39		-4.16		-3.89	-1.12E-05	-1.85	-2.76E-05	-2.21	
S94	-7.71E-06	-1.30	-4.52E-06	-1.85		-4.38		-1.74		-5.25	1.58E-05	-1.51	1.66E-05	-1.50	1.19E-05	-1.72		-1.13		-1.72	1.04E-05	-1.33		-1.78	

A95	-2.08	-1.91			-3.83	-2.72	-2.50	-3.47	1.42E-05	-2.89	-2.77	-8.20			-3.42						
D96	-2.03	-5.95E-06	-1.85	-1.95	-1.10E-05	-1.90	-1.73	1.27E-05	-1.78	1.01E-05	-1.74	1.03E-05	-1.72	-2.01	-2.06	-1.12E-05	-1.85	-8.55E-06	-1.77		
H97				-2.41					-1.78				-2.01	-2.06							
H98	0.22	0.05	0.36	0.05	1.02E-05	0.29		0.63		0.42	0.47		0.32	0.61					0.32		
G99	1.78E-05	-3.45	1.20E-05	-3.33	-3.78	-3.79	-3.57	1.35E-05	-3.44	1.03E-05	-3.86	9.54E-06	-3.64	1.22E-05	-3.55	1.21E-05	-3.29	1.21E-05	-3.75	1.10E-05	-3.55
H100	-7.28E-06	0.01		-2.69			-6.31	-7.16	-4.14	3.21E-05	-4.22	-2.20			-3.03E-05	-2.88				-3.32	
V101	2.12E-05	-9.17	2.89E-05	-10.05	-8.74	-9.43	-9.04	-9.73	-8.80		-9.37	-9.14	-10.23		-8.91					-9.85	
S102	-2.88	-2.94	-8.17	-3.41	-2.87	-3.05	-3.12	-3.12	-2.78	-2.86	-3.22	-3.04								-3.04	
T103	7.51E-06	-3.60	-5.54E-06	-3.87	-2.69	-3.03	-2.60	-2.94	-2.81	-3.18	-2.92	-3.51	6.92E-06	-2.68	-3.14					-3.14	
K104	-2.93E-05	0.10	-0.10	-2.36E-05	-0.42		-1.95	-1.89	-1.96	-1.95	-0.63	-1.07	-1.42E-05	-1.03	-1.29					-1.29	
E105	-1.38	-0.41	-1.32	-0.63	-0.88	1.58E-05	-1.51	-1.15	2.34E-05	-1.00	-0.85	-0.79	-8.18E-06	-0.99	-2.30E-05	-0.55				-0.55	
H106	-1.59E-05	-6.47	-5.15E-05	-6.92	-3.42	-3.16E-05	-5.33	1.73E-05	-4.62	3.54E-05	-5.89	-4.33	2.33E-05	-5.56	-5.13	-4.35E-05	-6.31	-4.50	-3.81E-05	-6.18	
H107	3.36E-05	-7.84					8.84E-05	-6.16	1.02E-05	-8.02	1.45E-05	-6.09	1.42E-05	-7.37	1.88E-05	-6.56	-8.66	-6.46		-7.76	
D108	-1.29E-04	-0.94	-8.29E-05	-1.11			-8.70	8.31E-05	-7.15	8.32E-05	-8.60	8.08E-05	-6.69	-7.59E-05	-4.84	-7.25E-05	-1.81	-9.14E-05	-8.59	-6.35E-05	-5.90
H109							1.43E-05														
D110	-1.72E-05	-0.14	-2.39E-05	-0.12	-1.24E-05	-0.71	-2.41E-05	-0.38	-0.17	1.04	0.01	0.30	-1.09E-05	-0.20	-1.43E-05	0.41	-1.04E-05	-0.33	-1.66E-05	0.32	
T111	-1.24E-05	0.93	0.74	-0.11	-0.12	0.76	1.15	0.57	0.51	0.40	0.57	0.31	0.34							0.34	
T112	3.29E-05	-3.13	-3.42	-1.91	-1.74	-1.96	-2.60	6.04E-06	-1.45	-1.78	-2.74	-3.91	1.21E-05	-1.50	-1.85					-1.85	
F113	-9.89E-06	-4.42	-4.26	-7.70E-06	-4.44	-3.88	8.11E-06	-3.52	-2.97	6.05E-06	-4.76	-4.57	-2.41E-05	-3.24	-2.02E-05	-3.00	-4.32			-4.11	
E114	-2.52	-1.03E-05	-2.94	-1.61	-1.83	-2.65			8.93E-06	-1.98	-2.24	-2.25E-05	-0.85	-2.60E-05	-1.18	-2.21	-2.30			-2.30	
E115	1.56E-05	-6.29	-6.27	-1.33	-1.69E-05	-7.16	1.19E-05	-6.05	-6.62	-6.29	-7.61	-6.35	-6.61	-6.63	-6.56					-6.56	
I116	-3.05	-2.97	-4.07	-9.66E-06	-3.94	-3.39	-3.72	-3.57	-3.42	-3.43	-3.46	-3.60	-3.26							-3.26	
I117	1.54E-05	-6.94	6.89E-06	-7.52	-6.85	-7.61	1.05E-05	-6.78	-7.53	-7.28	-7.70	9.80E-06	-7.04	-7.64	1.17E-05	-6.90	1.11E-05	-7.48		-7.48	
I118	-6.23E-05	-0.57	-3.93E-05	-1.21	-2.29E-05	-2.07	-3.50E-05	-1.87	-1.34	-1.12	-1.30	2.29E-05	-0.80	-2.81E-05	-0.25	-2.23E-05	0.15	-2.05E-05	-1.62	-2.29E-05	-1.30

