CHAPTER TWO

Assessing Coupled Protein Folding and Binding Through Temperature-Dependent Isothermal Titration Calorimetry

Debashish Sahu*, Monique Bastidas*, Chad W. Lawrence*,[1], William G. Noid*, Scott A. Showalter*,†,[2]

*Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania, USA
†Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania, USA
[2]Corresponding author: e-mail address: sas76@psu.edu

Contents

Abstract

Broad interest in the thermodynamic driving forces of coupled macromolecular folding and binding is motivated by the prevalence of disorder-to-order transitions observed when intrinsically disordered proteins (IDPs) bind to their partners. Isothermal titration calorimetry (ITC) is one of the few methods available for completely evaluating the thermodynamic parameters describing a protein–ligand binding event. Significantly, when the effective $\Delta H°$ for the coupled folding and binding process is determined by ITC in a

[1] Present address: Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99354.

Methods in Enzymology, Volume 567
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2015.07.032

temperature series, the constant-pressure heat capacity change (ΔC_p) associated with these coupled equilibria is experimentally accessible, offering a unique opportunity to investigate the driving forces behind them. Notably, each of these molecular-scale events is often accompanied by strongly temperature-dependent enthalpy changes, even over the narrow temperature range experimentally accessible for biomolecules, making single temperature determinations of $\Delta H°$ less informative than typically assumed. Here, we will document the procedures we have adopted in our laboratory for designing, executing, and globally analyzing temperature-dependent ITC studies of coupled folding and binding in IDP interactions. As a biologically significant example, our recent evaluation of temperature-dependent interactions between the disordered tail of FCP1 and the winged-helix domain from Rap74 will be presented. Emphasis will be placed on the use of publically available analysis programs written in MATLAB that facilitate quantification of the thermodynamic forces governing IDP interactions. Although motivated from the perspective of IDPs, the experimental design principles and data fitting procedures presented here are general to the study of most noncooperative ligand binding equilibria.

1. INTRODUCTION

Intrinsically disordered proteins (IDPs) lack a significant bias toward any particular spatially and temporally stable equilibrium structure under native conditions (Tompa, 2002; van der Lee et al., 2014). In contrast to IDP folding, their interactions appear to be more harmonious with the expectation of a funneled free energy surface that guides the IDP conformational ensemble toward a folded structure in the bound state (Csermely, Palotai, & Nussinov, 2010; Papoian & Wolynes, 2003). This result suggests there may be other parallels between folding of globular proteins and IDP interactions. Certain 3D folds are repeated in proteins with specific functions; similarly, some biological processes are optimally performed by IDPs, providing a strong incentive for better understanding the energetics of IDP interactions (Dunker, Silman, Uversky, & Sussman, 2008; Fuxreiter et al., 2008). The prevalence of IDP disorder-to-order transitions induced by interactions (Dyson & Wright, 2002; Tompa & Fuxreiter, 2008; Wright & Dyson, 2009), which often involve regions of the IDP possessing partially preformed structure (Mohan et al., 2006), has stimulated research into the thermodynamic driving forces coupling folding and binding (Ferreon & Hilser, 2004). The hypothesis emerging from this work is that native disorder provides a molecular pathway to high fidelity, yet reversible, association in processes including cellular signaling and gene transcription (Ganguly et al., 2012; Wright & Dyson, 2009). Accordingly, there is

considerable interest in quantification of the thermodynamic forces that govern IDP interactions.

When applicable, isothermal titration calorimetry (ITC) is one of the most robust methods available for thermodynamic studies (Ghai, Falconer, & Collins, 2012). ITC provides a direct measurement of the binding enthalpy ($\Delta H°$) and equilibrium association constant (K_A, more commonly reported as $K_D = {K_A}^{-1}$), enabling the calculation of the Gibbs free energy ($\Delta G°$) through the known temperature and parametric determination of the binding entropy ($\Delta S°$). Significantly, when $\Delta H°$ is determined by ITC in a temperature series, the constant-pressure heat capacity change (ΔC_p) associated with the binding process is also experimentally accessible. As a result, several temperature-dependent ITC data sets have been reported for IDPs binding to macromolecular partners (Cervantes et al., 2011; Cho et al., 2010; Drobnak et al., 2013; Ferreon & Hilser, 2004; Lacy et al., 2004; Spolar & Record, 1994; Xu, Oruganti, Gopalan, & Foster, 2012). For both unfolding and ligand dissociation events, large changes in $\Delta H°$ are often seen, even to the extent of producing a transition from endothermic to exothermic binding over narrow and near-room temperature ranges. Thus, measuring $\Delta H°$ at a single temperature is often less informative than typically assumed (Cooper, 2010) and the significance of experimentally determining ΔC_p when investigating coupled IDP folding and binding should not be underestimated. In particular, the influence of ΔC_p on complex stability and the insight it provides into, e.g., the role of forces including the hydrophobic effect in their binding mechanism is of special significance.

While this argument suggests that temperature-dependent studies are essential to an optimal ITC-based investigation of IDP folding and binding, the resulting data must be interpreted with caution. Specifically, it is clear that attributing too much mechanistic significance to entropy–enthalpy compensation is dangerous, because systematic under-reporting of ITC errors enriches the natural correlation that is already expected between these parameters when they are reported over narrow (experimentally accessible) temperature ranges (Chodera & Mobley, 2013). Perhaps more important than the simple presence of entropy–enthalpy compensation is that this phenomenon suggests interactions involving molecular disorder-to-order transitions should be relatively robust to temperature fluctuations, because the Gibbs free energy will experience relatively small temperature-induced changes over biologically tolerated temperature ranges as a result. This may be particularly significant where the interaction under investigation is formed through the accumulation of multiple cooperative and weak

interactions, which will give rise to the large observed ΔC_p (Cooper, 2010), as is nearly always the case for protein folding or coupled folding and binding.

As a model system to investigate the vital role of disordered proteins for regulating eukaryotic gene transcription (Tantos, Han, & Tompa, 2012), our laboratories have collaborated to study interactions between the intrinsically disordered C-terminal region of the TFIIF-associating CTD phosphatase (FCP1) and the cooperatively folded winged-helix domain from Rap74 (Kumar, Showalter, & Noid, 2013; Lawrence, Bonny, & Showalter, 2011; Lawrence, Kumar, Noid, & Showalter, 2014; Lawrence & Showalter, 2012; Wostenberg, Kumar, Noid, & Showalter, 2011). Briefly, FCP1 performs an essential function in transcription by dephosphorylating the hyperphosphorylated C-terminal domain of RNA Polymerase II in preparation for subsequent rounds of transcription. Biochemical studies suggest that the interaction between Rap74 and FCP1 facilitates this function. As illustrated in Fig. 1, our work has shown that the Rap74 binding surface of FCP1, which forms an amphipathic α-helix in complex (Kamada, Roeder, & Burley, 2003; Nguyen et al., 2003), is only partially α-helical in the unbound state (Lawrence et al., 2011). Given this structural knowledge, we set out to derive an easily generalized framework for investigating coupled folding and binding of IDPs. As described above, temperature-dependent ITC provides a thorough characterization of the effective thermodynamic properties of coupled folding and binding processes. Therefore, we have evaluated the temperature dependence of the FCP1–Rap74 interaction by ITC, using a general thermodynamic model in which the large observed change in ΔC_p is directly fit (Lawrence et al., 2014).

