THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

BIOPHYSICAL CHARACTERIZATION OF TRBP AND PDX1

DECLAN EVANS
SPRING 2017

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Biochemistry and Molecular Biology
and Mathematics
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Scott A. Showalter
Associate Professor of Chemistry
Associate Professor of Biochemistry and Molecular Biology
Thesis Supervisor

Wendy Hanna-Rose
Associate Professor of Biochemistry and Molecular Biology
Honors Adviser

*Signatures are on file in the Schreyer Honors College.
Abstract

Proteins are a diverse class of biomolecule with extreme importance to life. As the number of discovered proteins increases so too does the diversity of this class of molecule. Grand dogma that once governed our understanding of all proteins are being narrowed into specific classes. As a result old methods of studying proteins are becoming obsolete and new ones must be refined. This study focuses on two frontiers of protein chemistry; protein- RNA interactions and intrinsically disordered proteins.

The discovery of catalytic RNA in the 1980’s forever changed understanding of enzymes. The focus of enzymology shifted from exclusive study of proteins to include RNA and protein- RNA complexes. RNA is now known to play key roles in many different regulatory pathways. One widely studied example is RNA interference (RNAi), a type of post- transcriptional gene regulation. RNAi involves the use of a protein- RNA complex called the RNA induced silencing complex (RISC). The RISC is composed of microRNA (miRNA) and several proteins including Dicer and TRBP. While there is extensive data as to the catalytic nature of the RISC, little is known about the protein-RNA interface. This study examines the HIV-1 TAR RNA binding protein (TRBP) and the thermodynamics associated with its RNA binding event. Isothermal titration calorimetry (ITC) is used to determine this binding event is not coupled with any large- scaled distortion in the A formed geometry of the RNA.

Classical understanding of proteins has been rooted in our understanding of structure. Earliest crystalized proteins consisted of well formed domains of α-helices or β-sheets, and as a result, most analytic techniques are designed to probe these properties of proteins. Recently however, an
increasing number of proteins have been discovered that possess little to no secondary structure. Conventional methods have little success analyzing these proteins and novel methods must be created and refined. Here we examine the pancreatic and duodenal homeobox 1 (Pdx1), a member of the homeodomain family of transcription factors lacking large scale secondary structure outside of its DNA binding homeodomain. This protein is of particular interest due to its ties to diabetes. In this study we investigate the structure of Pdx1 and how changes can be linked to disease.
Table of Contents

List of Figures

List of Tables

Acknowledgements

1 Introduction
   1.1 TRBP .................................................. 2
      1.1.1 RNAi Pathway ........................................ 2
      1.1.2 Double Stranded RNA Binding Domains .............. 2
      1.1.3 RNA Bending ......................................... 3
      1.1.4 Significance ........................................ 3
   1.2 Pdx1 .................................................. 4
      1.2.1 Intrinsically Disordered Proteins ................. 4
      1.2.2 Pdx1 Structure ....................................... 5
      1.2.3 Pdx1 Function ....................................... 5
      1.2.4 Significance ........................................ 6

2 Materials and Methods .................................... 18
   2.1 Protein Expression ..................................... 19
   2.2 Protein Purification ................................... 19
   2.3 RNA Purification ....................................... 19
   2.4 Differential Scanning Calorimetry ..................... 20
   2.5 Isothermal Titration Calorimetry ...................... 20
   2.6 Campari Monte-Carlo Simulations ....................... 21

3 Results ..................................................... 22
   3.1 Differential Scanning Calorimetry ................. 23
   3.2 Isothermal Titration Calorimetry .................. 23
   3.3 Campari Monte-Carlo Simulations ................... 23

4 Discussion .................................................. 35
   4.1 TRBP .................................................. 36
   4.2 Pdx1 .................................................. 37
List of Figures

1.1 The miRNA maturation pathway ............................................. 7
1.2 dsRBDs in the miRNA maturation pathway ................................. 8
1.3 TRBP-dsRBD2 crystal structure ............................................. 9
1.4 Structure of pri-miR-16-1 miRNA ......................................... 10
1.5 RNA bending upon binding to dsRBDs ..................................... 11
1.6 Folding funnel of proteins .................................................. 12
1.7 Pdx1 homeodomain bound to DNA ....................................... 13
1.8 Structure of Pdx1 ............................................................ 14
1.9 The role of Pdx1 .............................................................. 15
1.10 Pdx1 mutations linked to diabetes ......................................... 16
1.11 A Visual Representation of PRE ............................................ 17

