THE RELATIONSHIP BETWEEN LATENT SECONDARY STRUCTURE WITHIN AN UNBOUND INTRINSICALLY DISORDERED PROTEIN AND THE THERMODYNAMICS OF COMPLEXATION WITH A TARGET MOLECULE

A Thesis in Chemistry

by

Anthony W. Zidell

© 2015 Anthony W. Zidell

Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science

May 2015
The thesis of Anthony W. Zidell was reviewed and approved* by the following:

Scott A. Showalter  
Associate Professor of Chemistry  
Thesis Advisor

David D. Boehr  
Associate Professor of Chemistry

William G. Noid  
Associate Professor of Chemistry

David S. Gilmour  
Professor of Molecular and Cell Biology

Kenneth S. Feldman  
Professor of Chemistry  
Graduate Program Chair

*Signatures are on file in the Graduate School.
Abstract

Many intrinsically disordered proteins (IDPs) undergo a disorder-to-order transition upon binding to their target molecules. Despite their plasticity, disordered proteins have often been observed in the unbound state to possess latent structure resembling the folded bound state. However, the role of this preformed structure and its effect on the thermodynamics of binding are currently not well understood. Our group has recently shown through temperature-series isothermal titration calorimetry that the folding-upon-binding of a model disordered protein, FCP1, to its target, RAP74, is equivalent to the cooperative folding of a globular protein. Changes in molar enthalpy upon formation of the FCP1–RAP74 complex were deconstructed into contributions from the primary driving force of that interaction, the hydrophobic effect, and other concomitant driving forces, such as those associated with helix formation.

In the current work, we attempt to parse the contribution of helix formation to stabilization of the IDP–target complex through the generation of a helicity series by rational mutagenesis of the disordered protein. In this way, a panel of FCP1 mutants is prepared with amino acid substitutions that confer modulated helicity without disrupting the hydrophobic binding face. The extent of latent helical structure within each FCP1 variant has been quantified using circular dichroism spectroscopy and shown to be linearly correlated with various thermodynamic parameters derived from isothermal titration calorimetry and urea denaturation of the FCP1–RAP74 complex. These observations allow us to better understand the contribution of latent secondary structure in an IDP to the energetics of IDP–target complex formation.
# Table of Contents

List of Figures ......................................................................................................................... vi
List of Abbreviations ................................................................................................................. vii
Acknowledgements .................................................................................................................. ix

Chapter 1: Introduction

I. General Introduction ................................................................................................................. 1
   Intrinsic Disordered Proteins ................................................................................................. 1
   A Model Disordered Protein: FCP1 .................................................................................... 2

II. Technical Introduction ......................................................................................................... 5
   AGADIR ................................................................................................................................. 5
   Circular Dichroism Spectroscopy ......................................................................................... 5
   Isothermal Titration Calorimetry ....................................................................................... 7

III. Thesis Overview ............................................................................................................... 12

Chapter 2: The Relationship between Latent Helicity and Binding Thermodynamics

I. Introduction ............................................................................................................................ 13

II. Materials & Methods ......................................................................................................... 14
   Rational Design of FCP1 Variants ..................................................................................... 14
   Preparation of FCP1 ............................................................................................................. 14
   Preparation of FCP1 Mutants ............................................................................................. 15
   Preparation of RAP74 .......................................................................................................... 15
   Shortened FCP1 Constructs ............................................................................................... 16
   Circular Dichroism Spectroscopy ....................................................................................... 16
   Isothermal Titration Calorimetry ....................................................................................... 16
   Statistics ............................................................................................................................... 16

III. Results & Discussion ........................................................................................................ 17
   AGADIR-Directed Mutation .............................................................................................. 17
   Thermal Denaturation of FCP1 .......................................................................................... 18
   Urea Denaturation of FCP1 ............................................................................................... 20
   Deconvoluting the Ellipticity Spectrum of the FCP1–RAP74 Complex .............................. 21
   Urea Denaturation of the FCP1–RAP74 Complex ............................................................... 22
   Isothermal Titration Calorimetry ....................................................................................... 24

IV. Conclusion ......................................................................................................................... 29
Chapter 3: Direction for Future Experiments

I. Introduction .................................................................................................................................................. 30
II. Preliminary Results & Proposed Experiments ......................................................................................... 30
   Differential Scanning Calorimetry .............................................................................................................. 30
   Generation of Structural Ensembles ........................................................................................................... 31
   NMR Spectroscopy of the FCP1–RAP74 Complex ................................................................................... 32
   Phosphorylation of Intrinsically Disordered Proteins ............................................................................... 33
List of Figures

FIGURE 1. Amino acid occurrence histogram and hydropathy plot of FCP1. ....................................................... 3

FIGURE 2. Solution structure model of FCP1 bound to RAP74. ............................................................................. 4

FIGURE 3. Conceptual decomposition of plane polarized light and elliptically polarized light ................. 6

FIGURE 4. Representative mean molar residue ellipticity spectrum of a primarily α-helical protein........................................................ ............................................................. 7

FIGURE 5. Schematic of an isothermal titration calorimeter .................................................................................. 8

FIGURE 6. Example isothermal titration calorimetry thermogram and isotherm ........................................ 11

FIGURE 7. Residue-level AGADIR prediction of fraction helicity for FCP1 variants .................................. 17

FIGURE 8. Representative mean molar residue ellipticity spectra of FCP1944–961 and FCP1944–961 at various temperature ........................................................ ........................................ 19

FIGURE 9. Mean molar residue ellipticity at 222 nm for FCP1944–961 variants as a function of temperature ........................................................................................................................................... 20

FIGURE 10. Mean molar residue ellipticity spectra of FCP1944–961 with various urea concentration ........................................................................................................................................... 21

FIGURE 11. Ellipticity spectrum of the FCP1–RAP74 complex ........................................................................ 22

FIGURE 12. Mean molar residue ellipticity at 222 nm as a function of admixed urea concentration for complexes between FCP1 variants and RAP74 ........................................................................................................................................... 23

FIGURE 13. Urea denaturation midpoint concentrations as a function of fraction helicity for variants of FCP1944–961 ........................................................................................................................................... 24

FIGURE 14. Representative thermograms and isotherms from isothermal titration calorimetry of the binding of RAP74 and FCP1944–961–K951R ........................................................................................................................................... 25

FIGURE 15. Change in molar enthalpy, molar Gibbs potential and molar entropy of injection as functions of temperature for the binding of RAP74 to the FCP1944–961 variants ........................................................................................................................................... 27

FIGURE 16. Change in molar enthalpy and temperature of endothermic-to-exothermic transition for FCP1944–961–RAP74 complexation as a function of fraction helicity ........................................................................................................................................... 28

FIGURE 17. Thermograms from differential scanning calorimetry of RAP74 with various mole ratio of FCP1944–961 ........................................................................................................................................... 31

FIGURE 18. Residual dipolar coupling data for FCP1930–961 .............................................................................. 32

FIGURE 19. HSQC spectra of isolated RAP74 and the FCP1–RAP74 complex as a function of urea concentration ........................................................................................................................................... 34
## List of Abbreviations