Here, we will document the procedures we have adopted in our laboratory for designing, executing, and globally analyzing temperature-dependent ITC studies of coupled folding and binding in IDP interactions.

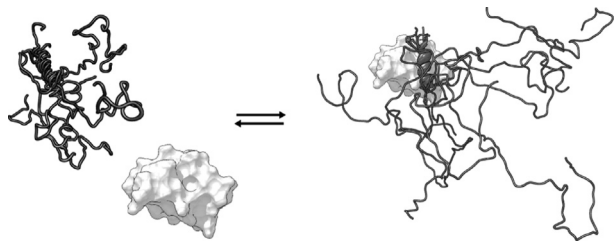

Figure 1 An illustration of the interaction between FCP1 (dark gray ribbon) and Rap74 (white surface model), which is the object of the temperature-dependent isothermal titration calorimetry study used to generate examples throughout the text.

Due to the nature of this volume, we will assume the reader is generally familiar with the ITC technique and therefore will not focus on a detailed protocol for conducting the individual titrations. Instead, our focus will be on avoiding the common pitfalls in analyzing temperature-dependent ITC data sets. Emphasis will be placed on the use of analysis programs we have written in MATLAB that are made publically available by the authors and which facilitate quantification of the thermodynamic forces that govern IDP interactions.

2. SAMPLE PREPARATION

In general, preparation of IDPs for ITC assays is no different from the preparation of cooperatively folding proteins for the same purpose. However, there are two general trends in designing purification strategies for high yield and high purity when working with IDPs that bear mentioning. Perhaps surprisingly, many IDPs seem to express well in *Escherichia coli*, without significant problems due to proteolysis. On the other hand, the yield of very small proteins from *E. coli* overexpression is often poor, which can make it necessary to express intrinsically disordered regions as fusion proteins incorporating a large domain, such as GST or maltose binding protein, in order to improve yield. Note, however, that this strategy was not necessary for the C-terminal region of FCP1, for which a representative purification protocol is presented below. Finally, caution must be exercised when selecting molecular weight cutoffs (MWCOs) for membranes used in dialysis or centrifugal concentration, because the nonglobular nature of IDP structure allows them to pass through membranes that are nominally restrictive enough to retain a protein with a particular IDP's mass. In general, selecting a MWCO that is less than half of the IDP mass is a good initial strategy that can be relaxed based on experience to improve membrane performance or purification speed and efficiency.

2.1 Plasmid Construction

A synthetic gene with *E. coli* optimized codon usage and also encoding a nonnative N-terminal tryptophan was purchased from Geneart to encode FCP1 (residues 878–961 of the human polypeptide sequence). The DNA encoding this FCP1 construct was cloned into the pET-47b (Novagen) expression vector to create the plasmid pET-47b-FCP1W, which was transformed into *E. coli* BL21 (DE3) competent cells. The same procedure was

repeated with an *E. coli* optimized synthetic gene encoding residues 434–517 of the human Rap74 polypeptide sequence, producing the plasmid pET-47b-Rap74.

2.2 Purification of FCP1 and Rap74

1. Inoculate LB media with *E. coli* bearing either the pET-47b-FCP1W or the pET-47b-Rap74 plasmid and then grow to an optical density between 0.8 and 0.9 at 600 nm. Induce expression through the addition of IPTG to a final concentration of 1.0 m*M*, followed by an additional 3–4 h of growth.
2. This step is performed at 4 °C. Harvest the cells by centrifugation for 20 min at 3300 times gravity. Suspend the cell pellet in lysis buffer (50 m*M* Tris, adjusted to pH 7.5 before use, 500 m*M* NaCl, 20 m*M* imidazole, 5 m*M* β-mercaptoethanol), 1 × protease inhibiting cocktail set V (EMD Millipore), and 1 × DNase (Promega). Lyse by sonication and then clarify the cell lysates by centrifugation for at least 30 min at 14,000 × *g*.
3. Equilibrate an immobilized metal affinity chromatography column containing a 5 mL settled bed volume of Ni^{2+}-NTA resin (Thermo Fisher), using 40 mL of lysis buffer. Pass the clarified cell lysate over the column, followed by passage of 40 mL wash buffer (lysis buffer plus 0.1% Triton X-100), and a final rinse with an additional 20 mL of lysis buffer. All steps may be performed using gravity flow.
4. Elute the His-tagged fusion protein in 20 mL of elution buffer (50 m*M* Tris, adjusted to pH 7.5 before use, 500 m*M* NaCl, 200 m*M* imidazole, 5 m*M* β-mercaptoethanol) and transfer the collected material to pres-oaked 1000 Da MWCO dialysis tubing. Add His-tagged 3C Protease to the tubing prior to closing and dialyze against 1.6 L of dialysis buffer (50 m*M* Tris, adjusted to pH 7.5 before use, 300 m*M* NaCl, 5 m*M* β-mercaptoethanol) overnight at 4 °C.
5. Achieve removal of the recombinant His-tag and His-tagged 3C Protease by passing the contents of the dialysis tubing back over a reequilibrated Ni^{2+}-NTA column using gravity flow. To ensure complete elution of FCP1 or Rap74, wash the column with an additional 20 mL of dialysis buffer.
6. Concentrate the protein using a 3000 Da MWCO centrifugal concentrator with a PES membrane (EMD Millipore) and store in lysis buffer until ready for ITC.

7. The day before ITC experiments are to be initiated, transfer FCP1 and Rap74 to separate pieces of 1000 Da MWCO dialysis tubing; codialyze FCP1 and Rap74 overnight against a shared reservoir of 50 m*M* cacodylate, pH 6.5, 50 m*M* KCl. Determine final concentrations of FCP1 and Rap74, either by absorbance at 280 nm or using a Direct Detect FT-IR (EMD Millipore).

3. SPECIAL CONSIDERATIONS FOR ITC OF INTRINSICALLY DISORDERED PROTEINS

3.1 Experimental Design

In general, the ITC experiment is set up by loading the ligand at high concentration into the injection syringe, which is then inserted into the sample cell that has been filled with a less concentrated solution containing the macromolecule. For one-to-one interactions, the assignment of the labels "ligand" and "macromolecule" is essentially semantic; thus, the investigator is at liberty to select which macromolecular species to inject from the syringe based on practical considerations. It must be stressed that all processes accompanied by the evolution or uptake of nonnegligible heat will be detected by the ITC experiment and so it is important to minimize contributions from off-target processes. Of relevance to the present discussion, many IDPs undergo reversible assembly state changes over the concentration ranges typically accessed upon loading into the syringe. Therefore, it is generally preferred that the IDP be loaded into the sample cell and the well-folded binding partner into the syringe, even if the IDP is the less massive species.