3.1 DSC results ....................................................................... 25
3.2 ITC results ....................................................................... 26
3.3 Rg vs Temperature of Simulation ......................................... 28
3.4 Rg Distribution ................................................................... 29
3.5 Asphericity of Pdx1c compared to R_g .................................. 30
3.6 Simulated PRE data- wild type Pdx1c ................................... 31
3.7 Simulated PRE data- wild type Pdx1c ................................... 32
3.8 PRE Plots Using Multiple NMR Methods .............................. 33

4.1 MD results ....................................................................... 38
List of Tables

3.1 ITC results ................................................................. 27
3.2 $\chi^2$ analysis of cutoff distances ................................... 34
Acknowledgements

I would first like to thank Dr. Showalter and every member of the Showalter Lab group for their support and guidance. The research conducted could not have been done without the support of the Erickson Discovery Grant. Lastly, I would like to thank my family for their unconditional support and my friends for making these last four years so memorable. Balder-D
Chapter 1

Introduction
1.1 TRBP

1.1.1 RNAi Pathway

Understanding of post-transcriptional gene regulation was forever altered in the mid 1990’s following the discovery of RNA silencing pathways, specifically RNAi. Since its discovery, RNAi has been shown to be vital to eukaryotic gene regulation. Manipulation of RNAi by invasive species is detrimental to human cells and has been linked to health issues including cancer and viral infections [1][2]. However, RNAi can also be used for a number of antiviral treatments and has shown promising results against the HIV virus[3]. Human RNAi employs a class of RNA called micro RNAs (miRNA) which consist of roughly 22 nucleotides when fully mature[4]. All canonical miRNA are transcribed in the nucleus as primary micro RNA (pri-miRNA) that mature to precursor micro RNA (pre-miRNA) through cleavage by Drosha[5]. pre-miRNA is transported out of the nucleus by Exportin-5 and is cleaved into mature miRNA by Dicer in the cytoplasm[5]. Dicer cleavage is affected by the presence of a cofactor TRBP[6][7]. Mature miRNA, Dicer and TRBP are then incorporated into the RISC, an enzyme complex that intercepts and degrades mRNA before translation[8]. A summary of this process is shown in Figure 1.1.

1.1.2 Double Stranded RNA Binding Domains

Not surprisingly, many of the proteins involved in the miRNA pathway contain at least one double stranded RNA binding domain (dsRBD). Figure 1.2 summarizes some of the dsRBDs found in several of the miRNA maturation proteins. A dsRBD is characterized by an $\alpha - \beta - \beta - \beta - \alpha$ motif, however amino acid sequence and loop size vary by dsRBD[9]. Most dsRBDs bind RNA along the minor groove [10] in a non-sequence specific manner[11] as seen in Figure 1.3. Interestingly, RNA binding affinity of each dsRBD varies greatly throughout the pathway. For example, TRBP-dsRBD2 has an unusually high affinity(see chapter 3) whereas there is no evidence to suggest TRBP-dsRBD3 and Drosha-dsRBD are capable of binding RNA[11]. This raises questions as
to the nature and mechanism for dsRBD binding.

1.1.3 RNA Bending

Previous findings as well as crystal structures (Figure 1.3) suggest dsRBDs generally recognize double-stranded RNA by shape rather than sequence[10]. Previous studies in our lab have shown dsRBD binding is confined to Watson-Crick basepairs and is excluded from any internal loops or bulges due to imperfections[11] features common in maturing miRNA (Figure 1.4). Specifically, several studies have shown dsRBDs recognize the A formed geometry of the RNA [12][13]. These observations conflict with crystal structures of several dsRBDs bound to RNA which suggest dsRBDs bind with large scale bending of the RNA. The earliest evidence of this is in the crystal structure of Xlrbpα-2 bound to a GC 10-mer dsRNA (Figure 1.5A) [14]. Analogously, there is strong evidence that shows the B formed geometry of DNA can bend up to 90° upon binding certain proteins [15][16]. There are additional crystal structures of some proteins, specifically *Aquifex aeolicus* RNase III and the two dsRBDs from ADAR2, that do not show any distortion of RNA A-form geometry upon binding [17] [18]. These conflicting observations complicates our understanding of the protein- RNA interface and should be investigated for clarity. Here we use TRBP-dsRBD2 binding a Watson-Crick 20-mer RNA as a model for this event. In this study, we use Isothermal Titration Calorimetry (ITC) as a primary means of determining thermodynamic parameters of this binding interaction. Measured thermodynamics will be used to infer conformational changes following a method first used in a study by Spolar and Record [19].