### Latin Alphabet

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{600}$</td>
<td>absorption at 600 nm</td>
</tr>
<tr>
<td>$b$</td>
<td>optical path length</td>
</tr>
<tr>
<td>$C$</td>
<td>molar concentration</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
<tr>
<td>$C_X$</td>
<td>molar concentration of ligand in syringe</td>
</tr>
<tr>
<td>$dP$</td>
<td>differential power applied to sample cell</td>
</tr>
<tr>
<td>DSC</td>
<td>differential scanning calorimetry; differential scanning calorimeter</td>
</tr>
<tr>
<td>E954D</td>
<td>FCP1 mutant with glutamic acid 954 substituted to aspartic acid</td>
</tr>
<tr>
<td>FCP1</td>
<td>TFIIF-associating RNA Pol II carboxyl-terminal domain phosphatase 1</td>
</tr>
<tr>
<td>$f_H$</td>
<td>fraction helix</td>
</tr>
<tr>
<td>$f_F$</td>
<td>fraction folded</td>
</tr>
<tr>
<td>$H$</td>
<td>mean net hydropathy</td>
</tr>
<tr>
<td>HRV</td>
<td>human rhinovirus</td>
</tr>
<tr>
<td>IDP</td>
<td>intrinsically disordered protein</td>
</tr>
<tr>
<td>ITC</td>
<td>isothermal titration calorimetry; isothermal titration calorimeter</td>
</tr>
<tr>
<td>$k$</td>
<td>injection number</td>
</tr>
<tr>
<td>$K_C$</td>
<td>equilibrium thermal denaturation constant</td>
</tr>
<tr>
<td>$K_D$</td>
<td>equilibrium binding dissociation constant</td>
</tr>
<tr>
<td>$K_U$</td>
<td>equilibrium urea unfolding constant</td>
</tr>
<tr>
<td>K951R</td>
<td>FCP1 mutant with lysine 951 substituted to arginine</td>
</tr>
<tr>
<td>LB</td>
<td>lysogeny broth</td>
</tr>
<tr>
<td>M</td>
<td>macromolecule</td>
</tr>
<tr>
<td>MoRF</td>
<td>molecular recognition fragment</td>
</tr>
<tr>
<td>$M_T$</td>
<td>total molar concentration of macromolecule</td>
</tr>
<tr>
<td>MX</td>
<td>macromolecule–ligand complex</td>
</tr>
<tr>
<td>$N$</td>
<td>number of ligand binding sites on macromolecule</td>
</tr>
<tr>
<td>$\langle N \rangle$</td>
<td>average number of ligands bound to macromolecule</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>$n_{MX}$</td>
<td>number of moles of macromolecule–ligand complex</td>
</tr>
<tr>
<td>$n_T$</td>
<td>total number of moles of ligand</td>
</tr>
<tr>
<td>NTA</td>
<td>nitrilotriacetic acid</td>
</tr>
<tr>
<td>N958Q</td>
<td>FCP1 mutant with asparagine 958 substituted to glutamine</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>$q$</td>
<td>excess heat of injection</td>
</tr>
<tr>
<td>$R$</td>
<td>molar Boltzmann constant</td>
</tr>
<tr>
<td>RAP74</td>
<td>74 kDa ribosome-associating protein</td>
</tr>
<tr>
<td>RNA Pol II</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecylsulfate</td>
</tr>
<tr>
<td>$t$</td>
<td>time</td>
</tr>
<tr>
<td>$T$</td>
<td>absolute temperature</td>
</tr>
<tr>
<td>TFIIF</td>
<td>transcription factor IIF</td>
</tr>
<tr>
<td>$V$</td>
<td>volume of cell</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type</td>
</tr>
</tbody>
</table>
List of Abbreviations (continued)

$X$ ligand
$X$ molar concentration of unbound ligand
$X_T$ total molar concentration of ligand

Greek Alphabet

$\alpha$ residue number
$\Delta C_P$ change in molar heat capacity at constant pressure
$\Delta G$ change in molar Gibbs potential
$\Delta H$ change in molar enthalpy
$\Delta S$ change in molar entropy
$\Delta V$ injection volume
$\theta$ mean molar residue ellipticity
$\Theta$ ellipticity
$\lambda$ wavelength
$\chi$ mole ratio of ligand to macromolecule
Acknowledgements

“Ah, if only I were a poet, I would rightly exalt you...”

D-503 of We by Yevgeny Zamyatin

Because I have been influenced for many years by innumerable people, I would like to apologize in advance for any inadvertent preterition on my part.

The work I have done at Penn State would not be possible without the support of my mentor, Dr. Scott Showalter, whose guidance often transformed my perspective on problem solving. I would also like to acknowledge the support of current and past members of the Showalter group. In particular, I appreciate the help of Dr. Debashish Sahu with NMR spectroscopic measurements and the help of Dr. Kaycee Quarles with molecular biology.

The biophysical techniques used in this work were made possible in part by the Huck Institutes of the Life Sciences Genomics Core and Automated Biological Calorimetry Facilities, and Nuclear Magnetic Resonance Spectroscopy Facility at Penn State.

To all the teachers who fostered my interest in learning, I am grateful. In particular, I would like to thank the late Mr. Ronald Hauck, from whom I learned meticulous, organized and explicit mathematical problem solving.

I would like to thank my roommate of three years, James Morse, for his immutable friendship and for our exhaustive scientific and philosophical dialogue.

For the support my family has given me, I am forever in debt. The unconditional love of my mother has provided a ubiquitous motivation to improve. I owe a special thanks to the source and stimulus of my scientific curiosity, my father, Dr. R. H. Zidell. Finally, for evoking an indelible imprint on my psyche, I would like to thank my uncle, to whom this work is dedicated.
For my late uncle, Dr. V. S. Zidell.
Chapter 1. Introduction

I. GENERAL INTRODUCTION

Intrinsically Disordered Proteins

The structure–function paradigm is a central theme in structural biology asserting that the function of a biological macromolecule is inherently dependent on its stable tertiary structure; precisely positioned moieties make specific contacts with the target molecule to effect some response. However, the current structure–function paradigm is often artificially limited in scope to the most well-studied proteins, i.e., those that fold into globular three-dimensional structure.

Within the past couple decades, it has become clear that 25–30% of proteins encoded by the human genome are entirely disordered and >50% contain long stretches of disordered regions. These intrinsically disordered proteins (IDPs) participate in numerous cellular processes, including signaling, transcription and replication, and many are either directly or indirectly involved in human disease.

IDPs tend to be enriched in polar and charged amino acids and depleted in large, nonpolar amino acids relative to globular proteins which evokes low sequence complexity and a tendency toward amino acid runs. These proteins have been characterized according to their low hydropathy and distinguished from well-folded proteins using a per residue index of mean net hydropathy versus mean net charge. IDPs tend to have negative hydropathy and high net charge relative to globular proteins. Because the dominant force in protein folding is the hydrophobic effect, the low hydropathy of IDPs contributes substantially to disorder.

The composition bias toward highly solvated amino acids in IDPs confers greater configurational entropy and a flatter free energy landscape. It has been proposed that heightened sampling of conformational space allows the IDP to more efficiently search for binding targets and to bind those targets with less specificity and lower affinity, which is especially important for rapid and reversible signaling within the cell.

Homologous binding regions of IDPs that experience similar binding-induced folding transitions are often referred to as MoRFs (molecular recognition fragments). Once in contact with the target, many IDPs undergo a disorder-to-order transition as structural heterogeneity within the MoRF becomes reduced, but likely remains prevalent. The diminished configurational entropy of the MoRF can be compensated in part by an increase in disorder within non-MoRF regions of the IDP and by favorable changes in solvation free energy when nonpolar residues in the IDP become buried upon binding.
Interestingly, it was recently demonstrated by our laboratory, in collaboration with the Noid group, that the binding of an IDP to a target is similar in nature to the cooperative folding of a single-chain globular domain of a well-folded protein.\(^{16}\) Whether the binding transition of any particular IDP follows a conformational capture model\(^{17}\) (only the pre-folded IDP may bind the target), an induced fit model\(^{18}\) (the IDP interacts with the target before completely folding), or some mixed model is still debated and cannot be answered with thermodynamic experiments.

A recent coarse-grained simulation using a derivative of the Dill HP model\(^{19-21}\) suggests IDPs may bind to their target in a stepwise manner\(^{22}\) while another study of IDP binding kinetics posits a conformational capture mechanism for at least that system.\(^{23}\) Kinetics studies show that IDP binding events, although fast, are not diffusion limited\(^{24}\) and, more importantly, are probably dependent on the lifetime of conformational states.\(^{17}\) Thus, it appears likely that IDPs indeed fill the spectrum of mixed binding models. Regardless of mechanism, what remain unclear are the driving forces for the formation of the IDP–target complex. There is a striking paucity of thermodynamic information describing folding-upon-binding transitions, so thorough investigation of these proteins is essential to understanding the means by which IDPs function without unique spatial or temporal structure.

**A Model Disordered Protein: FCP1**

Much work in the Showalter Group has focused on the intrinsically disordered region of FCP1\(^\dagger\) because it exhibits many of the common characteristics of IDPs,\(^{2,3,7}\) including low sequence complexity, enrichment in charged amino acids and low hydropathy (see Figure 1). The archetypal features of FCP1 make this protein an attractive model for the study of IDPs.

FCP1 is a dedicated phosphatase for a region of RNA polymerase II (Pol II) referred to as the CTD. Dephosphorylation of the CTD by FCP1 allows Pol II to be recycled for additional rounds of transcription.\(^{25-32}\) The phosphatase activity of FCP1 is enhanced through its direct interaction with the globular transcription factor, RAP74\(^\ddagger,33-42\) Upon associating with RAP74, the binding region within the C-terminal domain of FCP1 folds into an α-helix (see Figure 2).\(^{43,44}\) The MoRF of FCP1 has been thermodynamically characterized by our group using isothermal titration calorimetry\(^{45}\) and dynamically characterized using NMR spectroscopy\(^{46}\) and molecular dynamics simulation\(^{13}\) in both the bound and unbound states. Others have studied the structure of the FCP1–RAP74 complex using comparable techniques as well as X-ray crystallography.\(^{43,44}\)

\(^\dagger\) FCP1: TFIIF- (transcription factor IIF)-associating CTD (RNA polymerase II carboxyl-terminal domain) phosphatase. The intrinsically disordered C-terminal domain of FCP1 is studied in this work and will be referred to simply as FCP1 for brevity.