Locally, we have access to both a MicroCal, Inc. (now a part of Malvern) VP-ITC and a MicroCal AutoITC$_{200}$ for our work. While IDP binding interactions display just as wide an array of associated affinities as those measured between pairs of well-folded partners (see table 1 in Gibbs & Showalter, 2015), IDP interactions most commonly are accompanied by dissociation constants in the micromolar range. Therefore, the protocols discussed here will be presented under the assumption that a VP-ITC (or equivalent) is being used.

3.2 Calorimetric Parameters

1. Total injections—this is dependent on the volume of each titration, but typically on the order of 30–40 injections. The user can overestimate this value, without adversely affecting the titration; the instrument will report insufficient volume in the syringe and save the data. In addition,

the first injection is always set to a lower volume to account for inconsistencies in syringe filling (Pierce, Raman, & Nall, 1999) and to eliminate the known "backlash" anomaly (Mizoue & Tellinghuisen, 2004); but this initial point is not included in the analysis.

2. Cell temperature—the typical range of temperatures for these studies is 5–40 °C. If possible, determine the thermal denaturation temperature (e.g., by differential scanning calorimetry) for the cooperatively folded partner of the IDP, which will enable the final temperature range to be selected such that thermal denaturation of the folded protein can be avoided.
3. Reference power—the reference power is the constant power applied to equilibrate the temperatures between the reference and sample cells. Typically the value is set to 21 μJ/s (5 μCal/s in the MicroCal software), but for interactions yielding large heats, the reference power should be set to 42 μJ/s (10 μCal/s) or greater.
4. Initial delay—set to the minimum value of 60 s to establish a baseline.
5. Concentrations—the concentrations of the samples should be entered here, but if necessary, minor corrections can be made during analysis.
6. Injection parameters—typical settings include:
   a. Volume—the volume of each injection.
   b. Duration—the duration in seconds of each injection. As a practical value to establish a starting point, this parameter is usually set to produce an injection rate of 0.5 μL/s. For example, if the injection volume is 10 μL, the default setting for this parameter is 20 s.
   c. Spacing—after each injection, the differential heat signal should return completely to baseline. The necessary spacing between each injection will need to be determined based on the system, but typically 240 or 360 s is a good starting point.
   d. Filter period—this parameter alters the period in which a single data point is collected. For reactions occurring rapidly, 2 s is sufficient to represent a peak, which can be properly integrated to produce an accurate enthalpy.
7. Feedback mode/gain—The feedback mode dictates the response time of the instrument. For increased sensitivity when small heats of injection are expected, the passive (none) mode will provide the best data. For the quickest response time, a high response time is selected. This mode is recommended in the manual.
8. ITC equilibration options—typically, Fast Equil. and Auto are selected. This will ensure that the temperatures of the cell and sample are within

specified tolerances for the starting temperature. Note that these settings will dictate an automatic period of equilibration with stirring and final baseline equilibration prior to initiating the injection schedule.

4. DATA FITTING AND ANALYSIS

There are many commercial or open source software packages available for the analysis of ITC data, including Origin® (bundled with the MicroCal calorimeters we have used for this work and thus described here), SEDPHAT (Schuck, 2000), Affinimeter (Burnouf et al., 2012), or NanoAnalyze (TA Instruments). In our laboratory, we generally perform all of our ITC data fitting and analysis using in-house MATLAB macros that have features we will also discuss in this section. All MATLAB macros used for the analysis presented in this article are freely available through the Penn State ScholarSphere Data Repository at the following stable URL: https://scholarsphere.psu.edu/downloads/5712mr21f.

4.1 Standard Data Fitting in Origin

Standard ITC data analysis, as originally proposed by Wiseman et al., begins with fitting to the single class of sites model in Origin®, yielding apparent values for ΔH°, K_A, and the stoichiometry (n), as presented elsewhere in this volume (Wiseman, Williston, Brandts, & Lin, 1989). However, collecting calorimetric data at a single temperature can be problematic if the apparent binding enthalpy happens to be near zero at the temperature selected. For example, the raw ITC data generated for the present study of FCP1 binding to Rap74 (Fig. 2) clearly show a near-zero differential heat for each injection at 20 °C (Fig. 2B). Origin® analysis of nearly isoenthalpic data typically yields unreliable stoichiometries and association constants. Further, the uncertainty in the measured ΔH° is likewise increased, although this is rarely detrimental as determination that ΔH° was near zero is qualitatively insightful on its own.

Often, during the early design stages of the experiment, titrations will be performed at several temperatures in order to find the "optimal" temperature that yields both strong heats to measure and a robust determination of the equilibrium constant. As is clear from Fig. 3A, this search often reveals that the binding enthalpy varies dramatically with temperature and, in particular, demonstrates a large ΔC_p, which suggests that the binding event is driven by many, weak cooperative interactions. The conventional method for quantitative calculation of the ΔC_p associated with an interaction

