The Pennsylvania State University

The Graduate School

Eberly College of Science

# THERMODYNAMIC CHARACTERIZATION OF THE INTERACTION OF THE TRANSCRIPTION FACTOR PDX1 AND ITS REGULATORY PARTNER SPOP

A Thesis in

Chemistry

by

Krystal Diaz



Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science

December 2018

The thesis of Krystal Diaz was reviewed and approved* by the following:

Scott A. Showalter
Associate Professor of Chemistry and Biochemistry and Molecular Biology
Thesis Advisor
Graduate Program Chair

Philip C. Bevilacqua
Distinguished Professor of Chemistry and Biochemistry and Molecular Biology

Amie K. Boal
Assistant Professor of Chemistry and Biochemistry and Molecular Biology

*Signatures are on file in the Graduate School.

# Abstract

The homeodomain transcription factor Pdx1 is essential for pancreatic development, function, and maintenance. Mutations in *pdx1* lead to impairments in beta cell function, glucose tolerance, and insulin secretion that result in diabetic phenotypes. Most clinically documented Pdx1 sequence variants map to disordered regions of the protein whose function is not well-characterized. Our group has confirmed that the disordered C-terminus of Pdx1 participates in essential regulatory interactions and that mutations within this region disrupt binding.

The current work aims to elucidate mechanistic detail of the interaction of the Pdx1-C terminal domain with the MATH domain of the E3 ubiquitin ligase adaptor SPOP. This is achieved through the thermodynamic and structural characterization of the Pdx1-SPOP complex, which reveals binding affinities and atomistic details describing the binding event. The effects of multivalency have been assessed through calorimetric and spectroscopic binding assays using minimal peptides encompassing individual Pdx1 binding sites. The results herein provide a molecular level understanding of this interaction crucial to glucose homeostasis.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| a.a. | Amino Acid |
| $A_{600}$ | Absorbance at 600 nm |
| BACK | BTB and C-terminal Kelch domain |
| CD | Circular Dichroism |
| $\Delta C_p$ | Change in molar heat capacity |
| DNA | Deoxyribonucleic acid |
| DTT | Dithiothreitol |
| FT-IR | Fourier Transform Infrared Spectroscopy |
| GLUT2 | Glucose transporter 2 |
| GST | Glutathione-S-transferase |
| $\Delta H$ | Change in molar enthalpy |
| $His_6$ | 6X Histidine tag |
| HPLC | High Performance Liquid Chromatography |
| HSQC | Heteronuclear Single Quantum Coherence |
| IDR | Intrinsically disordered protein region |
| IDP | Intrinsically disordered protein |
| *ins* | Insulin gene |
| IPTG | isopropyl β-D-1-thiogalactopyranoside |
| ITC | Isothermal titration calorimetry |
| $K_a$ | Equilibrium binding association constant |
| $K_d$ | Equilibrium binding dissociation constant |
| LB | Lysogeny Broth |
| M | Macromolecule |
| MALDI | Matrix assisted laser desorption ionization |
| MATH | Meprin and Traf Homology Domain |
| MODY4 | Mature-onset diabetes of the young 4 |
| $M_{tot}$ | Total macromolecule concentration |
| MWCO | Molecular Weight Cutoff |
| MX | Concentration of macromolecule-ligand complex |
| n | Reaction stoichiometry |
| Ni-NTA | $Ni^{2+}$-nitrilotriacetic acid |
| NMR | Nuclear Magnetic Resonance |
| p300 | Histone acetyltransferase p300 |
| PDAC | Pancreatic ductal adenocarcinoma |
| Pdx1 | Pancreatic and duodenal homeobox 1 |
| Pdx1 S1 | Pdx1 binding site 1 |
| Pdx1 S2 | Pdx1 binding site 2 |

| | |
|---|---|
| Pdx1-C | C-terminus of Pdx1 |
| Pdx1-N | N-terminus of Pdx1 |
| PRE | Paramagnetic Resonance Enhancement |
| PUC | Puckered protein |
| $Q_{obs}$ | Observed fluorescence quenching |
| RING | Really Interesting New Gene Domain |
| SAXS | Small-angle X-ray scattering |
| SET7 | SET domain containing lysine methyltransferase 7 |
| SLiM | Short linear interaction motif |
| SPOP | Speckle-type POZ protein |
| WT | Wild-type |
| <X> | Average value of ligands bound |
| X | Ligand |
| $X_{free}$ | Concentration of free ligand |
| $\lambda_{EM,,max}$ | Wavelength of maximal fluorescence intensity |
| $\lambda_{emission}$ | Emission wavelength |
| π | Polar residue |
| Φ | Nonpolar residue |

# Chapter 1: Introduction

## 1.1 General Introduction

### *Intrinsically Disordered Proteins*

The study of protein chemical structure offers insight into protein biological function. The traditional hypothesis regarding protein function suggests that the formation of a unique tertiary structure is requisite for biological activity; thus, the research motivated by this hypothesis has limited structural studies to folded proteins for much of the twentieth century.[1–3] The importance of disordered protein regions was not widely understood until the mid-1990s, during which experimental and bioinformatics studies revealed that over 50% of human protein-coding genes encode for regions of disorder upwards of 30 amino acids in length.[4,5] Transcription factors are a class of proteins highly enriched in such disordered regions, which suggests that protein disorder has a distinct functional role. Relative to folded proteins, the amino acid composition of intrinsically disordered protein regions (IDRs) is enriched in polar and charged residues and lacking in the hydrophobic residues required for nucleation of protein folding.[6] As such, these protein regions are unable to spontaneously fold into a stable three-dimensional structure in their native state.[4,7]

The free energy landscape of an intrinsically disordered protein (IDP) and a stably folded protein differ drastically. A well-folded protein has a distinct global energy minimum that corresponds to its cooperatively folded state; in contrast, the potential energy surface of an unbound IDP is thought to be rugged with no discernible global minimum, and so the IDP adopts an ensemble of conformations with similar energies.[6] Intrinsically

disordered proteins commonly undergo coupled folding and binding, during which disordered regions fold upon association with their binding partners.[3,4] However, many IDPs do not fold upon binding and instead remain in an extended conformation. Even so, bound IDPs sample a much smaller conformational space than unbound IDPs and usually display a well-defined energy minimum corresponding to a stable bound structure. Many characteristics of IDPs draw parallels with well-folded proteins, further supporting the now widely accepted view that disorder imparts function and complements the function of structured protein regions.[2,3,4]

A common feature of IDRs is the presence of short linear motifs (SLiMs) that participate in interactions with multiple partners. To ameliorate an otherwise weak or potentially unfavorable interaction, IDPs often engage in multivalent interactions using multiple SLiMs in order to enhance overall binding affinity through avidity effects.[8,9] This behavior is advantageous in biological pathways because it allows for combinatorial amplification of the required signal, confers additional specificity to the binding interaction, and increases the overall affinity of the IDP for its target. Relative to folded proteins, the conformational diversity and increased flexibility of IDPs render them ideal candidates for the control of tightly regulated biological processes such as cellular signaling and transcription activation.

***Intrinsic Disorder in Transcription Factors***

Transcription initiation in eukaryotes is tightly controlled by interactions of transcription factors with proteins and DNA. Transcription factors are highly enriched in IDRs relative to other protein families.[10] These disordered regions are known to regulate DNA binding and mediate protein-protein interactions pivotal to transcription regulation;

variations within such regions often disrupt essential steps in biological processes and are associated with disease pathologies.[4,7,10,11] The prevalence of IDRs within transcription factors and the association of their aberrant expression with disease highlights their physiological importance and drives the hypothesis that structural disorder has a distinct functional role.[1,10] While the functions of many transcription factor IDRs are known, there remain many unanswered questions regarding the relation between the amino acid sequence and functional outputs of IDRs in transcriptional regulation. Basic mechanistic insight into sequence-dependent transcriptional regulation by IDRs is lacking, as these flexible regions are difficult to characterize by traditional methods.

The present study seeks to accomplish a complete biophysical characterization of the insulin-regulating transcription factor Pancreatic and Duodenal Homeobox 1 (Pdx1) and its complexes to increase understanding of its role in transcriptional regulation and the etiology of disease.

***Pdx1: Understanding the Role of IDRs in Transcription***

When blood glucose levels are elevated, glucose enters pancreatic β-cells and induces insulin production. The entry of glucose into the cell triggers a series of events that result in insulin secretion and the eventual reduction of blood glucose levels (Figure 1).[12–14] The insulin transactivation pathway, critical to glucose homeostasis, is under tight control through the action of complex, multi-step processes that allow for fine adjustments in insulin secretion in response to glucose metabolism.

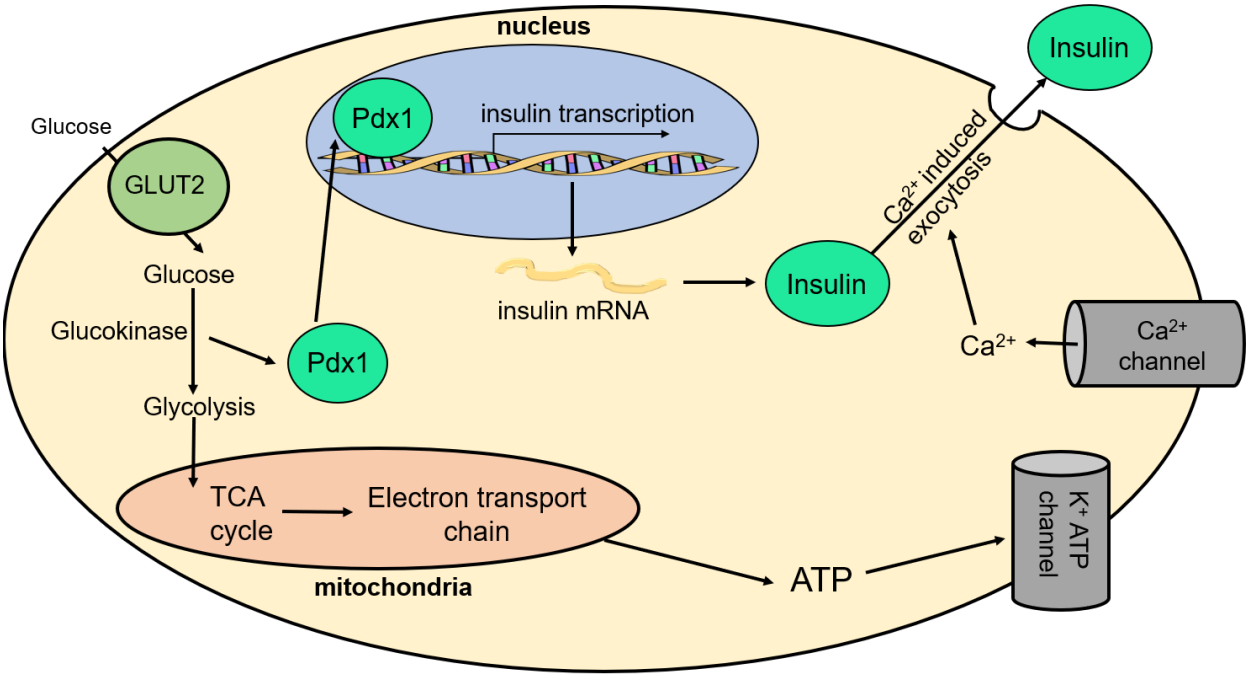



**Figure 1.** Schematic of pancreatic β-cell displaying some of the glucose-dependent processes that promote insulin production and secretion.

Pdx1 is essential for glucose-regulated pancreatic development, function, and maintenance. During embryonic development, Pdx1 is required for the differentiation of all pancreatic cell types. Its expression is later localized mostly to pancreatic β-cells, where it plays a broad role in the transcription and secretion of insulin and other β-cell-specific genes.[15–18] Pdx1 is hypothesized to play a role in glucose sensing, insulin biosynthesis, and insulin exocytosis.[19] Pdx1 activates the transcription of glucose transporter 2 (GLUT2) and glucokinase, which are key components in β-cell glucose sensing.[20,21] Under high glucose conditions, Pdx1 binds to *ins* A-box promoter elements with the core consensus sequence 5'-TAAT-3' and recruits various transcription factors to initiate insulin transcription.[14,22,23] For instance, Pdx1 mediates the recruitment of the lysine methyltransferase SET7 and the histone acetyltransferase p300 required for insulin transcription initiation.[20,24]

Though most common forms of type I and type II diabetes are polygenic, Pdx1 is also implicated in rare forms of monogenic adult-onset diabetes resultant from Pdx1 mutations that cause changes to its amino acid sequence.[25,26] Homozygous deletions of single *pdx1* nucleotides lead to frameshift mutations that result in the complete failure of pancreatic development, whereas heterozygous mutations in *pdx1* lead to impairments in β-cell function, glucose tolerance, and insulin secretion that can result in mature onset diabetes of the young 4 (MODY4).[23,26–29] Dysregulation of Pdx1 expression is linked to the pathogenesis of type II diabetes through improper maintenance of β-cell identity with eventual reduction in β-cell mass and insulin secretion.[16,17,19,30–32] Additionally, aberrant expression of Pdx1 is associated with the development of several cancers, namely pancreatic ductal adenocarcinoma (PDAC).[18,33,34] However, its role in the development and proliferation of cancer cells is not clearly understood as the current literature on this topic displays contradictory results. For example, many past studies suggest that Pdx1 is an oncogene, with its overexpression in PDAC cell lines linked to proliferation, invasiveness, and growth.[35,36] Contrarily, recent studies show that Pdx1 loss is associated with an aggressive subtype of PDAC, suggesting its role as a tumor surpressor.[18,29,37] In addition to the evidence presented here, there is overwhelming evidence that demonstrates Pdx1 has a distinct functional role in the development and prevention of disease. Because the biological function of Pdx1 is already well-known, it serves as a useful model system for the study of the role of IDRs in transcription. In addition to cellular studies, structural and biophysical characterization of Pdx1 and its interaction partners is required to develop a full understanding of the role of intrinsically disordered transcription factors in diabetes and pancreatic cancer.

### *Pdx1 Structure*

Human Pdx1 is a 283 amino acid, 31 kDa protein that contains a homeodomain flanked by two intrinsically disordered N- and C-terminal tails.[22,28] The central homeodomain binds to DNA and contains a nuclear localization signal.[26,38–40] The N-terminal tail (Pdx1-N) is involved in the recruitment of histone modification enzymes that promote open chromatin structure at the *ins* gene to make it accessible to other transcription factors and the transcriptional machinery.[24,28,41–43] The C-terminal tail (Pdx1-C) has recently been identified as a region required for regulation of Pdx1 stability.[28,40,44]

Crystallographic data for the homeodomain bound to DNA gives mechanistic insight into the interaction of Pdx1 with DNA (Figure 2), but structural characterization of

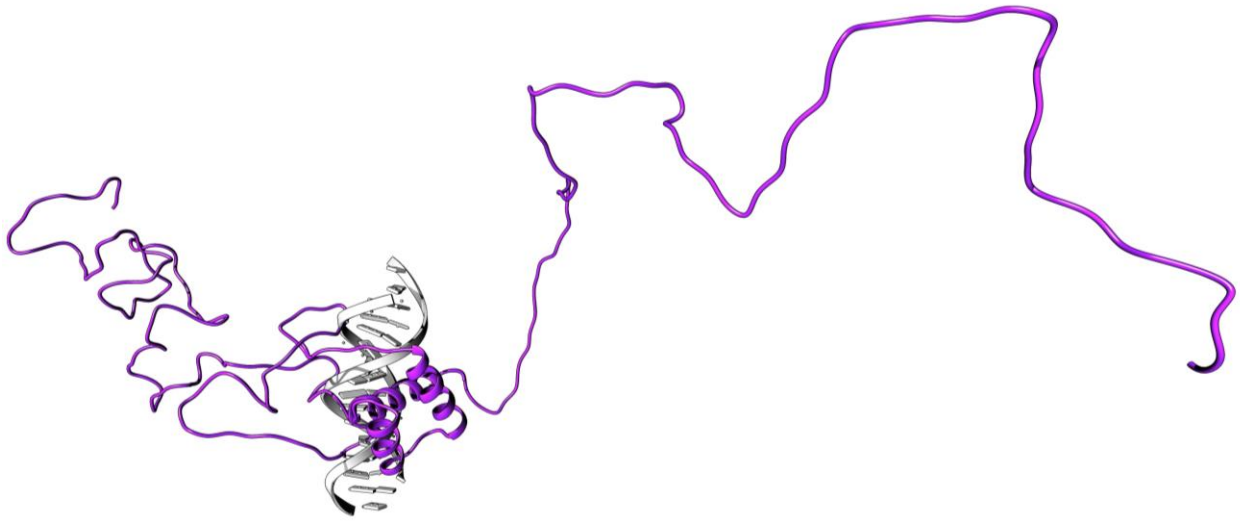


**Figure 2.** Structure of Pdx1 adapted from the Pdx1 homeodomain crystal structure (PDB ID 2H1K). Disordered tails were modeled using MODELLER.

the disordered termini is hindered by challenges associated with conventional methods of structure determination.[27] Disorder predictions suggest that the tails of Pdx1 are completely disordered (Figure 3), which was validated by the Showalter lab based on

nuclear magnetic resonance (NMR), circular dichroism (CD), and small-angle x-ray scattering (SAXS) structural studies of the protein (not shown).

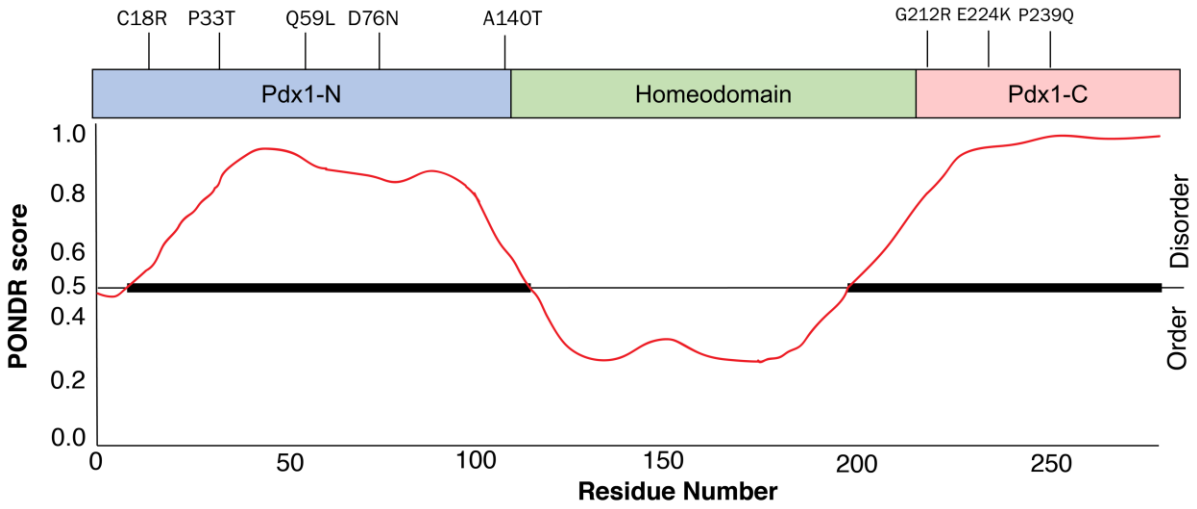



**Figure 3.** Domain architecture of Pdx1 and the diabetes-associated variants of its IDRs mapped onto PONDR disorder predictions. A PONDR score > 0.5 suggests disorder and a score < 0.5 suggests order.