\(^\ddagger\) RAP74: 74 kDa RNA polymerase II associating protein. The C-terminal domain of RAP74 that contains the binding site for FCP1 is studied in this work and will be referred to simply as RAP74 for brevity.
Figure 2 depicts the solution structure model of FCP1 (blue) bound to RAP74 (gray) computed from NMR spectroscopic data. The MoRF of FCP1 has been defined as residues 944–961 based on this model. It should be noted, however, that the binding face of FCP1 may extend N-terminally beyond residue 944 based on a 1 μs all atom, explicit solvent simulation of a FCP1

\[ \text{FCP1}_{930–961} \text{–RAP74} \]

**Figure 1.** Physical properties of FCP1 (residues 879–961) illustrated by (A) an amino acid occurrence histogram and (B) a hydropathic index plot. The amino acids in A are arranged from left to right in order of least to most commonly occurring in IDPs relative to globular proteins. Hydropathic index was computed using the Kyte–Doolittle method with a span of nine residues.
complex (Showalter, unpublished). The increased hydrophobic content within the MoRF (Figure 1) allows FCP1 to fold into an amphipathic α-helix and bind to a hydrophobic cleft on the surface of RAP74.

In the unbound state, FCP1 has been shown by our group to possess some residual helical structure. Circular dichroism spectroscopy of FCP1 with admixed 2,2,2-trifluoroethanol (TFE) demonstrated the IDP could be induced to adopt excess α-helical structure. From the atomic resolution of nuclear magnetic resonance spectroscopy, the enhanced helicity with the addition of TFE was confirmed to be specific to the binding region of FCP1. Moreover, NMR data demonstrated denaturation of latent helical content within the binding region with admixed urea. Residual structure has been observed in numerous other IDPs as well.\textsuperscript{48-50} Despite the prevalence of latent helicity in binding regions of disordered proteins, the relationship between residual secondary structure and thermodynamics of binding remains ambiguous. Helical biases within the MoRF of an unbound IDP are hypothesized here to play a crucial role in preparing the MoRF for binding, thereby enhancing the affinity of binding to its target.

To understand this relationship, the latent helicity in the model disordered protein, FCP1, will be modulated by amino acid substitution and quantified by circular dichroism spectroscopy. The helix-coil transition of each FCP1 variant will be monitored as a function of externally controlled variables,
such as temperature and osmolyte concentration. The thermodynamics of binding of the FCP1 variants will then be quantified using isothermal titration calorimetry.

II. TECHNICAL INTRODUCTION

AGADIR

A simple algorithm for predicting the helical tendency of a peptide was introduced as AGADIR in 1994 by Muñoz and Serrano. AGADIR uses a statistical thermodynamic approach to compute the average α-helix content within a sequence of amino acids by enumerating accessible configurations and determining their energies. The energy of a configuration is a sum of contributions from the sequence-dependent intrinsic energy of sampling helical dihedrals, the energy of forming backbone hydrogen bonds, the energy of side chain–side chain interactions between monomer α and α + 1,3 and 4, and the energy due to stabilization by N- and C-capping moieties. Each of these energies is derived from empirical data, such as those generated by Chakrabartty et al. In this work, AGADIR will be used as a guide for the rational design of mutants that will confer modulated helical content to the MoRF.

Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy is a common method of identifying and quantifying secondary structure in proteins. This technique relies on detecting changes in the ellipticity of incident plane polarized light. This type of radiation can be conceptually decomposed into right- and left-handed circular components (see Figure 3). Chiral species and those containing regular repeating structure tend to absorb one component to a greater extent than the other in the far-UV range. This differential absorption effects an elliptical nature to the incident light which is measured as a function of wavelength, thereby producing an ellipticity spectrum.

One such regular repeating structure is an α-helix of amino acids. α-Helices elicit a characteristic negative peak near 222 nm in the ellipticity spectrum (see Figure 4) that can be easily distinguished from the nearly null response of a random coil at that wavelength. Thus, the folding or unfolding of an α-helix of amino acids with externally controlled variables, like temperature or osmolyte concentration, can be easily monitored.
Figure 3. Conceptual decomposition of (A) plane polarized light (black vectors) and (B) elliptically polarized light (black vectors) into right- and left-handed circular components (gray vectors). In each step, the sum of the gray vectors is the black vector, which projects along the z-axis out of the page. The elliptical nature of the vectors in B is due to the smaller magnitude of the left-handed circular component.

Protein ellipticity spectra are generally normalized to the number of moles of residues per unit volume of sample and the path length of the sample. Molar residue ellipticity is defined as

$$\theta = \theta(\lambda) = \frac{\Theta}{NbC},$$

(1)

where $\Theta = \Theta(\lambda)$ is the background-corrected ellipticity, $N$ the number of residues in the protein, $b$ the optical path length of sample and $C$ the concentration of the active species. Conventional units for $\theta$ are deg cm$^2$ dmol$^{-1}$. 
FIGURE 4. Representative mean molar residue ellipticity spectrum of a primarily α-helical protein showing the negative peak near 222 nm characteristic of the α-helix (indicated by the vertical hatched line). This spectrum was generated at 273 K with RAP74436–517 in 20 mM potassium phosphate, pH = 7.0.

CD spectroscopy will be used in this work for the purpose of measuring the extent of perturbations to the FCP1 variants from varied temperature or osmolyte concentration. The unfolding of the RAP74–FCP1 variant complex will also be monitored as a function of osmolyte concentration to study the relationship between latent helicity in an IDP and the stability of the bound structure.

Isothermal Titration Calorimetry

The driving forces of IDP binding can be elucidated by determining equilibrium binding constants and the change in enthalpy, entropy and heat capacity due to association with a target. These thermodynamic parameters provide a wealth of knowledge about the nature of protein–protein interaction. A single isothermal titration calorimetry (ITC) experiment accesses the information necessary for the determination of change in enthalpy upon binding and the equilibrium binding dissociation constant, and therefore the changes in Gibbs potential and entropy upon binding.

In the ITC experiment, a solution of ligand is titrated into a solution of macromolecule at a fixed temperature. The macromolecule is contained in a coin-shaped cell surrounded by an adiabatic jacket (Figure 5). An adjacent cell contains no protein and serves as a reference. Both cells are maintained at the same temperature by applying some amount of power to each. Following injection, the differential power required to maintain the temperature of the sample relative to the reference is measured. The jacket allows sensitive measurement of the heat of injection, which is roughly equivalent to the heat of binding between the ligand and macromolecule. Thermodynamic parameters may be extracted from the data based on analysis with a simple binding model.
The following derivations rely on the framework of a one-to-one binding interaction between two species. This model is acceptable for many IDP–target interactions, including the binding of FCP1 and RAP74. It will be assumed that activity coefficients of each species are unity such that molar concentration may be used in this analysis. Furthermore, the partial molar volumes of each protein are assumed independent of concentration. Both assumptions are valid at the low concentrations used for ITC experiments.

A binding reaction between macromolecule M and ligand X to form a complex (MX) according to

\[ M + X \rightarrow MX \]

is governed by the equilibrium binding dissociation constant
\[ K_D = \frac{[M]X}{[MX]}, \quad (2) \]

where \( X \) is the molar concentration of unbound ligand and brackets denote molar concentration. The total molar concentration of macromolecule is \( M_T = [M] + [MX] \), so the average number \( \langle N \rangle \) of ligands bound to a macromolecule is the fraction of the total concentration of \( M \) that is in the complexed form, \textit{i.e.},

\[ \langle N \rangle = \frac{[MX]}{M_T} = \frac{X}{K_D + X}. \quad (3) \]

The rightmost expression is simplified from the definition of \( K_D \) (eq. 2) and is independent of \([M]\). Because distinguishing between the bound and unbound forms of \( X \) (and \( M \)) is experimentally challenging, it is desirable to determine \( \langle N \rangle \) as a function of the contrived total concentrations of those reagents.

The total molar concentration of ligand is \( X_T = X + [MX] \). From the relations in eq. 3 for both species comprising \( X_T \), the total molar concentration of ligand is a function of \( \langle N \rangle \) by

\[ X_T = \frac{K_D \langle N \rangle}{1 - \langle N \rangle} + \langle N \rangle M_T. \quad (4) \]

Solving for \( \langle N \rangle \) from the quadratic form of eq. 4 gives

\[ \langle N \rangle = \frac{1}{2M_T} \left( X_T + M_T + K_D - \sqrt{(X_T + M_T + K_D)^2 - 4X_T M_T} \right). \]

This \textit{ad hoc} expression for the average number of bound ligands is a function of the knowable total concentrations of \( X \) and \( M \), and is therefore more useful for analyzing ITC data.