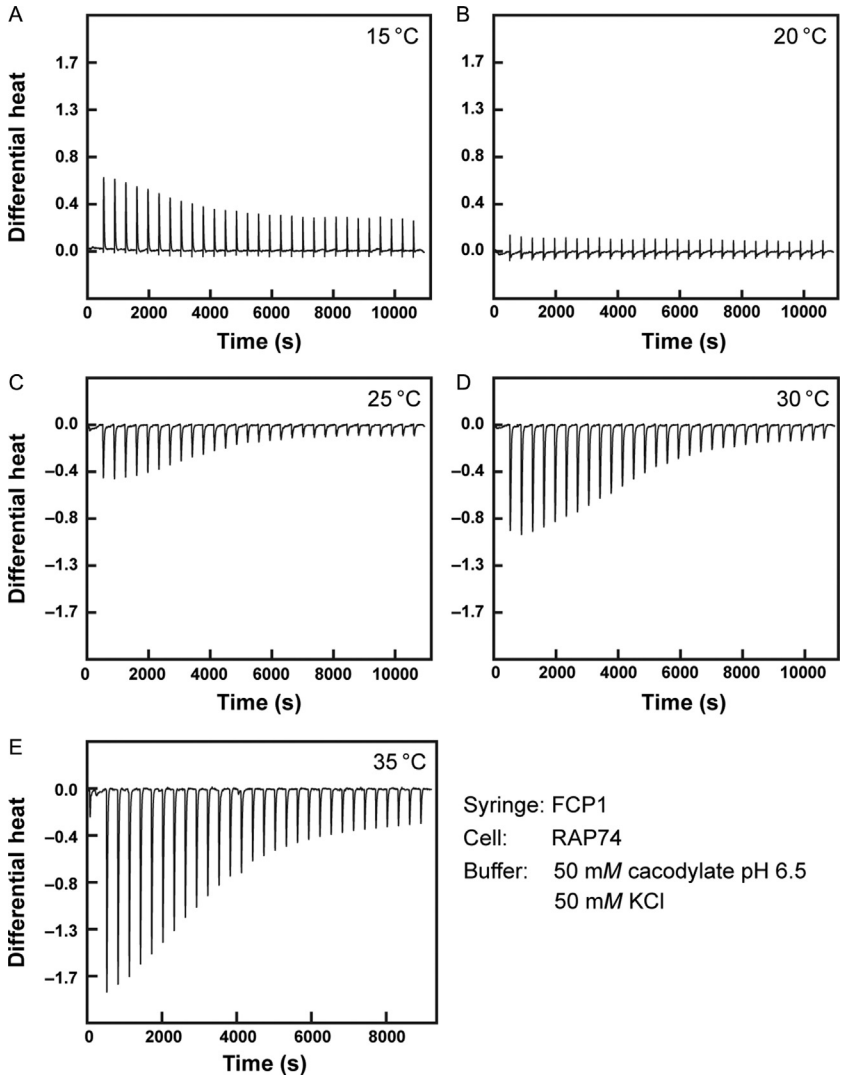

Figure 2 Isotherms generated by titration of FCP1 from the syringe into Rap74 in the cell of a VP-ITC at (A) 283 K, (B) 288 K, (C) 303 K, (D) 308 K, and (E) 308 K. The buffer conditions used for the ITC experiments were 50 m*M* cacodylate pH 6.5 and 50 m*M* KCl.

between ligand and macromolecule is to repeat the binding titration at different temperatures and obtain the apparent binding enthalpy as a function of temperature. A linear fit of the resulting $\Delta H^{\circ}(T)$ data series yields ΔC_{p} as the slope of the fit. For most macromolecular interactions, this procedure works well; there is a linear dependence of the enthalpic contribution to

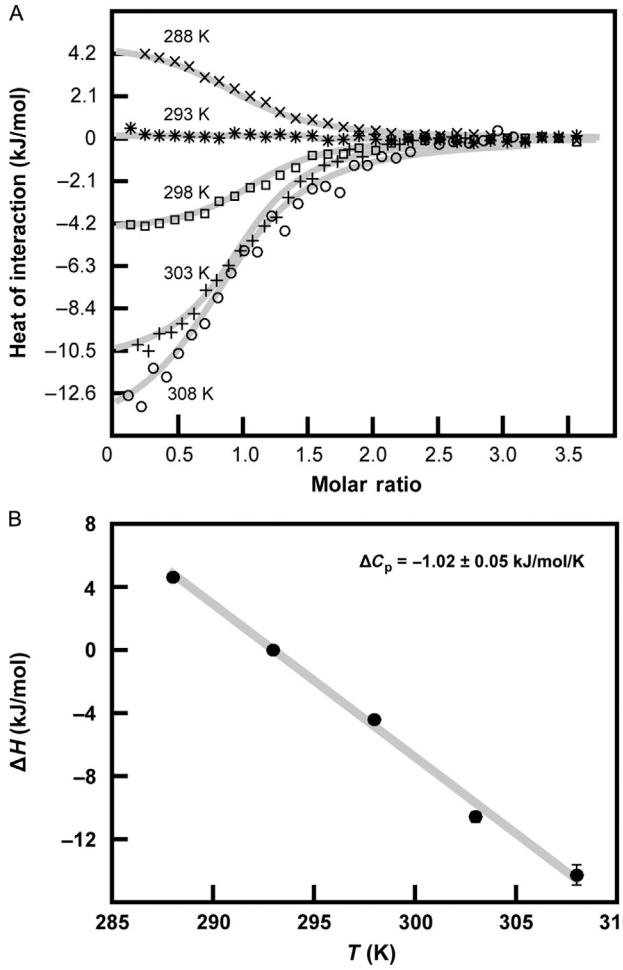

Figure 3 Serial analysis of temperature-dependent ITC data in Origin® yields an estimate of the constant-pressure heat capacity change (ΔC_p) from a linear fit to the apparent binding enthalpy (ΔH). (A) Overlay of individual fits (gray line) for each temperature, with the experimentally observed integrated heats of injection displayed at 288 K (×), 293 K (*), 298 K (□), 303 K (+), and 308 K (○). (B) Determination of ΔC_p from a linear fit (gray line) of the fitted ΔH (black circles) at each temperature, with the best-fit value provided as an inset.

ligand binding on temperature over the narrow temperature ranges experimentally accessed, unless a thermal unfolding transition for one or both of the biomolecules is approached. For example, the binding of FCP1 by Rap74 is accompanied by a linear temperature dependence to ΔH° over the range from 288 to 308 K (Fig. 3A). When the individual data sets are

fit in Origin®, ΔC_p can also be computed by plotting the fitted ΔH° as a function of temperature and performing a linear fit, yielding the heat capacity change $\Delta C_p = -1.02 \pm 0.05$ kJ/mol/K as the slope (Fig. 3B). Large negative ΔC_p values are typically interpreted to arise from changes in the amount of solvent exposed surface area upon binding (Dill, 1990), although there are many possible causes (Cooper, Johnson, Lakey, & Nollmann, 2001).

Protocol for data fitting

1. After each experiment, the data are manually integrated to minimize the error in the heat of injection.
2. Prior to fitting the data in Origin®, the baseline is corrected for heats of dilution of the buffer and/or the ligand. The baseline correction for the dilution of ligand into buffer is accounted for in one of two ways: a separate titration of ligand into buffer is collected or a 5-point baseline average is computed. For the analysis of FCP1–Rap74 interaction data, the raw data are manually integrated and a 4-point baseline average is used for baseline subtraction of the heats of dilution of Rap74 using the simple math function in Origin®.
3. Baseline corrected data are fit to a one-site binding model, where 100 iterations of fitting are completed until χ^2 is no longer reduced.
4. Repeat this procedure for each titration performed at the various temperatures.

4.2 Data Fitting in MATLAB

Generating accurate individual fits of ITC data to a single set of sites binding model in the Origin® software requires that the upper baseline approach nearly to zero without a sign inversion. Ideally, no heat should be released or absorbed during the final ~5 injections, but often this is not the case; various factors including buffer mismatch, dilution effects, or improper baseline subtraction can cause significantly nonzero heats for the final points or even transitions crossing from, e.g., exothermic to endothermic heats. This is problematic because the fitting function used in Origin® does not allow the best-fit line to cross zero (i.e., the fit is constrained to be semidefinite). The result of an improperly treated upper baseline will be artifacts in the fitted binding affinity, enthalpy, and number of binding sites. In order to circumvent this fitting issue, we relax the upper baseline constraints using in-house written MATLAB fitting macros that add an offset to the single class of sites binding model. This new method allows the user to directly take

the integrated heat data from the ITC with less extensive intervention and preprocessing.