Appendix B - Apparent pKas

Appendix B Table 1: Apparent pKas calculated from the NMR-monitored pH dependence of amide proton chemical shifts for wild-type (collected by Dr. Martin Smith), LLL, and I85L. Data were fit to one of two models; a Henderson-Hasselbach model that assumes one site of protonation, or one that is modified with a hill coefficient (See section: Analyzing NMR pH Titrations) to account for other proximate ionization events. Fitting to the models was executed in Origin software, and the reported errors are the standard error of fitting. Residues without values means that there was no assignment or that the fit did not satisfactorily converge.

Residue	WT Myr pKa		WT NM pKa		LLL Myr pKa		LLL NM pKa		I85L Myr pKa		I85L NM pKa	
	WT Myr pKa	Error	WT NM pKa	Error	LLL Myr pKa	Error	LLL NM pKa	Error	I85L Myr pKa	Error	I85L NM pKa	Error
G2	6.78	0.08	6.86	0.11	6.60	0.06						
N3	6.88	0.10	7.00	0.26	7.04	0.02			7.07	0.05		
R4	7.13	0.04	8.15	0.47	7.16	0.08						
A5	6.84	0.14	7.36	0.32	7.27	0.04						
F6			8.04	0.12					8.08	0.10	6.40	0.10
K7	6.06	0.17										
S8	6.58	0.14	6.80	0.09	6.08	0.04	6.34	0.03	5.81	0.06	6.09	0.08
H9	6.74	0.08										
H10	6.26	0.11	7.22	0.21	6.01	0.04	6.28	0.03	6.02	0.02	5.97	0.04
G11			8.48	0.16	5.89	0.14	6.49	0.13	6.18	0.07	5.75	0.27
H12			6.47	0.34					6.02	0.07	6.15	0.12
F13	6.70	0.07	7.10	0.04								
L14	6.54	0.07	7.58	0.15			7.37	0.02				
S15	7.02	0.05	7.40	0.27			7.26	0.02				
A16	6.49	0.06	7.48	0.06			6.83	0.06	7.33	0.02		
E17	6.33	0.15	6.78	0.07	6.48	0.05						
G18												
E19					6.37	0.03					6.94	0.17
A20	6.58	0.07	7.20	0.27	6.13	0.07	6.76	0.06	6.34	0.02	6.46	0.03
V21	7.06	0.07	6.52	0.10	6.41	0.06	6.73	0.04	6.34	0.01	6.34	0.04
K22	6.62	0.06	7.02	0.10	7.50	0.02			7.37	0.01	7.92	0.05
T23	6.66	0.07	7.52	0.16	6.68	0.06	6.89	0.02	6.53	0.03	6.66	0.03
H24	6.51	0.06			6.27	0.03	7.20	0.07	6.36	0.03		
H25												
G26												
H27												
H28												
D29												
H30												
H31												
T32												
H33	6.50		7.36	0.06			6.91	0.01	6.41	0.02		
F34	6.99	0.08	8.11	0.16								
H35												
V36	7.11	0.06	6.70	0.16			7.44	0.04	7.09	0.02		
E37	6.33	0.37	7.21	0.16			7.39	0.04			6.33	0.17
N38	6.09	0.16			6.54	0.04	6.62	0.03				
H39	5.98	0.23			6.22	0.13						
G40												
G41												
K42			6.81	0.11			6.36	0.07	6.23	0.04	6.48	0.03
V43	7.02	0.08	7.21	0.12	6.48	0.10	6.45	0.03			6.27	0.05