1.1.4 Significance

Manipulation of RNAi has incredible potential for improvements in a variety of fields from medicine and crop development [20][21][22]. Specifically in the miRNA pathway the activity of Dicer has varying effects on health from eyesight to cancer [23][24]. In order to fully realize the potential of RNAi, it is essential we better understand this process. While much is known about the catalytic event of mRNA degradation, the nature of the protein- RNA interface remains largely
unknown. Given that Dicer also recognizes dsRNA by shape[25], large-scale RNA conformational changes induced by TRBP binding could have profound effects on RNA processing. In this study we address the question of RNA bending induced by binding to TRBP in an effort to better understand the process in which miRNA develop.

1.2 Pdx1

1.2.1 Intrinsically Disordered Proteins

Since their discovery, proteins have been regarded as some of the most important molecules in living systems. Proteins are involved in nearly every cellular process and can serve a variety of purposes. Our current understanding of protein function is largely rooted in protein structure. Classically, it is argued that biologically active proteins adopt a single, well-folded structure that directly influence that protein’s function, however new evidence has challenged this understanding. Intrinsically disordered proteins (IDPs) are a recently discovered class of proteins that readily transition between many different conformations. Figure 1.6 compares characteristic folding funnels of well-folded and intrinsically disordered proteins. While classically well-folded proteins occupy a single lowest-energy conformation, IDPs can occupy and readily switch between several low-energy conformations.

Several studies have shown IDPs are biologically active in a variety of processes[26][27]. Additional studies have shown IDPs are relatively common proteins, with 33.0% of eukaryotic proteins exhibiting a disordered region of over 30 amino acids[27]. Transcription factors in particular tend to be especially disordered and most known transcription factors are predicted to exhibit disorder of 73 – 94%[28]. In this experiment, we study the function of Pdx1, an example of such a transcription factor.
1.2.2 Pdx1 Structure

Homeodomain transcription factors are an important aspect of both prokaryotic and eukaryotic gene expression[29]. Thousands of these proteins have been discovered across eukaryotic species and all share a conserved homeodomain, a roughly 60 amino acid sequence with a helix-turn-helix conformation[30]. This region is involved in DNA binding and allows the protein to activate gene expression as a transcription factor (Figure 1.7)[30]. Mammalian homeodomain proteins are expressed in a highly tissue specific manner and alterations in tissue expression produce remarkable phenotypes[31]. Several human diseases can be linked to proteins in this class of transcription factors. An example of this is the human pancreatic duodenal homeobox protein 1 (Pdx1, formerly known as the insulin promoter factor 1 or IFP1) which is expressed in the pancreas and largely controls insulin expression, linking mutations in this protein to diabetes[32]. Pdx1 contains the conserved and well structured homeodomain, however the flanking N and C terminus are long disordered regions with little to no secondary structure (Figure 1.8). While the ordered homeodomain can be crystalized with relative ease, the N- and C- termini do not fold cooperatively and as a result less is known about them.

1.2.3 Pdx1 Function

The primary role of Pdx1 in mature organisms is as an insulin transcription factor. Pdx1 is also essential for the development of the gut endoderm, playing key parts in both pancreatic development and $\beta$-cell maturation [33]. Specifically, Pdx1 is responsible for early embryonic pancreatic development [34] as homozygous knockout mice fail to develop a pancreas [35]. In $\beta$-cell formation, Pdx1 is responsible for the activation of insulin transcription[36]. A summary of this process is shown in Figure 1.9.
1.2.4 Significance

Given the importance of Pdx1 in pancreatic development, it is no surprise Pdx1 is linked to diseases such as diabetes. Previous studies have linked mutations in Pdx1 to maturity-onset diabetes of the young 4 (MODY4), a type of diabetes that typically onsets in the mid-thirties with earliest cases beginning at 17 years[37][38]. Figure 1.10 represents twelve sites on the structure of Pdx1 that are directly linked to this form of diabetes [39]. In this study, a mutation of the intrinsically disordered C termini (E224K) is examined for its effects on the local structure. Any structural changes observed will be investigated for possible ties to MODY4. Structures generated by Campari Monte-Carlo simulations are the primary focus of this study and will be used to predict paramagnetic relaxation enhancement (PRE) plots. PRE is a technique in which a MTSL group is covalently attached to a cystine residue. Residues in close proximity to the MTSL produce a notably lower signal than residues farther away (Figure 1.11). This makes PRE a helpful tool in determining three dimensional structures of proteins. Simulated data will be compared to experimental data collected by another member of the lab.
Overview of the miRNA maturation pathway. pri-miRNA is transcribed in the nucleus by RNA pol II. pri-miRNA is cleaved by Drosha and transported into the cytosol by XPO5. Once in the cytosol, pre-miRNA is cleaved by Dicer to form fully matured miRNA which is incorporated into the RISC. The RISC intercepts and degrades mRNA before it can be translated, thus acting as a gene-regulatory agent.
dsRBDs in the miRNA maturation pathway. Many proteins involved in the miRNA maturation pathway have at least one dsRBD. Each dsRBD has a characteristic $\alpha - \beta - \beta - \beta - \alpha$ motif, however amino acid and loop length vary by dsRBD. RNA binding affinity varies by dsRBD.
Figure 1.3: TRBP-dsRBD2 crystal structure