Interestingly, of the 11 clinically documented diabetes-associated Pdx1 sequence variants, only two variants map to the folded homeodomain. The remaining nine variants map to disordered regions of the protein, further highlighting the vital role these disordered regions have in proper Pdx1 function. Several mutations are found within a regulatory region of the disordered C-terminal domain (Figure 3) that contains segments required for maintenance of Pdx1 stability. The remainder of this thesis will focus on the biophysical characterization of Pdx1-C and its interaction with speckle-type POZ protein (SPOP), an E3 ubiquitin ligase adaptor involved in Pdx1 degradation.

***Pdx1 Interacts with Speckle-type POZ protein (SPOP) Leading to Ubiquitination***

Speckle-type POZ protein (SPOP) is a cullin3 RING ligase substrate adaptor that recruits substrates for ubiquitin-mediated proteasomal degradation. SPOP targets its substrates by recognition of primary sequence motifs of Φ-π-S-S/T- S/T (Φ nonpolar, π

polar) through its MATH domain, which binds with its amino acid residues Y87, F102, Y123, and F133.[9,45,46] Co-crystal structures of SPOP MATH domain bound to short peptide substrates show that target peptides bind to a narrow canonical groove on the MATH domain in an extended state (Figure 4).[46] The BTB domain of SPOP is responsible

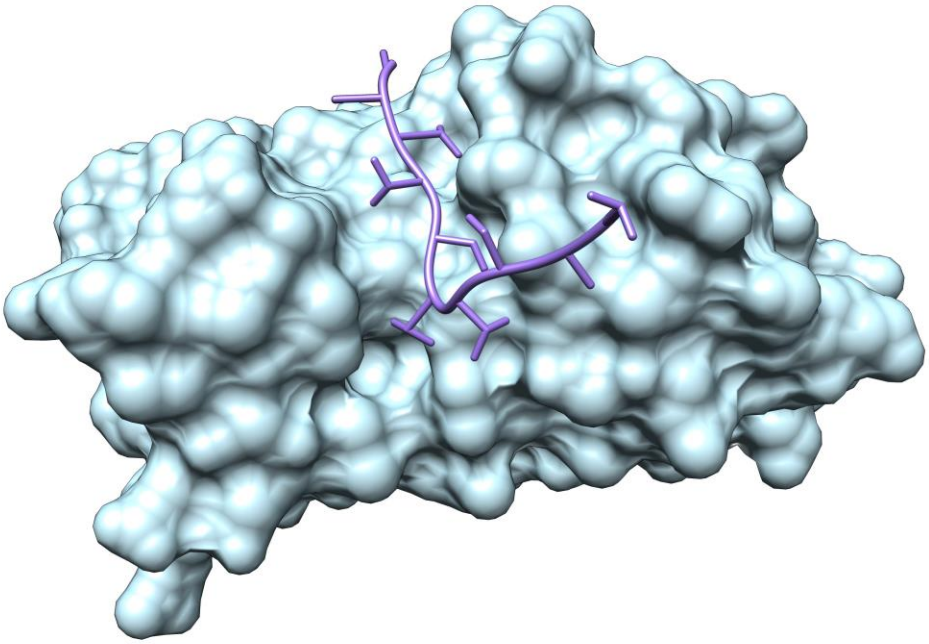


**Figure 4.** Crystal structure of SPOP MATH domain bound to a PUC peptide (PDB ID 3IVV).

for its dimerization, and its BACK domain facilitates further self-association into higher-order oligomers and subsequent aqueous phase separation into nuclear speckles.[9,47,48] In the presence of substrates, SPOP oligomers are known to cause aqueous phase separation into nuclear speckles.[47,48] The ability to form homo-oligomers allows for the binding of multiple MATH domains to the linear motifs of multivalent SPOP substrates. The combination of the weak individual interactions between SPOP and such motifs increases the overall affinity of the binding interaction through avidity effects. Additionally,

the shared multivalency of SPOP and its target substrates may confer additional specificity to their binding interaction and the regulation of substrate levels.[9]

SPOP mutations are associated with several cancers, including prostate and endometrial cancer.[9,49] A majority of SPOP mutations are centered on its canonical binding groove and disrupt substrate binding.[46,47,49] SPOP also has a role as a tumor suppressor in various cancers.[50,51] Understanding the mechanism of the SPOP:Pdx1 interaction will contribute to current SPOP literature by providing additional mechanistic detail into SPOP-substrate interactions.

***Pdx1-C as a Regulatory Region via Multivalent Interaction with SPOP***

Glucose homeostasis requires tight regulatory control of not only insulin expression, but also the expression of the transcription factors and transcription co-activators responsible for insulin production. When blood glucose levels are low, insulin transcription is downregulated by the ubiquitin-mediated proteasomal degradation of Pdx1.[28] The switch-like modulation of Pdx1 levels is an important facet of insulin regulation that facilitates the maintenance of glucose homeostasis. In order to develop a rationale for understanding the role of Pdx1 in insulin production, it is essential to understand Pdx1 degradation pathways and how their dysregulation is associated with disease.

Yeast-two hybrid screens identified SPOP in association with the Pdx1 C-terminus.[28] Co-transfection of Pdx1 with increasing amounts of SPOP showed that SPOP inhibits Pdx1 transcriptional activity in a dose-dependent manner.[28] Further study of this interaction found that transactivation by mouse, rat, human, and zebrafish Pdx1 were similarly reduced by SPOP, which led the authors to conclude that Pdx1-C may contain

conserved interaction motifs.[40] Current Pdx1 literature suggests that Pdx1-C is required for proper insulin transcription. Probing the molecular interactions between SPOP and Pdx1-C is therefore an interesting avenue to pursue due to the role of this interaction in glucose homeostasis.

Amino acids 210-238 of Pdx1 are conserved across multiple species, which suggests the presence of a molecular recognition motif at this site.[28,40] Co-transfection of Pdx1 and Pdx1Δ210-238 with SPOP found SPOP in association with full-length Pdx1 but not Pdx1Δ210-238, indicating that this conserved region is required for Pdx1-SPOP binding.[40] However, no additional confirmation of this binding site is present in the literature, which motivated our lab to verify this binding site through nuclear magnetic resonance (NMR) spectroscopy and thermodynamic binding assays.

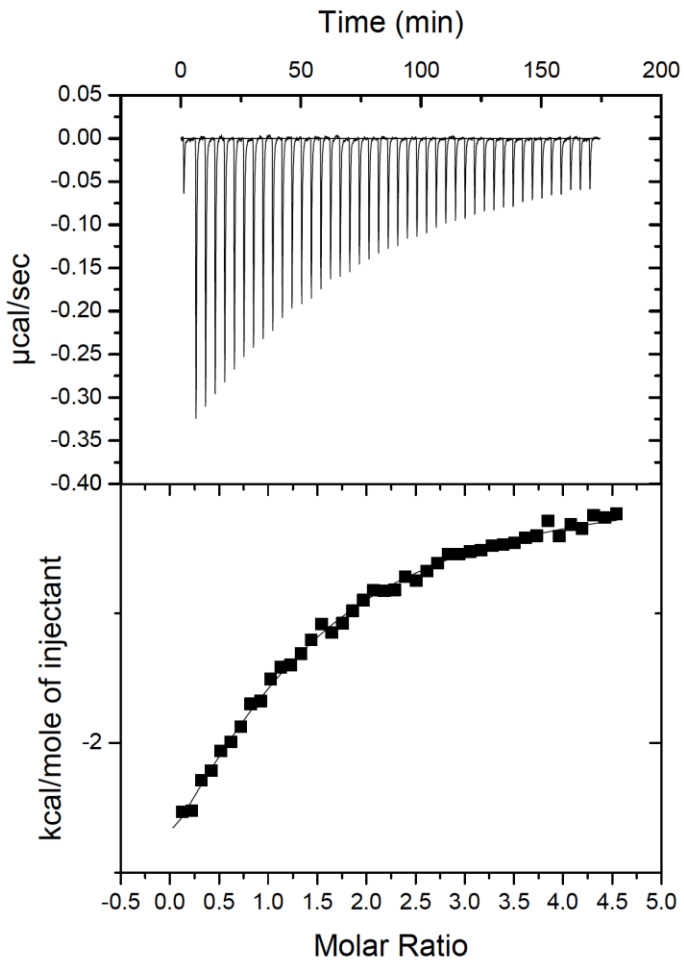



**Figure 5.** ITC thermogram and binding isotherm for titration of 500 μM Pdx1-C into 20 μM SPOP MATH 1-166 at 298 K. The estimated $K_d$ is 25 μM. From Monique Bastidas' Ph.D. thesis, 2015.[52]

The Showalter lab is well-suited for biophysical characterization of IDP interactions that were first identified in cells through the use of carbon-detect NMR and thermodynamic binding assays such as isothermal titration calorimetry (ITC). Former Showalter Lab graduate student Monique Bastidas confirmed that Pdx1-C directly binds to SPOP MATH 1-166 with a $K_d$ ~25 μM through isothermal titration calorimetry (Figure 5).[52] Monique then proceeded to confirm the $Pdx1_{210-238}$ interaction site in vitro by

collecting $^{15}N$,$^{13}C$-CON spectra of unbound $^{15}N$, $^{13}C$ Pdx1-C (Pdx1 a.a. 206-283) and the SPOP-Pdx1-C complex (Figure 6). Pdx1 residues involved in SPOP binding experience a change in chemical environment, which in turn perturbs Pdx1-C chemical shifts relative to unbound Pdx1-C.

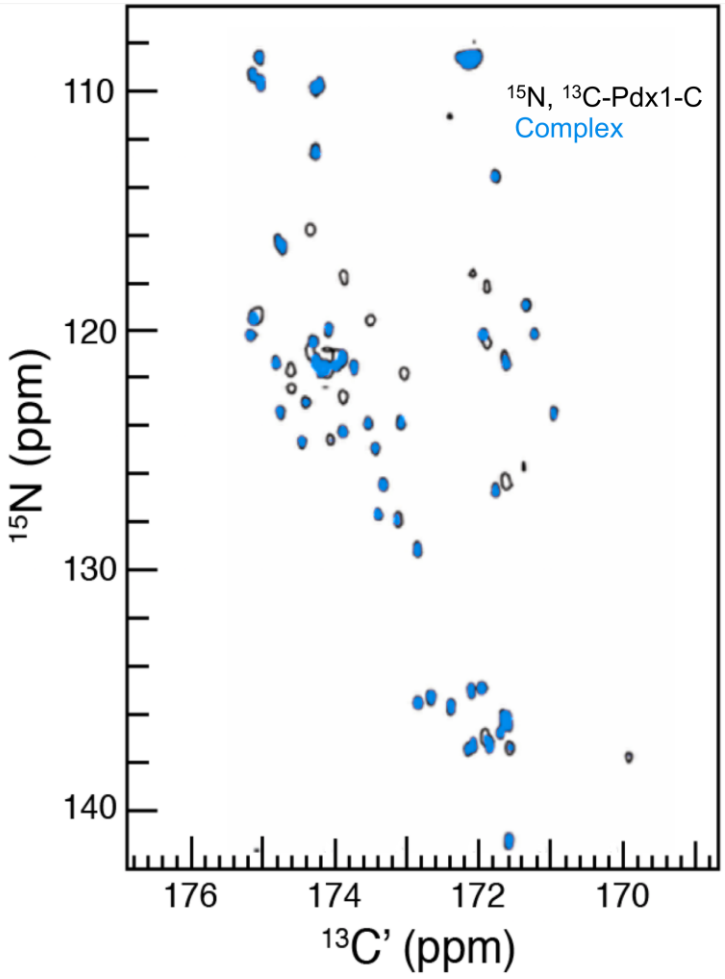



**Figure 6.** Overlay of $^{15}N$ Pdx1-SPOP $^{15}N$,$^{13}C$-CON spectra showing resonance loss upon binding. From Monique Bastidas' Ph.D. thesis, 2015.[52]

Resonance intensity loss upon binding (Figure 7) suggests that Pdx1 contains two SPOP binding sites comprising Pdx1 residues 224-236 and residues 270-275 or that Pdx1-C undergoes a conformational change upon association with SPOP. The binding interaction is in intermediate exchange and occurs on the same timescale of chemical

shifts, which results in complete resonance loss upon binding. This spectral feature allows for facile determination of Pdx1 binding sites, but precludes the generation of any additional structural insight into the Pdx1-C:SPOP complex.

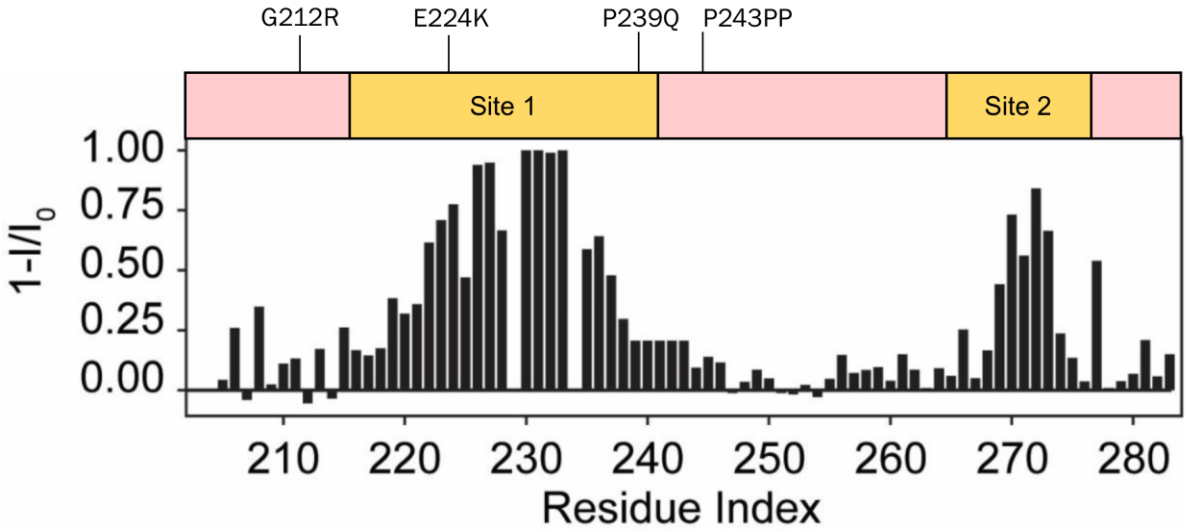



**Figure 7.** Pdx1-C clinical variant sites compared with percent resonance intensity loss for each residue, generated from $^{15}$N Pdx1-SPOP $^{15}$N,$^{13}$C-CON spectra. Locations with the highest percent resonance intensity loss correspond are potential sites of SPOP binding. From Monique Bastidas' Ph.D. thesis, 2015.[52]

Paramagnetic resonance enhancement (PRE) experiments conducted in the Showalter laboratory by Dr. Erik Cook confirm that Pdx1-C does not participate in any long-range interactions that could be broken upon SPOP binding (not shown). Therefore, the preliminary NMR data confirm a direct interaction between Pdx1 residues 220-236 (Pdx1 S1) and residues 270-275 (Pdx1 S2). Although Pdx1 S2 has not been identified in previous Pdx1-C studies, its identification is in agreement with the current literature consensus that SPOP substrates are multivalent.[9,47] Interestingly, Pdx1 S1 does not adhere to the SPOP consensus binding sequence of Φ-π-S-S/T- S/T (Φ nonpolar, π polar) very well, whereas Pdx1 S2 more tightly adheres to this consensus (Figure 7). This

leads us to hypothesize that Pdx1 may be a novel SPOP substrate, or that adherence to the consensus binding motif is not absolutely necessary for SPOP-substrate binding.

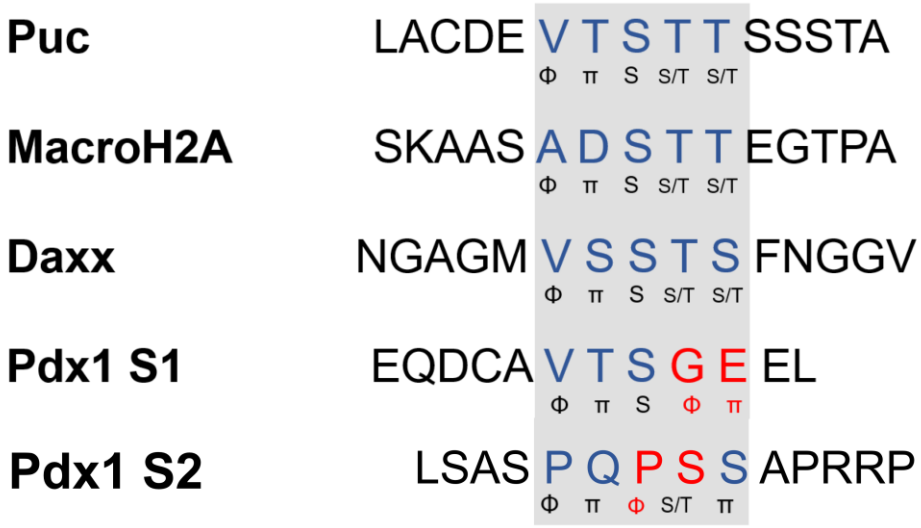



**Figure 8**. Known SPOP substrates and their consensus binding sequences compared with Pdx1 S1 and S2.

Cellular studies by our lab demonstrate that co-transfection of Pdx1 and SPOP results in colocalization into liquid-like nuclear speckles in HEK293T cells. The formation of membraneless organelles is a mechanism by which a protein sequesters its target, allowing for physical separation from the surrounding cellular milieu. Sequestration of Pdx1 by SPOP and subsequent degradation may be a molecular switch that allows SPOP to turn Pdx1 “off” under conditions of low glucose when insulin production is not required. This result agrees with the literature consensus that SPOP forms nuclear speckles in the presence of substrates.

Currently, there is no molecular-level description of the multivalent interaction between SPOP and Pdx1-C. Therefore, a rigorous structural and thermodynamic characterization of the complex is necessary in order to strengthen the conclusions of cellular studies and establish a molecular mechanism for functional Pdx1 regulation.

Motivated by the identification of two Pdx1-C binding sites, I hypothesize that Pdx1 stability is modulated by the multivalent interaction of individual weak motifs in Pdx1-C with SPOP MATH. The avidity effects conferred by this multivalency may provide a tunable signal that finely adjusts Pdx1 levels in response to changes in blood glucose. This hypothesis will be tested through the combination of biophysical and cellular studies of Pdx1 and the Pdx1-SPOP complex. Informed by the results of cellular studies, I will conduct binding assays that will provide the thermodynamic parameters that describe the formation of the Pdx1-SPOP complex and verify the affinity-enhancing effect of multivalency. This collaborative study aims to expand the breadth of knowledge on SPOP substrate recognition, Pdx1-C sequence-function relations, and, broadly, insulin regulation.

### 1.2 Technical Introduction

In the simplest case, biomolecular interactions involve two molecules recognizing and binding to each other. Study of the equilibrium thermodynamics of macromolecule-ligand interactions is important for biochemical processes, whose binding energies provide insight into their molecular mechanisms. Given the interesting thermodynamic profile that governs the conformational space IDPs sample, it is evident that the thermodynamics of their binding interactions are intimately related to their biological function. The techniques described below will be used to assess the thermodynamics surrounding Pdx1-SPOP complex formation and determine the enhancing effects of multivalency on binding affinity.