A44	6.56	0.14			6.72	0.07				6.23	0.13	
L45	7.86	0.14	6.45	0.09				6.08	0.03	6.15	0.05	
K46	6.94	0.06	7.53	0.07						6.88	0.07	
T47	6.26	0.10	5.91	0.45	6.03	0.13			6.21	0.08		
H48	7.48	0.08			8.07	0.07						
C49	8.59	0.68										
G50	7.47	0.05										
K51	7.78	0.10	7.03	0.06	7.12	0.07	7.30	0.03				
Y52	10.80	0.43	7.18	0.08								
L53	7.27	0.09	7.00	0.07	5.82	0.06	6.62	0.08	5.55	0.07	5.98	0.04
S54	6.48	0.12	6.17	0.10	6.48	0.02	6.60	0.04	6.27	0.02	6.28	0.02
I55	6.50		5.78	1.22			6.39	0.08				
G56	5.75	0.21	6.83	0.06	6.27	0.02	6.36	0.02	6.33	0.02	6.51	0.01
D57												
H58												
K59	6.50		7.58	0.12	6.20	0.03	6.31	0.03	6.12	0.03	6.29	0.02
Q60	6.50		7.04	0.19	6.26	0.04	6.33	0.03	6.23	0.03	6.28	0.03
V61	5.46	1.42			6.05	0.13						
Y62	7.79	0.20	9.05	0.31	6.14	0.04	5.78	0.12				
L63	7.70	0.08					6.03	0.05			6.25	0.03
S64	4.84	2.14					6.07	0.05	6.07	0.14		
H65												
H66	5.65	0.26	7.67	0.09	5.97	0.05	6.02	0.02	5.76	0.02	5.89	0.05
L67	6.50		7.01	0.12	7.51	0.11						
H68	5.56	0.41	6.40	0.05	6.04	0.02	6.12	0.02	6.03	0.02	6.02	0.05
G69			shape		5.97	0.04	5.91	0.06	6.09	0.02	6.39	0.02
D70	7.13	0.08			6.43	0.04	6.36	0.02	6.20	0.03	6.69	0.02
H71	6.50		7.32	0.11	6.14	0.03					6.72	0.20
S72	6.25	0.13	6.29	0.33	6.00	0.03	6.08	0.02	6.05	0.03		
L73	5.85	0.19	6.74	0.06	6.17	0.02	6.37	0.01	6.20	0.02	6.38	0.02
F74	7.18	0.11	7.12	0.07	6.50	0.04	6.70	0.03	6.72	0.03	6.53	0.05
H75	6.63	0.05	7.34	0.09	6.46	0.01	6.53	0.02	6.30	0.02	6.36	0.01
L76	7.08	0.05	6.45	0.07	6.81	0.07	6.95	0.03				
E77	6.45	0.09	7.26	0.18	6.37	0.03						
H78	6.88	0.06			6.85	0.02	7.05	0.03	6.46	0.03	6.91	0.03
H79	6.78	0.06			6.76	0.04	7.00	0.03	6.88	0.03	6.97	0.03
G80												
G81												
K82							7.57	0.02			7.71	0.03
V83	7.26	0.11	7.03	0.12	6.26	0.04	5.91	0.11	6.13	0.08	6.36	0.04
S84	6.98	0.04	7.11	0.06	6.74	0.03	6.84	0.03			6.70	0.03
I85	6.53	0.11	6.75	0.05	6.60	0.01	6.72	0.04	6.15	0.01	6.64	0.05
K86			6.24	0.40	6.56	0.02	6.72	0.02			6.49	0.02
G87	6.60	0.10	5.50	0.45	6.46	0.01	5.96	0.07	6.06	0.04	5.74	0.11
H88					6.44	0.04	6.99	0.07	7.34	0.11		