The change molar enthalpy \( \Delta H \) of the binding reaction of \( M \) and \( X \) is a function of the excess heat \( q \) of that reaction by

\[ \Delta H = \frac{q}{n_{MX}} = \frac{q}{\langle N \rangle M_T V}, \]

where \( n_{MX} = [MX]V \) is the number of moles of complex formed and \( V \) the volume of the system. The rightmost expression is simplified from eq. 3. Thus,

\[ q = q(K_D, \Delta H|M_T, X_T, V) = \frac{1}{2}V \Delta H \left( X_T + M_T + K_D - \sqrt{(X_T + M_T + K_D)^2 - 4X_T M_T} \right), \quad (5) \]

with \( M_T, X_T, V \) being externally controlled parameters in an ITC experiment.
In an ITC experiment, the entire active volume $V$ of an adiabatic cell is filled with the macromolecule at some initial total molar concentration $M_{T,0}$. The ligand is then titrated by injection through a needle into the cell in discrete increments. Each injection $k = 1, 2, 3, \ldots$ displaces a volume $\Delta V_k$ of liquid already in the cell to an inactive area (see Figure 5), thereby diluting $M$ to some new total molar concentration $M_{T,k}$. Assuming no amount of the $k$th injection is displaced as a result of its addition to the cell, the total molar concentration of $M$ following that injection will be

$$M_{T,k} = \frac{V - \Delta V_k}{V} M_{T,k-1}.$$  

Similarly, the addition of $X$ by injection will displace some amount of $X$ already present in the cell to the inactive area. So the total molar concentration of $X$ following injection $k$ will be

$$X_{T,k} = \frac{V - \Delta V_k}{V} X_{T,k-1} + \frac{\Delta V_k}{V} C_X,$$

where $C_X$ is the molar concentration of $X$ in the syringe. Thus, the mole ratio of $X$ to $M$ following injection $k$ is

$$\chi_k = \frac{X_{T,k}}{M_{T,k}},$$

which is used to produce an isotherm.

The differential power $d\mathcal{P}$ applied to the sample cell is plotted against time $t$ to yield a thermogram (Figure 6). Each injection may show a positive, negative or no peak in $d\mathcal{P}$ resulting from net endothermic, net exothermic or net athermic events, respectively, following that injection. Integration of peaks in the thermogram yields the excess heats $q$ of injection. Plotting $(\Delta q/\Delta n_T)$ versus $\chi$ (where $\Delta q = q_k - q_{k-1}$ and $\Delta n_T = (X_{T,k} - X_{T,k-1})V$) will produce an isotherm to which eq. 5 may be fit for the extraction of $\Delta H$ and $K_D$. Plotting data in this derivative form accentuates uncertainties in the data and provides a better visualization of the fitted data than a plot of $q$ versus $n_T$\textsuperscript{55,56} It is important to note that for a number $N > 1$ of indistinguishable binding sites on $M$, each $M_T$ term must be replaced by $NM_T$ to account for the multiple sites of interaction, so $q = q(K_D, \Delta H| N, M_T, X_T, V)$ provides an additional degree of freedom for fitting.
FIGURE 6. Example isothermal titration calorimetry thermogram (upper panel). Data corresponding to peak integrations (points) and the isotherm modeling those data (black curve) are shown in the lower panel. These data were acquired for the binding of the ligand RAP74 and macromolecule FCP1\textsubscript{1879–961} at 298 K.

The change in molar Gibbs potential $\Delta G$ for the macromolecule–ligand binding interaction is easily determined from its relation to the equilibrium binding dissociation constant, \textit{i.e.},

$$\Delta G = RT \ln K_D,$$

where $R$ is the molar Boltzmann constant and $T$ the absolute temperature at which the ITC experiment was conducted. The change in molar entropy $\Delta S$ can then be determined from the relation,

$$\Delta G = \Delta H - T \Delta S.$$  \hspace{1cm} (7)

Furthermore, a series of ITC experiments at various $T$ can provide the change in heat capacity $\Delta C_p$ upon binding through
\[ \Delta H = \int \Delta C_p dT. \]  

Recently, our laboratory, in collaboration with the Noid group, has shown that global fitting of isotherms collected at various temperature is a tractable method for extracting \( \Delta H \) and \( K_D \). \(^{16}\)

Each of the FCP1 variants will be studied by isothermal titration calorimetry to quantify the driving forces of the FCP1–RAP74 interaction at various temperatures. Those parameters will be analyzed to study the relationship between latent helical structure and the thermodynamics of binding.

### III. THESIS OVERVIEW

Chapter 2 summarizes and discusses the work the author has done to disambiguate the correlation of latent secondary structure in an IDP to the thermodynamics of IDP–target complexation. AGADIR is used in conjunction with the solution structure for the FCP1–RAP74 complex to rationally design mutants of FCP1 with modulated helical content within the MoRF. Circular dichroism spectroscopy is then used to identify and quantify helicity within wild-type FCP1 and helix-enhancing and helix-diminishing mutants. Urea denaturation and isothermal titration calorimetry are used to probe the thermodynamics of binding and stability of the FCP1–RAP74 complex.

Chapter 3 outlines future experiments that will provide further interrogation of the effects of latent bound-like structure of an IDP and the binding of that IDP to a target molecule. NMR spectroscopy will be a necessary addition to the panel of experiments proposed in this chapter. Characterizing changes in the FCP1 variants at the atomic level can be achieved by acquiring a suite of NMR spectroscopic data. Preliminary experiments have been conducted for unbound FCP1 and for RAP74 in complex with FCP1. Differential scanning calorimetry is a useful method for directly accessing heat capacity and can be used to generate additional useful data concerning the thermal denaturation of RAP74 in complex with FCP1 variants with modulated helicity. Preliminary DSC experiments reveal the stabilization of the complex relative to RAP74. Finally, experiments are proposed to study how the latent helicity of FCP1 is affected by phosphorylation, a post-translational modification predicted for FCP1 and common to IDPs. The same techniques used in Chapter 2 and those proposed for Chapter 3 can be used to study phospho-FCP1.
Chapter 2. The Relationship between Latent Helicity and Binding Thermodynamics

I. INTRODUCTION

Latent bound-like structure within binding regions or MoRFs (molecular recognition fragments) has been observed in many intrinsically disordered proteins (IDPs). Several methods exist for the identification of residual helicity, including circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopies. Because the prevalence of helicity in disordered proteins tends to be low, both of these spectroscopic techniques require the use of admixed osmolytes for the purpose of increasing or decreasing the helical content to convincingly demonstrate the presence of secondary structure. The molecular resolution of CD spectroscopy allows the recognition of changes in signal attributable to α-helices with osmolyte concentration, while the atomic resolution of NMR spectroscopy allows the identification of those residues undergoing changes in helicity. Residual structure within FCP1 was studied in this manner by our group using both spectroscopies in conjunction with admixed 2,2,2-trifluoroethanol to enhance helicity and admixed urea to diminish helicity.

In this work, the FCP1–RAP74 model system is used to study the effects of latent helical structure within an unbound IDP on binding thermodynamics of that IDP to a target molecule. Because the bound state of the FCP1–RAP74 complex is well-defined, we may assume that competent binding relies on precise positioning of the interaction faces of both molecules. For this reason, the use of admixed osmolytes to study the effects of latent helicity on binding thermodynamics is impractical because those same osmolytes are expected to affect the structure of the binding interface of RAP74. An alternate method is therefore required to investigate the connection between the latent structure of an IDP and the thermodynamics of IDP–target complexation.

The distinct helical propensity of each amino acid type can be taken advantage of to realize changes to helical propensity within the MoRF of FCP1. Amino acid substitutions to residues in FCP1 will allow the helicity of the IDP to be altered without affecting the structure of RAP74. The modulation of helical content within IDPs by mutation has been demonstrated by several other groups. Several nuances of this method must be considered carefully for the study of IDPs in this manner.

Ideally, for the purpose of studying latent helicity within the MoRF of an IDP, mutations will confer only modulated helical propensity. If this single variable is isolated, a “knob” is effectively created to dial the helicity and therefore, as we hypothesize, the thermodynamic parameters of
binding. It is necessary to avoid radical mutations because drastic changes to the chemical environment of the MoRF may lead to unexpected transient long range interactions affecting binding and convoluting data analysis. The rational design of mutants requires (i) a solved structure for the IDP–target complex and (ii) a viable tool for predicting helical content.

The solution structure model of the FCP1–RAP74 complex shown in Figure 2 details the binding face of the amphipathic FCP1 helix. Exposed side chains of FCP1 directly interact with RAP74 according to this model. Residues lowlighted in dark blue anchor the IDP by nesting the folded MoRF within the nonpolar cleft of RAP74 while the two side chains in light blue have been identified to interact via attractive electrostatic forces. Mutation of these residues would result in obfuscated perturbation to binding thermodynamics. Because these residues cannot be modified for the purpose of this work, the other solvated residues of FCP1 which play no apparent role in direct interaction with RAP74 are good candidates for mutation.