The code we provide for fitting temperature-dependent ITC data is written to implement the Levenberg–Marquardt (LM) algorithm (MATLAB function fit.m with LM options) for large-scale fitting of nonlinear functions. An alternative method that appears attractive is to implement the standard nonlinear numerical fitting algorithm present in the base MATLAB library (MATLAB function nlinfit.m), but in our experience, this algorithm tends to become trapped in local minima during the fitting of ITC data. Initial estimates of the apparent binding association equilibrium constant (K_A) and number of binding sites (n) must be provided, noting that quality initial guesses can improve fitting efficiency dramatically. It is not necessary for the user to provide initial guesses for $\Delta H°$ and the baseline offset, because the code is written to extract high-quality initial guesses for these parameters directly from the integrated heats provided as input. Finally, we note that as distributed our macro is designed to perform 10^6 LM iterations, which generally should be sufficient with reasonable initial estimates for the fitting parameters.

Our modified algorithm adds an additional fitting parameter (the upper baseline) compared to the simpler model implemented in Origin®, but in general the best-fit values for the thermodynamic parameters generated when our modified method is applied to well-collected data do not vary significantly with respect to those generated by the conventional method (Table 1). For example, when the binding of FCP1 by Rap74 is fit in MATLAB using ITCFit.m (Fig. 4), the fitted $\Delta C_p = -1.16 \pm 0.07$ kJ/mol/K, which is in good agreement with $\Delta C_p = -1.02 \pm 0.05$ cal/mol/K fitted in Section 4.1 using Origin® derived fits to the raw ITC data. Note that there are data sets for which the agreement between these two methods will be reduced. Specifically, a poorly defined upper or lower baseline in the data reduces the quality of the fit in both algorithms and differences are expected to arise. Notably, in cases where the upper baseline is poorly defined, Origin® will fix the upper baseline to zero and force the fitting of the observed heat released to conform with this choice, resulting in a poor convergence with the experimental data as well. In these cases, it is always advisable to repeat the experiments with increased concentration of macromolecule in the cell such that the "c value" (Wiseman et al., 1989) is improved.

Finally, we note that the addition of an extra parameter in the fitting algorithm does lead to increased values for the estimated uncertainties of the fitted parameters. This is of little concern because it has been noted that

Table 1 Summary of Best-Fit Parameters for FCP1 Binding to Rap74

T (K)	*N*	ΔH° (kJ/mol)	K_D (μ*M*)
Origin			
288	1.02 ± 0.02	4.8 ± 0.1	2.2 ± 0.2
293	N.B.D.	N.B.D.	N.B.D.
298	1.15 ± 0.02	−4.6 ± 0.1	1.4 ± 0.2
303	1.04 ± 0.02	−11.1 ± 0.3	1.9 ± 0.3
308	0.96 ± 0.03	−14.9 ± 0.7	3.1 ± 0.5
MATLAB			
288	1.03 ± 0.02	5.1 ± 0.2	2.7 ± 0.3
293	N.B.D.	N.B.D.	N.B.D.
298	1.18 ± 0.02	−4.9 ± 0.2	1.9 ± 0.3
303	1.08 ± 0.02	−12.7 ± 0.5	3.1 ± 0.5
308	1.01 ± 0.04	−17.7 ± 1.4	5.3 ± 1.2

These parameters produce the gray best-fit lines in Figs. 3A and 4A. For the 293 K data heats were not sufficient to constrain the fit and the results are reported as no binding detected (N.B.D.)

the true error in ITC fitting parameters is probably systematically underestimated (Chodera & Mobley, 2013). It is our view that the simpler data handling in our MATLAB-driven fitting procedure minimizes user bias and removes potential artifacts, making an inflation of the error estimates a price well worth paying.

Protocol for data fitting

1. At the end of an ITC experiment, the user will initially integrate the heat released or absorbed in each injection, as outlined in Section 4.1.
2. Following integration, the data series containing injection volume (v_i), ligand concentration (X_T), macromolecule concentration (M_T), molar ratio (X_T/M_T), and heat of injection (ndh) in tabular form can be accessed in Origin® through the "Windows" menu tab. These data are saved in a standard ASCII text format that is directly used by our provided MATLAB scripts as inputs. Macromolecule and ligand refer here to the contents in the cell and the syringe, respectively.
3. The file generated in step 2, along with the active cell volume (V_o) and syringe concentration of ligand (X_o), are used as inputs into the MATLAB fitting macro (ITCFit.m), which fits the data to a modified

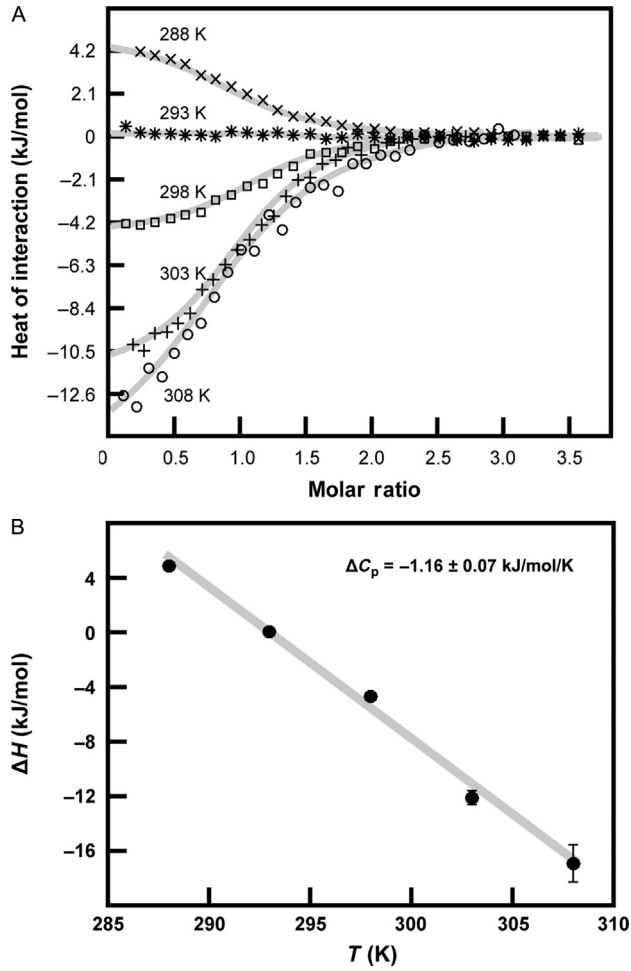

Figure 4 Serial analysis of temperature-dependent ITC data in MATLAB yields an estimate of the constant-pressure heat capacity change (ΔC_p) from a linear fit to the apparent binding enthalpy (ΔH). (A) Overlay of individual fits (gray line) for each temperature, with the experimentally observed integrated heats of injection displayed at 288 K (×), 293 K (*), 298 K (□), 303 K (+), and 308 K (○). (B) Determination of ΔC_p from a linear fit (gray line) of the fitted ΔH (black circles) at each temperature, with the best-fit value provided as an inset.

single class of sites binding model as discussed above. Initial estimates of the apparent binding association equilibrium constant (K_A) and number of binding sites (n) must be provided. If the fit does not converge, it is recommended to repeat the fitting procedure using different initial estimates of K_A and n.