### *Isothermal Titration Calorimetry*

The understanding of IDP binding energetics requires a precise characterization of IDP-target complex formation. Enthalpy and entropy changes associated with binding reflect the magnitude and nature of the conformational changes coupled to binding. The determination of the binding association constant is especially important in the case of IDPs, whose weak interactions are often governed by large dissociation constants.[53] Isothermal titration calorimetry (ITC) is currently the only technique that simultaneously and directly yields information describing the reaction stoichiometry, enthalpy change upon binding, and equilibrium binding constant in a single experiment.

The ITC instrument contains a reference cell and a sample cell in an adiabatic environment, maintained at a constant and equal temperature slightly offset from that of the bath (Figure 8). Stepwise addition of ligand to the macromolecule induces a heat flow that requires a countervailing application or reduction of power to the sample cell in order to keep both cells at the same temperature. Small aliquots of ligand are titrated into the sample and the resultant thermogram is generated by plotting the applied differential power against time. Integration of the area under each peak produces data points that can be fit to a particular binding model to construct a binding isotherm (Figure 9) from

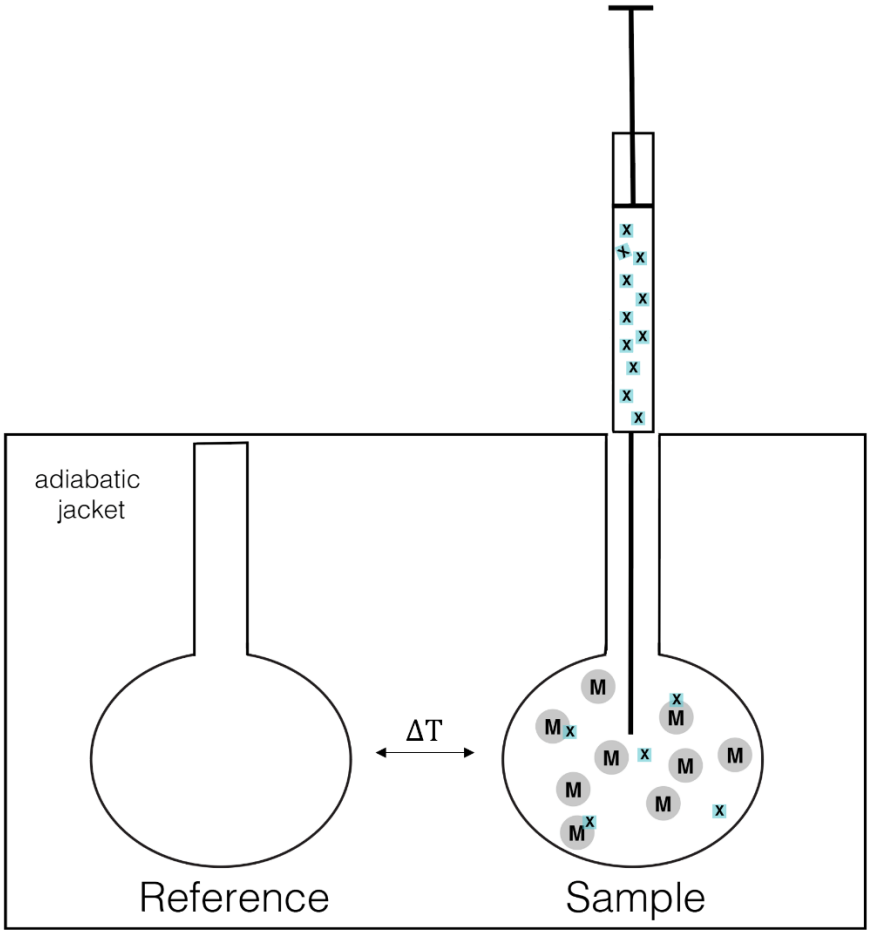



**Figure 9.** Diagram of a typical ITC instrument. The sample cell is filled with the macromolecule solution (M) and the injection syringe is filled with the ligand (X) solution. The source of signal in an ITC experiment is the differential power required to maintain constant temperature in both cells following each injection of titrant.

which the change in molar binding enthalpy (ΔH), the equilibrium binding association constant ($K_a$) and the reaction stoichiometry (n) can be determined. The data obtained from titrations at several different temperatures can be used to calculate the heat capacity change ($\Delta C_P$) associated with a binding interaction using the following relation:

$$\Delta H = \int \Delta C_P \, dT. \qquad (1)$$

Where dCp is assumed to remain constant over the relevant temperature range. Given that the temperature is known, the overall change in free energy (ΔG) is related to the equilibrium binding association constant through the expression

$$\Delta G = -RT \ln K_A = RT \ln K_D , \quad (2)$$

which allows for subsequent determination of the binding entropy using the relation

$$\Delta G = \Delta H - T\Delta S. \quad (3)$$

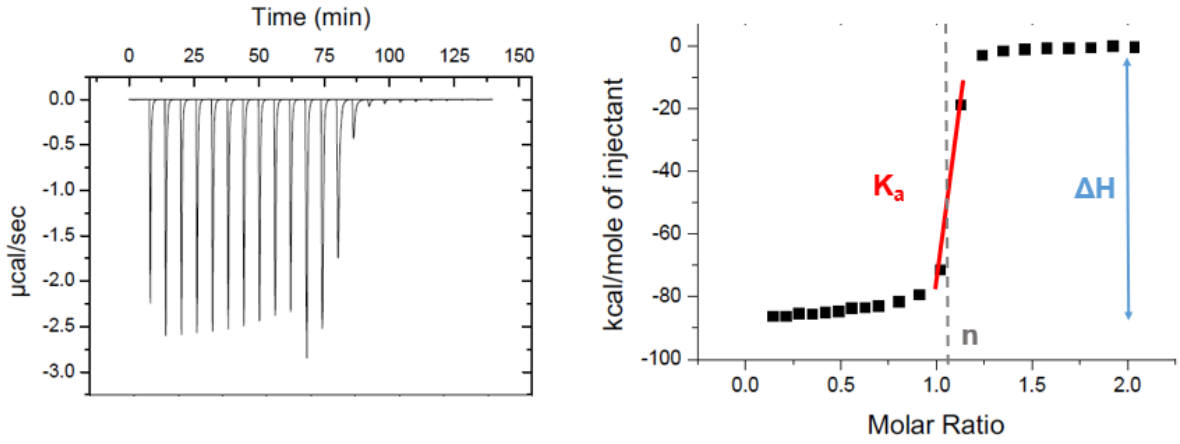



**Figure 10.** Example ITC thermogram and resultant heat curve from which the thermodynamic parameters $K_a$, n, and ΔH are determinable. These data were obtained by titration of 75 µM ssDNA into 5 µM of its complementary strand at 298 K.

### *Equilibrium Fluorescence Titrations*

Tryptophan fluorescence quenching experiments are relatively simple and consume minimal amounts of potentially expensive protein reagents. Initial ITC experiments predict that individual Pdx1 binding sites have very weak affinities for SPOP. Although the affinities of these sites may lie within the measurable range of binding affinities by ITC, preliminary results suggest that titration of minimal Pdx1 peptides into SPOP requires large and potentially unobtainable amounts of peptide in order to reach

full saturation. The increased sensitivity, ease of use, and minimal sample requirements afforded by tryptophan fluorescence quenching qualify this method, in parallel to ITC, as a robust label-free binding assay for the determination of binding affinities in protein-protein interactions.

Intrinsic protein fluorescence is a result of the excitation of predominantly tryptophan residues by 280 nm light. Intrinsic fluorescence is typically very weak or nonexistent in some systems.[54] In order to remedy this problem, extrinsic fluorophores are often incorporated into proteins by attachment of the fluorophore to native or engineered cysteine residues. The availability of commercial fluorophores has greatly stimulated their use, with a large percentage of biochemical binding assays being based on the use of extrinsic fluorophores.[55,56] Although the use of extrinsic fluorophores has yielded progress in fluorescence binding assays, their use has several shortcomings that can negatively affect the outcome of an experiment. Most notably, the covalent modification of proteins can result in structural and functional changes that affect the energetics of an interaction.[54,55] Due to their extended nature, intrinsically disordered proteins often participate in long-range interactions with their binding partners and themselves.[6] Extrinsic fluorophore labeling of IDPs may disrupt this complex network of interactions and abrogate the binding interaction altogether. To avoid such shortcomings, the experiments in this study do not rely on the use of extrinsic fluorophores and rather use intrinsic tryptophan fluorescence as a means to track binding.

This study uses intrinsic tryptophan fluorescence quenching to monitor binding of non-fluorescent minimal Pdx1 peptides to SPOP MATH. There are two accessible tryptophan residues on SPOP that can be selectively targeted by excitation at 295 nm,

with maximal emission occurring at ~350 nm. The excitation of the fluorophore is accomplished using a spectrofluorometer, wherein the fluorophore is placed in a reaction cell and titrated with ligand. Emission scans are collected for macromolecule sans ligand and following each addition of ligand. During an emission scan, the excitation monochromator is set to a fixed wavelength and the emission monochromator is scanned across a range of wavelengths. The design of this experiment requires care when selecting the fluorophore concentration and excitation/emission wavelength bandwidth.

Preliminary measurements of binding affinity can aid in the determination of optimal concentrations of ligand and macromolecule required for full saturation. Maximal saturation of the macromolecule with ligand ensures that experimental points display an even distribution when plotted over a logarithmic scale of titrant concentration at each data point.[54] Photobleaching of fluorophores is a commonly encountered problem in fluorescence experiments that is caused by exposure to high excitation energy for an extended period and leads to a decrease in fluorescence intensity over time.[54,55] High excitation bandwidths increase the excitation energy and should thus be avoided by selecting the narrowest possible excitation slit width setting.[54,57] The titration of ligand into a macromolecule results in dilution of both species, and data analysis must account for dilution effects on the total concentration of macromolecule. Preliminary determination of binding affinities is accomplished by plotting the change in signal intensity against total peptide concentration. However, a more robust analysis is achieved by conducting titrations at different macromolecule concentrations, fitting the data, and constructing a true binding isotherm.[54,58]

### *Determination of the Equilibrium Binding Association Constant*

The binding of a ligand (X) to a macromolecule (M) with one binding site can be defined by the reaction scheme

$$\mathrm{M} + \mathrm{X} \rightleftharpoons \mathrm{MX}$$

and is described by the equilibrium binding association constant, $K_a$:

$$\mathrm{K_a} = \frac{[\mathrm{MX}]}{[\mathrm{M}][\mathrm{X}]} \tag{4}$$

where $[\mathrm{M}]$, $[\mathrm{X}]$, and $[\mathrm{MX}]$ are the equilibrium concentrations of free macromolecule, free ligand, and macromolecule-ligand complex, respectively. For convenience, the free ligand concentration will be expressed as $[\mathrm{X}]_{\mathrm{free}}$ throughout the remainder of this discussion. The solution to Eq. 4 is possible through the determination of $[\mathrm{M}]$, $[\mathrm{X}]_{\mathrm{free}}$, and $[\mathrm{MX}]$; however, it is often difficult to simultaneously quantify the molar concentration of each species in solution. The determination of the binding constant is most convenient when utilizing the fractional quantity described by Eq. 5:

$$\langle \mathrm{X} \rangle = \frac{[\mathrm{MX}]}{[\mathrm{M}] + [\mathrm{MX}]} \tag{5}$$

where $\langle \mathrm{X} \rangle$ is the average value of ligands bound per macromolecule. For a macromolecule that possesses one binding site for X, Eqs. 4 and 5 transform into

$$\langle \mathrm{X} \rangle = \frac{\mathrm{K_a}[\mathrm{X}]_{\mathrm{free}}}{1 + \mathrm{K_a}[\mathrm{X}]_{\mathrm{free}}}. \tag{6}$$

The equilibrium binding model presented here can be used to fit titration curves using data from fluorescence binding assays. However, the experimental determination of $K_a$ is only achievable if $[\mathrm{X}]_{\mathrm{free}}$ and $\langle \mathrm{X} \rangle$ are calculable and the relationship between the observable signal change and $\langle \mathrm{X} \rangle$ is known. For a monomeric macromolecule that does

not undergo self-association, $\langle X \rangle$ is exclusively a function of $[X]_{free}$ as per Eq. 6. Additionally, for this same set of conditions, the observed fluorescence quenching ($Q_{obs}$) is an intrinsic property of the system determined only by $\langle X \rangle$.[54,58,59] Therefore, if $Q_{obs}$ is constant, both $\langle X \rangle$ and $[X]_{free}$ remain constant as well. The known variables in these titrations are the total macromolecule concentration $[M]_{tot}$ and the total ligand concentration $[X]_{tot}$. These variables are related to $\langle X \rangle$ and $[X]_{free}$ through the expression

$$[X]_{tot} = [X]_{free} + \langle X \rangle [M]_{tot} \tag{7}$$

in Eq. 7. The parameters $\langle X \rangle$ and $[X]_{free}$ can be determined by conducting at least two titrations at different total macromolecule concentrations and constructing a plot of $Q_{obs}$ against $[X]_{tot}$ for each separate titration. The points in the titration curve where $Q_{obs}$ is constant should result in the same value of $\langle X \rangle$ for any pair of concentrations $[X]_{tot}$ and $[M]_{tot}$.[58] The unknown variables can then be calculated using Eqs. 8 and 9:

$$\langle X \rangle = \frac{[X_2]_{tot} - [X_1]_{tot}}{[M_2]_{tot} - [M_1]_{tot}} \tag{8}$$

$$[X]_{free} = \frac{[M_1]_{tot}[X_2]_{tot} - [M_2]_{tot}[X_1]_{tot}}{[M_1]_{tot} - [M_2]_{tot}} \tag{9}$$

A true binding isotherm can be constructed by repeating this calculation for multiple $Q_{obs}$ values and plotting $\langle X \rangle$ vs $[X]_{tot}$ (or $[X]_{free}$).

## 1.3 Thesis Overview

Chapter 2 of this thesis details the experimental framework used to elucidate the mechanistic details of the interaction between the transcription factor Pdx1 and the E3 ubiquitin ligase substrate adaptor SPOP. Calorimetric and spectroscopic binding assays are used to determine the effects of multivalency on this protein-protein interaction essential for proper insulin production. Isothermal titration calorimetry and tryptophan fluorescence assays are used to determine thermodynamic parameters describing this binding event and the results herein inform the experiments proposed in Chapter 3.

Chapter 3 discusses additional experiments conducted by the author that will motivate future experiments into the thermodynamics involved in the Pdx1-SPOP binding interaction. First, the collection of triple resonance protein NMR experiments allows for the assignment of the SPOP HSQC spectrum and will prove useful for future NMR characterization of the SPOP-Pdx1 complex. Differential scanning calorimetry (DSC) results offer some insight into the entropic forces that may drive this interaction. Experiments involving Pdx1 clinical variants are proposed as a means to fully understand the role of the Pdx1-C terminal domain in Pdx1-SPOP binding and, ultimately, insulin regulation.

# Chapter 2: Examining the role of the Pdx1 C-terminal domain in Pdx1 degradation

## 2.1 Introduction

The transcription factor pancreatic duodenal homeobox 1 (Pdx1) plays a crucial role in pancreatic development and insulin production in β-cells. Clinically documented amino acid substitutions in conserved regions of the C-terminal IDR of Pdx1 (Pdx1-C) are associated with the development of maturity-onset diabetes of the young (specifically, MODY4) and various forms of adult-onset type 2 diabetes. Current Pdx1 literature suggests that Pdx1-C is required for proper *ins* transcription and regulation of Pdx1 stability, however, the molecular mechanisms describing these processes are currently unclear. In order to fully understand the role of Pdx1-C in insulin regulation, it is imperative to understand atomistic details of its role in Pdx1 degradation. Under low glucose conditions, Pdx1 is downregulated via SPOP-mediated proteasomal degradation through interaction with Pdx1-C. This event reduces the production of insulin within the β-cell. The aforementioned amino acid substitutions within Pdx1-C can diminish the efficiency of this process, which can result in insulin overproduction leading to the insulin resistance characteristic of type II diabetic phenotypes. Relative to wild-type Pdx1, the binding interaction of SPOP with clinical Pdx1 variants may result in the alteration or complete disruption of crucial atomic interactions at the Pdx1 binding site. Therefore, a complete thermodynamic and structural characterization of the Pdx1-SPOP interaction is necessary to fully understand the effects of Pdx1 amino acid substitutions on this important regulatory binding event.

Preliminary NMR studies conducted by our lab have identified a previously unknown second binding site for the E3 ubiquitin ligase adaptor SPOP (Figure 10). Although this secondary binding site was not identified in vivo, the existence of multiple linear binding motifs is a characteristic of Pdx1 that is shared with all currently known SPOP substrates. Interestingly, the residues that comprise this binding site coincide with known sites of Pdx1 glucose-dependent phosphorylation**.** The Pdx1-SPOP binding interaction may therefore be modulated in some way by phosphorylation of Pdx1. Investigation into the Pdx1-C:SPOP interaction may reveal mechanistic details of glucose-dependent, SPOP-mediated proteasomal degradation modulated by phosphorylation events.

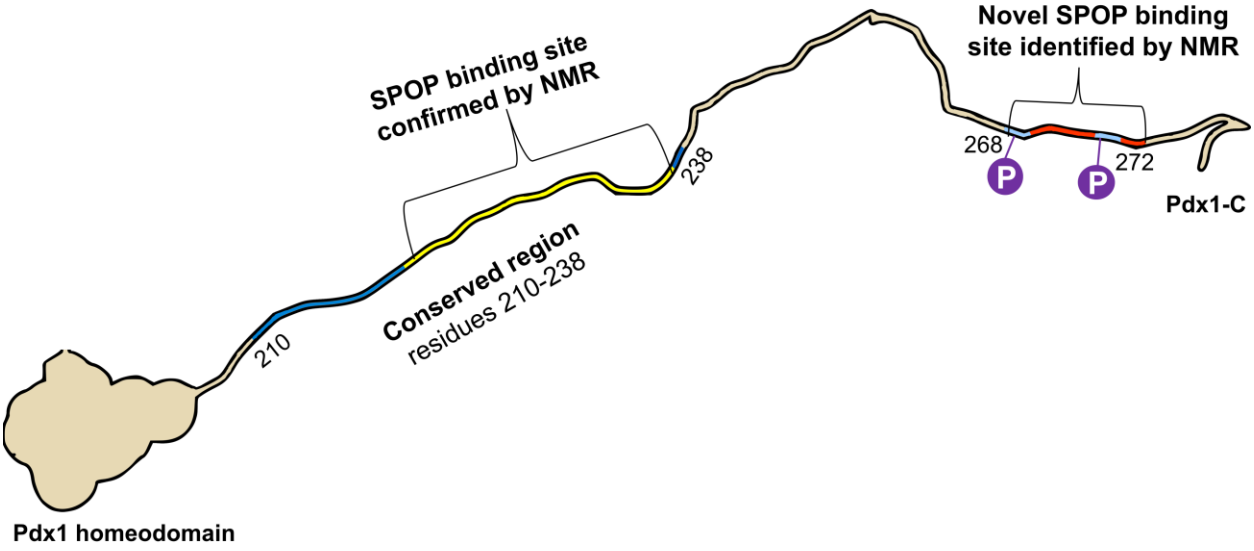



**Figure 11.** Cartoon summary of the conclusions informed by preliminary data. The blue region on Pdx1-C was previously identified as a conserved region required for maximal Pdx1 transactivation and SPOP binding. The region in yellow is the SPOP binding site verified by the Showalter lab through NMR titrations. The region in red encompasses Pdx1 residues 265-275 and is the novel SPOP binding site identified through NMR titrations. Pdx1 undergoes phosphorylation at the residues in light blue with attached purple phosphorylation marks.