AGADIR is an online calculator for helical propensity in peptides (see Chapter 1 for additional detail). This calculator will be used as a guide to choose amino acid substitutions that will confer substantially modulated helicity. Because AGADIR is ignorant to tertiary contacts, such as those that stabilize helices, artificially introduced long range interaction will go undetected by the calculator. Thus, again, conservative amino acid substitutions to residues with similar chemical properties but distinct helical propensity (e.g., glutamate to aspartate or lysine to arginine) are the most sensible for the study of the relationship between latent helicity and binding energetics. It is important to note that because of the limited number of residues available for mutation, enhancement of helical content is challenging without drastic alteration to the chemical properties of the peptide by less conservative mutation.

II. MATERIALS & METHODS

Rational Design of FCP1 Variants

The online helical propensity predictor AGADIR was used to identify those conservative mutations that may confer the widest breadth of helicity in the MoRF of FCP1. Calculation was performed on a per-residue basis with no N- or C-terminal protection for FCP1_{879–961} at 298 K, an ionic strength of 0.1 M, and pH = 7.00.

Preparation of FCP1

Competent BL21(DE3) Escherichia coli cells hosting a pET-47b(+) plasmid containing a synthetic gene encoding FCP1_{879–961} were allowed to proliferate in the presence of 30 μg mL^{-1}
kanamycin in Miller LB at 37°C to an optical density $A_{600 \text{ nm}} = 0.7–0.9$ measured with a 1 cm path length. His-tagged FCP1$_{879–961}$ was subsequently overexpressed at 30°C in the presence of 5 mM isopropyl β-D-1-thiogalactopyranoside for 3–4 h. Cells were harvested by centrifugation at $3400 \times g$ for 10 min at 4°C and lysed by sonication in an ice bath. The lysate was clarified by centrifugation at $14200 \times g$ for 15–20 min at 4°C. His-tagged FCP1$_{879–961}$ was purified from the lysate by immobilized metal affinity chromatography on a Ni-NTA column. The protein was eluted and the His-tag cleaved from FCP1$_{879–961}$ by His-tagged HRV 3C protease, leaving a non-native N-terminal Gly–Pro–Gly tripeptide. FCP1$_{879–961}$ was purified from the excised His-tag and 3C on the Ni-NTA column and was subsequently concentrated and buffer-exchanged at 4°C to 20 mM potassium phosphate, pH = 7.0, in a 3 kDa nominal molecular mass cutoff centrifugal filter. Purity was assessed by SDS-PAGE and concentration was determined on a DirectDetect FTIR spectrometer (Millipore) to ±5% relative uncertainty or better.

**Preparation of FCP1 Mutants**

FCP1$_{879–961}$ mutants were prepared from a pET-47b(+) plasmid hosting the fcp1$_{879–961}$ synthetic gene using a QuikChange Lightning site-directed mutagenesis kit (Agilent) with codon usage optimized for expression in *E. coli*. Once the mutated DNA was cloned into a new pET-47b(+) vector and that vector transformed into competent BL21(DE3) *E. coli* cells, purification proceeded as described above for wild-type FCP1$_{879–961}$.

**Preparation of RAP74**

Competent BL21(DE3) *E. coli* cells hosting a pET-47b(+) plasmid containing a synthetic gene coding for RAP74$_{436–517}$ were allowed to proliferate in the presence of 30 μg mL$^{-1}$ kanamycin in Miller LB at 37°C to an optical density $A_{600 \text{ nm}} = 0.7–0.9$ measured with a 1 cm path length. His-tagged RAP74$_{436–517}$ was subsequently overexpressed at 30°C in the presence of 5 mM isopropyl β-D-1-thiogalactopyranoside for 3–4 h. Cells were harvested by centrifugation at $3400 \times g$ for 10 min at 4°C and lysed by sonication in an ice bath. The lysate was clarified by centrifugation at $14200 \times g$ for 10–20 min at 4°C. His-tagged RAP74$_{436–517}$ was purified from the lysate by immobilized metal affinity chromatography on a Ni-NTA column. The His-tag was cleaved from RAP74$_{436–517}$ by His-tagged HRV 3C protease, leaving a non-native N-terminal Gly–Pro–Gly tripeptide. RAP74$_{436–517}$ was purified from the excised His-tag and 3C on the Ni-NTA column and was subsequently concentrated and buffer-exchanged at 4°C to the required buffer in a 3 kDa nominal molecular mass cutoff centrifugal filter. Purity was assessed by SDS-PAGE and concentration was determined by UV-vis spectroscopy to ±0.6% relative uncertainty or better using the Gill–von Hippel method.$^{58}$
Shortened FCP1 Constructs

Solid state synthesized FCP1\textsubscript{944–961} variants were obtained as lyophilized powders from the Penn State Hershey Macromolecular Core Facility. The peptides were gravimetrically massed and dissolved in the necessary buffer to achieve desired concentrations to $\pm 5\%$ relative uncertainty or better. Phosphate concentration, urea concentration and buffer pH were precise to $\pm 0.7\%$, $\pm 1.0\%$ and $\pm 0.7\%$ relative uncertainty or better, respectively.

Circular Dichroism Spectroscopy

300 $\mu$L samples were prepared with the proteins at desired concentrations (generally $\sim 100 \ \mu$M) and were allowed to equilibrate in an ice bath for 8–16 h. Ellipticity spectra were acquired using a J-1500 spectropolarimeter (Jasco). The xenon lamp was allowed to warm up for $>30 \text{ min}$ with a nitrogen flow rate of 15–30 ft$^3$ h$^{-1}$. Samples were contained in a quartz cuvette with an optical path length of 1.00 mm and housed in a Peltier jacket capable of maintaining temperature to $\pm 0.02 \text{ K}$ throughout measurement. Samples were incubated within the jacket for 2–3 min for experiments conducted at 273 K, and 4–10 min for experiments conducted at greater temperature. Each spectrum was acquired with four accumulations over the wavelength domain, 245–205 nm, with a scanning rate of 100–200 nm min$^{-1}$ and with a slit width of 2 nm. A 0.025 nm step resolution was used to generate 1601 points for each spectrum.

Isothermal Titration Calorimetry

RAP74\textsubscript{436–517} and the FCP1\textsubscript{944–961} variants were extensively co-dialyzed in 50 mM potassium chloride/20 mM potassium phosphate, pH = 7.0, to obtain exact buffer match. Concentrations of potassium chloride and potassium phosphate were precise to $\pm 0.7\%$ and $\pm 0.5\%$, respectively, and pH was precise to $\pm 0.03 \text{ pH}$ units. A 3 kDa nominal molecular mass cutoff regenerated cellulose membrane was used for RAP74\textsubscript{436–517} and a 0.1–0.5 kDa nominal molecular mass cutoff cellulose ester membrane was used for the FCP1\textsubscript{944–961} variants (Spectrum Labs). ITC data were acquired using a VP-ITC isothermal titration calorimeter (MicroCal). Following an initial delay of 60 s to establish the 5 $\mu$cal s$^{-1}$ reference power baseline and a 2.5 $\mu$L initial injection, 15 $\mu$L aliquots of 500 $\mu$M RAP74\textsubscript{436–517} were injected into 35–40 $\mu$L FCP1\textsubscript{944–961} variant at 0.5 $\mu$L s$^{-1}$ with 300–480 s spacing between injections. A stirring rate of 307 rpm was used to maintain homogeneity within the sample cell. Data were analyzed using the Origin 7 data acquisition software (OriginLabs).

Statistics

In all cases where replicate experiments exist, uncertainty of any value $Y$ computed from measurements $x_i$ was propagated using standard deviation $s$ according to
\[ s_Y^2 = s_{x_1}^2 \left( \frac{\partial Y}{\partial x_1} \right)^2_{x_2, x_3, \ldots} + s_{x_2}^2 \left( \frac{\partial Y}{\partial x_2} \right)^2_{x_1, x_3, \ldots} + s_{x_3}^2 \left( \frac{\partial Y}{\partial x_3} \right)^2_{x_1, x_2, \ldots} + \cdots. \]

The true mean \( \mu \) of that value \( Y \) is then reported as the 95% confidence interval about its experimental mean \( \bar{Y} \), i.e.,

\[ \mu = \bar{Y} \pm \frac{t s_Y}{\sqrt{n}}, \]

where \( t = t(n) \) is Student’s t-statistic for the 95% confidence interval with a number \( n \) of replicate experiments. It should be noted that the 95% confidence interval is \( 1.84s_Y \) for the typical \( n = 3 \) replicate experiments. Error bars on all plots will reflect the 95% confidence interval rather than standard deviation.