Crystal structure of TRBP-dsRBD2 bound to a 10mer of double stranded RNA suggesting TRBP-dsRBD2 recognizes the shape of the RNA rather than the sequence (PDB 3ADI).
**Figure 1.4: Structure of pri-miR-16-1 miRNA**

![miRNA Structure](image)

**Structure of pri-miR-16-1 miRNA.** Most miRNA contain several Watson-Crick imperfections. Shown is the sequence of a biologically active miRNA pri-miR-16-1. Bases shown in red correspond to Watson-Crick imperfections which exclude binding by TRBP.
RNA bending upon binding to dsRBDS Crystal structures of Xlrbpa2 (left, PDB ID: 1DI2) indicate bending of the A-formed helical geometry of the RNA upon binding. Contrasting this observation is the crystal structure of ADAR2-dsRBM2 (right, PDB ID: 2I3C) which shows less deformation of the RNA helix.
Folding funnel of proteins. Well-folded proteins exhibit a deep folding funnel with one lowest energy conformation (left). Proteins in this well require large amounts of energy to change conformations. IDPs have a much more shallow folding funnel with multiple lowest energy conformations (right). As a result, IDPs are much more capable of changing conformation and do so freely in solution.
Figure 1.7: Pdx1 homeodomain bound to DNA

Pdx1 homeodomain bound to DNA. Crystal structure showing the isolated Pdx1 homeodomain bound to DNA (PDB ID: 2H1K). Unlike the RNA binding proteins mentioned earlier, Pdx1 binds DNA in a sequence-specific manner [40]
Structure of Pdx1. A structural representation of Pdx1 in its entirety. The core homeodomain (grey) is the only region of the protein which exhibits secondary structure. The N- and C- termini (blue) are largely unstructured and can take a number of conformations.
The role of Pdx1. Taken with permission from [36]. Pdx1 has several important roles in pancreatic development. Shown is the effect of Pdx1 on β cell maturation.
Pdx1 mutations linked to diabetes. A diagram of the Pdx1 homeodomain bound to DNA. The N and C termini are not involved in DNA binding. Grey circles represent eleven locations throughout Pdx1 with mutations linked to diabetes. Note here mutations are present on both the homeodomain as well as the N and C termini, indicating their importance in insulin production.
**Figure 1.11: A Visual Representation of PRE**

A Visual Representation of PRE. Two PDBs of Pdx1-C generated by EOM simulations. The cystine 227 residue is highlighted in red for both. Amino acids in close proximity to this cystine (shown in the red ball) give no signal by PRE due to the added MTSL. By varying the radius of the sphere in simulations (15Å in this figure) we can get data to more closely fit the observed data.
Chapter 2

Materials and Methods
2.1 **Protein Expression**

BL21 cells expressing the desired constructs were grown in a solution of 10 mL LB 10 µL Kanamycin at 37 °C for 8 hours. The cells were then added to a 50 mL LB 50 µL Kanamycin growth at 37 °C for 16 hours. This solution was then split into 4 cultures each of 1000 mL LB 1 mL Kanamycin and incubated at 37 °C. Optical density was measured every 30 min with fixed wavelength of 600 nm to assess cell growth. Once the cells reached an absorbance of 0.8, IPTG was added to achieve a final concentration of 0.5 mM and the temperature was dropped to 23 °C for 16 hours.