4.3 Global Analysis of Temperature-Dependent Data Sets

The procedure outlined in the previous two sections is robust and effective, but the elegance of ITC data is that it directly measures heats, granting unique access to fundamental thermodynamic quantities capable of producing direct mechanistic insights. In our view, it is advantageous to consider global fitting of temperature-dependent ITC series, as we will describe in this final section. One clear advantage of a global data analysis is that it minimizes the impact of uncertainty inherent to data sets that happen to have been collected near the isoenthalpic temperature for the process under investigation. A second advantage of the method is that it directly estimates ΔC_p. Global data analysis inherently requires model building, noting that the single set of sites model discussed above will typically need to be discarded, because the equilibrium association constant need not be independent of temperature. The greatest strength of this approach is that the models used can rely entirely on fundamental thermodynamic equations, with only the interpretation depending on system-specific assumptions, making the results reasonably model independent.

A global model must determine both ΔH° and K_A at each temperature T for which ITC experiments were performed. Our central assumption is that ΔC_p is constant over the relevant temperature range. The standard enthalpy and entropy then vary according to

$$\Delta H^\circ(T) = \Delta C_p(T - T_H) + \Delta H_0 \quad (1)$$

$$\Delta S^\circ(T) = \Delta C_p \ln(T/T_S) + \Delta S_0 \quad (2)$$

where $\Delta H_0 = \Delta H^\circ(T_H)$ and $\Delta S_0 = \Delta S^\circ(T_S)$. From these expressions, the association constant is determined from $K_A(T) = \exp[-\Delta G^\circ(T)/RT]$ with $\Delta G^\circ(T) = \Delta H^\circ(T) - T\Delta S^\circ(T)$. Note that T_H and T_S are not necessarily the same and that these expressions are exact over the range for which ΔC_p is constant.

Previous studies have demonstrated that the interaction between FCP1 and Rap74 is mediated by a large nonpolar interface with only one direct salt-bridge interaction (Kamada et al., 2003). This extensive nonpolar interface, as well as the large, negative heat capacity that is inferred from independently fitting the ITC data at each temperature, suggests that hydrophobic forces may drive the coupled folding–binding interaction between FCP1 and Rap74. Moreover, these considerations suggest that the large, negative heat capacity primarily reflects the desolvation of nonpolar surface area. Accordingly, we interpreted the measured thermodynamic parameters

with a simple hydrophobic model described by Baldwin (1986). In this model, $T_H = 295$ K is the temperature at which the enthalpy of hydrophobic interactions vanishes, while $T_S = 386$ is the temperature at which the entropy of hydrophobic interactions vanishes. Within the context of this model, ΔC_p is assumed to describe hydrophobic effects. In contrast, ΔH_0 and ΔS_0 are interpreted as temperature-independent contributions to $\Delta H°(T)$ and $\Delta S°(T)$ that are *not* due to hydrophobic effects; e.g., ΔH_0 and ΔS_0 are interpreted to describe contributions from electrostatic interactions.

When $T_H = 295$ K and $T_S = 386$ K are predetermined by the Baldwin model, the global fit depends upon three parameters. Using the MATLAB script Fcn_Fit_Cp_Ndil.m (distributed through the Penn State ScholarSphere as described elsewhere; see *Protocol for data fitting* below for details), we determined $\Delta C_p = -1.01 \pm 0.04$ kJ/mol/K, $\Delta H_0 = -2.24 \pm 0.23$ kJ/mol, and $\Delta S_0 = -0.17 \pm 0.01$ kJ/mol/K (Fig. 5). Indeed, these parameters are quite consistent with the interpretation of the Baldwin model. The best-fit ΔC_p is in good agreement with the relation proposed by Spolar and Record (1994) for relating solvated surface areas to heat

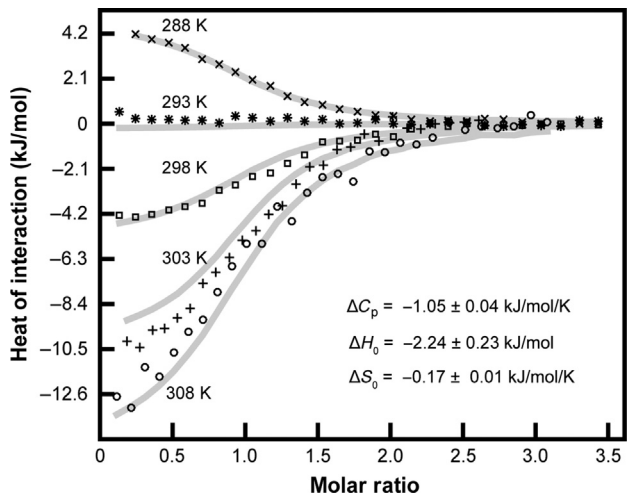

Figure 5 Global fit of temperature-dependent ITC titrations in MATLAB yields a direct estimate of the constant-pressure heat capacity change (ΔC_p), as well as temperature-independent contributions to the overall enthalpy change (ΔH_0) and entropy change (ΔS_0). The output best-fit parameters are displayed as an inset. Calculated isotherms resulting from the global fit are displayed as a gray line for each temperature, while the experimentally observed integrated heats of injection are displayed at 288 K (×), 293 K (*), 298 K (□), 303 K (+), and 308 K (○).

capacity changes. The best-fit ΔH_0 is consistent with the presence of a small number of favorable polar interactions at the protein–protein interface. Finally, the best-fit ΔS_0 is consistent with the expectation of the rotational and translational entropy loss due to the association of the two proteins, although ΔS_0 is somewhat smaller than that estimated by Spolar and Record. We previously suggested that this difference may be due to the "fuzziness" (Tompa & Fuxreiter, 2008) that we have previously observed in the Rap74–FCP1 protein complex (Lawrence & Showalter, 2012; Wostenberg et al., 2011).

In closing, we note that although the results of the global fit are quite consistent with the Baldwin model, the fitting procedure does not fundamentally depend upon this model. Rather, the Baldwin model provides a molecular interpretation of the parameters that are determined by the global analysis. In particular, the Baldwin model relates ΔC_p to hydrophobic effects, while attributing ΔH_0 and ΔS_0 to other sources. However, the only essential assumption of the global fitting procedure we describe is that ΔC_p is constant over the experimental temperature range, which is identical to the assumption made for serial fitting as described in Sections 4.1 and 4.2. The definition of $T_H = 295$ K and $T_S = 386$ K only results in trivial shifts in ΔH_0 and ΔS_0, while having no effect on ΔC_p. Thus, our thermodynamic model can be readily applied for interactions that are not dominated by hydrophobic interactions.