### III. RESULTS & DISCUSSION

**AGADIR-Directed Mutation**

The design of FCP1\textsubscript{879–961} mutation was guided by AGADIR for the generation of variants that may confer the widest breadth of helical propensity (Figure 7). N958Q and K951R mutants of FCP1\textsubscript{879–961} were subsequently used for circular dichroism spectroscopy to identify secondary structure and perturbations to that structure by varied temperature. Following the recognition of a need for shorter constructs (\textit{vide infra}), E954D, N958Q, K951R and wild-type variants of FCP1\textsubscript{944–961} were obtained for the purpose of studying the isolated binding region of FCP1.

![Graph](image_url)

**Figure 7.** Residue-level AGADIR prediction of fraction helicity for wild-type FCP1 and several mutants spanning a range of helical content within the MoRF. Calculations were performed for the peptides with no N- or C-terminal protection, and at 298 K, 0.1 M ionic strength and pH = 7.00.
Thermal Denaturation of FCP1

The thermal denaturation of latent helical structure within the binding region of FCP1\textsubscript{879–961} has been monitored with circular dichroism spectroscopy. Experimental temperature was sufficiently precise to limit overall precision by only that of the experimental method. Figure 8A demonstrates that with advancing temperature, the minimum near 222 nm erodes to a featureless curve. The diminution of that peak is expected because it reflects the abolition of α-helical structure. However, contrary to intuition, the magnitude of the signal increases as the peptide becomes thermally denatured. This trend has also been observed for the mutants of FCP1\textsubscript{879–961}.

Random coils elicit a negative peak near 195 nm which is much stronger and broader than that of α-helices. Because CD spectroscopy measures the mean response of an entire molecule, the signal is dominated by the primarily disordered region of FCP1\textsubscript{879–961} (see Figure 7) which obfuscates the characteristic peaks of α-helical structure. For this reason, shortened constructs (FCP1\textsubscript{944–961} and its mutant derivatives) containing only the MoRF\textsuperscript{43} are expected to disambiguate the helix–coil equilibrium of FCP1 in the unbound state.

The mean molar residue spectra for FCP1\textsubscript{944–961} in Figure 8B show a more defined local minimum at 222 nm indicating that the shortened construct is indeed primarily α-helical. Furthermore, the temperature dependence of spectra of FCP1\textsubscript{944–961} follows the expected trend of decreasing magnitude of signal as the peptide becomes thermally denatured. The E954D, N958Q and K951R mutants of FCP1\textsubscript{944–961} behave similarly.

The magnitude of the local minimum at 222 nm measures the extent of helical structure in the FCP1\textsubscript{944–961} variants, so those data are used to compute a fraction helical content $f_\text{H}$ assuming the two-state model,

$$M_\text{H} \rightleftharpoons M_\text{C},$$

where the peptide transitions between a helix state ($M_\text{H}$) and a thermally denatured state ($M_\text{C}$). To adequately define the signal for completely helical and completely coil species, the thermal denaturation of a well-folded and primarily α-helical protein is best used as a standard. The mean molar residue ellipticity for RAP74 is used to calibrate the data for the FCP1\textsubscript{944–961} variants. RAP74 and each of the variants converge to an experimentally indistinguishable mean molar residue ellipticity at high temperature, which is assumed to be $f_\text{H} = 0$. The signal observed for RAP74 at 273 K is assumed to be the mean molar residue ellipticity for completely folded α-helices, \textit{i.e.}, $f_\text{H} = 1$. 


Figure 8. Representative mean molar residue ellipticity spectra of (A) FCP1_{879–961} and (B) FCP1_{944–961} at various temperature. The vertical hatched line at 222 nm indicates the wavelength at which the local minimum attributable to α-helical peptides exists.

The fraction helix is used to determine the equilibrium constant for this model as

\[ K_C = \frac{1 - f_H}{f_H}. \]

The change in molar Gibbs potential for thermal denaturation is then

\[ \Delta G_C = -RT \ln K_C. \]  

(9)
Figure 9. (A) Fraction helix for FCP1_{944−961} variants and (B) the change in molar Gibbs potential for thermal denaturation assuming a two-state model. For reference to a well-folded primarily α-helical protein, the hatched red line in A follows equivalent data for isolated RAP74. With all three variants, it is apparent that some degree of helicity is possessed at low temperature and that structure is melted out as heat is added to the system. All data are based on mean molar residue ellipticity at 222 nm and were collected at the stated temperature with a buffer of 20 mM potassium phosphate, pH = 7.0. Error bars represent the 95% confidence interval.

The fraction helicity and change in Gibbs potential for each of the FCP1_{944−961} variants are plotted as a function of temperature in Figure 9. Data for the N958Q mutant have been omitted due to an apparent lack of modulated helicity, i.e., the signal for this peptide overlays with wild-type FCP1_{944−961}. In agreement with the AGADIR-predicted trend, it is apparent that the K951R mutation confers a modest enhancement in latent helical content within the MoRF of FCP1, and the E954D mutation greatly diminishes that structure.

Urea Denaturation of FCP1

The unfolding of latent helical structure within the MoRF of FCP1 by admixed urea has also been monitored with circular dichroism spectroscopy. Gravimetric preparation of urea solutions was sufficiently precise to limit overall precision by only that of the experimental method. Similar to the thermal denaturation of FCP1_{944−961}, unfolding of the peptide by urea decreases the magnitude of the mean molar residue ellipticity signal (Figure 10). These data will be essential for the analysis of the urea denaturation of the FCP1−RAP74 complex.
Deconvoluting the Ellipticity Spectrum of the FCP1–RAP74 Complex

The blue curve in Figure 11 is a representative circular dichroism spectrum collected at 273 K for a sample containing 767 μM FCP1_{944–961} and 121 μM RAP74. This type of spectrum is inherently difficult to interpret because the observed signal is a linear combination of signals from unbound FCP1 (M), unbound RAP74 (X) and the complex (MX).

A rigorous method for deconvoluting the spectrum produced by the sample used to study the FCP1–RAP74 complex is required for the analysis of data and quantitative comparison with the spectrum of isolated RAP74. To disambiguate the spectrum, we will first assume that a >4× molar excess of FCP1 will force all available RAP74 to assume the bound state. Previous differential scanning calorimetry experiments with variable mole ratios demonstrate this assumption to be valid (see Figure 17). Consequently, the observed ellipticity \( \Theta_{\text{obs}} \) is comprised only of contributions from unbound FCP1 and the complex and not from unbound RAP74 (i.e., \( \Theta_X = 0 \)) such that

\[
\Theta_{\text{obs}} = \Theta_M + \Theta_{MX}.
\]

This assumption also requires that the concentration of complex is equivalent to the total concentration of RAP74, i.e.,

\[
[MX] \equiv X_T.
\]

The concentration of unbound FCP1 is then the difference between the total concentrations of FCP1 and RAP74, or

\[
[FCP1] = [X_T] - [MX].
\]

Figure 10. Mean molar residue ellipticity spectra of FCP1_{944–961} at 273 K over a range of urea concentration. The vertical hatched line at 222 nm indicates the wavelength at which the local minimum attributable to \( \alpha \)-helical peptides exists.
\[ [M] = M_T - X_T. \]  

(12)

Thus, the molar residue ellipticity of the complex is

\[ \theta_{MX} = \frac{\theta_{\text{obs}} - \Theta_M}{N_{MX} b X_T}, \]  

(13)

where \( N_{MX} = 103 \) is the number of residues in the FCP1–RAP74 complex. Ellipticity measurements of FCP1 in isolation allow the back calculation of \( \Theta_M \) for any concentrations \( M_T \) and \( X_T \) using eq. 1.

\[ \text{Figure 11. Ellipticity spectrum of the FCP1–RAP74 complex (blue line) at 273 K from 121 \mu M \text{ RAP74 and 767}\ \mu M \text{ FCP1. Ellipticity spectra for FCP1 and RAP74 were back-calculated for those concentrations from spectra acquired for the isolated proteins (dotted and hatched gray lines, respectively). The sum of the spectra of isolated proteins (solid gray line) differs substantially from the spectrum of the complex and thereby demonstrates the increase in helical content in one or both proteins upon binding.} \]

**Urea Denaturation of the FCP1–RAP74 Complex**

Spectra collected for the FCP1–RAP74 complex at various concentrations of admixed urea must also be disambiguatated. With increasing concentration of urea, the complex will dissociate and the structure of the two proteins will diminish. Although the concentration of complex varies in this manner, we are interested in the fate of the entire original complex, such that the denominator of eq. 13 will be general for analyses at all concentrations of urea. Because the extent of secondary structure in each species is dependent upon the concentration of urea, \( \theta_{\text{obs}} \) and \( \Theta_M \) are both functions of this externally controlled parameter. This method requires that the mean molar residue ellipticity of FCP1 be known at each concentration of urea used to study the unfolding of the complex.