2.2 **Protein Purification**

Cultures were removed from incubation, all following processes were carried out at 4 °C. The cells were centrifuged at 4200 × g for 20 min and the supernatant was removed. The pellets were resuspended in lysis buffer (50 mM Tris pH 7.5, 500 mM NaCl) and lysed by sonication (1 sec on/ 2 sec off for 2 total min at 50% amplitude). The lysates were centrifuged at 11,500 × g for 1 hour. The supernatant was then treated with 750 µL PEI and centrifuged again at 11,500 × g for 1 hour. The supernatant was mixed with 12.65 g ammonium sulfate and centrifuged again at 11,500 × g for 1 hour. The supernatant was decanted over a Ni-His column. The proteins were eluted with elution buffer (50 mM Tris pH 7.5, 500 mM NaCl, 2000 mM imidazole). The elutents were dialyzed with 3C protease in dialysis buffer (50 mM Tris pH 7.5, 300 mM NaCl) at 4 °C for 16 h. The samples were then again passed over a Ni-His column. Purity was measured by SDS-PAGE gel electrophoresis.

2.3 **RNA Purification**

A synthetic self-complementary RNA of sequence 5′-GCGCGCGCGCGCGCGCGC-3′ (GC20) was ordered from Dharmacon (GE Healthcare, Chicago, IL). RNA was deprotected upon
arrival following a protocol set by the manufacturer. RNA concentration and purity were determined by UV spectroscopy using an extinction coefficient of 165,000 L · mol$^{-1}$ · cm$^{-1}$. RNA duplexation was performed by heat denaturation at 90 °C for 45 s, followed by fast annealing at 4 °C for 5 min.

### 2.4 Differential Scanning Calorimetry

In order to determine the optimal temperature range for ITC studies, thermal stability of each species, namely; TRBP-RBD2, and GC20 RNA, was monitored by Differential Scanning Calorimetry (DSC) using a Microcal VP-Capillary DSC (Malvern, Westborough, MA). In each experiment, each sample was equilibrated in ITC buffer (50 mM sodium cacodylate pH 7.0, 50 mM potassium glutamate) and diluted to working concentrations (80 µM TRBP-dsRBD2 and 25 µM GC20 RNA). Experiments were conducted in triplicate alongside a buffer reference with temperature varying from 30 °C to 80 °C at a scan rate of 90 °C hr$^{-1}$. The reference buffer run was subtracted from the experimental data and peaks were determined using the cursor initiated non-two state program. The data was then fitted individually using the Origin 7.0 DSC software. Fitted data was then plotted using in-house Matlab scripts.

### 2.5 Isothermal Titration Calorimetry

Before each titration, TRBP-dsRBD2 and GC20 RNA were codialized in ITC buffer overnight. Each species was diluted to working conditions (80 µM TRBP-dsRBD2 and 25 µM GC20 RNA). The GC20 RNA was added to the sample cell and the TRBP-dsRBD2 was added to the syringe. Each titration was carried out on a Microcal VP-ITC DSC (Malvern, Westborough, MA). Experiments were conducted in triplicate at temperatures 10 °C, 15 °C, 20 °C, and 25 °C and. Each set of experiments was averaged using a best-fit parameter and analyzed using the one-set-of-sites Matlab functions. Thermodynamic values directly measured were the equilibrium association constant ($K_A$), change in enthalpy ($\Delta H$), and binding stoichiometry (n). These values were used to
calculate dissociation constant \( K_D = K_A^{-1} \), Gibbs free energy of binding \( \Delta G = RT \ln(K_D) \) and change in entropy \( \Delta G = \Delta H - T\Delta S \). Errors were calculated using standard error propagation.

### 2.6 Campari Monte-Carlo Simulations

Campari Monte-Carlo simulations were generated using the Campari molecular modeling software (v2) developed by Rohit V. Pappu and Andreas Vitalis. Random pdb files were generated for the amino acid sequence for Pdx1c (residues 202-283) for twenty different temperatures from 230 K to 420 K at 10 K increments. Ten thousand pdb files were generated for each temperature. Generated pdb files were processed using in-house Matlab scripts to calculate the desired results (see Campari Monte-Carlo Simulations (results)).
Chapter 3

Results
3.1 Differential Scanning Calorimetry

Calorimetry experiments were able to successfully determine an appropriate temperature range for ITC titrations. Figure 3.1 shows the average values of the heat capacity ($C_p$) as the samples were heated from 30 °C to 80 °C. Melting temperature ($T_m$) was determined and reported in Figure 3.1. Positive values for $\Delta C_p$ correspond to thermally denatured species [41].

3.2 Isothermal Titration Calorimetry

In our study, heat capacity change due to binding interactions ($\Delta C_p$) was measured using a set of temperature-dependent ITC titrations. These titrations produced various thermodynamic values associated with this event and the results are summarized in Table 3.1. Most notably, the change in enthalpy $\Delta H$ was observed for the various temperatures. The recorded values of $\Delta H$ for each time are recorded in Figure 3.2. This information was used to calculate $\Delta C_p$ using the established thermodynamic relation $\Delta C_p = (\partial \Delta H / \partial T)_T$. Calculation revealed a $\Delta C_p$ value of $-70 \pm 40 \text{cal} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$.