Protocol for data fitting

1. At the end of an ITC experiment, the user will initially integrate the heat released or absorbed in each injection, as outlined in Section 4.1.
2. Following integration, the data series containing injection volume (v_i), ligand concentration (X_T), macromolecule concentration (M_T), molar ratio (X_T/M_T), and heat of injection ($n$dh) tabular form can be accessed in Origin® through the "Windows" menu tab. These data are saved in a standard ASCII text format that is directly used by our provided MATLAB scripts as inputs after the following changes are made. The first line must be commented out if headers are saved through Origin®; the last line must also be commented out in order to discard the last point. Macromolecule and ligand here refer to the contents in the cell and the syringe, respectively.
3. The ITC data files generated in step 2, along with the active cell volume (V_0), syringe concentration of ligand (X_0), experiment temperature (T) in K, and number of initial points to exclude (npd, which will typically be at least 1), are provided as inputs into the MATLAB fitting macro

Table 2 An Example of the Master Input File That Is Provided to the Macro Fcn_Fit_Cp_Ndil.m for Global Fitting of Temperature-Dependent ITC Data Sets

File_name.DAT	T	V_0	X_0	npd
15C.DAT	288	1427.47	0.3	2
20C.DAT	293	1427.47	0.3	1
25C.DAT	298	1427.47	0.3	1
30C.DAT	303	1427.47	0.3	2
35C.DAT	308	1427.47	0.3	1

(Fcn_Fit_Cp_Ndil.m) in a separate user written, tab delimited text file. A sample of what this file will look like is provided as Table 2. Running this protocol will produce best-fit values for ΔC_p, ΔH_0, and ΔS_0.

5. SUMMARY

There is significant interest in establishing general theories to describe IDP interactions (Levy, Cho, Onuchic, & Wolynes, 2005; Liu, Faeder, & Camacho, 2009; Whitford, Sanbonmatsu, & Onuchic, 2012). Quantitative analysis of temperature-dependent ITC data sets offers an unequaled opportunity to evaluate structure–function relationships for IDPs. Through these studies, it is possible to establish the energetic consequences of intrinsic disorder and to define the driving forces for disordered protein interactions, which is especially interesting when binding is coupled to a folding transition in the IDP. This chapter provides a framework for designing, executing, and analyzing temperature-dependent ITC studies in IDP interactions. The analysis stage of the research protocol documented here is largely achieved using MATLAB macros to facilitate either serial or global fitting of the temperature-dependent data series. All of the MATLAB scripts needed for this protocol are freely available through the stable URL: https://scholarsphere.psu.edu/downloads/5712mr21f.

In the closing section of this chapter, we have provided a specific example of a procedure implemented in MATLAB for global fitting of temperature-dependent ITC data that has the benefit of directly estimating the apparent ΔC_p for the transition(s) being monitored in the titration. Although our case study involved coupled folding and binding, the thermodynamic framework presented is not limited to describing interactions of this

nature; however, we do note that the functional form derived is only rigorously applicable to one-to-one binding. With this caveat in mind, it is our general finding that the global fitting protocol we have established here is preferable to serial fitting of temperature-dependent data. The principle strength of this method is its reliance on directly fitting to fundamental thermodynamic equations, from which any desired equation of state may be calculated. Extension of the model to incorporate additional thermodynamic observables (e.g., pressure effects, the influence of added osmolytes, or macromolecular crowding) is straightforward. Of practical significance, this fitting method reliably incorporates data from (nearly) isoenthalpic titrations, making these data sets valuable contributors to the overall constraint of the parameters. In order to optimally utilize our protocol, we recommend balancing data collection over a wide temperature range and the need to sample where ΔC_p for the transition(s) is constant. In practice, this means that the upper temperature limit should be ~10 °C below the thermal denaturation midpoint of any folded biomolecules. For best results, we find that collecting data at a minimum of 4–6 temperatures provides a robust fit, with passage through the isoenthalpic point preferred if possible. Although we have only shown a single titration at each temperature for clarity in the figures presented here (Fig. 5), we generally find fitting of triplicate titrations at each selected temperature to provide a more robust sampling of the errors associated with the measurements.

For the case study we present, our experiments demonstrate that the FCP1–Rap74 interaction is readily reversed and characterized by a modest equilibrium dissociation constant of 2–5 μ*M* under nearly all assayed conditions. This suggests that intrinsic disorder may provide an important functional advantage by making IDP interactions more robust with respect to certain local perturbations, although this appears somewhat in contrast to the notion that IDPs function as molecular rheostats. Moreover, our ITC protocol reveals a large and negative ΔC_p associated with this interaction. Without making any assumptions regarding the chemical details of the binding interface, this large and negative ΔC_p suggests that the FCP1–Rap74 interaction is driven by hydrophobic forces resulting from desolvating the binding surface. Furthermore, as a consequence of this large and negative ΔC_p, the binding interaction switches from entropically driven to enthalpically driven near room temperature. This emphasizes the importance of ITC experiments at multiple temperatures because ITC experiments performed at a single temperature would fail to reveal this transition. This transition is particularly important, as it is generally believed

that the entropic penalty associated with folding IDPs plays a critical role in regulating their moderate binding affinity. Consequently, we assert that conducting variable-temperature ITC experiments provides a powerful framework for elucidating the thermodynamic forces that drive IDP interactions, as well as for revealing the unique biophysical properties and functional advantages associated with intrinsic disorder.

ACKNOWLEDGMENTS

We would like to thank the Penn State Automated Biomolecular Calorimetry Facility and its staff for the vital role they play in our research program and for providing helpful discussion during the preparation of this manuscript. We also thank Dr. Sushant Kumar for assistance validating the MATLAB scripts presented. This work was funded in part by an NSF CAREER award (MCB-0953918) to S.A.S. and a NIH predoctoral fellowship (F31GM101936) to M.B. This work was also supported by fellowships from the Sloan Foundation and the Penn State Institute for Cyberscience awarded to WGN. We thank the Penn State ScholarSphere team (http://scholarsphere.psu.edu) for its help in establishing permanent and stable web hosting for the products of our research program.

REFERENCES

Baldwin, R. L. (1986). Temperature dependence of the hydrophobic interaction in protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, *83*, 8069–8072.

Burnouf, D., Ennifar, E., Guedich, S., Puffer, B., Hoffmann, G., Bec, G., et al. (2012). kinITC: A new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *Journal of the American Chemical Society*, *134*, 559–565.

Cervantes, C. F., Bergqvist, S., Kjaergaard, M., Kroon, G., Sue, S. C., Dyson, H. J., et al. (2011). The RelA nuclear localization signal folds upon binding to IkappaBalpha. *Journal of Molecular Biology*, *405*, 754–764.