A denaturation curve for RAP74 in complex with each FCP1 variant is constructed from the mean molar residue ellipticity data at 222 nm (Figure 12). The data are normalized such that the fraction
folded in the absence of urea is \( f_F = 1 \) and the fraction denatured is \( f_U = 1 - f_F = 0 \). The hatched red line follows equivalent data for isolated RAP74. Each of the variants in complex with RAP74 shift the denaturation midpoint to higher concentrations of urea relative to RAP74 in isolation.

The change in molar Gibbs potential for the denaturation of each complex can be computed with the assumption that the process follows a two-state model, i.e.,

\[
(MX)_F \rightleftharpoons (MX)_U,
\]

where \((MX)_F\) is the folded state of the complex and \((MX)_U\) the unfolded state. The change in molar Gibbs potential is then

\[
\Delta G_U = -RT \ln \frac{f_U}{f_F},
\]

(14)

where \( R \) is the molar Boltzmann constant and \( T \) the absolute temperature of the experiment. \( \Delta G_U \) data about the denaturation midpoint for each of the complexes and isolated RAP74 are shown in Figure 12B.

**Figure 12.** (A) Denaturation curves for RAP74 in complex with FCP1\(_{944-961}\) variants and (B) the change in molar Gibbs potential for unfolding of those complexes as a function of urea concentration assuming a two-state model. The hatched red line in A follows equivalent data for isolated RAP74. The denaturation midpoint is shifted to greater concentration of urea for the complex with all three FCP1 variants relative to isolated RAP74. Furthermore, complexation with the helix-enhancing mutant, K951R, has the greatest midpoint, followed by wild-type FCP1, and finally the helix-diminishing mutant, E954D. This trend can be more clearly seen by the change in molar Gibbs potential at a given concentration of urea. All data are based on mean molar residue ellipticity at 222 nm and were collected at 273 K for samples with >4× molar excess of FCP1 variant in 20 mM potassium phosphate, pH = 7.0, and the stated concentration of urea.
Whereas binding the helix-enhancing K951R mutant drives the complex to the greatest denaturation midpoint, binding the helix-diminishing E954D mutant produces a noticeably lesser increase relative to isolated RAP74. Interestingly, data for binding with wild-type FCP1 falls between the two mutants producing a trend that follows the observations from CD spectroscopy of the unbound FCP1 variants. In fact, the denaturation midpoint concentration of urea correlates well with the fraction helical content for each of the variants (Figure 13). Although the coefficient of determination $R^2$ is not a meaningful statistic for data with large uncertainties in abscissae, it is nevertheless shown to demonstrate the linearity of the mean data. This finding is of particular interest because it implies that the stability of an IDP–target complex can be predicted based solely on the latent helical content in the binding region of the IDP.

![Figure 13](image)

**Figure 13.** Urea denaturation midpoint concentrations as a function of fraction helicity at 273 K for variants of FCP1$_{944-961}$. The denaturation midpoint, and therefore the stability of the complex, is a linear function of latent helical content within the binding region of FCP1. Error bars represent the 95% confidence interval.

**Isothermal Titration Calorimetry**

The binding of RAP74 and FCP1$_{944-961}$ variants has been thermodynamically quantified using isothermal titration calorimetry. Following integration of peaks in each thermogram, eq. 5 was modified to include a parameter for a number $N$ of indistinguishable binding sites and fit to the data to yield $N$, the equilibrium binding dissociation constant $K_D$ and the change in molar enthalpy $\Delta H$. Experiments were performed at various temperature to provide a more complete description of the thermodynamics of binding between RAP74 and each FCP1 variant. Representative thermograms
FIGURE 14. Representative thermograms (upper panels) and resultant isotherms (lower panels) from isothermal titration calorimetry of the binding of RAP74 and FCP1944−961−K951R (green) and FCP1944−961−E954D (brown) at (A) 283 K and (B) 313 K. Isotherms were fit (solid lines in lower panels) using eq. 5. The experimental buffer for all experiments was 50 mM potassium chloride/20 mM potassium phosphate, pH = 7.0.
and isotherms are shown in Figure 14. At lower temperature, peaks in the thermogram are positive, indicating a net endothermic process following injection. In contrast, at higher temperature, peaks in the thermogram are negative, indicating a net exothermic process.

The fit parameters, $K_D$ and $\Delta H$, for each titration allow the change in molar Gibbs potential $\Delta G$ and the change in molar entropy $\Delta S$ according to eqs. 6 and 7. Each of the three energies, $\Delta G$, $\Delta H$ and $-T\Delta S$, are plotted as functions of temperature in Figure 15. As qualitatively identified in Figure 14, $\Delta H$ for each of the variants tends to decrease with advancing temperature. Interestingly, this decrease is compensated by an increase in $-T\Delta S$, to produce an experimentally indistinguishable change in $\Delta G$ for all FCP1 variants over the experimental range of temperature. This observation has been made by our group for the interaction of wild-type FCP1$_{879-961}$ and RAP74$_{485-517}$–A485W.$^{16}$

The most obvious departure from conformity between the three FCP1$_{944-961}$ variants is the temperature $T_H$ at which binding transitions from a net endothermic to a net exothermic processes. Wild-type FCP1 transitions at 295 K, while the helix-enhancing mutant (K951R) and helix-diminishing mutant (E954D) transition at 296 and 286 K, respectively.

The slope of a linear fit to the $\Delta H$ versus $T$ data is the change in molar heat capacity $\Delta C_p$ at constant pressure according to eq. 8. From a linear regression analysis of each data set, $\Delta C_p$ for wild-type FCP1 and the K951R and E954D mutants is found to be experimentally indistinguishable as $-0.19 \pm 0.05$, $-0.15 \pm 0.02$ and $-0.15 \pm 0.03$ kcal mol$^{-1}$ K$^{-1}$, respectively, where parenthetical values denote the 95% confidence interval. Similar values have been obtained from ITC measurements of the interaction of wild-type FCP1$_{879-961}$ and RAP74$_{485-517}$–A485W.$^{16}$ Large negative changes in molar heat capacity, such as those found here, are attributable to burial of nonpolar surface area for proteins. Because FCP1 folds into an amphipathic helix upon binding the nonpolar site of interaction on RAP74, it is not surprising that large negative $\Delta C_p$ are observed for these systems.

Due to the binding-induced folding transition of FCP1, contributions from the changes in solvation of the backbone of that protein would also be expected to be substantial. Each of the variants are then expected to have unique contributions as a result of their distinct helical propensities. Interestingly, however, the $\Delta C_p$ are indistinguishable for all three variants of FCP1$_{944-961}$. This observation implies that the change in heat capacity for the binding of RAP74 and any FCP1 variant is primarily a result of burial of hydrophobic surface at the binding interface of the interacting proteins and not the disorder-to-order transition of FCP1.
FIGURE 15. Change in (A) molar enthalpy, (B) molar Gibbs potential and (C) molar entropy of injection from isothermal titration calorimetry as functions of temperature for the binding of RAP74 to the FCP1<sub>944-961</sub> variants. The hatched lines in A are linear fits to the data for K951R and E954D. Error bars represent the 95% confidence interval.
The data for the change in molar enthalpy of injection follow the trend of fraction helix determined by circular dichroism spectroscopy. Similar to the data for the midpoint concentration of urea for the denaturation of the complex, a plot of the change in molar enthalpy as a function of fraction helix produces a linear trend (Figure 16A). Likewise, the temperature at which the binding reaction transitions from endothermic to exothermic is a linear function of fraction helix (Figure 16B). As noted previously for the urea denaturation curves, these observations imply that thermodynamic parameters for IDP–target complexation can be predicted if the extent of latent helical content in the binding region of the IDP is known.

![Figure 16](image.png)

**Figure 16.** (A) Change in molar enthalpy of injection at 283 K (upper data) and 313 K (lower data) as a function of fraction helical content for the binding interaction of FCP1$_{944-961}$ and RAP74. (B) Temperature of transition from endothermic to exothermic processes for that binding interaction as a function of fraction helical content.

Each of the FCP1 mutants was designed to host conservative amino acid substitutions such that the unbound forms possess distinct helical tendencies, while the bound forms assume the same folded state as the wild-type peptide. Thus, the change in molar enthalpy of transitioning from the largely unfolded state of the E954D mutant to the mostly helical state of the K951R mutant is simply the difference between the molar enthalpies for those proteins. With the generation of additional FCP1 mutants, this concept will allow us to determine the contribution of latent helicity in the disordered protein to the thermodynamics of binding a target molecule.
IV. CONCLUSION

FCP1 and its interaction with RAP74 have been investigated in this work because the properties and behavior of this protein are typical for IDPs. FCP1 is disordered in the unbound state and folds upon binding RAP74 into an α-helix. Unbound FCP1 exhibits a substantial fraction of latent helical content, which has been shown previously by our laboratory to exist in the binding region. The current work has demonstrated that the extent of latent helicity determines the stability of the FCP1–RAP74 complex.