H89	6.30	0.07			6.64	0.03	6.70	0.02	6.49	0.01	6.57	0.01
H90	6.67	0.09	7.60	0.12	7.46	0.06	7.41	0.04	6.74	0.03		
H91	6.20	0.10	6.94	0.72	6.54	0.01	6.59	0.01	6.32	0.01	6.46	0.02
Y92	6.42	0.11	6.84	0.06			6.56	0.02			6.51	0.02
I93	6.50		7.86	0.13			6.44	0.08			6.40	0.05
S94	6.50	0.18										
A95			5.55	0.29			7.36	0.05	6.95	0.03		
D96	7.20	0.08					7.51	0.02	7.43	0.11		
H97												
H98	6.84	0.05	6.63	0.06	6.63	0.03	6.74	0.03	6.55	0.03	6.59	0.04
G99	6.75	0.14	7.48	0.13	6.27	0.04	6.33	0.03	6.06	0.03	6.29	0.03
H100					7.16	0.02			5.96	0.13		
V101	7.06	0.05	7.02	0.05	7.06	0.02	7.09	0.03	6.99	0.02	6.95	0.03
S102	7.97	0.24			6.13	0.09	6.23	0.19				
T103	7.68	0.07	7.82	0.28	6.60	0.04	6.63	0.03	6.58	0.02		
K104			8.67	0.28	6.70	0.03	6.86	0.27				
E105	6.83	0.21					6.63	0.03				
H106	6.89	0.06	7.16	0.07	7.24	0.03	7.20	0.02	7.06	0.02	7.15	0.01
H107	6.79	0.06	6.85	0.10	6.47	0.03	6.02	0.10			6.19	0.03
D108	5.68	0.21			6.32	0.02	6.49	0.03	6.27	0.01	6.33	0.01
H109												
D110	6.07	0.12	6.52	0.09	6.35	0.01	6.38	0.02	6.19	0.01	6.25	0.01
T111	5.90	0.40	6.21	0.18	6.24	0.03	6.16	0.02	6.02	0.02	6.02	0.02
T112	5.37	0.52	6.53	0.10	6.46	0.04	6.50	0.07	6.34	0.02	6.66	0.03
F113					6.23	0.11	6.60	0.07	6.57	0.04	6.50	0.03
E114	7.17	0.14	7.42	0.16	6.26	0.05	6.44	0.07			6.30	0.03
E115	7.52	0.07					7.51	0.04				
I116	7.80	0.06	8.44	0.20			6.96	0.07	8.26	0.12	7.12	0.10
I117	7.43	8.44	6.28	0.29	7.37	0.05	7.02	0.05	7.56	0.08		
I118	6.75	0.07			7.17	0.06	6.61	0.03	6.28	0.02	6.86	0.06

Appendix C

Pulse sequences.

1D Excitation Sculpting

```
;zgesgp
;avance-version (02/05/31)
;1D sequence
;water suppression using excitation sculpting with gradients
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
; Series A 112 275-279 (1995)
```

```
prosol relations=<triple>
```

```
#include <Avance.incl>
#include <Grad.incl>
```

```
"p2=p1*2"
```

```
"d12=20u"
```

```
1 ze
2 30m
d12 p1:f1 BLKGRAD
d1
p1 ph1
```

```
50u UNBLKGRAD
p16:gp1
d16 p10:f1
(p12:sp1 ph2:r):f1
```

```
4u
d12 pl1 : f1

p2 ph3

4u
p16 : gp1
d16
50u
p16 : gp2
d16 pl0 : f1
(p12 : sp1 ph4 : r) : f1
4u
d12 pl1 : f1
```

```
p2 ph5
```

```
4u
p16 : gp2
d16
```

```
go=2 ph31
30m mc #0 to 2 F0(zd)
4u BLKGRAD
exit
```

```
ph1=0
ph2=0 1
ph3=2 3
ph4=0 0 1 1
ph5=2 2 3 3
ph31=0 2 2 0
```

```
;pl0 : 120dB
;pl1 : f1 channel – power level for pulse (default)
;sp1 : f1 channel – shaped pulse 180 degree
```

```
;p1 : f1 channel – 90 degree high power pulse
;p2 : f1 channel – 180 degree high power pulse
;p12: f1 channel – 180 degree shaped pulse (Squa100.1000) [2 msec]
;p16: homospoil/gradient pulse
;d1 : relaxation delay; 1–5 * T1
;d12: delay for power switching [20 usec]
;d16: delay for homospoil/gradient recovery
;NS: 8 * n, total number of scans: NS * TD0
;DS: 4
```

```
;use gradient ratio: gp 1 : gp 2
;                   31 : 11
```

```
;for z-only gradients:
;gpz1: 31%
;gpz2: 11%
```

```
;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
```

```
;$Id: zgesgp,v 1.4 2002/06/12 09:05:21 ber Exp $
```

1H, 15H HSQC

```
;invifp3gpsi
;avance-version
;2D H-1/X correlation via double inept transfer
;  using sensitivity improvement
;phase sensitive using Echo/Antiecho-TPPI gradient selection
;with decoupling during acquisition
;using f3 - channel
;using flip-back pulse
;A.G. Palmer III , J. Cavanagh, P.E. Wright & M. Rance, J. Magn.
;  Reson. 93, 151-170 (1991)
;L.E. Kay, P. Keifer & T. Saarinen, J. Am. Chem. Soc. 114,
;  10663-5 (1992)
;J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky,
;  S.J. Glaser, O.W. Sorensen & C. Griesinger, J. Biomol. NMR 4,
;  301-306 (1994)
;S. Grzesiek & A. Bax, J. Am. Chem. Soc. 115, 12593-12594 (1993)
;modified by jgaspar to include delay prior to aquisition
;
```

```
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>
```

```
"p2=p1*2"
```

```
"p22=p21*2"
```

```
"d0=3u"
```

```
"d11=30m"
```

```
"d26=1s/(cnst4*4)"
```