As motivated by the identification of this additional binding site, I hypothesize that the Pdx1 C-terminal IDR forms multivalent contacts with SPOP, and that this multivalency increases the affinity of this interaction. Additionally, I demonstrate that strict adherence to the SPOP binding consensus motif is not required for SPOP-substrate binding. Based on our preliminary data, I propose three possible models for Pdx1-SPOP binding (Figure 11). In this work, minimal Pdx1 peptides corresponding to Pdx1 S1 (a.a. 224-236) and Pdx1 S2 (a.a. 265-278) will be complexed with SPOP MATH to determine their individual binding affinities for SPOP. Full-length Pdx1-C (a.a. 206-283) will also be complexed with SPOP MATH to investigate the affinity-enhancing effects of multivalency. Isothermal titration calorimetry and equilibrium fluorescence binding assays will be used to determine affinities and to confirm the correct binding model for this interaction.

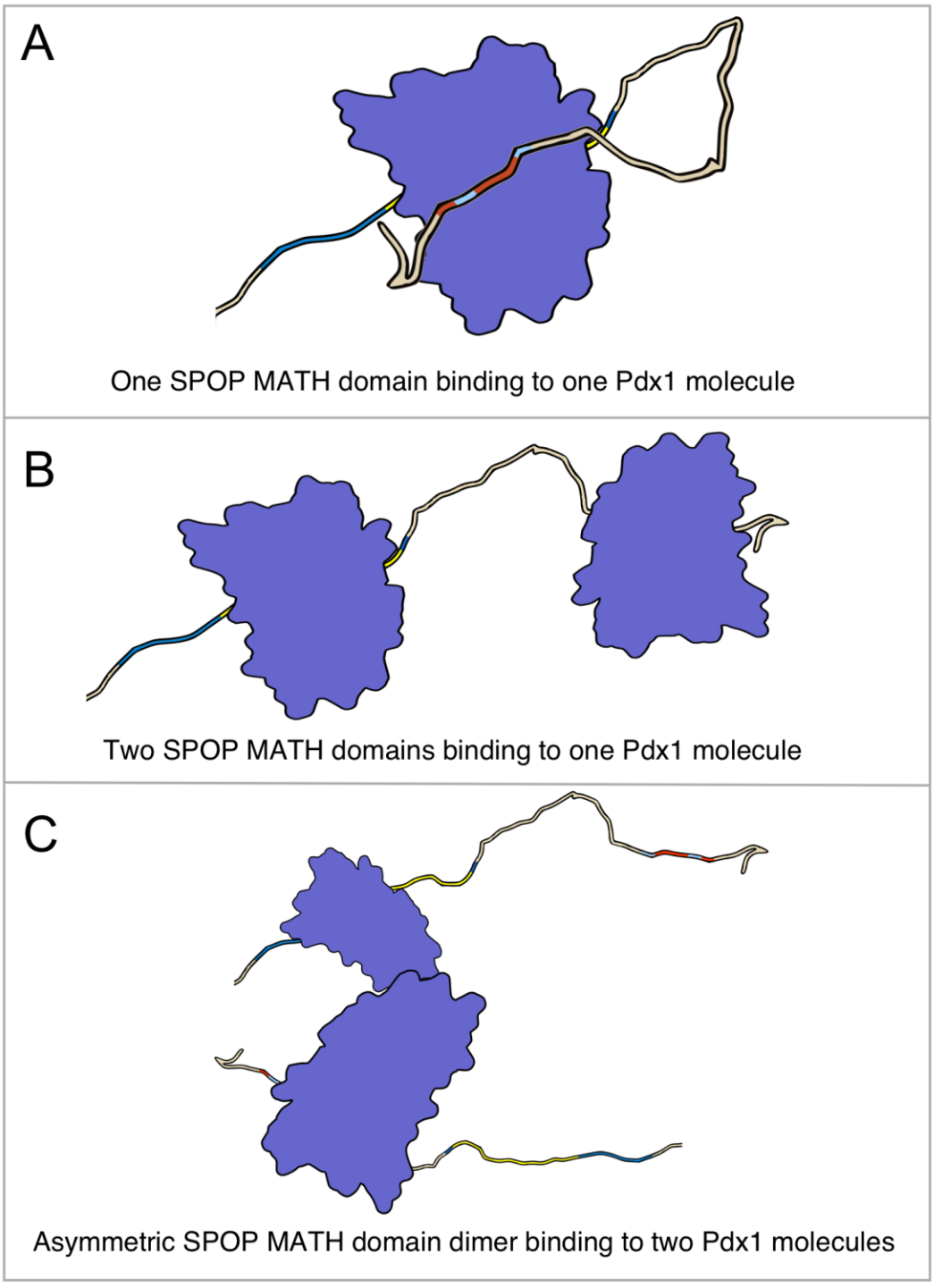



**Figure 12**. Proposed models for Pdx1-C:SPOP interaction based on preliminary data presented in Chapter 1.

## 2.2 Materials and Methods

### *Preparation of Pdx1-C*

A pet-49b(+) plasmid containing GST-$His_6$-Pdx1-C (a.a. 206-283) was transformed into competent BL21-Gold (DE3)pLysS cells for protein overexpression. Cells were allowed to grow in LB medium with 30 µg $µL^{-1}$ kanamycin at 37 ºC to an optical density of $A_{600}$ = 0.7-0.9 and were induced with 500 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3 h. The cells were then harvested by centrifugation at 3400 x *g* for 30 mins at 4 ºC and subsequently lysed via sonication under ice. The cell lysate was clarified by centrifugation at 14200 x *g* for 30 min. The cleared lysate was then poured over Ni-NTA resin and purified by affinity chromatography to yield GST-$His_6$-Pdx1-C. The GST and $His_6$ tags were cleaved from Pdx1-C by overnight incubation with His-tagged HRV 3C protease at 4 ºC. The cleaved Pdx1-C/His-3C protease mixture was placed over Ni-NTA resin and Pdx1-C was collected as flowthrough. Pdx1-C was purified further by size-exclusion chromatography using a HiPrep 26/60 Sephacryl S-200 HR column (GE Life Sciences) in 50 mM tris-HCl pH 7.5, 100 mM potassium chloride, and 3 mM β-mercaptoethanol. The protein was then concentrated and buffer-exchanged into the appropriate buffer using a 3kDa MWCO centrifugal filter. Protein concentration was determined by FT-IR spectroscopy.

### *Preparation of SPOP MATH*

A pet-49b(+) plasmid containing GST-$His_6$-SPOP MATH (a.a. 28-166) was transformed into competent BL21-Gold (DE3)pLysS cells for protein overexpression. Cells were allowed to grow in LB medium with 30 µg $µL^{-1}$ kanamycin at 37 ºC to an optical

density of $A_{600}$ = 0.6-0.7 and were induced with 600 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) overnight at 24 ºC. The cells were then harvested by centrifugation at 3400 x *g* for 30 mins at 4 ºC and subsequently lysed via sonification under ice. The cell lysate was clarified by centrifugation at 14200 x *g* for 1 h. The cleared lysate was then poured over Ni-NTA resin and purified by affinity chromatography to yield GST-$His_6$-SPOP MATH. The GST and $His_6$ tags were cleaved from SPOP MATH by overnight incubation with His-tagged HRV 3C protease at 4 ºC. The cleaved SPOP MATH/His-3C protease mixture was again placed over Ni-NTA resin and SPOP MATH was collected as flowthrough. SPOP MATH was purified further by size-exclusion chromatography using a HiPrep 26/60 Sephacryl S-200 HR column (GE Life Sciences) in 50 mM tris-HCl pH 7.5, 100 mM potassium chloride, and 3 mM β-mercaptoethanol. The protein was then concentrated and buffer-exchanged into the appropriate buffer using a 3kDa or 10kDa MWCO centrifugal filter. SPOP MATH aliquots were denatured in 6 M guanidine hydrochloride and subject to a UV-Vis spectroscopy absorption wavelength scan from 220-320 nm. Protein concentration was calculated using the Beer-Lambert law with a molar extinction coefficient of 22180 $M^{-1}$ $cm^{-1}$ at 280 nm and a molecular weight of 16300 g/mol. The molar extinction coefficient was calculated based on amino acid sequence by using the ProtParam tool at www.expasy.org.

***Synthetic peptides***

$Pdx1_{224-236}$, $Pdx1_{262-278}$, and PUC SBC1 were obtained as lyophilized powders from the Tufts University Core Facility. The peptides were extensively purified by high-performance liquid chromatography (HPLC) and purity was assessed using MALDI mass spectrometry. Pure peptides were dissolved in the appropriate buffer and protein

concentration was confirmed by FT-IR spectroscopy. Table 1 describes the amino acid sequence of each synthetic peptide.

**Table 1.** Synthetic peptide sequences for titration experiments

| Peptide | Sequence |
|---|---|
| $Pdx1_{224-236}$ | EQDCAVTSGEELL |
| $Pdx1_{262-278}$ | LSASPQPSSAPR |
| PUC SBC1 | LACDEVTSTTSSSTA |

***Truncated Pdx1-C constructs***

A recombinant plasmid of $Pdx1_{206-256}$ was generated from $Pdx1_{206-283}$ by mutation of the AAA codon encoding for Lys 257 to the stop codon TAA using a Quikchange Lightning site-directed mutagenesis kit (Agilent). Recombinant plasmids of Pdx1-CW, Pdx1-$CW_{202-247}$ and Pdx1$CW_{248-283}$ were also generated through PCR amplification of desired regions using Q5 site-directed mutagenesis. Purification protocols for truncated constructs are identical to those described for Pdx1-C.

***Nuclear Magnetic Resonance***

Standard protein NMR heteronuclear single quantum coherence experiments were performed at 298 K on a 500 or 600 MHz BrukerAvance III spectrometer. All samples were buffer exchanged into 100 mM potassium phosphate pH 6.5, 100 mM potassium chloride, and 5 mM β-mercaptoethanol. Samples also contained 10% $D_2O$ (w/v) for a lock

signal and 0.01% sodium azide to prevent microbial growth. HSQC spectra of 300-500 μM $^{15}$N SPOP MATH in the absence and presence of 500-1000 μM Pdx1 S1 or Pdx1 S2 were collected. The relative intensity change was calculated and used to colorize the SPOP MATH surface for qualitative visualization of the SPOP binding surface for Pdx1. NMR spectra were processed in TOPSPIN and analyzed with SPARKY.

***Isothermal Titration Calorimetry***

Pdx1-C and SPOP MATH were purified as described in previous sections. Purified Pdx1-C and SPOP MATH were buffer-matched by buffer exchange over PD-10 resin into 100 mM potassium phosphate pH 6.5, 100 mM potassium chloride, and 1 mM β-mercaptoethanol. Synthetic peptides were diluted with ITC buffer and were not buffer-exchanged. A MicroCal VP-ITC calorimeter was used to titrate 500 μM Pdx1-C into 30 μM SPOP MATH at 25 ºC with a stirring rate of 307 rpm. The reference power was set to 5-10 μCal/s. The initial delay was set to 60 s for most experiments, but was set as high as 600 s in cases where it was difficult to quickly obtain a reference power baseline. The initial injection volume was set to 4 μL to account for errors in syringe loading and the spurious heat often seen upon first injection. This data point is excluded from the analysis and subsequent calculation of thermodynamic parameters. Following the first injection, 25 injections of 10 μL aliquots of a Pdx1 construct were titrated into SPOP MATH at a rate of 0.5 μL/s. Spacing between each injection was set to 400 s to ensure each measurement returned to baseline. Origin 7 was used for standard data fitting and subsequent analysis of ITC data. The injection data points were manually integrated to ensure each measurement returned to baseline. A 5-point baseline average was

computed and subtracted from each heat of injection to account for any heats associated with dilution. The data underwent 100 simplex iterations of fitting to a one-site binding model until the corresponding chi-squared value was fully reduced.

***Intrinsic Tryptophan Fluorescence Quenching***

Fluorescence intensity measurements were collected on a Varian Cary Eclipse Fluorescence spectrophotometer. All samples were diluted with 50 mM tris(hydroxymethyl)aminomethane hydrochloride pH 7.5, 50 mM potassium chloride, and 5 mM diothiothreitol (DTT). After powering on the instrument, the lamp was allowed to warm up for at least 15 minutes to reduce ambient temperature fluctuations during the experiment. The fluorescence spectrophotometer temperature was kept at 25 C. The excitation slit width was 5 nm and the emission slit width was 10-20 nm, dependent on the concentration of the sample. The emission slit width was increased for samples with a concentration < 2 μM. A SPOP MATH solution with a concentration in the range of 0.5 to 5 μM was placed into a quartz cuvette with a 0.4 mm pathlength. Stock solutions of 0.50 to 14 mM substrate peptide were titrated into SPOP according to spreadsheet designs similar to that in Table 2. The excitation wavelength was set to 278 nm. After a 60 s equilibration delay, the sample was excited and an emission wavelength scan from 300-400 nm was collected. Data points were collected using a wavelength scan rate of 600 nm/min for macromolecule only ($i = 0$) and following each addition of ligand, where each data point $i$ corresponds to the emission spectrum collected upon the $i$th addition of ligand. The measurements were repeated 3-5 times to ensure that the signal remained stable. The titration continued until the total peptide concentration reached 300 μM, or when the dilution of SPOP reached 10%, whichever occurred first.

**Table 2.** Sample fluorescence titration experimental design spreadsheet for titration of 1 μM SPOP with Pdx1 S1 in 50 mM tris(hydroxymethyl)aminomethane hydrochloride pH 7.5, 50 mM potassium chloride, and 5 mM diothiothreitol (DTT).

| ***i*** | **[peptide] in cell (μM)** | **[titrant stock] (mM)** | **volume of titrant (μL)** | **total titrant volume (μL)** |
|---|---|---|---|---|
| 0 | 0.00 | N/A | 0.00 | 0.00 |
| 1 | 0.35 | 0.50 | 0.70 | 0.70 |
| 2 | 0.70 | 0.50 | 0.70 | 1.40 |
| 3 | 1.05 | 0.50 | 0.70 | 2.10 |
| 4 | 1.40 | 0.50 | 0.70 | 2.80 |
| 5 | 1.74 | 0.50 | 0.70 | 3.50 |
| 6 | 3.14 | 2.00 | 0.70 | 4.20 |
| 7 | 4.54 | 2.00 | 0.70 | 4.90 |
| 8 | 5.94 | 2.00 | 0.70 | 5.60 |
| 9 | 8.74 | 4.00 | 0.70 | 6.30 |
| 10 | 11.54 | 4.00 | 0.70 | 7.00 |
| 11 | 14.34 | 4.00 | 0.70 | 7.70 |
| 12 | 17.14 | 4.00 | 0.70 | 8.40 |
| 13 | 19.94 | 4.00 | 0.70 | 9.10 |
| 14 | 22.74 | 4.00 | 0.70 | 9.80 |
| 15 | 25.54 | 4.00 | 0.70 | 10.50 |

### *Data Analysis for Tryptophan Fluorescence Quenching Experiments*

The total fluorescence intensity of a mixture of $i$ fluorescent species can be described by

$$\mathrm{F} = \sum \mathrm{F_i}[\mathrm{X}]_\mathrm{i} \qquad (10)$$

Individual plots of wavelength against fluorescence intensity for each titration point $i$ were plotted onto the same figure for a qualitative assessment of fluorescence

quenching. This plot also aided in the observation of shifts in the emission wavelength at which fluorescence intensity is maximal, $\lambda_{EM,max}$.

The observed change in relative fluorescence quenching can be written as

$$\mathrm{Q_{obs}} = \mathrm{Q_{max}} \frac{\mathrm{K_{obs}[X]_{free}}}{1 + \mathrm{K_{obs}[X]_{free}}} . \quad (11)$$

Direct measurement of $[X]_{free}$ is not possible and the preferred approach for binding constant determination is to express $[X]_{free}$ in terms of $[X]_{tot}$ as described in Eq. 12:

$$[X]_{free} = [X]_{tot} - [MX] \quad (12)$$

For determination of the equilibrium binding dissociation constant, the above can be expanded and ultimately transformed into the quadratic equation

$$[MX]^2 - (\mathrm{[M]_{tot}} + \mathrm{[X]_{tot}} + \mathrm{K_D})\mathrm{[MX]} + \mathrm{[M]_{tot}[X]_{tot}} , \quad (13)$$

which has one real solution describing the concentration of macromolecule (SPOP) in complex with the ligand (Pdx1 peptide):

$$[MX] = \frac{1}{2}(\mathrm{[M]_{tot}} + \mathrm{[X]_{tot}} + \mathrm{K_D}) - \sqrt{(\mathrm{[M]_{tot}} + \mathrm{[X]_{tot}} + \mathrm{K_D})^2 - 4\mathrm{[M]_{tot}[X]_{tot}}} . \quad (14)$$

Equation 14 is related to $\mathrm{Q_{obs}}$ by the expression in Equation 14:

$$[MX] = \tfrac{1}{2}\mathrm{Q_{obs}}\,[1 + ((\mathrm{[M]_{tot}} + \mathrm{[X]_{tot}} + \mathrm{K_D}) - \sqrt{(\mathrm{[M]_{tot}} + \mathrm{[X]_{tot}} + \mathrm{K_D})^2 - 4\mathrm{[M]_{tot}[X]_{tot}}})] + Q_{obs,min} \quad (15)$$

Values of normalized $\mathrm{Q_{obs}}$ from a window of 10 nm were summed and plotted against the concentration of peptide present in the sample cell. The resultant data were fit to equation 15, which produced a binding curve that allowed for extraction of the equilibrium dissociation binding constant. For a more rigorous analysis, additional titrations at various concentrations of SPOP MATH were conducted in order to determine $\mathrm{[X]_{free}}$ and $\langle \mathrm{X} \rangle$. A Langmuir isotherm was generated from these data according to the description in Chapter 1.2.

## 2.3 Results and Discussion

### *Nuclear Magnetic Resonance Binding Experiments*

The collection of the $^{15}N$,$^{13}C$-CON experiments described in Chapter 1 allowed us to monitor the binding interaction from the perspective of isotopically enriched Pdx1-C. In order to establish whether Pdx1-C binds in the canonical SPOP MATH binding groove, HSQC spectra of $^{15}N$-enriched SPOP in the absence and presence of Pdx1-C were collected. I went on to complete this suite of experiments by collecting the standard triple resonance experiments used to assign the HSQC of SPOP MATH (see Chapter 3.2). The assignment of the SPOP MATH HSQC spectrum allowed us to identify the location of the binding site for Pdx1 on the SPOP MATH surface. Upon assignment, we were able to conclude that the previous NMR titration of Pdx1-C into $^{15}N$-labeled SPOP MATH displayed chemical shift perturbations not only in regions corresponding to the canonical SPOP binding groove, but also around the circumference of the SPOP surface (Figure 12). This led to the working hypothesis that Pdx1-C may behave as a unique SPOP substrate that wraps around the entire circumference of the MATH domain.