Helical content within wild-type FCP1_{944–961}, the helix-enhancing mutant (K951R), and the helix-diminishing mutant (E954D) was quantified using circular dichroism spectroscopy. Urea titrations of RAP74 in isolation and in complex with the FCP1_{944–961} variants then showed that the stability of the complex scales linearly with latent helicity. Surprisingly, isothermal titration calorimetry revealed that the change in molar Gibbs potential for binding was independent of temperature within experimental uncertainty. This observation means that although the IDP and target molecule have reached a stage in evolution at which complexation is insensitive to fluctuations in temperature, the IDP–target complex may still be sensitive to chaotropic denaturation.

Understanding the relationship between latent helical structure in the binding region of an IDP and the thermodynamics of its interaction with a target molecule is imperative for the realization of methods for predicting and interpreting IDP properties. A course of proposed experiments are outlined in Chapter 3 to provide broader perspectives for understanding these proteins.
Chapter 3. Direction for Future Experiments

I. INTRODUCTION

Several avenues stemming from this work need to be investigated more thoroughly to fully understand the nature of intrinsically disordered proteins. Nuclear magnetic resonance spectroscopy will provide a swath of useful data reporting on structure to complement the results of circular dichroism spectroscopy reported in Chapter 2. Differential scanning calorimetry of RAP74 in complex with each of the FCP1 variants will yield thermodynamic information from a different perspective than isothermal titration calorimetry. Preliminary and proposed experiments using these techniques are discussed in the following section.

II. PRELIMINARY RESULTS & PROPOSED EXPERIMENTS

Differential Scanning Calorimetry

The denaturation of a structured biological polymer, such as a protein, is accompanied by an increase in heat capacity attributed primarily to the exposure of the nonpolar core to solvent. Careful analysis of the heat signature of the thermal denaturation process can yield valuable information describing the energetics of the phase transition from the folded state to the unfolded state. Differential scanning calorimetry (DSC) accesses this thermodynamic information directly by measuring the excess heat capacity of a solution at variable temperature.

DSC has been used by the author to measure thermodynamic properties of RAP74 in isolation and in complex with FCP1\textsubscript{879–961} (Figure 17). While isolated RAP74 has a phase transition midpoint near 327 K, isolated FCP1\textsubscript{879–961} exhibits no change in excess heat capacity over the experimental temperature range. At saturating mole ratios of FCP1 (M) to RAP74 (X), a single peak is observed for the complex near 340 K. Due to the large equilibrium binding dissociation constant, a bimodal distribution is observed at subsaturating concentrations of FCP1. The shift in transition midpoint to greater temperature for the complex indicates the complex is substantially more stable than isolated RAP74. Furthermore, the peak for the complex is greater in magnitude, \textit{i.e.}, more heat is required to thermally denature the complex, so the change in enthalpy for the transition is greater than that for isolated RAP74 (reference eq. 8).

This work may lay the foundation for future studies of RAP74 in complex with wild-type FCP1 and the helix-enhancing and -diminishing mutants. This method has the potential to provide
additional thermodynamic data to more completely understand the effects of latent helical structure within an IDP on binding to its partner.

**Figure 17.** Heat capacity profiles for 1 mg mL\(^{-1}\) (105 µM) RAP74 (X) with various mole ratios of FCP1\(_{879-961}\) (M) from differential scanning calorimetry. Isolated RAP74 shows a peak in molar heat capacity near 327 K, while the fully complexed form of RAP74 peaks near 340 K. Unbound FCP1 (teal) exhibits no heat signature and therefore does not contribute to the heat capacity profile of the FCP1–RAP74 complex. All experiments were performed in a buffer of 50 mM potassium chloride/50 mM sodium cacodylate, pH = 6.5.

**Generation of Structural Ensembles**

Structural ensembles can be constrained by numerous data, including those from nuclear magnetic resonance (NMR) spectroscopy. In particular, NMR spectroscopy can be used to measure dipolar couplings within biological macromolecules when the motion of those molecules is sufficiently restrained that the effects of global tumbling no longer average those couplings to zero.

Several media have been used to restrain motion and effect partial alignment. One such alignment medium commonly used for proteins is compressed polyacrylamide gel. This medium has been chosen by our lab for the measurement of residual dipolar couplings (RDCs) because it properly aligns many of the proteins we study while maintaining sufficiently weak interactions to prevent alteration to structure of those proteins. However, a simple polyacrylamide alignment medium was proven incompetent for RDC measurement for a construct of FCP1 containing residues 930–961.

Several resonances of FCP1\(_{930-961}\) vanish or shift dramatically in frequency when subjected to alignment by compressed polyacrylamide. Presumably, the peptide at least interacts transiently with the alignment medium. Like many IDPs, the extreme C-terminus of FCP1 is highly negatively charged.
The author has taken advantage of this physicochemical property to mitigate interactions by developing a novel charged alignment medium specifically for the purpose of measuring RDCs in FCP1_{930–961}. Charged gels are generated using the same protocol for preparing acrylamide/bisacrylamide gels for SDS-PAGE, but substituting half of the normal acrylamide for 2-acrylamido-2-methylpropane sulfonic acid. This medium has since been used to align other IDPs in our laboratory. This unique polyacrylamide-derivative medium introduces a negatively charged moiety to effectively repel FCP1_{930–961} and thereby reduce interaction. The amide nitrogen–amide proton residual dipolar couplings within FCP1_{930–961} could then be measured (Figure 18).

![Residual dipolar coupling data for FCP1_{930–961} (teal). For comparison to the previously studied construct including the full disordered C-terminus, data have been included for FCP1_{879–961} (gray; Showalter, unpublished). Breaks in the data for FCP1_{879–961} are due to immeasurable peaks.](image)

Additional NMR spectroscopic data must be acquired to constrain an ensemble for FCP1_{930–961}. Because resonance assignments for this peptide were based on data for FCP1_{879–961}, it is best to verify the peaks by performing experiments to independently assign FCP1_{930–961}. Ideally, chemical shifts for the amide proton and nitrogen, carbonyl carbon, and aliphatic carbons and protons will be measured to provide data to constrain an ensemble. In addition to chemical shifts, three bond scalar couplings are especially useful for ensemble generation.

**NMR Spectroscopy of the FCP1–RAP74 Complex**

The stability of the complex formed by RAP74 and several FCP1_{944–961} variants was studied by circular dichroism spectroscopy and isothermal titration calorimetry in Chapter 2. NMR spectroscopy may also be used to interrogate these systems. Measuring changes in the chemical shifts
of resonances in a heteronuclear single quantum coherence (HSQC) spectrum with admixed urea provides an additional means of monitoring the unfolding of an IDP–target complex.

The author has previously acquired a series of HSQC spectra for urea titrations of $^{15}$N isotopically enriched RAP74 in isolation and in complex with FCP1$_{879-961}$ at urea concentrations of 0–4.4 M (Figure 19). From the denaturation curves in Figure 12, it is evident that with 4.4 M urea, isolated RAP74 and the FCP1–RAP74 complex should both be nearly completely unfolded. In agreement with the results of Chapter 2, a qualitative inspection of the HSQC spectra indicates that RAP74 denatures at lower concentrations of urea than the complex.

**Phosphorylation of Intrinsically Disordered Proteins**

The carboxyl-terminus of FCP1 is predicted to be post-translationally modified at residues adjacent to the RAP74 binding site by phosphorylation of Ser942 and Ser944. Phosphorylation of these residues *in vitro* effects an increased affinity of the peptide for RAP74. However, the effects of phosphorylation of FCP1 on latent helical structure are currently unknown. From the results of Chapter 2, it is reasonable to predict that phosphorylation of FCP1 enhances helicity within the binding region, which subsequently produces the increase in affinity for RAP74.

This hypothesis may be easily tested with the techniques used in Chapter 2. Comparing circular dichroism spectra of phosphorylated and non-phosphorylated FCP1$_{944-961}$ will allow the fraction helical content to be quantified for both species. Solid state synthesized phosphopeptides are commercially available from numerous sources, including the Penn State Hershey Macromolecular Core Facility from which the FCP1$_{944-961}$ variants of Chapter 2 were obtained. Isothermal titration calorimetry can then measure the thermodynamics of phosphorylated FCP1 binding to RAP74 to further quantify the correlation between phosphorylation-induced changes in helical propensity and complexation.
Figure 19. HSQC spectra of (A) isolated RAP74 and (B) the FCP1–RAP74 complex at various concentration of urea. All experiments were performed at 298 K and used a buffer of 50 mM potassium chloride/50 mM sodium cacodylate, pH = 6.5.
References