"DELTA=p16+d16+p2+d0*2"

"DELTA1=p16+d16+8u"

"CEN_HN1=(p21-p1)/2"

"CEN_HN2=(p22-p2)/2"

"I3=(td1/2)"

1 ze

d31

d11 p16 : f3

2 d1 do : f3

6m

3 d11

18m

4 (p1 ph1)

d26 p13 : f3

(CEN_HN2 p2 ph2) (p22 ph6) : f3

d26 UNBLKGRAD

(p1 ph2)

4u p10 : f1

(p11:sp1 ph1:r) : f1

4u

p16:gp1

d16 p11 : f1

(p21 ph3) : f3

d0

p2 ph7

d0

p16:gp2*EA

d16

(p22 ph4) : f3

DELTA

(CEN_HN1 p1 ph1) (p21 ph4) : f3

d24

```
(CEN_HN2 p2 ph1) (p22 ph1):f3
d24
(CEN_HN1 p1 ph2) (p21 ph5):f3
d26
(CEN_HN2 p2 ph1) (p22 ph1):f3
d26
(p1 ph1)
DELTA1
(p2 ph1)
4u
p16:gp3
d16 pl16:f3
4u BLKGRAD
go=2 ph31 cpd3:f3
d1 do:f3 wr #0 if #0 zd
3m ip5 igrad EA
3m ip5
lo to 3 times 2
d11 id0
3m ip3
3m ip3
3m ip6
3m ip6
3m ip31
3m ip31
lo to 4 times l3
exit
```

```
ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=1 1 3 3
ph6=0
ph7=0 0 2 2
ph31=2 0 0 2
```

```

;p10 : 120dB
;p11 : f1 channel – power level for pulse (default)
;p13 : f3 channel – power level for pulse (default)
;p116: f3 channel – power level for CPD/BB decoupling
;sp1: f1 channel – shaped pulse 90 degree
;p1 : f1 channel – 90 degree high power pulse
;p2 : f1 channel – 180 degree high power pulse
;p11: f1 channel – 90 degree shaped pulse
;p16: homospoil/gradient pulse [1 msec]
;p21: f3 channel – 90 degree high power pulse
;p22: f3 channel – 180 degree high power pulse
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1–5 * T1
;d11: delay for disk I/O [30 msec]
;d16: delay for homospoil/gradient recovery
;d24: 1/(4J)YH for YH
; 1/(8J)YH for all multiplicities
;d26: 1/(4J(YH))
;d31: delay between experiments in multizg
;cnst4: = J(YH)
;l3: loop for phase sensitive 2D using E/A method : l3 = td1/2
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 1 * n
;DS: >= 16
;td1: number of experiments
;MC2: echo–antiecho
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel – 90 degree pulse for decoupling sequence

;use gradient ratio: gp 1 : gp 2 : gp 3
; 50 : 80 : 20.1 for C–13
; 50 : 80 : 8.1 for N–15

;for z–only gradients:
;gpz1: 50%

```

```
;gpz2: 80%  
;gpz3: 20.1% for C-13, 8.1% for N-15
```

```
;use gradient files:
```

```
;gpnam1: SINE.100
```

```
;gpnam2: SINE.100
```

```
;gpnam3: SINE.100
```

15N-Edited NOESY - HSQC

```
;noesyhsqcfpf3gpsi3d
;avance-version (03/06/18)
;NOESY-HSQC
;3D sequence with
;  homonuclear correlation via dipolar coupling
;  dipolar coupling may be due to noe or chemical exchange
;  H-1/X correlation via double inept transfer
;    using sensitivity improvement
;phase sensitive (t1)
;phase sensitive using Echo/Antiecho-TPPI gradient selection (t2)
;with decoupling during acquisition
;using flip-back pulse
;using f3 - channel
;(use parameterset NOESYHSQCFFP3GPSI3D)
;
;O. Zhang, L.E. Kay, J.P. Olivier & J.D. Forman-Kay,
;  J. Biomol. NMR 4, 845 - 858 (1994)
;A.G. Palmer III, J. Cavanagh, P.E. Wright & M. Rance, J. Magn.
;  Reson. 93, 151-170 (1991)
;L.E. Kay, P. Keifer & T. Saarinen, J. Am. Chem. Soc. 114,
;  10663-5 (1992)
;J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky,
;  S.J. Glaser, O.W. Sorensen & C. Griesinger, J. Biomol. NMR 4,
;  301-306 (1994)
```

```
prosol relations=<triple>
```

```
#include <Avance.incl>
```

```
#include <Grad.incl>
```

```
#include <Delay.incl>
```