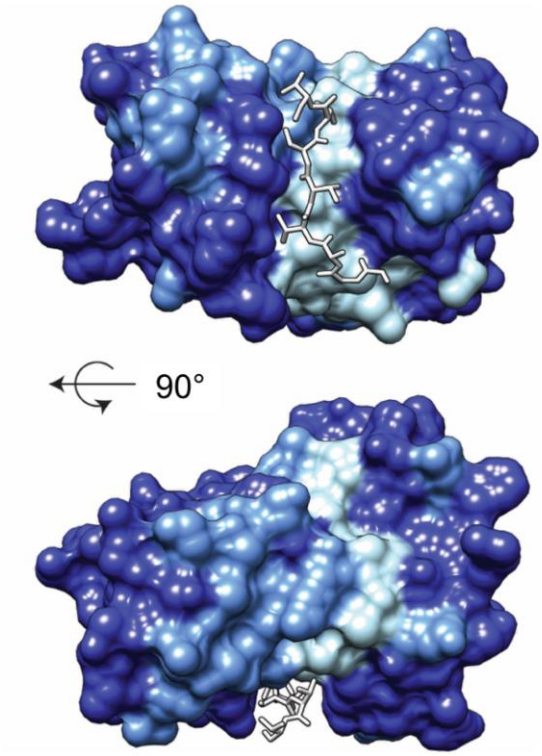



**Figure 13.** The collection of 3D NMR experiments allowed for SPOP MATH HSQC assignment and subsequent analysis of the SPOP MATH-Pdx1 binding experiments collected by Monique Bastidas. The SPOP MATH surface is colorized at residues corresponding to resonance intensity loss upon binding. Dark blue areas correspond to regions with minimal intensity change upon binding, and lighter blue areas correspond to regions with high intensity change upon binding. Pdx1-C binds in the canonical SPOP MATH binding groove. Rotation of the surface by 90 degrees shows chemical shift perturbation alongside the entire circumference of the SPOP MATH domain.

There are alternative binding models that can account for the chemical shift perturbations seen in the NMR titration of Pdx1-C into SPOP. Based on preliminary NMR data, I propose 3 models describing Pdx1-SPOP binding in Figure 11. The model in Figure 11A suggests that an extended conformation of Pdx1 binds to SPOP at sites Pdx1 S1 and Pdx1 S2. In this model, Pdx1 binds to SPOP MATH at multiple sites that wrap around the MATH domain. Figure 11B describes a binding event that involves two SPOP MATH domains binding to one Pdx1 molecule, with each MATH domain binding at Pdx1 S1 and Pdx1 S2. Figure 11C involves an asymmetric SPOP MATH dimer binding to two

separate Pdx1 molecules, with one monomeric subunit binding to Pdx1 S1 on one Pdx1 molecule, and the other subunit binding to Pdx1 S2 on another Pdx1 molecule.

To simplify the results from our NMR titration of full-length Pdx1-C, I conducted NMR titrations of Pdx1 S1 and Pdx1 S2 into $^{15}$N SPOP MATH. An overlay of HSQC spectra corresponding to each binding event showed resonance intensity loss upon binding for both Pdx1 S1 and S2 (Figure 13). This confirmed that both sites do indeed bind to SPOP MATH. Pdx1 S1 caused greater intensity losses in SPOP MATH, which suggests that this binding event may be stronger than that involving Pdx1 S2. Fractional intensity loss was again mapped onto the SPOP MATH surface and colorized according to the degree of intensity loss upon binding (Figure 14).

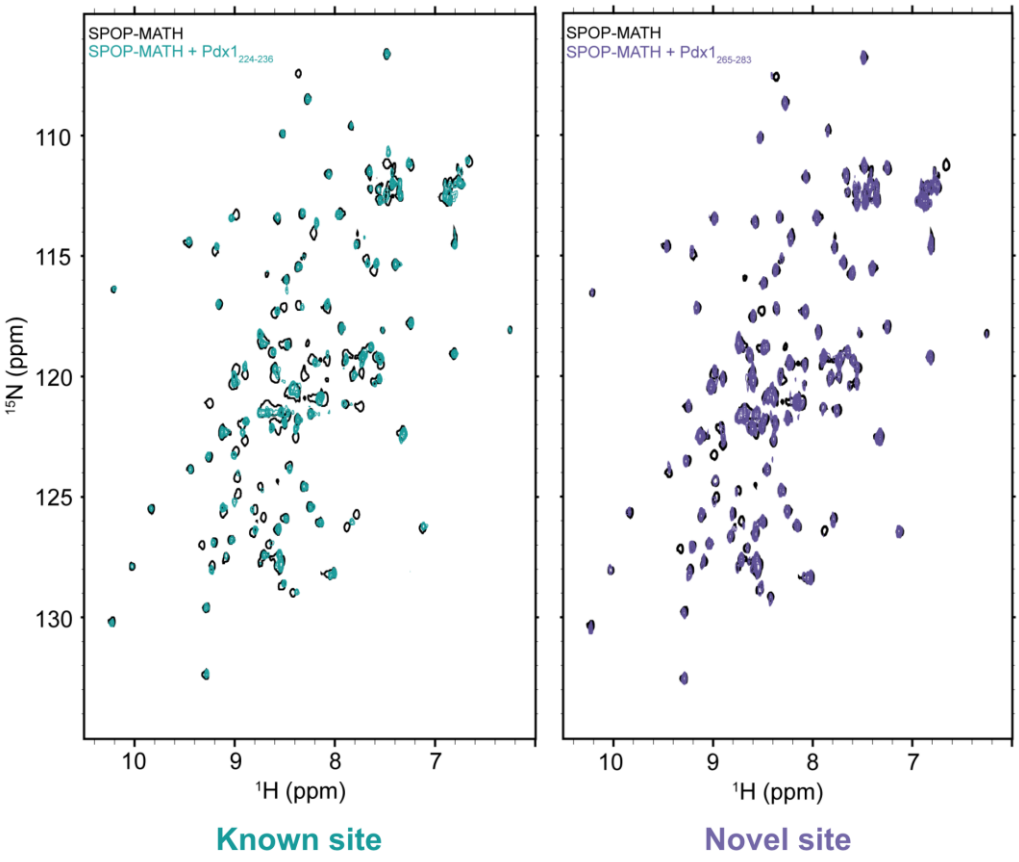



**Figure 14.** Loss of SPOP MATH resonance intensity upon Pdx1 S1 and S2 binding confirms that both sites directly bind to SPOP.

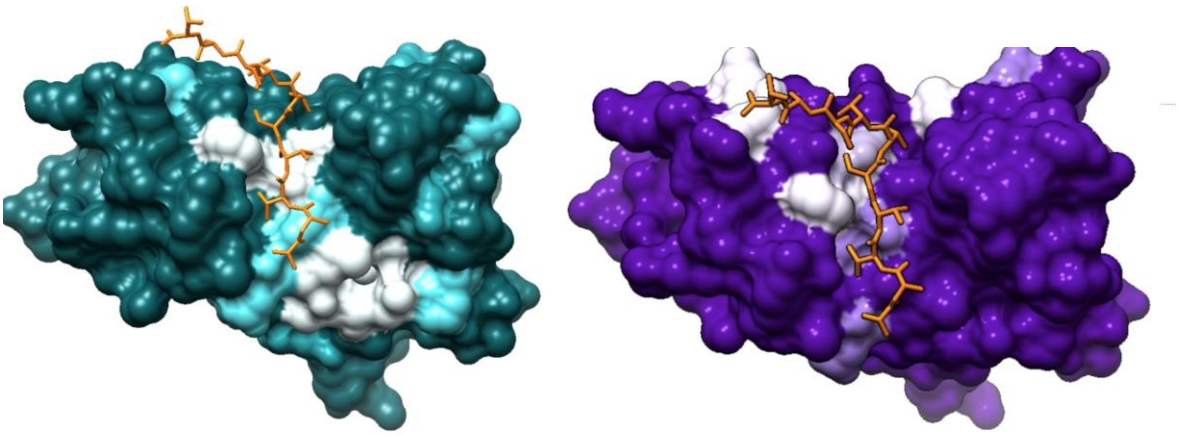


**Figure 15.** Colorized SPOP MATH surfaces generated from NMR titration of Pdx1 S1 (right, teal) and Pdx1 S2 (left, purple). Figures generated by Grace Usher.

Both SPOP MATH peptides bind in the canonical SPOP MATH binding groove, with Pdx1 S1 displaying some chemical shift perturbation wrapping around behind the canonical binding groove (not shown). Pdx1 S2 showed little to no chemical shift perturbation outside the canonical SPOP MATH binding groove. Again, Pdx1 S2 displays weaker chemical shift perturbations that Pdx1 S1, which suggests that this site has a lower binding affinity for SPOP. However, additional thermodynamic binding assays were necessary to quantitatively determine the binding affinities at each site.

***Isothermal Titration Calorimetry***

According to our previous NMR experiments, SPOP binds to multiple sites on the Pdx1 C-terminus. Given that SPOP substrates often contain multiple binding motifs that participate in a multivalent interaction with SPOP, I aimed to test the hypothesis that multivalency has an enhancing effect on Pdx1-C:SPOP binding. Pdx1-C, $Pdx1_{224\text{-}236}$ (Pdx1 S1), and $Pdx1_{265\text{-}278}$ (Pdx1 S2) were titrated into dilute SPOP to determine the binding dissociation constant and demonstrate the enhancing effects of multivalency. An additional peptide corresponding to a SPOP binding motif on the phosphatase Puckered

protein (PUC) was also titrated into SPOP as a control. The dissociation constant for PUC binding to SPOP has been reported to lie in the range of 4-8 μM.

Zhang *et. al* conducted binding studies of different SPOP MATH domain-peptide interactions and found that these substrates displayed a range of affinities spanning from 4 μM to 1 mM.[46] Despite the divergent SPOP-binding consensus motif of Pdx1 S1, Monique Bastidas confirmed weak binding of SPOP MATH to Pdx1-C by ITC (Figure 5). The equilibrium dissociation constant ($K_d$) was determined to be 25 ± 1 μM, which lies within the range of affinities for SPOP-substrate interactions. However, the system did not reach saturation and the resultant binding isotherm did not contain a lower baseline.

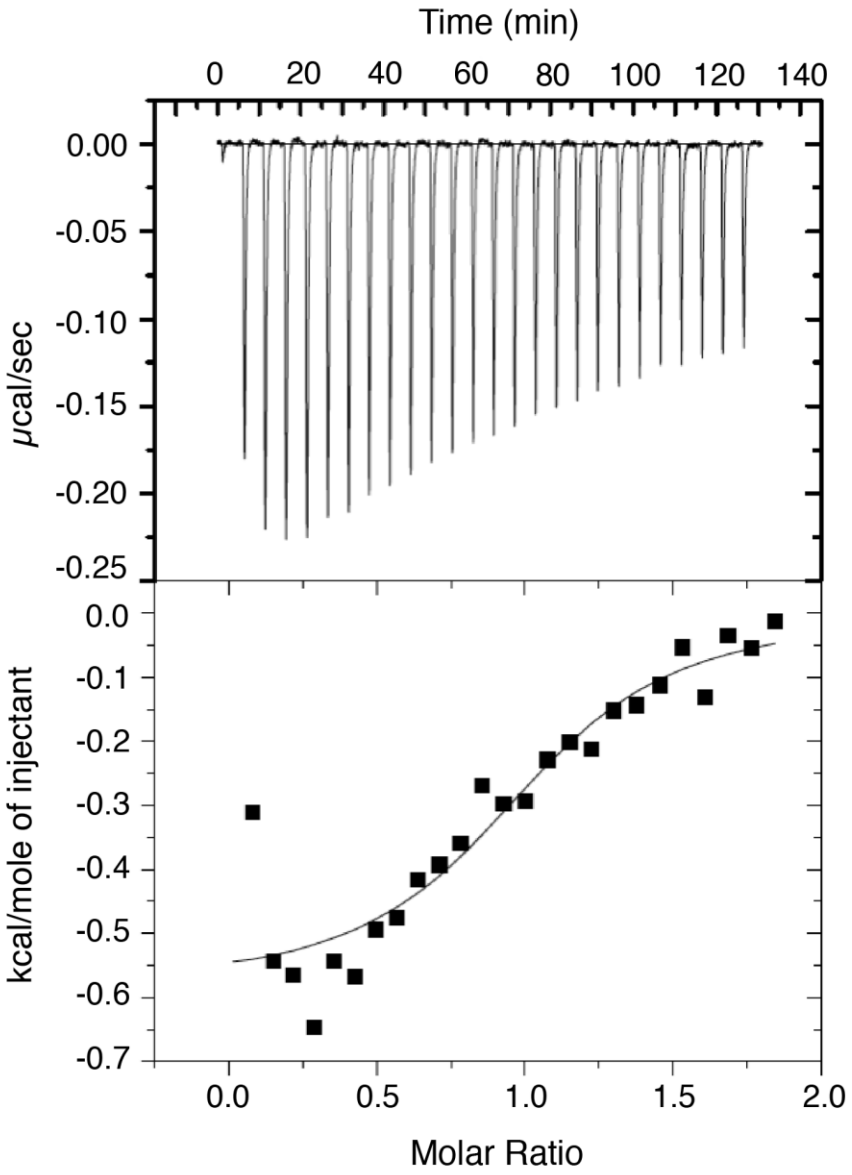



**Figure 16.** Titration of 500 µM Pdx1-C into 30 µM dimeric $SPOP_{28-337}$ at 298 K. The dissociation constant is $K_D$ ~4.8 µM.

I repeated the titration of Pdx1-C into SPOP MATH 28-166 in the pursuit of obtaining a lower baseline, however, my results were similar to Monique's and I could not generate a lower baseline. To demonstrate the enhancing effects of SPOP multivalency and oligomerization, Pdx1-C was also titrated into a dimeric form of SPOP, $SPOP_{28-337}$ (Figure 16). This titration was repeated various times, but had very low heats and never reached full saturation. The binding curve suggests a $K_d$ of 4.8 µM and a stoichiometry of n=1,

which would reject the model proposed in Figure 11B and shows a relative affinity increase compared to titration of Pdx1 into the monomeric SPOP $MATH_{28-166}$. However, due to the featureless binding isotherm these values may not be accurate.

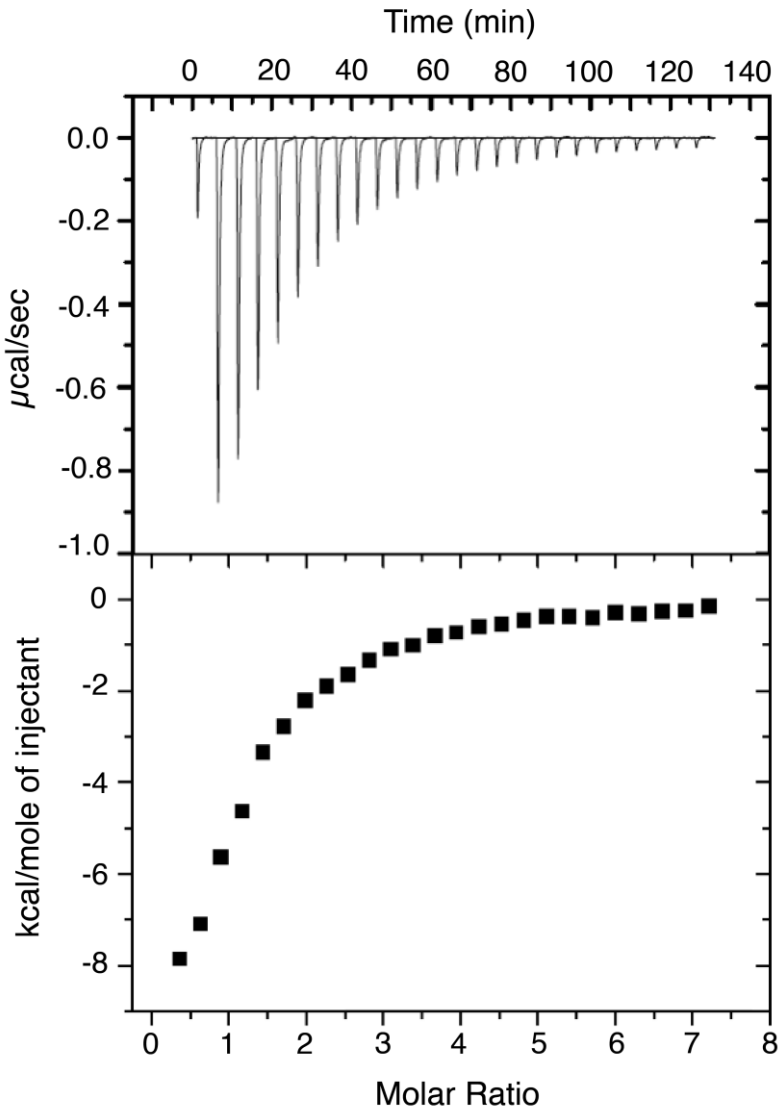



**Figure 17.** Titration of 500 µM PUC into 30 µM SPOP MATH at 298 K.

A control titration of 500 µM PUC peptide into 30 µM SPOP MATH showed that this system reached at least 90% saturation (Figure 17). The dissociation constant was calculated as 6 $\pm$ 1 µM, which agrees with the literature value of 7.7 $\pm$ 0.3 µM.[46] This

titration was performed with the same biological sample used in the Pdx1-C:SPOP experiment but displayed a markedly different binding curve. This suggests that the dissociation constant of Pdx1-C for SPOP may be lower than what we had originally found, or that there are additional complexities to the interaction that prevent SPOP MATH from fully saturating upon titration with Pdx1-C.

Titrations of minimal Pdx1-C peptides corresponding to the individual binding sites $Pdx1_{224-236}$ and $Pdx1_{265-278}$ were conducted to simplify our experimental ITC results. Titration of 500 µM Pdx1 S1 into 30 µM SPOP MATH displayed an entirely featureless isotherm that resembled a straight line (Figure 17). The calculated $K_d$ was 18 µM, but due to the linear binding isotherm, our confidence in this value was lower than that for full length Pdx1-C.