3.3 Campari Monte-Carlo Simulations

Campari simulations were useful in predicting structure models for the C-terminus of Pdx1. These structures were validated by calculating radius of gyration $R_g$ for each temperature for both the mutant and wild type Pdx1-C. Figure 3.3 shows the calculated $R_g$ for each temperature. As expected, $R_g$ increases with temperature indicating some expansion of the coil. $R_g$ does not vary linearly with temperature, however, and a lower-limit of 15 Å is observed in both the mutant and wild type protein. This same limit is observed in the distribution of $R_g$ for each temperature (Figure 3.4). Interestingly, for temperatures 230 K to 280 K, $R_g$ of the mutant is noticeably larger than that of the wild type. Additionally, shape was quantified as asphericity and plotted against $R_g$ for several temperatures (Figure 3.5. This data follows the trend that indicates the peptide is
expanding at higher temperatures.

Paramagnetic relaxation enhancement (PRE) plots were also simulated for the temperature replicate series (Figures 3.6 3.7) as well as for varying NMR experiments (Figure 3.8). Simulated PRE data relied heavily on a "Cutoff Distance" variable which indicated the distance (in Å) to which PRE signal of a given amino acid was affected by proximity to the cystine residue. Statistical analysis for a best fit model is given by chi-squared analysis is shown in table 3.2.
Figure 3.1: DSC results

DSC results. Taken (with permission) from [42]. Heat capacity plotted against temperature for TRBP-dsRBD2 and GC20-RNA. Each curve represents the average value obtained from triplicate runs of each species heated independently. Positive values of $\Delta C_p$ indicate thermal denaturation. This data was used to determine an appropriate temperature range for ITC titrations.
ITC results. Taken (with permission) from [42]. A Shows heat released due to binding of the titrations. Heat curves of four different temperatures are overlaid and shown in this figure. B Represents the change in enthalpy ($\Delta H$) of titrations for the four temperatures used. Change in heat capacity ($\Delta C_p$) was calculated from this result using the relation $\Delta C_p = \left( \frac{\partial \Delta H}{\partial T} \right)_p$. 
Table 3.1: ITC results

<table>
<thead>
<tr>
<th>T (K)</th>
<th>[TRBP-dsRBD2] (µM)</th>
<th>ΔH (kcal·mol⁻¹)</th>
<th>ΔG (kcal·mol⁻¹)</th>
<th>K_D (µM)</th>
<th>n (sites)</th>
</tr>
</thead>
<tbody>
<tr>
<td>283</td>
<td>74</td>
<td>−6.90 ± 0.20</td>
<td>−8.00 ± 0.05</td>
<td>0.69 ± 0.07</td>
<td>4.86 ± 0.09</td>
</tr>
<tr>
<td>288</td>
<td>60</td>
<td>−6.10 ± 0.10</td>
<td>−8.38 ± 0.06</td>
<td>0.45 ± 0.05</td>
<td>4.79 ± 0.08</td>
</tr>
<tr>
<td>293</td>
<td>60</td>
<td>−5.62 ± 0.09</td>
<td>−8.70 ± 0.05</td>
<td>0.33 ± 0.03</td>
<td>4.86 ± 0.06</td>
</tr>
<tr>
<td>298</td>
<td>62</td>
<td>−5.87 ± 0.09</td>
<td>−8.81 ± 0.05</td>
<td>0.35 ± 0.03</td>
<td>4.93 ± 0.06</td>
</tr>
</tbody>
</table>

Average value for thermodynamic constants determined by ITC along with their associated errors.
Figure 3.3: $R_g$ vs Temperature of Simulation

*Rg vs Temperature of Simulation.* Average Rg along with associated error calculated using in-house Matlab scripts plotted versus temperature of the simulation.
**Figure 3.4: Rg Distribution**

The distribution of Rg over several temperatures. Figures 3.4 A, B refer to 230 K, 3.4 C, D refer to 330 K, and 3.4 E, F refer to 420 K.
Asphericity of Pdx1c compared to $R_g$. To relate size and shape, asphericity (the quantification of how similarly a peptide resembles a sphere) is plotted for each simulation. Figures 3.5 A, B refer to 230 K, 3.5 C, D refer to 330 K, and 3.5 E, F refer to 420 K.
Simulated PRE data- wild type Pdx1c. Predicted plots for PRE experiments of wild type Pdx1c at 230 K (Figure 3.6A), 330 K (Figure 3.6B), 420 K (Figure 3.6C). Solid lines correspond to simulated results and plotted points correspond to preliminary data obtained from a member of our lab.
Figure 3.7: Simulated PRE data- E224K Pdx1c