Cho, S., Swaminathan, C. P., Bonsor, D. A., Kerzic, M. C., Guan, R., Yang, J., et al. (2010). Assessing energetic contributions to binding from a disordered region in a protein–protein interaction. *Biochemistry*, *49*, 9256–9268.

Chodera, J. D., & Mobley, D. L. (2013). Entropy–enthalpy compensation: Role and ramifications in biomolecular ligand recognition and design. *Annual Review of Biophysics*, *42*, 121–142.

Cooper, A. (2010). Protein heat capacity: An anomaly that maybe never was. *Journal of Physical Chemistry Letters*, *1*, 3298–3304.

Cooper, A., Johnson, C. M., Lakey, J. H., & Nollmann, M. (2001). Heat does not come in different colours: Entropy–enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions. *Biophysical Chemistry*, *93*, 215–230.

Csermely, P., Palotai, R., & Nussinov, R. (2010). Induced fit, conformational selection and independent dynamic segments: An extended view of binding events. *Trends in Biochemical Sciences*, *35*, 539–546.

Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry*, *29*, 7133–7155.

Drobnak, I., De Jonge, N., Haesaerts, S., Vesnaver, G., Loris, R., & Lah, J. (2013). Energetic basis of uncoupling folding from binding for an intrinsically disordered protein. *Journal of the American Chemical Society*, *135*, 1288–1294.

Dunker, A. K., Silman, I., Uversky, V. N., & Sussman, J. L. (2008). Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology*, *18*, 756–764.
Dyson, H. J., & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Current Opinion in Structural Biology*, *12*, 54–60.
Ferreon, J. C., & Hilser, V. J. (2004). Thermodynamics of binding to SH3 domains: The energetic impact of polyproline II (PII) helix formation. *Biochemistry*, *43*, 7787–7797.
Fuxreiter, M., Tompa, P., Simon, I., Uversky, V. N., Hansen, J. C., & Asturias, F. J. (2008). Malleable machines take shape in eukaryotic transcriptional regulation. *Nature Chemical Biology*, *4*, 728–737.
Ganguly, D., Otieno, S., Waddell, B., Iconaru, L., Kriwacki, R. W., & Chen, J. (2012). Electrostatically accelerated coupled binding and folding of intrinsically disordered proteins. *Journal of Molecular Biology*, *422*, 674–684.
Ghai, R., Falconer, R. J., & Collins, B. M. (2012). Applications of isothermal titration calorimetry in pure and applied research—Survey of the literature from 2010. *Journal of Molecular Recognition*, *25*, 32–52.
Gibbs, E. B., & Showalter, S. A. (2015). Quantitative biophysical characterization of intrinsically disordered proteins. *Biochemistry*, *54*, 1314–1326.
Kamada, K., Roeder, R. G., & Burley, S. K. (2003). Molecular mechanism of recruitment of TFIIF-associating RNA polymerase C-terminal domain phosphatase (FCP1) by transcription factor IIF. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 2296–2299.
Kumar, S., Showalter, S. A., & Noid, W. G. (2013). Native-based simulations of the binding interaction between RAP74 and the disordered FCP1 peptide. *Journal of Physical Chemistry B*, *117*, 3074–3085.
Lacy, E. R., Filippov, I., Lewis, W. S., Otieno, S., Xiao, L. M., Weiss, S., et al. (2004). p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nature Structural & Molecular Biology*, *11*, 358–364.
Lawrence, C. W., Bonny, A., & Showalter, S. A. (2011). The disordered C-terminus of the RNA Polymerase II phosphatase FCP1 is partially helical in the unbound state. *Biochemical and Biophysical Research Communications*, *410*, 461–465.
Lawrence, C. W., Kumar, S., Noid, W. G., & Showalter, S. A. (2014). Role of ordered proteins in the folding-upon-binding of intrinsically disordered proteins. *Journal of Physical Chemistry Letters*, *5*, 833–838.
Lawrence, C. W., & Showalter, S. A. (2012). Carbon-detected N-15 NMR spin relaxation of an intrinsically disordered protein: FCP1 dynamics unbound and in complex with RAP74. *Journal of Physical Chemistry Letters*, *3*, 1409–1413.
Levy, Y., Cho, S. S., Onuchic, J. N., & Wolynes, P. G. (2005). A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. *Journal of Molecular Biology*, *346*, 1121–1145.
Liu, J., Faeder, J. R., & Camacho, C. J. (2009). Toward a quantitative theory of intrinsically disordered proteins and their function. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 19819–19823.
Mizoue, L. S., & Tellinghuisen, J. (2004). The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. *Analytical Biochemistry*, *326*, 125–127.
Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K., et al. (2006). Analysis of molecular recognition features (MoRFs). *Journal of Molecular Biology*, *362*, 1043–1059.
Nguyen, B. D., Abbott, K. L., Potempa, K., Kobort, M. S., Archambault, J., Greenblatt, J., et al. (2003). NMR structure of a complex containing the TFIIF subunit RAP74 and the RNA polymerase II carboxyl-terminal domain phosphatase FCP1. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 5688–5693.

Papoian, G. A., & Wolynes, P. G. (2003). The physics and bioinformatics of binding and folding-an energy landscape perspective. *Biopolymers*, *68*, 333–349.
Pierce, M. M., Raman, C. S., & Nall, B. T. (1999). Isothermal titration calorimetry of protein–protein interactions. *Methods*, *19*, 213–221.
Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical Journal*, *78*, 1606–1619.
Spolar, R. S., & Record, M. T., Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. *Science*, *263*, 777–784.
Tantos, A., Han, K. H., & Tompa, P. (2012). Intrinsic disorder in cell signaling and gene transcription. *Journal of Molecular and Cellular Endocrinology*, *348*, 457–465.
Tompa, P. (2002). Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, *27*, 527–533.
Tompa, P., & Fuxreiter, M. (2008). Fuzzy complexes: Polymorphism and structural disorder in protein–protein interactions. *Trends in Biochemical Sciences*, *33*, 2–8.
van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., et al. (2014). Classification of intrinsically disordered regions and proteins. *Chemical Reviews*, *114*, 6589–6631.
Whitford, P. C., Sanbonmatsu, K. Y., & Onuchic, J. N. (2012). Biomolecular dynamics: Order–disorder transitions and energy landscapes. *Reports on Progress in Physics*, *75*, 076601.
Wiseman, T., Williston, S., Brandts, J. F., & Lin, L. N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry*, *179*, 131–137.
Wostenberg, C., Kumar, S., Noid, W. G., & Showalter, S. A. (2011). Atomistic simulations reveal structural disorder in the RAP74-FCP1 complex. *Journal of Physical Chemistry B*, *115*, 13731–13739.
Wright, P. E., & Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology*, *19*, 31–38.
Xu, Y., Oruganti, S. V., Gopalan, V., & Foster, M. P. (2012). Thermodynamics of coupled folding in the interaction of archaeal RNase P proteins RPP21 and RPP29. *Biochemistry*, *51*, 926–935.