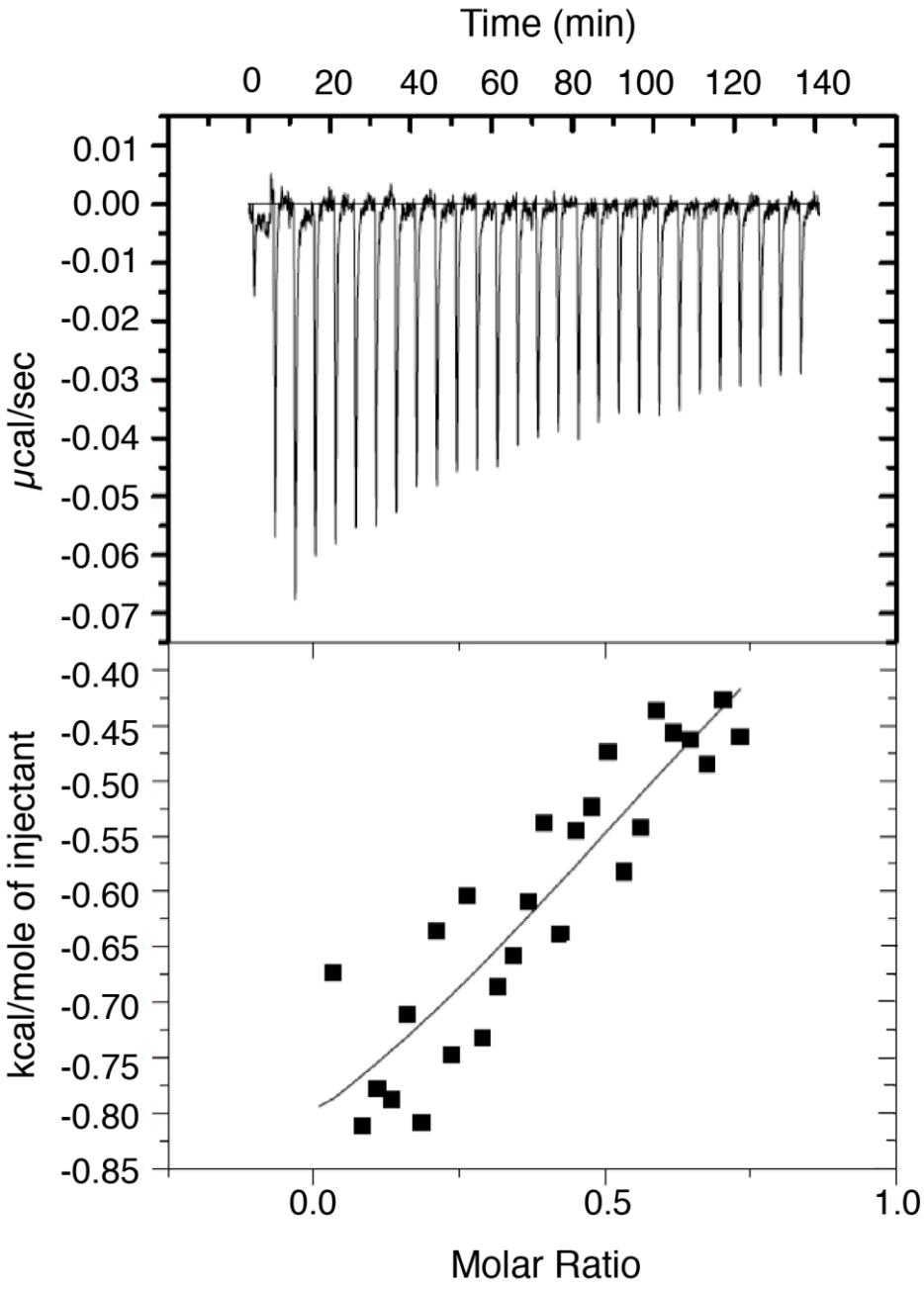



**Figure 18.** Titration of 500 µM Pdx1 S1 into 30 µM SPOP MATH at 298 K.

The lack of saturating conditions for titration of Pdx1 peptides into SPOP MATH prompted me to increase the concentrations of both SPOP MATH and Pdx1 peptide in order to achieve full saturation. A titration of 900 µM Pdx1 S2 into 70 µM SPOP MATH also yielded a noisy thermogram displaying low heats and a binding isotherm completely

absent of the features necessary to quantify the thermodynamic parameters describing this binding event (not shown).

The unusual results from my ITC experiments required me to employ an alternative technique allowing for determination of the dissociation constant. Fluorescence binding assays were selected as a next step due to their sensitivity and minimal sample concentration requirements. Intrinsic tryptophan fluorescence quenching reduces experimental costs because the amount of synthetic peptide required is greatly reduced relative to ITC.

### *Equilibrium Fluorescence Binding Assays*

The SPOP MATH domain contains two accessible tryptophan surface residues capable of monitoring SPOP-substrate binding. An initial control titration of PUC into SPOP MATH was conducted to verify intrinsic tryptophan fluorescence quenching as a viable technique for our system. Titrations of a PUC peptide into 0.5 and 1 μM SPOP MATH displayed reasonable tryptophan fluorescence quenching of about 10% (Figure 18). The fluorescence intensities were normalized and corrected for dilution effects. A binding isotherm for PUC titration into varying concentrations of SPOP MATH was generated (not shown), yielding a dissociation constant of $K_d$ = 6.0 $\pm$ 1 μM, which is in close agreement with established literature values.[9,46]

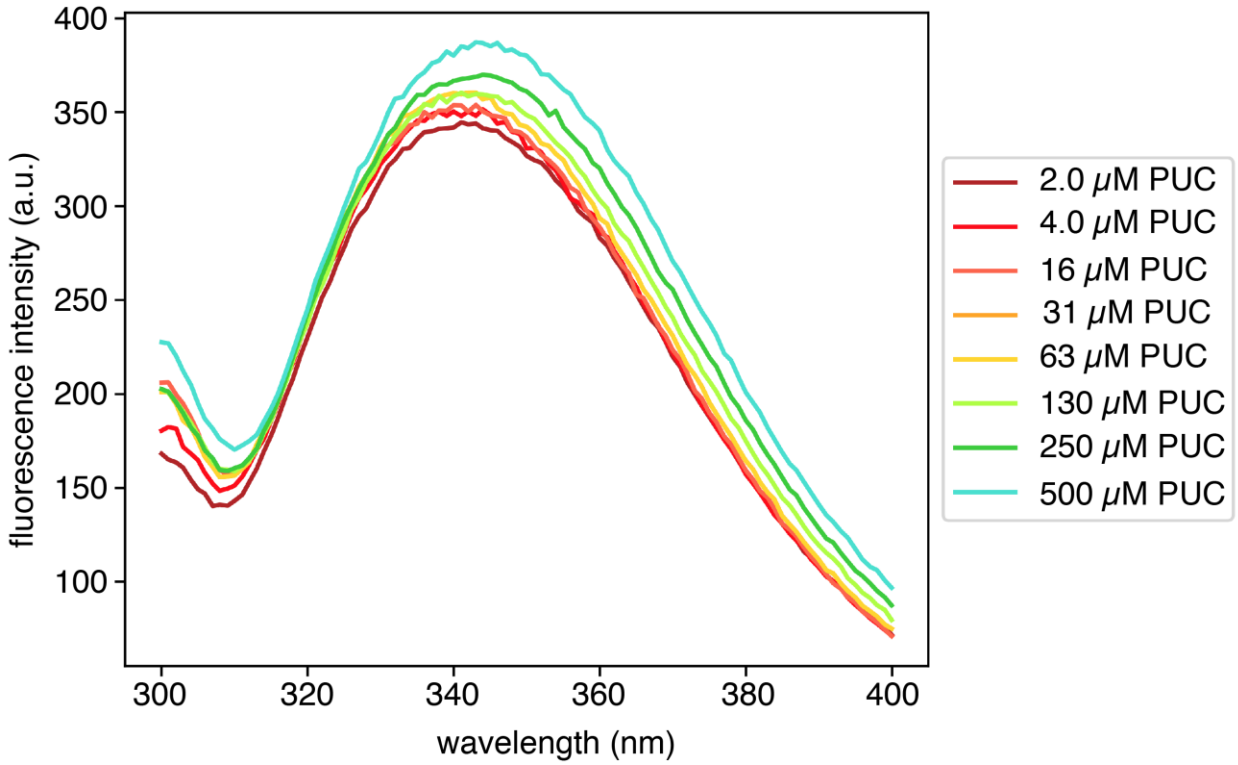



**Figure 19.** Representative overlay of emission spectra from titration of PUC peptide into 1 μM SPOP MATH. $K_d$ ~ 6.0 μM.

The reproducible results from the PUC titrations deemed fluorescence binding assays as a viable technique for characterization of Pdx1-SPOP interactions. Titration of Pdx1 S1 into SPOP displayed unexpected results that we initially attributed to unsuccessful titrations. The expected result for the titration of a peptide into a fluorescent macromolecule is that the relative fluorescence intensity should decrease with increasing peptide concentration, as tryptophan fluorescence should be quenched as a response to substrate binding. There are, however, systems that deviate from this expected behavior due to various factors including choice of solvent and the occurrence of multiple binding events. Repeated titrations of Pdx1 S1 into 1 μM SPOP MATH showed a reproducible blue shift upon increasing peptide concentration (Figure 19). Tryptophan emission is typically maximal at $\lambda_{emission}$ = 350 nm. The initial fluorescence intensity for this titration

was maximal at ~350 nm, however, upon an increase in peptide concentration, the $\lambda_{EM,max}$ blue shifted to a final value of $\lambda_{EM,max}$ = 309 nm. Additionally, the fluorescence intensity in the region of the emission spectra from 310-320 nm steadily increased upon each addition of ligand. Conversely, the fluorescence intensity in the region from 370-380 nm decreased upon each addition of ligand. These unusual results can be attributed to the occurrence of multiple binding events or the exposure of a previously buried tryptophan residue upon Pdx1 S1 binding.

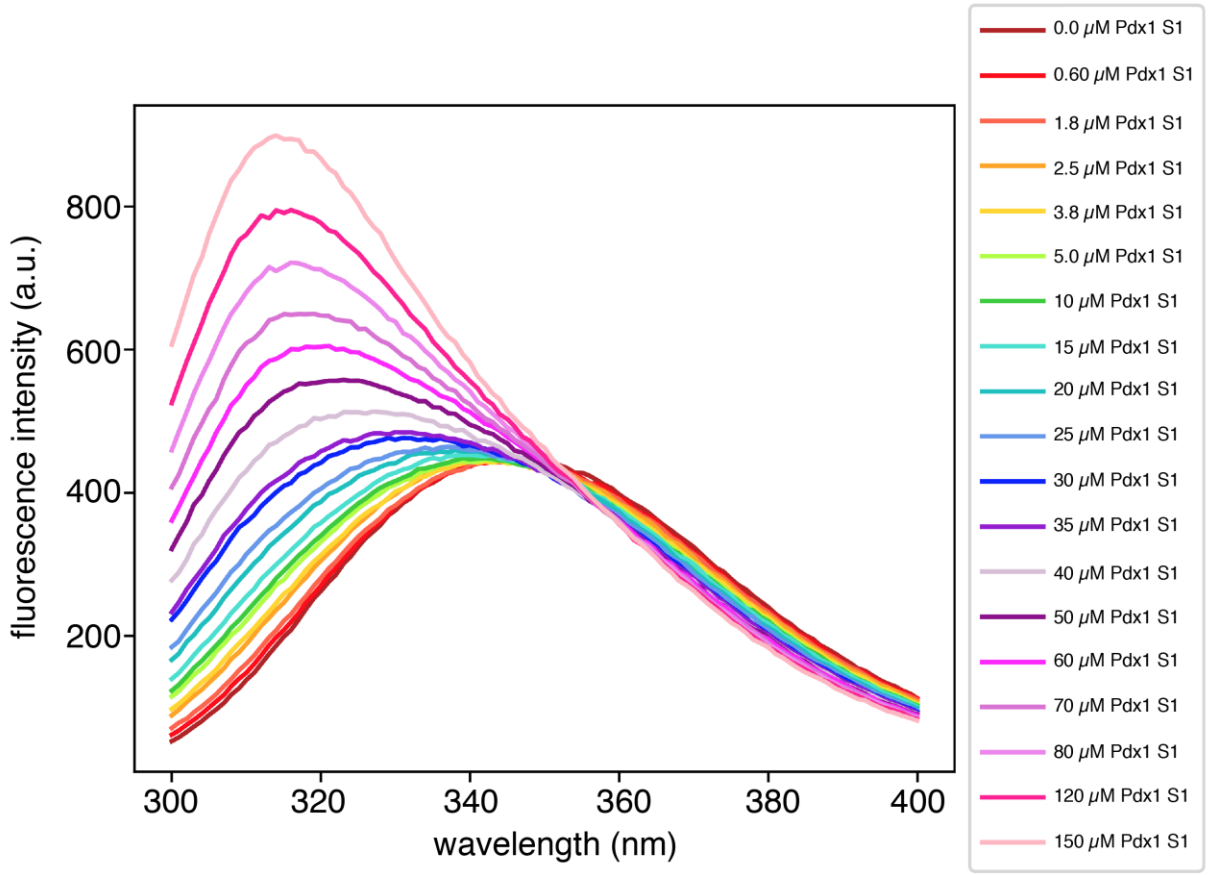



**Figure 20.** Representative spectral overlay of emission spectra from titration of 1 μM SPOP MATH with Pdx1 S1. The spectra show a blue shift and increase in fluorescence intensity with increasing peptide concentration, which suggests that two tryptophan residues on SPOP MATH are reporting on separate binding events.

SPOP MATH contains two accessible tryptophan residues on its surface that are capable of monitoring binding interactions (Figure 20). There is also a single tryptophan residue buried within the surface of SPOP that may be exposed upon conformational changes of SPOP MATH upon binding. Additional fluorescence experiments will need to be conducted to fully understand the source of the signal and its unusual response to an increase in peptide concentration.

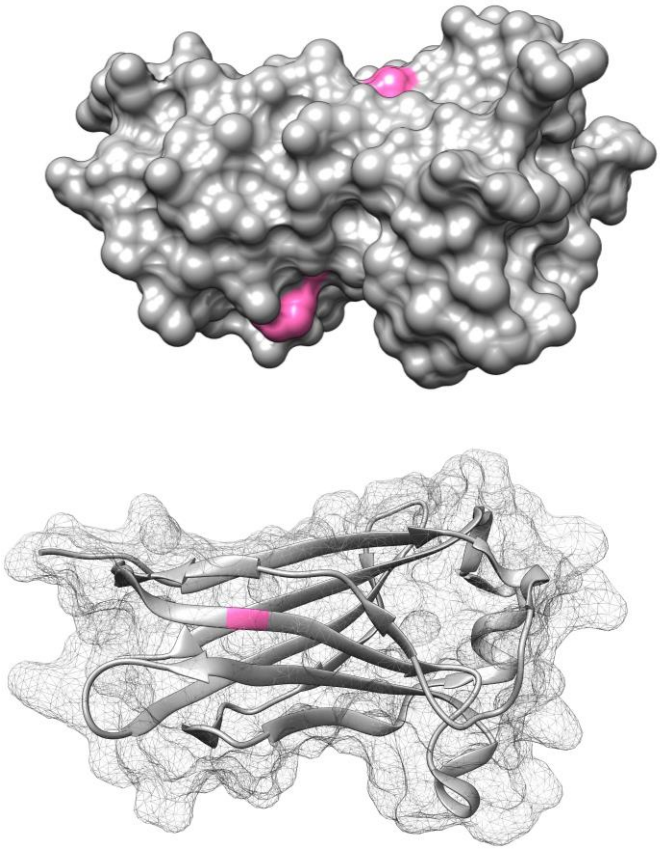


**Figure 21.** Locations of accessible surface tryptophan residues Trp 36 and Trp 131 (top solid surface) and the buried Trp 67 residue (bottom mesh surface) on the interior of SPOP MATH.

Operating under the premise that an additional binding event is causing the fluorescence increase and blue shift, the dissociation constants for each binding event were calculated. Upon inspection of the overlaid emission spectra, it appears that there

is an initial binding event that achieves saturation, as evinced by the overlapping spectra in the 370-380 nm region. The equilibrium binding dissociation constant describing this first binding event was calculated as $K_d$ = 30 $\pm$ 5 μM (Figure 23). This binding event has a lower SPOP affinity compared to Pdx1-C, which was expected based on my initial hypothesis that the presence of multiple linear motifs with individually weak affinities can enhance overall binding affinity through avidity effects.

The appearances of the spectra in the 310-320 nm region suggest the occurrence of a secondary binding event much weaker than the first. This secondary binding event does not appear to reach saturation, even after the addition of 200 μM Pdx1 S1. The $K_d$ for this second binding event was very low at 400 $\pm$ 72 μM (Figures 19 and 23). Due to extremely weak binding, the concentration of Pdx1 S1 needs to be increased to achieve full saturation in this second binding event, however, this may not be feasible as the reaction may be so weak that unobtainable amounts of peptide may be necessary for maximal saturation.

Titration of Pdx1 S2 into SPOP MATH displayed fluorescence quenching < 10%, which suggests weak binding at this site (Figure 21). A blue shift also occurs with this site, however, the range of $\lambda_{EM,max}$ is negligible relative to that seen in Pdx1 S1 titrations. The equilibrium binding dissociation constant was calculated to be $K_d$ = 50 $\pm$ 17 μM (Figure 23), which is substantially weaker than the dissociation constant of full-length Pdx1-C.

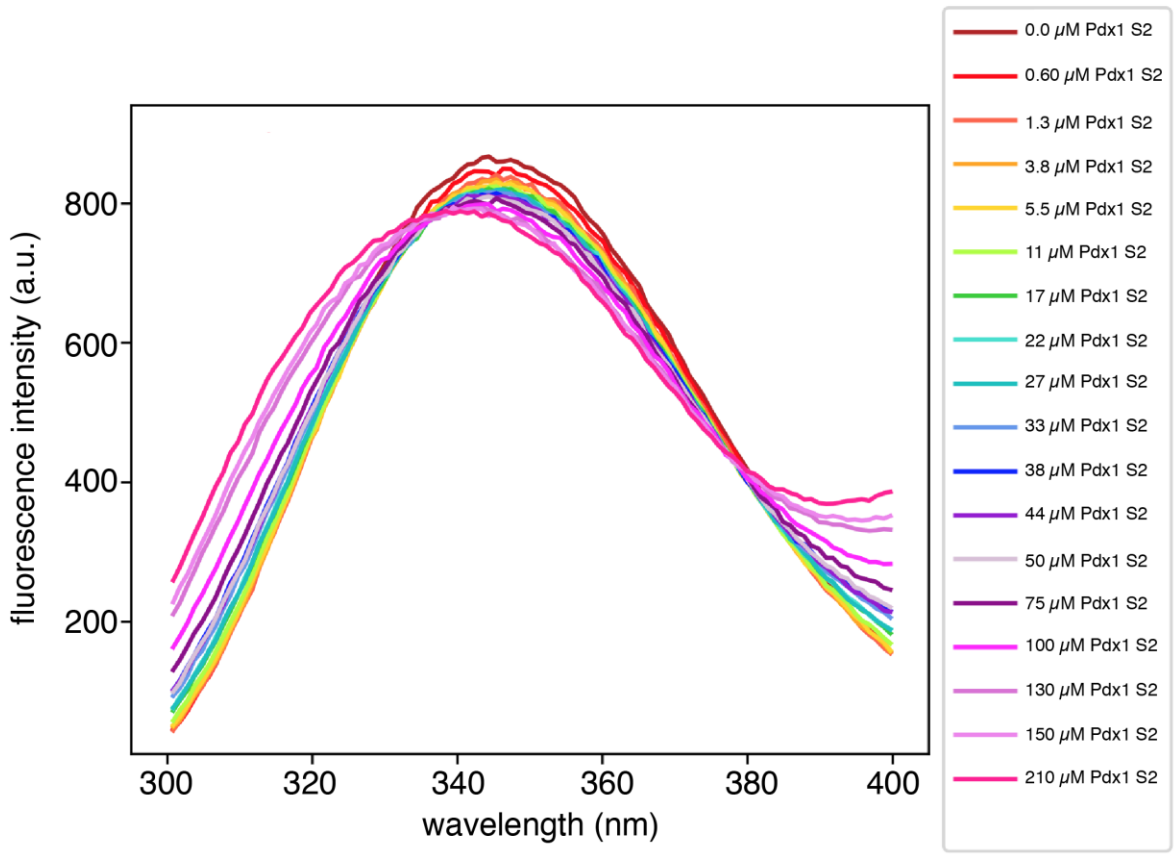



**Figure 22.** Representative spectral overlay for titration of Pdx1 S2 into 1 µM SPOP.

Figure 23 displays a summary of the calculated $K_d$ values from this data set. The preliminary binding isotherms in Figure 22 were generated by plotting ligand concentration against normalized fluorescence intensity.