"p2=p1*2"

"p22=p21*2"

"d0=3u"

"d10=3u"

"d11=30m"

"d12=20u"

"d13=4u"

"d26=1s/(cnst4*4)"

"p19=300u"

"DELTA1=d13+p19+d16+4u"

"DELTA3=d24-p16-d16"

"DELTA4=d26-p16-d16"

ifdef LABEL_CN

"DELTA=p16+d16+larger(p2,p8)+d10*2"

"DELTA2=larger(p8,p22)+d0*2"

else

"DELTA=p16+d16+p2+d10*2"

"DELTA2=p22+d0*2"

endif /*LABEL_CN*/

"TAU=d8-p16-d16"

aqseq 321

```

1 ze
d11 pl16:f3
2 d11 do:f3
3 d12
d1
(p1 ph8)
DELTA2 pl3:f3
(p2 ph9)
d0

#   ifdef LABEL_CN
(center (p8:sp13 ph1):f2 (p22 ph1):f3 )
#   else
(p22 ph1):f3
#   endif /*LABEL_CN*/

d0
(p1 ph10)
TAU UNBLKGRAD
p16:gp1
d16
(p1 ph11)
d26
(center (p2 ph1) (p22 ph6):f3 )
d26
(p1 ph2)

4u pl0:f1
(p11:sp1 ph12):f1
4u
p16:gp2
d16 pl1:f1

(p21 ph3):f3
d10

#   ifdef LABEL_CN
(center (p2 ph7) (p8:sp13 ph1):f2 )

```

```

#   else
(p2 ph7)
#   endif /*LABEL_CN*/

d10
p16:gp3*EA
d16
(p22 ph4):f3
DELTA
(center (p1 ph1) (p21 ph4):f3 )
p16:gp4
d16
DELTA3
(center (p2 ph1) (p22 ph1):f3 )
DELTA3
p16:gp4
d16
(center (p1 ph2) (p21 ph5):f3 )
p16:gp5
d16
DELTA4
(center (p2 ph1) (p22 ph1):f3 )
DELTA4
p16:gp5
d16
(p1 ph1)
DELTA1
(p2 ph1)
d13
p19:gp6
d16 pl16:f3
4u BLKGRAD
go=2 ph31 cpd3:f3
d11 do:f3 mc #0 to 2
F1PH(rd10 & ip8 & ip9 , id0)
F2EA(igrad EA & ip5*2, id10 & ip3*2 & ip6*2 & ip31*2)
exit

```

```

ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=1 1 3 3
ph6=0
ph7=0 0 2 2
ph8=0 0 0 0 2 2 2 2
ph9=1 1 1 1 3 3 3 3
ph10=0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
ph11=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph12=2 2 2 2 2 2 2 2 0 0 0 0 0 0 0 0
ph31=0 2 2 0 2 0 0 2 2 0 0 2 0 2 2 0
2 0 0 2 0 2 2 0 0 2 2 0 2 0 0 2

```

```

;p10 : 120dB
;p11 : f1 channel – power level for pulse (default)
;p13 : f3 channel – power level for pulse (default)
;p16: f3 channel – power level for CPD/BB decoupling
;sp1 : f1 channel – shaped pulse 90 degree
;sp13: f2 channel – shaped pulse 180 degree (adiabatic)
;p1 : f1 channel – 90 degree high power pulse
;p2 : f1 channel – 180 degree high power pulse
;p8 : f2 channel – 180 degree shaped pulse for inversion (adiabatic)
;p11: f1 channel – 90 degree shaped pulse
;p16: homospoil/gradient pulse [1 msec]
;p19: homospoil/gradient pulse [300 usec]
;p21: f3 channel – 90 degree high power pulse
;p22: f3 channel – 180 degree high power pulse
;d0 : incremented delay (F1 in 3D) [3 usec]
;d1 : relaxation delay; 1–5 * T1
;d8 : mixing time
;d10: incremented delay (F2 in 3D) [3 usec]
;d11: delay for disk I/O [30 msec]
;d12: delay for power switching [20 usec]