Simulated PRE data- E224K Pdx1c Predicted plots for PRE experiments of E224K Pdx1c at 230 K (Figure 3.7A), 330 K (Figure 3.7B), 420 K (Figure 3.7C). Solid lines correspond to simulated results and plotted points correspond to preliminary data obtained from a member of our lab.
Figure 3.8: PRE Plots Using Multiple NMR Methods

**PRE Plots Using Multiple NMR Methods.** Different NMR methods used to collect PRE data can yield different results. Figure 3.8A shows the predicted plot generated by the CON method and Figure 3.8B shows the predicted plot generated by the HACACON method.
<table>
<thead>
<tr>
<th>Cutoff Distance (Å)</th>
<th>$\chi^2$ CON</th>
<th>$\chi^2$ HACACON</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.4783109</td>
<td>9.9882150</td>
</tr>
<tr>
<td>2</td>
<td>6.4783109</td>
<td>9.9882150</td>
</tr>
<tr>
<td>3</td>
<td>6.4783109</td>
<td>9.9882150</td>
</tr>
<tr>
<td>4</td>
<td>5.0620067</td>
<td>8.2320967</td>
</tr>
<tr>
<td>5</td>
<td>4.9828692</td>
<td>8.1285370</td>
</tr>
<tr>
<td>6</td>
<td>4.1302328</td>
<td>7.1170977</td>
</tr>
<tr>
<td>7</td>
<td>6.9724801</td>
<td>6.3289259</td>
</tr>
<tr>
<td>8</td>
<td>3.2799897</td>
<td>5.4698376</td>
</tr>
<tr>
<td>9</td>
<td>3.5859388</td>
<td>4.7876851</td>
</tr>
<tr>
<td>10</td>
<td>17.840672</td>
<td>8.8687448</td>
</tr>
<tr>
<td>11</td>
<td>2.8589424</td>
<td>3.3667463</td>
</tr>
<tr>
<td>12</td>
<td>4.8049505</td>
<td>3.3959896</td>
</tr>
<tr>
<td>13</td>
<td>28.353450</td>
<td>12.069819</td>
</tr>
<tr>
<td>14</td>
<td>4.4756021</td>
<td>2.2909217</td>
</tr>
<tr>
<td>15</td>
<td>10.249786</td>
<td>4.0387226</td>
</tr>
<tr>
<td>16</td>
<td>49.877107</td>
<td>19.354737</td>
</tr>
<tr>
<td>17</td>
<td>1622.6962</td>
<td>735.67824</td>
</tr>
<tr>
<td>18</td>
<td>46.508646</td>
<td>18.993235</td>
</tr>
<tr>
<td>19</td>
<td>269.12100</td>
<td>110.16089</td>
</tr>
<tr>
<td>20</td>
<td>2465.1208</td>
<td>1425.8546</td>
</tr>
</tbody>
</table>

$\chi^2$ analysis of cutoff distances. Cutoff distance is directly related to how close an amino acid to the labeled cystine. Represented is a statistical analysis of how closely simulated PRE plots match preliminary data.
Chapter 4

Discussion
4.1 TRBP

Although TRBP lacks the catalytic properties of Dicer, its presence has been shown to increase the accuracy and precision of the cleavage event[43][44]. Likewise in the microprocessor complex, DGCR8 has been shown to greatly affect Drosha’s ability to process pri-miRNA[45]. These effects of double stranded RNA binding proteins on miRNA maturation promote questions as to the nature of their binding mechanism in an effort to better understand miRNA maturation as a whole. Here, we examined the hypothesis that dsRBDs induce a strong bend in dsRNA upon binding. We experimentally measured the thermodynamics of TRBP-dsRBD2 binding a 20-mer dsRNA as a model system for this event. Temperature dependent ITC experiments yielded a near-zero positive $\Delta C_p$, suggesting large-scale conformational changes are unlikely attributed to this event. Additionally, these results were compared to MD simulations (performed by another member of our lab) which suggest there are no appreciable changes in major and minor groove withs in the dsRNA upon binding (Figure 4.1) [42].