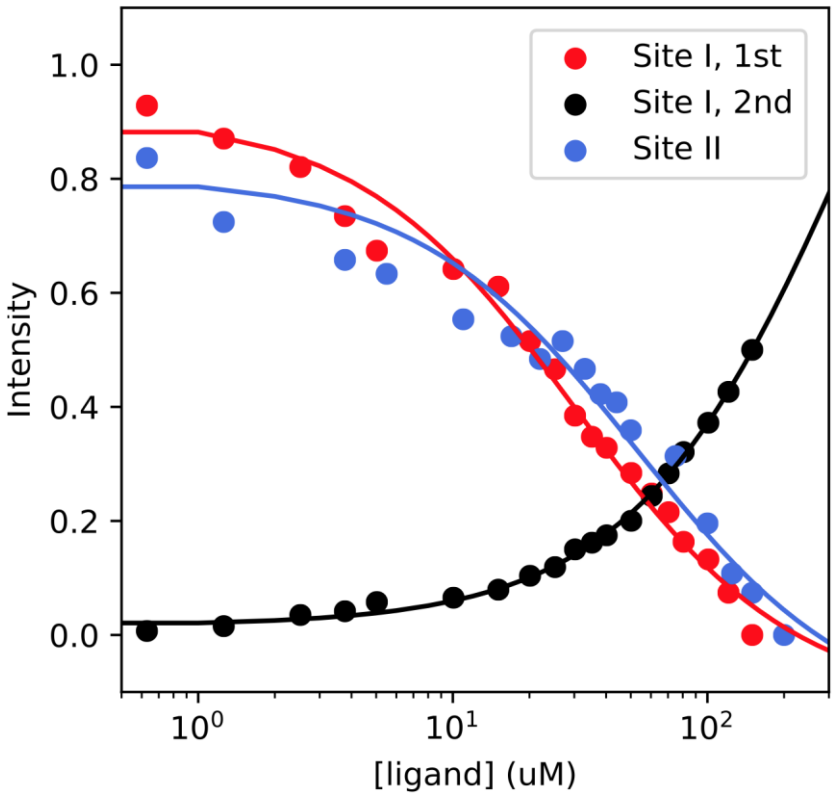
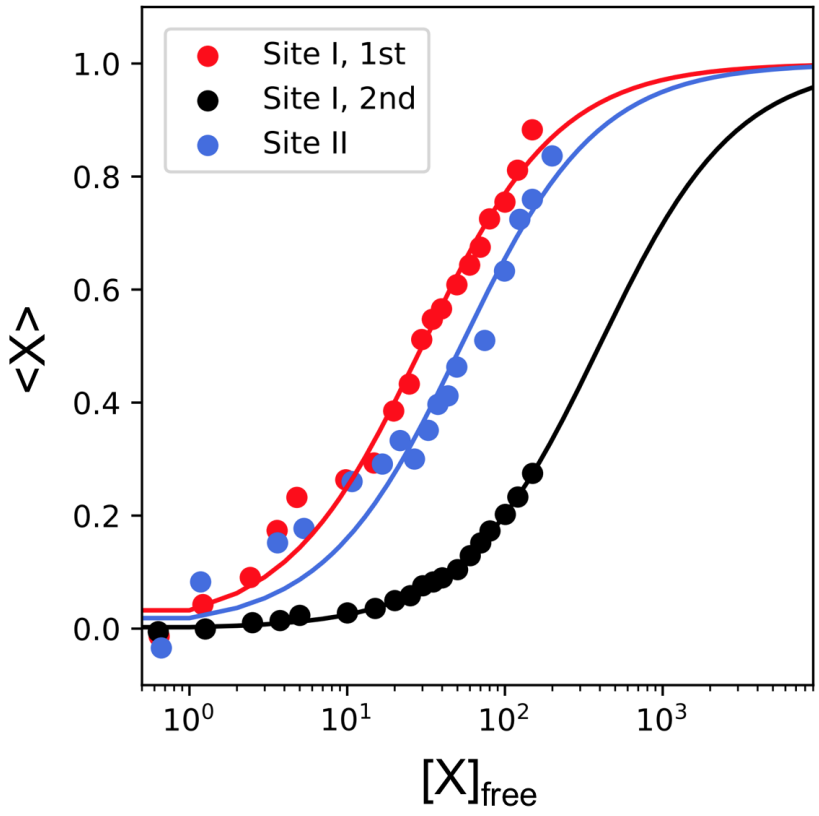



**Figure 23.** Binding curves generated from plotting normalized fluorescence intensity vs total ligand concentration and estimated $K_d$ values.




| Peptide | $K_d$ (μM) | N |
|---|---|---|
| Pdx1 S1, event 1 | 30 ± 5 | 5 |
| Pdx1 S1, event 2 | 400 ± 72 | 5 |
| Pdx1 S2 | 50 ± 17 | 5 |
| PUC | 6 ± 1 | 3 |

**Figure 24.** Langmuir isotherm generated from data describing the titration of Pdx1 peptides into 1.0 and 2.0 uM SPOP and the calculated $K_d$ values.

***Conclusions***

The binding interaction of Pdx1 with SPOP was studied to elucidate the role of the disordered C-terminus of Pdx1 in its degradation. Our lab has identified two binding sites on Pdx1-C that directly bind to the MATH domain of SPOP, which is consistent

with literature findings suggesting that SPOP substrates are multivalent. In order to confirm that this multivalency model applies to Pdx1, I generated recombinant Pdx1 truncation constructs corresponding to individual binding sites and compared their binding affinities to that of full-length Pdx1-C. Experimental results suggest that both Pdx1 sites bind in the canonical SPOP MATH binding groove with weaker affinities than full length Pdx1-C.

The disordered Pdx1-C binds to SPOP MATH in a canonical groove even though its binding motifs do not tightly adhere to the SPOP consensus binding sequence. The binding affinity of full-length Pdx1-C and peptides corresponding to individual SPOP binding sites was measured using ITC. This technique did not reliably probe the thermodynamics of our system due to weak binding. Equilibrium fluorescence titrations offered an alternative way to determine the equilibrium dissociation constant for this system. The intrinsic tryptophan fluorescence of SPOP was measured in the absence and the presence of increasing concentrations of peptides. The Pdx1 binding motif most divergent from the SPOP consensus binding sequence displayed two binding events in its emission spectra, with an initial binding event with $K_d$ of 30 $\pm$ 5 µM and a substantially weaker secondary binding event with a $K_d$ of 400 $\pm$ 72 µM. The identification of an additional binding event may provide reasoning for the extensive binding surface seen in NMR titrations of Pdx1-C into SPOP. In addition to the weak binding displayed by Pdx1 S1, the secondary binding event likely convoluted ITC data further, which made analysis difficult. The Pdx1 S2 binding site was determined to have a $K_d$ of 50 $\pm$ 17 µM. The affinities for these two individual sites were weaker than those of full-length Pdx1-C complex with dimeric SPOP. These results support my hypothesis

that the multivalency of Pdx1 for SPOP increases the overall affinity of their binding interaction through avidity effects. However, additional experiments are necessary to thoroughly test this hypothesis.

# Chapter 3: Additional Experiments and Future Directions

## 3.1 Introduction

This work provides a preliminary assessment of the affinity-enhancing results of multivalency in Pdx1-SPOP complex formation. Given that the two Pdx1 binding motifs appear to bind in similar locations on the canonical SPOP MATH binding groove, the possibility of a competitive binding event must be explored. Additional experiments are therefore necessary to fully understand the importance of multivalency in this interaction. Preliminary NMR and DSC data are shown in section 3.2.

The effect of Pdx1 phosphorylation on SPOP binding is an avenue yet to be explored in great detail. Upon the development of a reproducible phosphorylation protocol, the suite of experiments presented in Chapter 2 can be used to assess the effects of phosphorylation on Pdx1-SPOP binding.

Now that the characterization of the interaction of wild-type Pdx1-C with SPOP is nearly complete, we have a frame of reference for functional Pdx1-SPOP binding. The effects of clinical Pdx1 variants on SPOP binding can now be studied by comparison to experimental results from WT Pdx1-SPOP data. Experimental results will provide atomic-level detail into the Pdx1-SPOP interaction and can inform future SPOP-substrate and diabetes studies.

## 3.2 Additional Experiments and Preliminary Results for Future Directions

### *Nuclear Magnetic Resonance Assignment of SPOP MATH*

Two SPOP binding sites on Pdx1-C were previously identified by Monique Bastidas using the $^{15}N$, $^{13}C$-CON experiment. In order to assess whether Pdx1 binds in the canonical SPOP binding groove, previous Showalter lab graduate students collected HSQC spectra where Pdx1-C was titrated into $^{15}N$ SPOP MATH and compared to the HSQC of unbound SPOP MATH. I went on to collect the backbone triple resonance experiments used to assign the HSQC of SPOP MATH. This standard set of experiments includes the three-dimensional HNCO, HN(CA)CO, CBCA(CO)NH, and HNCACB. This set of experiments produces backbone carbon resonances for residue *i-1* on the amino acid chain and the resonances for the carbonyls on residues *i* and *i-1* corresponding to each N-H group. Through the generation of strip plots, one can “daisy chain” their way along the protein backbone and assign the resonances seen in the HSQC spectrum (Figure 24).

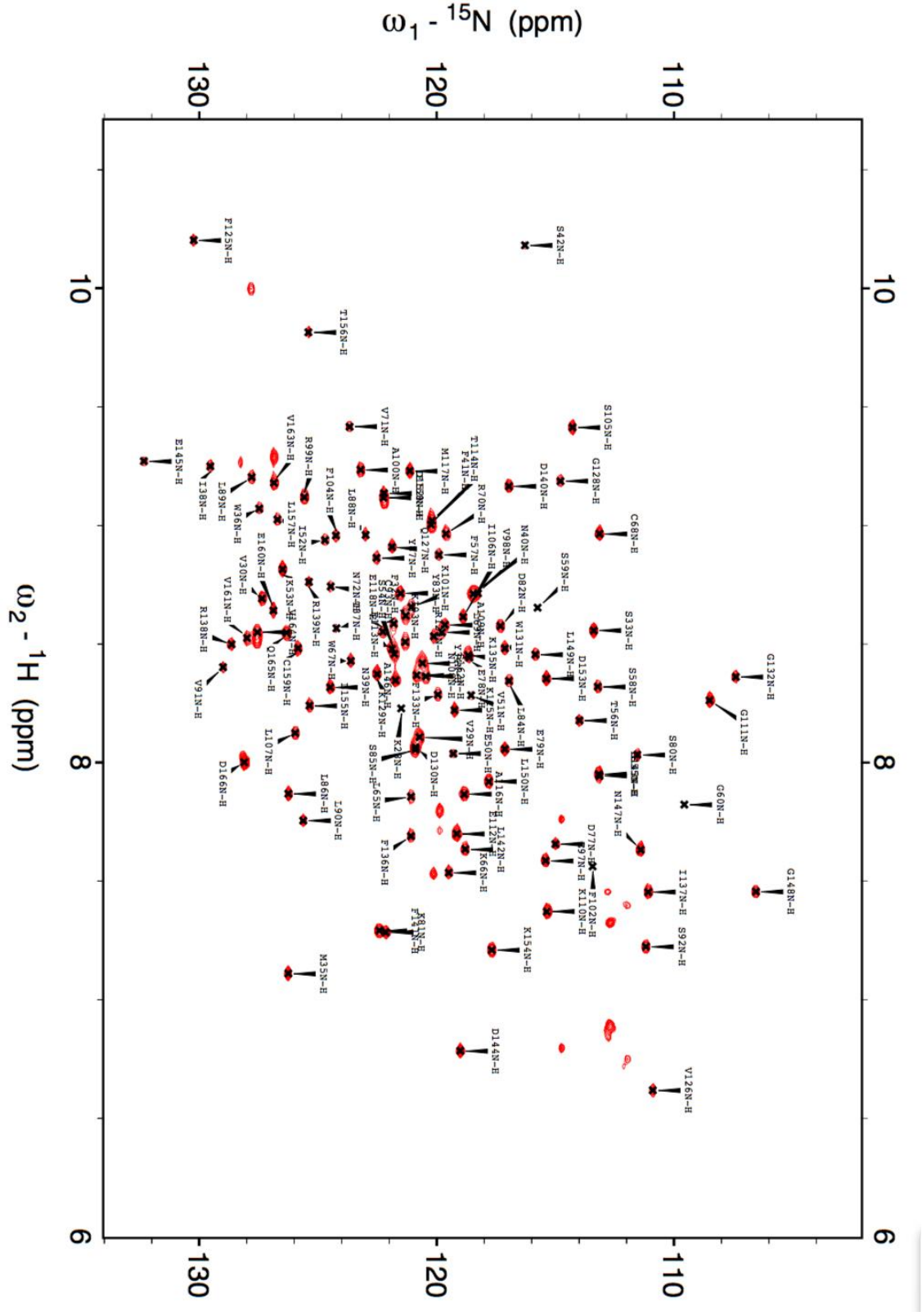



**Figure 25.** Assigned HSQC for SPOP MATH 28-166. Assigned by SAS.

***Differential Scanning Calorimetry***

The molecular recognition of Pdx1 by SPOP may involve transitional conformational changes that offer insight into the mechanisms that describe this binding interaction. Differential scanning calorimetry is a powerful technique complimentary to ITC that provides thermodynamic information describing conformational transitions associated with ligand binding. This technique is useful in the study of systems with very tight or very weak binding constants. These systems, similar to Pdx1-SPOP, are difficult to study by ITC without difficult modifications. In this technique, a biological macromolecule is denatured and the excess heat capacity associated with denaturation is measured across a temperature range of approximately 80 ˚C.[60] If a ligand tightly binds to a folded protein, complex formation should be favored, and the resultant macromolecule-ligand complex MX should have a higher thermal stability relative to each free species. Upon denaturation, the complex should have a significantly higher melting temperature than the macromolecule alone. If a suitable thermodynamic model of the binding process is known, DSC can be used to determine the binding energies of very tight or weak interactions.

Grace Usher and I performed DSC experiments on SPOP MATH alone and SPOP MATH in the presence of varying concentrations of Pdx1 S1 or Pdx1 S2 peptide in an effort to determine the equilibrium dissociation constant (Figure 26). Unbound SPOP undergoes a phase transition at $T_m$ = 65 ˚C and does not display a significant increase in $T_m$ with increasing Pdx1 S1 concentration. Unexpectedly, the molar heat capacity decreased with an increase in Pdx1 S1 concentration. Titration of Pdx1 S2 into SPOP MATH displayed a melting transition temperature that remained fairly constant upon each

addition of Pdx1 S2. The constant melting temperature of each titration suggests that Pdx1 binding does not stabilize nor destabilize native SPOP MATH. The interaction is likely not driven by the stabilizing effects of complex formation and may instead be entropically driven. The data from these DSC experiments were not suitable for the extraction of any quantitative thermodynamic information but do offer some qualitative insight into the driving forces of this interaction.

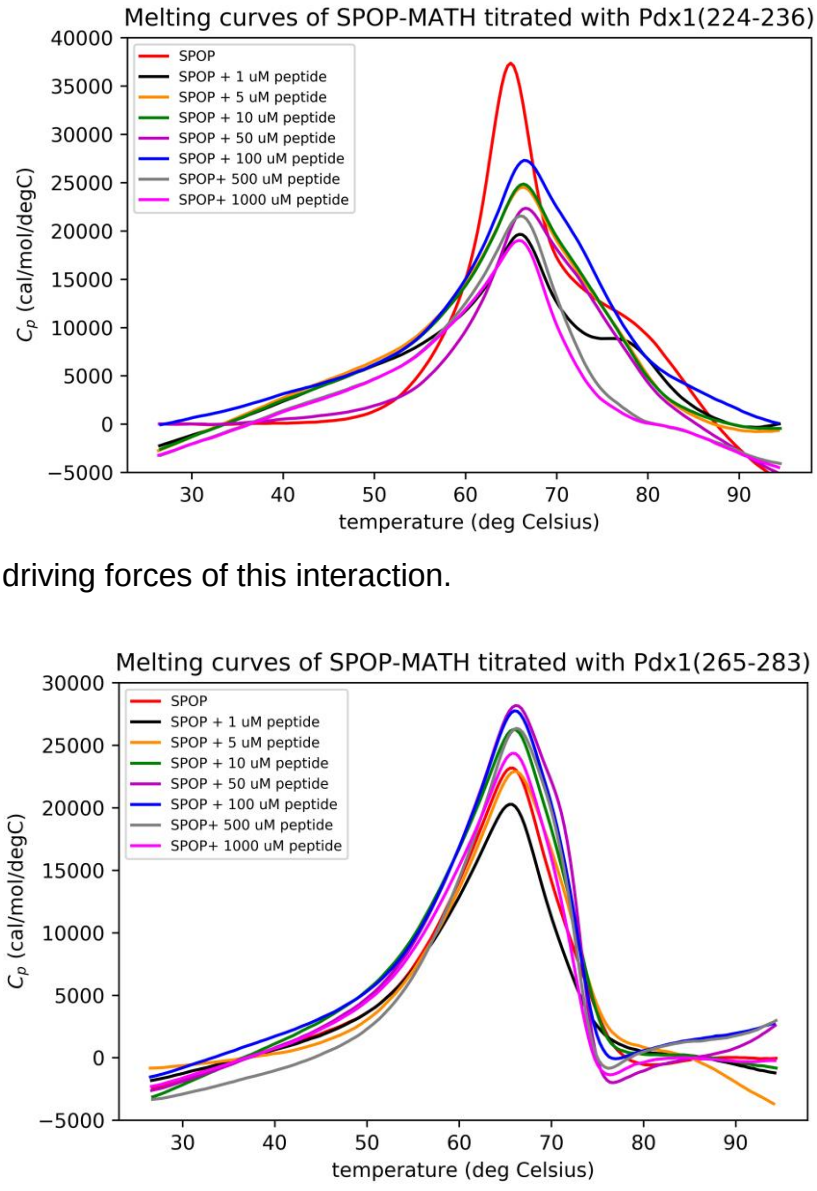



**Figure 26.** DSC data from titration of Pdx1 S1 (top panel) or Pdx1 S2 (bottom panel) into SPOP MATH.

### 3.3 Future Directions

### *Post-translational Modification of Pdx1-C*

Phosphorylation induces significant structural changes in IDRs that alter their functionality, thus modulating the strength of their protein-protein interactions.[61,62] Under low glucose conditions, Pdx1-C residues S268 and S272 are known to be phosphorylated, which is hypothesized to promote the proteasomal degradation of Pdx1.[63] Interestingly, we have identified a novel interaction site (Figure 5) for SPOP that includes these target residues, leading to the hypothesis that phosphorylation at S268 and S272 enhances SPOP binding to Pdx1-C.

This hypothesis can be tested through ITC and NMR experiments of phosphorylated wild-type Pdx1-C in complex with SPOP. Site-directed mutagenesis can be used to generate loss-of-function mutants that can be used to assess the role of S268 and S272 in Pdx1 regulation by SPOP. Dr. Erik Cook has designed Pdx1 S268A, S272A, and a double mutant S268A/S272A as a means to inhibit phosphorylation of Pdx1-C. Titrations of pPdx1-C and loss-of-function mutants into SPOP will test the hypothesis that phosphorylation of Pdx1-C enhances SPOP binding. Additionally, the phosphorylation-mimicking mutants S268E and S272E can be generated to simulate serine phosphorylation if phosphorylation of Pdx1-C proves difficult. Cellular viability assays must be conducted to ensure that the mutated constructs do not have any deleterious effects in cells.