```

```

;d13: short delay [4 usec]
;d16: delay for homospoil/gradient recovery
;d24: 1/(4J)YH for YH
;      1/(8J)YH for all multiplicities
;d26: 1/(4J(YH))
;cnst4: = J(YH)
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;in10: 1/(2 * SW(X)) = DW(X)
;nd10: 2
;NS: 8 * n
;DS: >= 16
;td1: number of experiments in F1
;td2: number of experiments in F2
;FnMODE: States-TPPI (or TPPI) in F1
;FnMODE: echo-antiecho in F2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;use gradient ratio:   gp 1 : gp 2 : gp 3 : gp 4 : gp 5 : gp 6
;                      30 : -50 : 80 : 11 : 5 : 8.1*x

;   with x = p16/p19

;for z-only gradients:
;gpz1: 30%
;gpz2: -50%
;gpz3: 80%
;gpz4: 11%
;gpz5: 5%
;gpz6: 27% (8.1% *p16(=1 ms)/p19(=300 us))

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
;gpnam3: SINE.100
;gpnam4: SINE.100

```

```
;gpnam5: SINE.100  
;gpnam6: SINE.50
```

```
;preproce$\beta$or-flags-start  
;LABEL_CN: for C-13 and N-15 labeled samples start experiment with  
;          option -DLABEL_CN (eda: ZGOPTNS)  
;preproce$\beta$or-flags-end
```

```
;$Id: noesyhsqcfpf3gpsi3d,v 1.1.2.3 2003/06/18 15:27:56 ber Exp $
```

15N-Edited TOCSY - HSQC

```
;dipsiif3gpsi3d
;avance-version (00/10/05)
;TOCSY-HSQC
;3D sequence with
;  homonuclear Hartman-Hahn transfer using DIPSI2 sequence
;    for mixing
;  H-1/X correlation via double inept transfer
;    using sensitivity improvement
;phase sensitive (t1)
;phase sensitive using Echo/Antiecho-TPPI gradient selection (t2)
;using trim pulses in inept transfer
;using f3 - channel
;A.G. Palmer III , J. Cavanagh, P.E. Wright & M. Rance, J. Magn.
;  Reson. 93, 151-170 (1991)
;L.E. Kay, P. Keifer & T. Saarinen, J. Am. Chem. Soc. 114,
;  10663-5 (1992)
;J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky,
;  S.J. Glaser, O.W. Sorensen & C. Griesinger, J. Biomol. NMR 4,
;  301-306 (1994)
```

```
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>
```

```
"p2=p1*2"
```

```
"p22=p21*2"
```

```
"d0=3u"
```

```
"d10=3u"
```

"d11=30m"

"d12=20u"

"d13=4u"

"d26=1s/(const4*4)"

"DELTA=p16+d16+p2+d10*2"

"DELTA1=d13+p16+d16+4u"

"DELTA2=p22+d0*2"

"CEN_HN1=(p21-p1)/2"

"CEN_HN2=(p22-p2)/2"

"FACTOR1=(d9/(p6*115.112))/2+0.5"

"l1=FACTOR1*2"

aqseq 321

1 ze

d11 pl16:f3

2 d11 do:f3

3 d12 pl9:f1

d1 cw:f1 ph29

4u do:f1

d12 pl1:f1

(p1 ph8)

DELTA2 pl3:f3

(p2 ph9)

d0

(p22 ph1):f3
d0

(p1 ph10)
d20 pl10:f1

;begin DIPSI2
7 p6*3.556 ph23
p6*4.556 ph25
p6*3.222 ph23
p6*3.167 ph25
p6*0.333 ph23
p6*2.722 ph25
p6*4.167 ph23
p6*2.944 ph25
p6*4.111 ph23

p6*3.556 ph25
p6*4.556 ph23
p6*3.222 ph25
p6*3.167 ph23
p6*0.333 ph25
p6*2.722 ph23
p6*4.167 ph25
p6*2.944 ph23
p6*4.111 ph25

p6*3.556 ph25
p6*4.556 ph23
p6*3.222 ph25
p6*3.167 ph23
p6*0.333 ph25
p6*2.722 ph23
p6*4.167 ph25
p6*2.944 ph23
p6*4.111 ph25

p6*3.556 ph23

p6*4.556 ph25
p6*3.222 ph23
p6*3.167 ph25
p6*0.333 ph23
p6*2.722 ph25
p6*4.167 ph23
p6*2.944 ph25
p6*4.111 ph23
lo to 7 times l1
;end DIPS12

d21 p11 :f1
(p1 ph11)