Our findings are somewhat at odds with the literature. There is extensive evidence that indicate all protein- dsRNA binding interactions are accompanied by a large conformational change in the protein or the RNA [46][47][48]. However recent findings suggest some protein- dsRNA binding follows a lock-and-key model rather than an induced fit model[17][18][49]. As more is known about these dsRBDs and their interactions with dsRNA, it may become obvious no one domain can serve as an adequate model for the whole system.

Regarding TRBP-dsRBD2 specifically, our results are consistent with previous studies in our lab which demonstrated TRBP-dsRBD2 preferentially binds the A formed geometry of dsRNA over deformations[11]. As helical deformations (such as Watson-Crick mismatches and loops) are common for pre-miRNA, this property of TRBP-dsRBD2 is likely more significant in miRNA recognition than any small- scaled conformational changes upon dsRNA binding. Future studies of this event should be targeted at how manipulation of this binding affects Dicer processing of pre-miRNA.
4.2 Pdx1

The lack of structure of Pdx1 should not detract from our understanding of the protein. In this study, we were able to predict attributes of $R_g$ and asphericity from simulated pdb files in order to better understand the properties of both mutant and wild type Pdx1. The average $R_g$ undergoes an expected increase with temperature indicating the protein unwinds at higher temperatures. The lower-limit of $R_g$ is also as follows due to the finite volume of the protein itself. Interestingly, at temperatures 230-280 K the $R_g$ for mutant Pdx1 is notably larger than that of the wild type, indicating the mutation may influence the proteins ability to compact. This observation is consistent with low temperature plots of $R_g$ (Figure 3.4) and asphericity (Figure 3.5).

Simulated PRE plots successfully predicted data that was consistent with preliminary results. Temperature series plots varied as expected with temperature. Simulations at lower temperatures show more residues are closer to the labeled cystine indicating the protein size decreases as temperature decreases. Additionally for varying NMR experiments, statistical analysis was able to predict the best-fit parameters. As expected, HACACON experiments yield a wider trough than CON experiments. The impact of these plots will not be evident until more complete experimental data can be obtained.

Current results cannot definitively state the effect of the E224K mutation on diabetes. While it appears this mutation affects the protein’s ability to compact at low temperatures we must still draw the connection between this observation and any phenotypic consequence. As the C terminus of Pdx1 interacts with the Pdx1 C terminus-interacting factor-1 (Pcif1, also known as SPOP)\cite{50}\cite{51}, additional studies will be focused on the impact of this mutation on this interaction.
MD results. Taken (with permission) from [42]. MD results conducted in tandem with the ITC experiments of this study by a member of our lab. These MD simulations measured the groove with before and after binding. No appreciable changes were observed, supporting the claim that binding by TRBP does not bend the RNA in any significant way.
Bibliography


[50] Kathryn C Claiborn, Mira M Sachdeva, Corey E Cannon, David N Groff, Jeffrey D Singer,

Education

**B.S. With Honors in Biochemistry and Molecular Biology (2017)**
Undergraduate degree with honors from the Pennsylvania State University in Biochemistry and Molecular Biology with a focus in biochemistry.

**B.S. in Mathematics (2017)**
Undergraduate degree from the Pennsylvania State University in Mathematics.

**Minor in Physics (2017)**
Undergraduate minor from the Pennsylvania State University in Physics.

Professional Experience

**Showalter Lab Group, Pennsylvania State University, PA**
**Undergraduate Researcher**

01.2016 - 05.2017
Designed and conducted biophysical studies of the disordered regions of the Pdx1 transcription factor and their relation to diabetes.

**Regeneron Pharmaceuticals, Rensselaer, NY**
**Intern, QC Assay Validation**
05.2016 - 08.2016
Involved in assay validation for the QC chemistry department. Wrote and carried out protocols in a GMP setting to test for repeatability, linearity, and robustness of quality control procedures. Assays learned include: vapor-pressure osmometry, high-performance liquid chromatography, peptide mapping.

**UpGrade Tutoring, State College, PA**
**Tutor**
01.2017 - 07.2017
Privately tutored students through the UpGrade tutoring company in Calculus 1 & 2 as well as in General Physics: Mechanics.
Honors and Awards

Erickson Discovery Grant (2015)
I wrote, submitted and was awarded the Erickson Discovery Grant for research the summer of 2015 to continue my work on the microRNA project in the Showalter lab. I presented the work I conducted with this grant at the Penn State Undergraduate Research Poster Presentation in the spring of 2016.

Eagle Scout (2013)
I received my Eagle Scout award from the Boy Scouts of America in May of 2013. My project involved the planning and construction of a shed for a community garden in my area.

Publications