Our lab has successfully completed chemical shift assignment of Pdx1-C using $^{15}$N,$^{13}$C-CON experiments (Figure 5) and determined that its conformation in solution is that of a random coil (not shown). A suite of three-dimensional NMR experiments can be used to assess the solution conformation of phosphorylated Pdx1-C (pPdx1-C). The

difference between the of the $C_\alpha$ and $C_\beta$ chemical shifts observed for pPdx1-C and those predicted for a random coil will offer insight into its local secondary structure and test the hypothesis that phosphorylation induces structural changes in Pdx1-C that facilitate SPOP binding.

***Thermodynamic Characterization of Clinically Documented Pdx1 Variants***

The prevalence of clinically documented diabetes-associated mutations mapping to the disordered tails of Pdx1 underscores the importance of studying these regions. Clinical mutants that map to Pdx1-C can be generated by site-directed mutagenesis and titrated into SPOP MATH in ITC and equilibrium fluorescence binding assays. The characterization of the interaction between wild-type Pdx1-C and SPOP MATH is nearly complete. This data set provides a frame of reference for functional wild-type Pdx1 activity and can be used to interpret results from assays involving Pdx1 diabetes-associated variants. The comparison of experimental results from assays of wild-type Pdx1-C and Pdx1 clinical variants will provide mechanistic insight into the Pdx1-C:SPOP interaction by identifying the key residues that participate in binding and will aid in the confirmation or alteration of the SPOP binding consensus motif.

The proposed experiments seek to generate insights that will improve our understanding of how SPOP modulates Pdx1 stability in a phosphorylation-mediated and glucose-dependent manner. The results from these experiments will improve our knowledge of SPOP-substrate interactions, Pdx1 regulation, and on a broader scope, insulin regulation.

# References


(1) Van Der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; et al. Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **2014**, *114* (13), 6589–6631.

(2) Wright, P. E.; Dyson, H. J. Intrinsically Unstructured Proteins: Re-Assessing the Protein Structure-Function Paradigm. *J. Mol. Biol.* **1999**, *293* (2), 321–331.

(3) Dyson, H. J. Making Sense of Intrinsically Disordered Proteins. *Biophys. J.* **2016**, *110* (5), 1013–1016.

(4) Wright, P. E.; Dyson, H. J. Intrinsically Disordered Proteins in Cellular Signalling and Regulation. *Nat. Rev. Mol. Cell Biol.* **2014**, *16* (1), 18–29.

(5) Jaeken, L. The Neglected Functions of Intrinsically Disordered Proteins and the Origin of Life. *Prog. Biophys. Mol. Biol.* **2016**, *126*, 31–46.

(6) Gibbs, E. B.; Showalter, S. A. Quantitative Biophysical Characterization of Intrinsically Disordered Proteins. *Biochemistry* **2015**, *54* (6), 1314–1326.

(7) Babu, M. M.; van der Lee, R.; de Groot, N. S.; Gsponer, J. Intrinsically Disordered Proteins: Regulation and Disease. *Curr. Opin. Struct. Biol.* **2011**, *21* (3), 432–440.

(8) Hayama, R.; Sparks, S.; Hecht, L. M.; Dutta, K.; Karp, J. M.; Cabana, C. M.; Rout, M. P.; Cowburn, D.; Allewell, N. M. Thermodynamic Characterization of the Multivalent Interactions Underlying Rapid and Selective Translocation through the Nuclear Pore Complex. *J. Biol. Chem.* **2018**, *293* (12), 4555–4563.

(9) Pierce, W. K.; Grace, C. R.; Lee, J.; Nourse, A.; Marzahn, M. R.; Watson, E. R.; High, A. A.; Peng, J.; Schulman, B. A.; Mittag, T. Multiple Weak Linear Motifs Enhance Recruitment and Processivity in SPOP-Mediated Substrate Ubiquitination. *J. Mol. Biol.* **2016**, *428* (6), 1256–1271.

(10) Liu, J.; Perumal, N. B.; Oldfield, C. J.; Su, E. W.; Uversky, V. N.; Dunker, A. K. Intrinsic Disorder in Transcription Factors. *Biochemistry* **2006**, *45* (5), 6873–6888.

(11) Aguzzi, A.; Altmeyer, M. Phase Separation: Linking Cellular Compartmentalization to Disease. *Trends Cell Biol.* **2016**, *26* (7), 547–558.

(12) Handorf, A. M.; Sollinger, H. W.; Alam, T. Genetic Engineering of Surrogate β Cells for Treatment of Type 1 Diabetes Mellitus. *J. Diabetes Mellit.* **2015**, *5* (November), 295–312.

(13) Macfarlane, W. M.; Campbell, S. C.; Elrick, L. J.; Oates, V.; Bermano, G.; Lindley, K. J.; Aynsley-Green, A.; Dunne, M. J.; James, R. F. L.; Docherty, K. Glucose Regulates Islet Amyloid Polypeptide Gene Transcription in a PDX1- and Calcium-Dependent Manner. *J. Biol. Chem.* **2000**, *275* (20), 15330–15335.

(14) Wu, H.; MacFarlane, W. M.; Tadayyon, M.; Arch, J. R.; James, R. F.; Docherty, K. Insulin Stimulates Pancreatic-Duodenal Homoeobox Factor-1 (PDX1) DNA-Binding Activity and Insulin Promoter Activity in Pancreatic Beta Cells. *Biochem. J.* **1999**, *344 Pt 3*, 813–818.

(15) Claiborn, K.; Sachdeva, M. Pcif1 Modulates Pdx1 Protein Stability and Pancreatic β Cell Function and Survival in Mice. *J. Clin.* **2010**, *120* (10), 3713–3721.

(16) Sachdeva, M. M.; Claiborn, K. C.; Khoo, C.; Yang, J.; Groff, D. N.; Mirmira, R. G.; Stoffers, D. A. Pdx1 (MODY4) Regulates Pancreatic Beta Cell Susceptibility to ER

Stress. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (45), 19090–19095.
(17) Gao, T.; McKenna, B.; Li, C.; Reichert, M.; Nguyen, J.; Singh, T.; Yang, C.; Pannikar, A.; Doliba, N.; Zhang, T.; et al. Pdx1 Maintains β Cell Identity and Function by Repressing an α Cell Program. *Cell Metab.* **2014**, *19* (2), 259–271.
(18) Roy, N.; Takeuchi, K. K.; Ruggeri, J. M.; Bailey, P.; Chang, D.; Li, J.; Leonhardt, L.; Puri, S.; Hoffman, M. T.; Gao, S.; et al. PDX1 Dynamically Regulates Pancreatic Ductal Adenocarcinoma Initiation and Maintenance. *Genes Dev.* **2016**, *30* (24), 2669–2683.
(19) Johnson, J. D.; Ahmed, N. T.; Luciani, D. S.; Han, Z.; Tran, H.; Fujita, J.; Misler, S.; Edlund, H.; Polonsky, K. S. Increased Islet Apoptosis in Pdx1+/- Mice. *J. Clin. Invest.* **2003**, *111* (8), 1147–1160.
(20) Mosley, A. L.; Corbett, J. A.; Özcan, S. Glucose Regulation of Insulin Gene Expression Requires the Recruitment of P300 by the β-Cell-Specific Transcription Factor Pdx-1. *Mol. Endocrinol.* **2004**, *18* (9), 2279–2290.
(21) Handorf, A. M.; Sollinger, H. W.; Alam, T. Genetic Engineering of Surrogate β Cells for Treatment of Type 1 Diabetes Mellitus. *J. Diabetes Mellit.* **2015**, *5* (November), 295–312.
(22) Bastidas, M.; Showalter, S. A. Thermodynamic and Structural Determinants of Differential Pdx1 Binding to Elements from the Insulin and IAPP Promoters. *J. Mol. Biol.* **2013**, *425* (18), 3360–3377.
(23) Andrali, S. S.; Sampley, M. L.; Vanderford, N. L.; Özcan, S. Glucose Regulation of Insulin Gene Expression in Pancreatic β-Cells. *Biochem. J.* **2008**, *415* (1), 1–10.
(24) Tuano, N. K.; Okabe, J.; Ziemann, M.; Cooper, M. E.; El-Osta, A. Set7 Mediated Interactions Regulate Transcriptional Networks in Embryonic Stem Cells. *Nucleic Acids Res.* **2016**, *44* (19), 9206–9217.
(25) Caetano, L. A.; Santana, L. S.; Costa-Riquetto, A. D.; Lerario, A. M.; Nery, M.; Nogueira, G. F.; Ortega, C. D.; Rocha, M. S.; Jorge, A. A. L.; Teles, M. G. PDX1 - MODY and Dorsal Pancreatic Agenesis: New Phenotype of a Rare Disease. *Clin. Genet.* **2017**, No. April, 1–5.
(26) Reis, A. F.; Ye, W. Z.; Dubois-Laforgue, D.; Bellanne-Chantelot, C.; Timsit, J.; Velho, G. Mutations in the Insulin Promoter Factor-1 Gene in Late-Onset Type 2 Diabetes Mellitus. *Eur. J. Endocrinol.* **2000**, *143* (4), 511–513.
(27) Longo, A.; Guanga, G. P.; Rose, R. B. Structural Basis for Induced Fit Mechanisms in DNA Recognition by the Pdx1 Homeodomain. *Biochemistry* **2007**, *46* (11), 2948–2957.
(28) Liu, A.; Desai, B. M.; Stoffers, D. A. Identification of PCIF1, a POZ Domain Protein That Inhibits PDX-1 (MODY4) Transcriptional Activity. *Mol. Cell. Biol.* **2004**, *24* (10), 4372–4383.
(29) Bailey, P.; Chang, D. K.; Nones, K.; Johns, A. L.; Patch, A. M.; Gingras, M. C.; Miller, D. K.; Christ, A. N.; Bruxner, T. J. C.; Quinn, M. C.; et al. Genomic Analyses Identify Molecular Subtypes of Pancreatic Cancer. *Nature* **2016**, *531* (7592), 47–52.
(30) Weir, G. C.; Sharma, A.; Zangen, D. H.; Bonner-Weir, S. Transcription Factor Abnormalities as a Cause of Beta Cell Dysfunction in Diabetes: A Hypothesis. *Acta Diabetol.* **1997**, *34* (3), 177–184.
(31) Fujimoto, K.; Polonsky, K. S. Pdx1 and Other Factors That Regulate Pancreatic β-

Cell Survival. *Diabetes, Obes. Metab.* **2009**, *11*, 30–37.
(32) Fujimoto, K.; Hanson, P. T.; Tran, H.; Ford, E. L.; Han, Z.; Johnson, J. D.; Schmidt, R. E.; Green, K. G.; Wice, B. M.; Polonsky, K. S. Autophagy Regulates Pancreatic Beta Cell Death in Response to Pdx1 Deficiency and Nutrient Deprivation * □. **2009**, *284* (40), 27664–27673.
(33) Yuan, Y.; Hartland, K.; Boskovic, Z.; Wang, Y.; Walpita, D.; Lysy, P. A.; Zhong, C.; Young, D. W.; Kim, Y. K.; Tolliday, N. J.; et al. A Small-Molecule Inducer of Pdx1 Expression Identified by High-Throughput Screening. *Chem. Biol.* **2013**, *20* (12), 1513–1522.
(34) Naqvi, A. A. T.; Hasan, G. M.; Hassan, M. I. Investigating the Role of Transcription Factors of Pancreas Development in Pancreatic Cancer. *Pancreatology* **2018**, *18* (2), 184–190.
(35) Liu, S.; Ballian, N.; Belaguli, N.; Patel, S.; Li, M.; Templeton, N.; Gingras, M.-C.; Gibbs, R.; Fisher, W.; Brunicardi, F. PDX-1 Acts as a Potential Molecular Target for Treatment of Human Pancreatic Cancer. *Pancreas* **2008**, *37* (2), 210–220.
(36) Heller, R. S.; Stoffers, D. A.; Bock, T.; Svenstrup, K.; Jensen, J.; Horn, T.; Miller, C. P.; Habener, J. F.; Madsen, O. D.; Serup, P. Improved Glucose Tolerance and Acinar Dysmorphogenesis by Targeted Expression of Transcription Factor PDX-1 to the Exocrine Pancreas. *Diabetes* **2001**, *50* (18), 1553–1561.
(37) Vinogradova, T. V.; Sverdlov, E. D. PDX1: A Unique Pancreatic Master Regulator Constantly Changes Its Functions during Embryonic Development and Progression of Pancreatic Cancer. *Biochem.* **2017**, *82* (8), 887–893.
(38) Meng, R.; Al-Quobaili, F.; Müller, I.; Götz, C.; Thiel, G.; Montenarh, M. CK2 Phosphorylation of Pdx-1 Regulates Its Transcription Factor Activity. *Cell. Mol. Life Sci.* **2010**, *67* (14), 2481–2489.
(39) Lu, M.; Miller, C. P.; Habenerg, F. Functional Regions of the Homeodomain Protein IDX-1 Required for Transactivation of the Rat Somatostatin Gene. *Endocrinology* **1996**, *137* (7), 2959–2967.
(40) Liu, A.; Oliver-Krasinski, J.; Stoffers, D. A. Two Conserved Domains in PCIF1 Mediate Interaction with Pancreatic Transcription Factor PDX-1. *FEBS Lett.* **2006**, *580* (28–29), 6701–6706.
(41) Wang, W.; Shi, Q.; Guo, T.; Yang, Z.; Jia, Z.; Chen, P.; Zhou, C. PDX1 and ISL1 Differentially Coordinate with Epigenetic Modifications to Regulate Insulin Gene Expression in Varied Glucose Concentrations. *Mol. Cell. Endocrinol.* **2016**, *428*, 38–48.
(42) Babu, D. A.; Deering, T. G.; Mirmira, R. G. A Feat of Metabolic Proportions: Pdx1 Orchestrates Islet Development and Function in the Maintenance of Glucose Homeostasis. *Mol. Genet. Metab.* **2007**, *92* (1–2), 43–55.
(43) Cissell, M. A.; Zhao, L.; Sussel, L.; Henderson, E.; Stein, R. Transcription Factor Occupancy of the Insulin Gene in Vivo. *J. Biol. Chem.* **2002**, *278* (2), 751–756.
(44) Claiborn, K. C.; Sachdeva, M. M.; Cannon, C. E.; Groff, D. N.; Singer, J. D.; Stoffers, D. A. Pcif1 Modulates Pdx1 Protein Stability and Pancreatic β Cell Function and Survival in Mice. *J. Clin. Invest.* **2010**, *120* (10), 3713–3721.
(45) Mani, R. S. The Emerging Role of Speckle-Type POZ Protein (SPOP) in Cancer Development. *Drug Discov. Today* **2014**, *19* (9), 1498–1502.
(46) Zhuang, M.; Calabrese, M. F.; Liu, J.; Waddell, M. B.; Nourse, A.; Hammel, M.;

Miller, D. J.; Walden, H.; Duda, D. M.; Seyedin, S. N.; et al. Structures of SPOP-Substrate Complexes: Insights into Molecular Architectures of BTB-Cul3 Ubiquitin Ligases. *Mol. Cell* **2009**, *36* (1), 39–50.
(47) Marzahn, M. R.; Marada, S.; Lee, J.; Nourse, A.; Kenrick, S.; Zhao, H.; Ben-Nissan, G.; Kolaitis, R.; Peters, J. L.; Pounds, S.; et al. Higher-order Oligomerization Promotes Localization of SPOP to Liquid Nuclear Speckles. *EMBO J.* **2016**, *35* (12), 1254–1275.
(48) Nagai, Y.; Kojima, T.; Muro, Y.; Hachiya, T.; Nishizawa, Y.; Wakabayashi, T.; Hagiwara, M. Identification of a Novel Nuclear Speckle-Type Protein, SPOP. *FEBS Lett.* **1997**, *418* (1–2), 23–26.
(49) Blattner, M.; Liu, D.; Robinson, B. D.; Huang, D.; Poliakov, A.; Gao, D.; Nataraj, S.; Deonarine, L. D.; Augello, M. A.; Sailer, V.; et al. SPOP Mutation Drives Prostate Tumorigenesis In Vivo through Coordinate Regulation of PI3K/MTOR and AR Signaling. *Cancer Cell* **2017**, *31* (3), 436–451.
(50) Guo, Z. Q.; Zheng, T.; Chen, B.; Luo, C.; Ouyang, S.; Gong, S.; Li, J.; Mao, L. L.; Lian, F.; Yang, Y.; et al. Small-Molecule Targeting of E3 Ligase Adaptor SPOP in Kidney Cancer. *Cancer Cell* **2016**, *30* (3), 474–484.
(51) Gan, W.; Dai, X.; Lunardi, A.; Li, Z.; Inuzuka, H.; Liu, P.; Varmeh, S.; Zhang, J.; Cheng, L.; Sun, Y.; et al. SPOP Promotes Ubiquitination and Degradation of the ERG Oncoprotein to Suppress Prostate Cancer Progression. *Mol. Cell* **2015**, *59* (6), 917–930.
(52) Bastidas, M. Structural and Functional Characterization of the Transcription Factor Pdx1 and Its Interactions with DNA and Proteins, The Pennslyvania State University, 2015.
(53) Guerra-Castellano, A.; Díaz-Moreno, I.; Velázquez-Campoy, A.; De La Rosa, M. A.; Díaz-Quintana, A. Structural and Functional Characterization of Phosphomimetic Mutants of Cytochrome c at Threonine 28 and Serine 47. *Biochim. Biophys. Acta - Bioenerg.* **2016**, *1857* (4), 387–395.
(54) Kozlov, A.; Galletto, R.; Lohman, T. SSB–DNA Binding Monitored by Fluorescence Intensity and Anisotropy. In *Methods in Molecular Biology*; 2012; Vol. 922, pp 55–83.
(55) Ghisaidoobe, A.; Chung, S. Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *Int. J. Mol. Sci.* **2014**, *15* (12), 22518–22538.
(56) Moon, C. P.; Fleming, K. G. *Using Tryptophan Fluorescence to Measure the Stability of Membrane Proteins Folded in Liposomes*, 1st ed.; Elsevier Inc., 2011; Vol. 492.
(57) Klostermeier, D.; Rudolph, M. G. *Biophysical Chemistry*; CRC Press, Taylor & Francis Group, 2017.
(58) Lohman, T. M.; Bujalowski, W. Thermodynamic Methods for Model-Independent Determination of Equilibrium Binding Isotherms for Protein-DNA Interactions: Spectroscopic Approaches to Monitor Binding. *Methods Enzymol.* **1991**, *208*, 258–290.
(59) Lin, J. B.; Lucius, A. L. *Analysis of Linked Equilibria*, 1st ed.; Elsevier Inc., 2015; Vol. 562.
(60) Barone, G.; Catanzano, P.; Vecchio, P. Del; Giancola, C.; Graziano, G.

Differential Scanning Calorimetry as a Tool to Study Protein-Ligand Interactions. *Pure Appl. Chem.* **1995**, *67* (11), 1867–1872.
(61) Nishi, H.; Hashimoto, K.; Panchenko, A. R. Phosphorylation in Protein-Protein Binding: Effect on Stability and Function. *Structure* **2011**, *19* (12), 1807–1815.
(62) Bah, A.; Forman-Kay, J. D. Modulation of Intrinsically Disordered Protein Function by Post-Translational Modifications. *J. Biol. Chem.* **2016**, *291* (13), 6696–6705.
(63) Humphrey, R. K.; Yu, S. M.; Flores, L. E.; Jhala, U. S. Glucose Regulates Steady-State Levels of PDX1 via the Reciprocal Actions of GSK3 and AKT Kinases. *J. Biol. Chem.* **2010**, *285* (5), 3406–3416.