d26
(CEN_HN2 p2 ph1) (p22 ph6):f3
d26 UNBLKGRAD
p28 ph1
d13
(p1 ph2)
3u
p16:gp1
d16
(p21 ph3):f3
d10
p2 ph7
d10
p16:gp2*EA
d16
(p22 ph4):f3
DELTA
(CEN_HN1 p1 ph1) (p21 ph4):f3
d24
(CEN_HN2 p2 ph1) (p22 ph1):f3
d24
(CEN_HN1 p1 ph2) (p21 ph5):f3
d26
(CEN_HN2 p2 ph1) (p22 ph1):f3

```

d26
(p1 ph1)
DELTA1
(p2 ph1)
d13
p16:gp3
d16 pl16:f3
4u BLKGRAD
go=2 ph31 cpd3:f3
d11 do:f3 mc #0 to 2
F1PH(rd10 & ip8 & ip9 & ip29 , id0)
F2EA(igrad EA & ip5*2, id10 & ip3*2 & ip6*2 & ip31*2)
exit

```

```

ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=1 1 3 3
ph6=0
ph7=0 0 2 2
ph8=0 0 0 0 2 2 2 2
ph9=1 1 1 1 3 3 3 3
ph10=2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
ph11=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph23=1
ph25=3
ph29=0
ph31=0 2 2 0 2 0 0 2 2 0 0 2 0 2 2 0
2 0 0 2 0 2 2 0 0 2 2 0 2 0 0 2

```

```

;pl0 : 120dB
;pl1 : f1 channel – power level for pulse (default)
;pl3 : f3 channel – power level for pulse (default)
;pl9 : f1 channel – power level for presaturation

```

```

;p110: f1 channel – power level for TOCSY–spinlock
;p116: f3 channel – power level for CPD/BB decoupling
;p1 : f1 channel – 90 degree high power pulse
;p2 : f1 channel – 180 degree high power pulse
;p6 : f1 channel – 90 degree low power pulse
;p16: homospoil/gradient pulse [1 msec]
;p21: f3 channel – 90 degree high power pulse
;p22: f3 channel – 180 degree high power pulse
;p28: f1 channel – trim pulse [1 msec]
;d0 : incremented delay (F1 in 3D) [3 usec]
;d1 : relaxation delay; 1–5 * T1
;d9 : TOCSY mixing time
;d10: incremented delay (F2 in 3D) [3 usec]
;d11: delay for disk I/O [30 msec]
;d12: delay for power switching [20 usec]
;d13: short delay [4 usec]
;d16: delay for homospoil/gradient recovery
;d20: first z–filter delay [10 usec]
;d21: second z–filter delay [10 usec]
;d24: 1/(4J)YH for YH
; 1/(8J)YH for all multiplicities
;d26: 1/(4J(YH))
;cnst4: = J(YH)
;l1: loop for DIPSI cycle: ((p6*115.112) * l1) = mixing time
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;in10: 1/(2 * SW(X)) = DW(X)
;nd10: 2
;NS: 8 * n
;DS: >= 16
;td1: number of experiments
;td2: number of experiments in F2
;FnMODE: States–TPPI (or TPPI) in F1
;FnMODE: echo–antiecho in F2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel – 90 degree pulse for decoupling sequence

```

```
;use gradient ratio:   gp 1 : gp 2 : gp 3
;                       50 :   80 :   8.1       for N-15
;                       50 :   80 :  20.1       for C-13

;for z-only gradients:
;gpz1: 50%
;gpz2: 80%
;gpz3: 8.1% for N-15, 20.1% for C-13

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
;gpnam3: SINE.100

;set pl9 to 120dB when presaturation is not required
;  use 70 - 80dB to reduce radiation damping

;$ld: dipsiif3gpsi3d ,v 1.7.2.1 2001/09/11 10:14:36 ber Exp $
```

1D 1H Presaturation

```
;zgpr
;avance-version (02/05/31)
;1D sequence with f1 presaturation
```

```
#include <Avance.incl>
```

```
"d12=20u"
```

```
1 ze
2 30m
d12 pl9:f1
d1 cw:f1 ph29
4u do:f1
d12 pl1:f1
p1 ph1
go=2 ph31
30m mc #0 to 2 F0(zd)
exit
```

```
ph1=0 2 2 0 1 3 3 1
ph29=0
ph31=0 2 2 0 1 3 3 1
```

```
;pl1 : f1 channel – power level for pulse (default)
;pl9 : f1 channel – power level for presaturation
;p1 : f1 channel – 90 degree high power pulse
;d1 : relaxation delay; 1–5 * T1
;d12: delay for power switching [20 usec]
;NS: 1 * n, total number of scans: NS * TD0
```

; \$Id: zgpr,v 1.7 2002/06/12 09:05:24 ber Exp \$